

RACIST ALGORITHMS OR SYSTEMIC PROBLEMS?

Risk Assessments and Racial Disparities

GINA M. VINCENT 

University of Massachusetts Medical School

JODI L. VILJOEN 

Simon Fraser University

As recent and historical events attest, racial and ethnic disparities are widely engrained into the justice system. Recently, scholars and policymakers have raised concerns that risk assessment instruments may exacerbate these disparities. While it is critical that risk instruments be scrutinized for racial bias, some concerns, though well-meaning, have gone beyond the evidence. This article explains what it means for an instrument to be “biased” and why instruments should not all be painted with the same brush (some will be more susceptible to bias than others). If some groups get apprehended more, those groups will score higher on non-biased, well-validated instruments derived to maximize prediction of recidivism because of mathematics. Thus, risk instruments shine a light on long-standing systemic problems of racial disparities. This article concludes with suggestions for research and for minimizing disparities by ensuring that systems use risk assessments to avoid unnecessary incarceration while allowing for structured discretion.

Keywords: criminal justice; risk assessment; race; ethnicity; decision-making

Risk assessment instruments are often used by personnel in justice settings to inform decisions at various points in the legal process in which risk to public safety is relevant (e.g., Starr, 2015; Wachter, 2015). National data indicate most U.S. justice systems use a pretrial risk assessment instrument to make determinations about risk to community safety (<https://pretrialrisk.com/>). The value of risk instruments generally is that they are more accurate and reliable in their estimates of one’s likelihood of reoffending than unstructured hunches (e.g., Ægisdóttir et al., 2006; Grove et al., 2000). Moreover, there is meta-analytic evidence that when properly implemented and paired with evidence-based practice, risk instruments can reduce incarceration without an increased threat to public safety (Viljoen et al., 2019). However, risk instruments recently have been subject to substantial concerns (e.g., Robinson & Koepke, 2019).

AUTHORS’ NOTE: *The authors wish to thank Lauren McDowell, M.A., for her assistance with this article. Correspondence concerning this article should be addressed to Gina M. Vincent, Department of Psychiatry, Law & Psychiatry Program and Implementation Science & Practices Advances Research Center, University of Massachusetts Medical School, 222 Maple Ave, Shrewsbury, MA 01545; e-mail: Gina.Vincent@umassmed.edu.*

CRIMINAL JUSTICE AND BEHAVIOR, 202X, Vol. XX, No. X, Month 2020, 1–9.

DOI: 10.1177/0093854820954501

Article reuse guidelines: sagepub.com/journals-permissions

© 2020 International Association for Correctional and Forensic Psychology

Critics are concerned that risk instruments may exacerbate racial disparities. This issue gained increased attention when Former Attorney General Holder (2014/2015) discussed the propensity for use of risk assessment tools in sentencing decisions to “. . . exacerbate unwarranted and unjust disparities that are already far too common in our criminal justice system and in our society.,” which was reaffirmed when Senators Booker & Schatz (2018) stated that there are valid concerns that some algorithms may exacerbate racial disparities. This statement is partially in reaction to a recent study of the COMPAS (Angwin et al., 2016) and writings suggesting risk instruments sentence people based on poverty and send a “noxious expressive message” (Starr, 2015, p. 206). Today there is an entire website displaying pretrial risk assessment tools, referred to as “RATS,” as a method to inflate racial biases while masquerading as “science” (<https://pretrialrisk.com/>), and the Pretrial Justice Institute (2020) recently stated that pretrial risk instruments can no longer be “part of our solution for building equitable pretrial justice systems” (p. 1).

The notion that racial and ethnic disparities exist in the justice system is indisputable (e.g., Hockenberry & Puzanchera, 2020). It is unlikely that risk instruments, or any other single approach, will be powerful enough to completely eradicate these disparities, but in accordance with our ethical standards (American Educational Research Association [AERA], American Psychological Association [APA], National Council on Measurement in Education [NCME], & Joint Committee on Standards for Educational and Psychological Testing (U.S.), 2014) instruments certainly should not exacerbate them. In fact, a primary intent of many risk instruments was to contribute to efforts to decrease disparities. It is critical for researchers and instrument developers to make a concerted effort to carefully attend to disparities and adequately validate and avoid biases in these instruments. It is equally critical that legal scholars and advocates ask themselves “what is the alternative?” and be careful to ensure that their calls to stop using instruments do not worsen problems. This article reviews the evidence about racial bias in risk instruments, and provides recommendations for research and for minimizing bias without eliminating risk instruments.

RISK INSTRUMENTS ARE NOT ALL CREATED THE SAME

The majority of criticisms raised over the last few years referring to “racist algorithms” (Schwartzapfel, 2019) germinated primarily from concerns about pretrial risk instruments. It is quite possible some of these instruments are, in fact, biased but we must be careful not to generalize these criticisms to all risk instruments. Although critics often lump all tools into the same category, they are not all the same.

INSTRUMENTS DIFFER IN THEIR PURPOSE

Instruments are different depending on the purpose for which they were designed (e.g., pretrial, dispositional planning; Grisso, 2005). Pretrial instruments generally need to be brief and efficient because of the large volume of defendants and the need for quick decisions. In the United States there are limits to what personnel are permitted to ask defendants due to concerns about self-incrimination. As such, most pretrial instruments tend to rely almost solely on one’s criminal history (e.g., number of prior felony convictions) as a consequence of not having access to interview defendants. Scholars have raised concerns that these criminal history items are a proxy for race (Harcourt, 2015).

INSTRUMENTS DIFFER IN THE WAY RISK LEVEL IS DETERMINED

Not all risk instruments use an algorithm or score to determine one's level of risk. There are two primary frameworks used to determine risk level that have equivalent predictive accuracy (Yang et al., 2010); actuarial and structured professional judgment (SPJ). Actuarial decision-making means the determination of one's risk level is formulated based on a non-discretionary algorithm generated by ratings on a fixed set of risk factors (Monahan, 2008). Risk level classifications are generated by identifying the optimal cut scores to differentiate who is more or less likely to recidivate. SPJ instruments do not use algorithms but instead permit some discretion by combining structure with professional judgment. Trained professionals (e.g., probation officers) weigh the relevance of ratings on validated risk factors to the outcome in question (e.g., re-arrest, violence) for a particular individual and make an individualized determination of whether the individual is low, moderate, or high risk.

INSTRUMENTS DIFFER IN THE TYPES OF ITEMS INCLUDED AND METHODS OF CONSTRUCTION

Short actuarial risk instruments, like pretrial instruments, focus primarily on static items like criminal history. These are in sharp contrast to the plethora of instruments that include both static and dynamic risk factors (i.e., antisocial associates, aggressive tendencies). These instruments, which includes all SPJ and static/dynamic actuarial instruments, require practitioners (e.g., probation officers) to interview individuals and gather collateral information. The focus is on risk management or "prevention" rather than on maximizing "prediction." In justice settings, instruments that contain both static and dynamic items have become known as risk and needs assessment instruments. In forensic settings, we simply refer to them as risk assessment instruments because there is widespread recognition that dynamic risk factors are just as, if not more, important as historical factors in predicting reoffending (Esienberg et al., 2019; Vincent et al., 2011). The more information we have about an individual, the more accurate we are at determining their risk (M. A. Campbell et al., 2009).

Some instruments are created actuarially, meaning that researchers examine a particular sample of people and conduct analyses or use machine-learning (systems or algorithms that detect patterns in large datasets to develop prediction models for occurrences of high uncertainty; Rohwer et al., 1994) to determine which risk factors predict official recidivism within that sample. Most dynamic instruments were created by selecting risk factors validated in multiple studies to be associated with both official records and self-reported illegal activity across multiple jurisdictions. This is a critical distinction. Self-reported violence and offending has significantly less racial disparity than criminal records (Loeber et al., 2015). Indeed, studies have found no significant differences between Black and White youth on items of an SPJ instrument that counts prior illegal acts and conduct as opposed to only what is on one's official record (e.g., Chapman et al., 2006; Perrault et al., 2017).

WHAT DOES IT MEAN FOR AN INSTRUMENT TO BE RACIALLY BIASED?

There has been considerable disagreement, and perhaps even some misunderstanding, regarding what constitutes racial bias in an instrument. An instrument is not necessarily racially biased if one group (e.g., people of color) simply scores higher, on average, than another group (e.g., White people). There are ethical standards in the Psychology and

Education fields (but not yet in Criminology) regarding how to validate our instruments (AERA, APA, NCME, & Joint Committee on Standards for Educational and Psychological Testing [U.S.], 2014). These standards state that test bias (e.g., racial bias) is present when scores function differently for different groups of people. For example, men score higher on risk assessments than women, on average. Does that mean all risk instruments are biased against men? Not necessarily, because men also engage in more crime than women, on average; men have more risk factors. Thus, men's higher scores simply reflect true differences in reoffending and the risk instrument is doing its job. Researchers have tested questions of racial bias and fairness within three areas.

TEST BIAS AND PREDICTIVE ACCURACY

As Skeem and Lowenkamp (2016) explained, risk assessment scores should statistically relate to the outcome the instrument was designed to detect and should do so in the same way regardless of group membership: Each group should have a similar probability of recidivism at each score on the instrument. Thus, we must compare the "functional form" (does an average risk score of X relate to an average recidivism rate of Y for different racial groups?) of test scores. Skeem and Lowenkamp made this comparison using moderated hierarchical regression to examine potential interactions between race and risk scores in the prediction of recidivism.

In our review of peer-reviewed, published studies since 2000, we found only nine studies with nonoverlapping samples that evaluated a risk instrument's differential predictive accuracy by race or culture using a method to compare the functional form of risk scores or structured risk judgments.¹ Eight tested actuarial instruments (C. Campbell et al., 2018; Flores et al., 2016; Lowder et al., 2019; Perrault et al., 2017; Schwalbe et al., 2004, 2007; Skeem & Lowenkamp, 2016), only one of which was a pretrial risk instrument (Cohen & Lowenkamp, 2018), and three tested SPJ instruments (Lowder et al., 2019; Muir et al., 2020; Perrault et al., 2017). Four of the 10 found significant racial differences in predictive accuracy. Two indicated the difference favored Black youth (C. Campbell et al., 2018; Schwalbe et al., 2004), one indicated differences were present only when gender also was considered (Muir et al., 2020), one found overprediction for Hispanic adults (Cohen & Lowenkamp, 2018), and one found overprediction for Black youth, which appeared to be a result of the criminal history items (Schwalbe et al., 2007). None of the SPJ instrument studies test bias on the structured risk judgments.

ERROR (FALSE POSITIVE) RATES

Other studies have defined bias as a difference in error classification or false positive rates, meaning a risk instrument may falsely classify one group as high risk at a higher rate than another group, which could easily result in unfair, harsher treatment of the group with more misclassifications. This approach was taken by a few non-peer-reviewed studies that reported a bias against Black adults (e.g., Angwin et al., 2016; Lason et al., 2016 [both were discredited, see Flores et al., 2016]), and a few peer-reviewed studies since 2000 (e.g., Muir et al., 2020; Rembert et al., 2014; Skeem & Lowenkamp, 2020), which found differences in false positive rates were small to nonsignificant or favored the particular group of color (meaning more individuals who were not categorized as high-risk recidivated), with one exception (see Dressel & Farid, 2018).

Cohen and Lowenkamp (2018) argued that evaluating and maximizing predictive accuracy is of “paramount importance” (p. 238) over evaluating error classification within justice settings due to the great need to optimally assess likelihood of danger to the community. They explained it is impossible to improve an instrument’s error classifications without also decreasing predictive accuracy except in circumstances where prediction is either perfect or the recidivism base rates between groups are equal (Kleinberg et al., 2016). Skeem and Lowenkamp (2020) rigorously demonstrated this in the context of a risk instrument using multiple algorithms, concluding like other scientists (Berk et al., 2017; Chouldechova, 2017) that it is impossible to satisfy both fairness in predictive accuracy and in error classifications at the same time.

DISPARATE IMPACT

The biggest concern by critics is that significant mean score or error rate differences on instruments by race will result in harsher system-related responses (Skeem & Lowenkamp, 2016), like incarceration. Because many studies have found such differences on risk instruments (e.g., Olver et al., 2014), this concern warrants considerable attention.² For score-related differences to have an impact on decisions, these must translate into significant differences in the proportions of individuals categorized at low-to-high risk. To determine whether risk instruments are having a disparate impact, we need to uncover evidence that they are leading to greater system disparity than the traditional approach of relying on hunches. There is currently no adequate evidence of that; however, there are few studies on this issue and most have poor research designs (Viljoen et al., 2019).

CONCLUSIONS AND RECOMMENDATIONS

In sum, by our ethical standards, there is currently no valid evidence that instruments in general are biased against individuals of color. Where bias has been found, it appears to have more to do with the specific risk instrument. On balance, the advent of risk instruments has led to a modest reduction in incarceration (Viljoen et al., 2019) without increased risk to public safety. The fact structured risk instruments improve accuracy over hunches about who is likely to be a danger has been well established (e.g., Grove et al., 2000; Lin et al., 2020). Moreover, risk “assessments” by judges and personnel still occur in the absence of actual instruments, and justice personnel are not immune to biases when they are not using instruments (see Bridges & Steen, 1998; Graham & Lowery, 2004). Judgments by personnel at different points in the system, which are generally guided by offense severity, have led to considerable racial disparities in the United States. Thus, reverting to professionals relying on their intuition is unlikely to fare better than using instruments (Picard et al., 2019).

Critics should be aware that not all risk instruments use an algorithm or include items based solely on one’s criminal history. Instrument developers should be aware that when we create tools that solely comprise these items, we are at risk of carrying forward the system’s racial disparities to build tools that have intrinsic racial bias (Hart, 2016). For risk instruments created actuarially, regardless of whether one used traditional statistical modeling or “machine-learning” methods, the underlying issue is the same. Due to simple mathematics, we must expect that if Black defendants have a higher rate of official recidivism than White defendants, and an algorithm is highly predictive of or well calibrated to those outcomes,

the algorithm will classify a greater proportion of Black defendants as high risk; and therefore, a greater proportion of Black defendants who ultimately do not recidivate will have been classified as high risk (Skeem & Lowenkamp, 2020). In short, we are confounding the question of who is likely to engage in illegal and potentially harmful conduct with who is likely to get apprehended, and we are shining a light on the long-standing problem of systemic injustices.

Based on the current evidence, we provide some suggestions for promising approaches to minimizing bias and unfairness in justice settings and areas in need of research.

Practitioners, justice agencies and courts should:

- Never make decisions based solely on score-based classifications of risk, period. If it is impossible to have an algorithm with both high predictive accuracy and low error classifications—we should not base a person’s life on it. Generally, every court decision involves a human decision-maker who must weigh the relevance of the evidence. Rather than eliminate risk instruments, decision-makers should use them but “think beyond the algorithm” (Picard et al., 2019). One approach could be to borrow strategies from the structured professional judgment method whereby decision-makers presented with results of risk instruments could be trained to consider the relevance of the risk factors to the individual (e.g., multiple prior convictions may be less relevant for defendants of color than for White defendants) before making their final decision about risk to community safety. Another approach may be the adjusted actuarial method (see Picard et al., 2019).
- Allow time for completion of instruments that contain dynamic risk factors when facing court decisions where incarceration is a distinct possibility. These instruments can guide more meaningful strategies for mitigating risk and avoiding unnecessary incarceration. In the pre-trial context, the time involved in conducting better risk assessments could be manageable if courts instituted valid criteria to first “screen out” the majority of individuals who should not be “detention eligible” (see Picard et al., 2019; Robinson & Koepke, 2019).
- Seek training and experience in cultural competence to gain awareness of potential biases in decisions and work to prevent these (Hart, 2016).

Risk assessment researchers should:

- Evaluate all risk assessment instruments to determine if they predict equally well across racial and ethnic groups using the appropriate methods as described earlier or latent test theory methods (Hart, 2016) in samples from multiple jurisdictions.
- Examine whether the use of actuarial versus SPJ instruments in practice actually reduce or increase disparities in incarceration at different decision-points (e.g., pretrial, sentencing, and release) as compared with unstructured judgments.
- Conduct both experimental (e.g., vignette studies) and applied research designs to evaluate whether the rater discretion permitted in SPJ instruments introduces more racial bias than actuarial instruments within the same samples of evaluatees.
- Examine whether contextual or individual factors affect the likelihood of bias in rater discretion; such as whether the race/ethnicity of the evaluator matches the evaluatee, whether it is a high-stakes case, and so on.

ORCID iDs

Gina M. Vincent  <https://orcid.org/0000-0002-9315-9584>

Jodi L. Viljoen  <https://orcid.org/0000-0002-1609-0202>

NOTES

1. Most studies have instead compared the strength of instruments' predictive accuracy by race using Areas Under the Receiver Operating Curve (AUCs). Most have not found significant differences, including a meta-analysis of the Level of Service Inventory Scales (Olver et al., 2014), but there has been some variability, with exceptions typically being studies that test AUCs by gender and race subsamples (e.g., C. Campbell et al., 2018).

2. As an aside, we must be careful to not automatically interpret such differences as a result of biased instruments because categorical differences in risk shown in any study conducted in one jurisdiction may merely reflect a difference in who gets arrested in that jurisdiction.

REFERENCES

- Ægisdóttir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., Cook, R., Nichols, C. N., Lampropoulos, G. K., Walker, B. S., Cohen, G., & Rush, J. D. (2006). The meta-analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *The Counseling Psychologist, 34*(3), 341–382. <https://doi.org/10.1177/0011000006295149>
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational and Psychological Testing (U.S.). (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). Machine bias: There's software used across the country to predict future criminals, and it's biased against Blacks. *ProPublica*. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- Berk, R., Heidari, H., Jabbari, S., Kearns, M., & Roth, A. (2017, May 28). *Fairness in criminal justice risk assessments: The state of the art*. arXiv. <https://arxiv.org/abs/1703.09207>
- Bridges, G. S., & Steen, S. (1998). Racial disparities in official assessments of juvenile offenders: Attributional stereotypes as mediating mechanisms. *American Sociological Review, 63*(4), 554–570. <https://doi.org/10.2307/2657267>
- Campbell, C., Papp, J., Barnes, A., Onifade, E., & Anderson, V. (2018). Risk assessment and juvenile justice: An interaction between risk, race, and gender. *Criminology & Public Policy, 17*, 1–22. <https://doi.org/10.1111/1745-9133.12377>
- Campbell, M. A., French, S., & Gendreau, P. (2009). The prediction of violence in adult offenders: A meta-analytic comparison of instruments and methods of assessment. *Criminal Justice and Behavior, 36*(6), 567–590. <https://doi.org/10.1177/0093854809333610>
- Chapman, J. F., Desai, R. A., Falzer, P. R., & Borum, R. (2006). Violence risk and race in a sample of youth in juvenile detention: The potential to reduce disproportionate minority confinement. *Youth Violence and Juvenile Justice, 4*(2), 170–184. <https://doi.org/10.1177/1541204006286316>
- Chouldechova, A. (2017). Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big Data, 5*(2), 153–163. <https://doi.org/10.1089/big.2016.0047>
- Cohen, T. H., & Lowenkamp, C. T. (2018). Revalidation of the federal PTRAs: Testing the PTRAs for predictive biases. *Criminal Justice and Behavior, 46*, 234–260. <https://doi.org/10.1177/0093854818810315>
- Booker, C. A., & Schatz, B. (2018). *Letter to Attorney General Sessions, Judge Pryor, and Sentencing Commission*. Booker. Senate.gov. <https://www.scribd.com/document/370800709/02052018-Booker-Schatz-Letter-to-DOJ-and-USSC>
- Dressel, J., & Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism. *Science Advances, 4*(1), 1–5. <https://doi.org/10.1126/sciadv.aao5580>
- Eisenberg, M. J., van Horn, J. E., Dekker, J. M., Assink, M., van der Put, C. E., & Hendriks, G. (2019). Static and dynamic predictors of general and violent criminal offense recidivism in the forensic outpatient population: A meta-analysis. *Criminal Justice and Behavior, 46*, 732–750. <https://doi.org/10.1177/0093854819826109>
- Flores, A. W., Bechtel, K., & Lowenkamp, C. T. (2016). 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.” *Federal Probation, 80*(2), 38–46.
- Graham, S., & Lowery, B. S. (2004). Priming unconscious racial stereotypes about adolescent offenders. *Law and Human Behavior, 28*, 483–504. <https://doi.org/10.1023/B:LAHU.0000046430.65485.1f>
- Grisso, T. (2005). Why we need mental health screening and assessment in juvenile justice programs. In T. Grisso, G. Vincent, & D. Seagrave (Eds.), *Mental health screening and assessment in juvenile justice* (pp. 3–21). Guilford Press.
- Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment, 12*(1), 19–30. <https://doi.org/10.1037/1040-3590.12.1.19>
- Harcourt, B. E. (2015). Risk as a proxy for race: The dangers of risk assessment. *Federal Sentencing Reporter, 27*(4), 237–243. <https://doi.org/10.1525/fsr.2015.27.4.237>
- Hart, S. D. (2016). Culture and violence risk assessment: The case of Ewert v. Canada. *Journal of Threat Assessment and Management, 3*(2), 76–96. <https://doi.org/10.1037/tam0000068>

- Hockenberry, S., & Puzanchera, C. (2020, April). *Juvenile court statistics 2018*. National Center for Juvenile Justice.
- Holder, E. (2015). *Attorney General Eric Holder Speaks at the National Association of Criminal Defense Lawyers 57th Annual Meeting and 13th State Criminal Justice Network Conference* [Speech transcript]. <http://www.justice.gov/opa/speech/attorney-general-eric-holder-speaks-national-association-criminal-defense-lawyers-57th> (Original work published 2014)
- Kleinberg, J., Mullainathan, S., & Raghavan, M. (2016, September 19). *Inherent trade-offs in the fair determination of risk scores*. arXiv. <https://arxiv.org/abs/1609.05807>
- Lason, J., Mattu, S., Kirchner, L., & Angwin, J. (2016, May 23). How we analyzed the COMPAS recidivism algorithm. *ProPublica*. <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>
- Lin, Z., Jung, J., Goel, S., & Skeem, J. (2020). The limits of human predictions of recidivism. *Science Advances*, 6(7), 1–8. <https://doi.org/10.1126/sciadv.aaz0652>
- Loeber, R., Farrington, D. P., Hipwell, A. E., Stepp, S. D., Pardini, D., & Ahonen, L. (2015). Constancy and change in the prevalence and frequency of offending when based on longitudinal self-reports or official records: Comparisons by gender, race, and crime type. *Journal of Developmental Life Course Criminology*, 1, 150–168. <https://doi.org/10.1007/s40865-015-0010-5>
- Lowder, E. M., Desmarais, S. L., Rade, C. B., Johnson, K. L., & Van Dorn, R. A. (2019). Reliability and validity of START and LSI-R assessments in mental health jail diversion clients. *Assessment*, 26, 1347–1361. <https://doi.org/10.1177/1073191117704505>
- Monahan, J. (2008). Structured risk assessment of violence. In R. I. Simon & K. Tardiff (Eds.), *Textbook of violence assessment and management* (pp. 17–33). American Psychiatric Publishing.
- Muir, N. M., Viljoen, J. L., Jonnson, M. R., Cochrane, D. M., & Rogers, B. J. (2020). Predictive validity of the Structured Assessment of Violence Risk in Youth (SAVRY) with Indigenous and Caucasian female and male adolescents on probation. *Psychological Assessment*, 32, 594–607. <https://doi.org/10.1037/pas0000816>
- Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2014). Thirty years of research on the Level of Service Scales: A meta-analytic examination of predictive accuracy and sources of variability. *Psychological Assessment*, 26(1), 156–176. <https://doi.org/10.1037/a0035080>
- Perrault, R. T., Vincent, G. M., & Guy, L. S. (2017). Are risk assessments racially biased? Field study of the SAVRY and YLS/CMI in probation. *Psychological Assessment*, 29(6), 664–678. <https://doi.org/10.1037/pas0000445>
- Picard, S., Watkins, M., Rempel, M., & Kerodal, A. (2019). *Beyond the algorithm: Pretrial reform, risk assessment, and racial fairness*. https://www.courtinnovation.org/sites/default/files/media/document/2019/Beyond_The_Algorithm.pdf
- Pretrial Justice Institute. (2020, February 7). *Updated position on pretrial risk assessment tools*. <https://www.pretrial.org/wp-content/uploads/Risk-Statement-PJI-2020.pdf>
- Rembert, D. A., Henderson, H., & Pirtle, D. (2014). Differential racial/ethnic predictive validity. *Youth Violence and Juvenile Justice*, 12(2), 152–166. <https://doi.org/10.1177/1541204013485606>
- Robinson, D. G., & Koepke, L. (2019, December). *Civil rights and pretrial risk assessment instruments*. John D. and Catherine T. Macarthur Foundation. <http://www.safetyandjusticechallenge.org/wp-content/uploads/2019/12/Robinson-Koepke-Civil-Rights-Critical-Issue-Brief.pdf>
- Rohwer, R., Wynne-Jones, M., & Wyszotzki, F. (1994). Neural networks. In D. Michie, D. J. Spiegelhalter, & C. C. Taylor (Eds.), *Machine learning, neural and statistical classification (Ellis Horwood series in artificial intelligence)* (pp. 84–106). Ellis Horwood. <http://www1.maths.leeds.ac.uk/~charles/statlog/>
- Schwalbe, C. S., Fraser, M. W., & Day, S. H. (2007). Predictive validity of the Joint Risk Matrix with juvenile offenders: A focus on gender and race/ethnicity. *Criminal Justice and Behavior*, 34(3), 348–361. <https://doi.org/10.1177/0093854806292244>
- Schwalbe, C. S., Fraser, M. W., Day, S. H., & Arnold, E. M. (2004). North Carolina Assessment of Risk (NCAR): Reliability and predictive validity with juvenile offenders. *Journal of Offender Rehabilitation*, 40, 1–22. https://doi.org/10.1300/J076v40n01_01
- Schwartzapfel, B. (2019, July 1). Can racist algorithms be fixed? The Marshall Project. <https://www.themarshallproject.org/2019/07/01/can-racist-algorithms-be-fixed>
- Skeem, J. L., & Lowenkamp, C. T. (2016). Risk, race, and recidivism: Predictive bias and disparate impact. *Criminology*, 54(4), 680–712. <https://doi.org/10.1111/1745-9125.12123>
- Skeem, J. L., & Lowenkamp, C. T. (2020). Using algorithms to address trade-offs inherent in predicting recidivism. *Behavioral Sciences & the Law*, 38(3), 259–278. <https://doi.org/10.1002/bsl.2465>
- Starr, S. B. (2015). The risk assessment era: An overdue debate. *Federal Sentencing Reporter*, 27(4), 205–206. <https://doi.org/10.1525/fsr.2015.27.4.205>
- Viljoen, J. L., Jonnson, M. R., Cochrane, D. M., Vargen, L. M., & Vincent, G. M. (2019). Impact of risk assessment instruments on rates of pretrial detention, postconviction placements, and release: A systematic review and meta-analysis. *Law and Human Behavior*, 43(5), 397–420. <https://doi.org/10.1037/lhb0000344>

- Vincent, G. M., Chapman, J., & Cook, N. E. (2011). Risk-needs assessment in juvenile justice: Predictive validity of the SAVRY, racial differences, and the contribution of needs factors. *Criminal Justice and Behavior*, 38(1), 42–62. <https://doi.org/10.1177/0093854810386000>
- Wachter, A. (2015). *Statewide risk assessment in juvenile probation*. *JGPS StateScan*. National Center for Juvenile Justice.
- Yang, M., Wong, S. C., & Coid, J. (2010). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin*, 136(5), 740–767. <https://doi.org/10.1037/a0020473>

Gina M. Vincent is an associate professor in the Department of Psychiatry and Co-Director of the Law & Psychiatry Program at the University of Massachusetts Medical School. Her research is in risk assessment and its implementation in justice settings, adolescent offending, mental health symptoms, and substance abuse.

Jodi L. Viljoen is a professor of clinical and forensic psychology in the Department of Psychology at Simon Fraser University. Her research focuses on risk assessment, risk management, and adolescent offending.