2023

The Need for Fairness and Accuracy for Women in Sentencing: Surmounting Challenges to Gender-Specific Statistical Risk Assessment Tools

Elizabeth E. Wainstein
University of California - Berkeley

THE NEED FOR FAIRNESS AND ACCURACY FOR WOMEN IN SENTENCING: SURMOUNTING CHALLENGES TO GENDER-SPECIFIC STATISTICAL RISK ASSESSMENT TOOLS

Elizabeth E. Wainstein*

States across the country have increasingly adopted statistical risk assessment tools in multiple stages of their criminal legal systems with the hope of reducing incarceration without increasing crime. These tools use various characteristics to estimate an individual's future risk of recidivism, and judges consider the results of these assessments when determining levels of custody or community supervision for convicted individuals. Despite much debate amongst academics and activists on the utility and fairness of these tools, one critique seems beyond debate: the tools are built for men, not women. These tools are based on criteria, statistics, and theory drawn from the experiences of men and thereby result in inaccurate and inequitable sentencing when applied to women. When women are sentenced according to the higher rates of violence and recidivism that are associated with men, they are often incarcerated or under supervision longer than justified by their gender-specific risk to society. The unfairness of these assessments is specifically concerning when one considers that, as of 2019, 1.2 million women in the United States were under the supervision of the criminal legal system, with approximately fifty-eight percent of them leaving at least one minor child at home without a mother's care and guidance.

Separate risk assessment tools for men and women can combat the inaccurate sentencing of women. While many commentators have argued for

* J.D. Candidate, University of California Berkeley School of Law, 2023. Thank you to Professor Andrea Roth for invaluable insight and guidance throughout the research and writing process; to Professor Katerina Linos, Professor Catherine Fisk, and the participants of the 2022 Note Publishing Workshop for providing feedback on drafts; and to the student editors of the Journal of Criminal Law and Criminology Online. Finally, thanks to my father for the many rounds of proofreading and to my entire family for their support throughout the various iterations of this project.

separate tools for men and women, they have not sufficiently addressed how such an approach would survive legal, theoretical, and policy hurdles. This Comment argues (1) that gender-specific assessments could survive an equal protection challenge; (2) that such assessments for women should be implemented despite the need for further research and work on the conflation of sex and gender and the utilization of a gender binary in the United States criminal legal system; and (3) that they could be adapted for women defendants without opening the floodgates to a demand for assessments designed for every conceivable category of criminal defendant.

INTRODUCTION 33
I. RECIDIVISM AND THE ISSUES WITH STATISTICAL RISK ASSESSMENT TOOLS 35
A. Background 35
B. The Problem with Statistical Risk Assessment Tools in Sentencing 37
II. A POSSIBLE SOLUTION—GENDER-SPECIFIC RISK ASSESSMENT TOOLS 41
III. ANTICIPATING OBJECTIONS TO HAVING SEPARATE TOOLS FOR MEN AND WOMEN 42
A. Would the use of separate statistical risk assessment tools for men and women in sentencing violate the Equal Protection Clause? 43
1. Equal Protection Objection 44
2. Governmental Objectives Served by Gender-Based Classification 47
a. Deploying Separate Tools for Men and Women Would Increase Accuracy 47
b. Appropriate Sentencing Better Protects the Community 48
3. Tailoring to Important Governmental Objectives 50
B. Are there theoretical objections to identifying individuals in sentencing by gender? 52
1. The Conflation of Sex and Gender 52
2. Gender as a Spectrum 52
C. Policy Concerns with Having Different Tools Based on Gender 55
CONCLUSION 57

INTRODUCTION

There appears to be widespread consensus that the American criminal legal system suffers systematic failures which results in over-punishment.[1] In 2019, 2,086,600 people were incarcerated in the United States. [2] With 810 out of every 100,000 adult residents being held in prisons or jails, the United States incarcerated a larger proportion of its population than any other country for which data is available. This has resulted in people along the political spectrum calling for criminal legal reform.[3]

In the wake of this movement to reduce mass incarceration, statistical risk assessment algorithms were presented as a potential means of reducing incarceration without increasing crime, by more precisely targeting defendants at a high risk of reoffending.[4] States across the country are increasingly adopting these tools throughout their criminal legal systems.[5]

Despite much debate on whether these tools are appropriate and fair, one critique seems beyond contention: the tools are built for men, not women. These statistical risk assessment tools, which are based on risk factors and theories of crime that focus on men, contribute to inaccuracies and inequalities in sentencing between men and women in the criminal legal system.[6] In particular, because the tools are trained using data sets involving men, and because men on average have higher base rates of reoffending, the predictions of risk for women are artificially inflated.

The potential for these tools to falsely label women as high risk is specifically concerning because the number of incarcerated women increased

---

[1] *See* Maggie Astor, *Left and Right Agree on Criminal Justice: They Were Both Wrong Before*, N.Y. TIMES (May 16, 2019), https://www.nytimes.com/2019/05/16/us/politics/criminal-justice-system.html [https://perma.cc/6B4G-M98B].

[2] Todd D. Milton, Lauren G. Beatty, & Zhen Zeng, *Correctional Populations in the United States, 2019 – Statistical Tables*, BUREAU OF JUST. STAT. 1 (July 2021), https://bjs.ojp.gov/sites/g/files/xyckuh236/files/media/document/cpus19st.pdf [https://perma.cc/R4PJ-TVSS].

[3] John Gramlich, *America's Incarceration Rate Falls to Lowest Level Since 1995*, PEW RSCH. CTR. (Aug. 16, 2021), https://www.pewresearch.org/fact-tank/2021/08/16/americas-incarceration-rate-lowest-since-1995 [https://perma.cc/KRM7-Q3P8]; *see generally* E. Ann Carson, *Prisoners in 2018*, BUREAU OF JUST. STAT. (Apr. 2020), https://www.bjs.gov/content/pub/pdf/p18.pdf [https://perma.cc/4AFZ-LZ93].

[4] *See* Jennifer L. Skeem, John Monahan, & Christopher T. Lowenkamp, *Gender, Risk Assessment, and Sanctioning: The Cost of Treating Women Like Men*, 40 PUB. L. & LEGAL THEORY RSCH. PAPER SERIES 500, 584–93 (Feb. 2016).

[5] Rebecca Foxwell, *Risk Assessments and Gender for Smarter Sentencing*, 3 VA. J. CRIM. L. 435, 438–39 n.10 (2015).

[6] Elizabeth Wainstein, *The Use of Gender in Risk Assessment Tools* 61 (July 15, 2020) (MPhil dissertation, University of Cambridge) (on file with author).

from 26,326 to 152,854 between 1980 and 2020, more than a 475% increase.[7] The overall rate of growth for incarcerating women in the United States has been twice as high as that of men since 1980 and, as of 2020, 1 million women were under the supervision of the criminal legal system.[8] The staggeringly-high social cost of mass sentencing disparities is only exacerbated by the fact that 58% of the women in state prisons have at least one child under the age of 18.[9] The number of children under 18 years old with a mother in prison has more than doubled between 1991 and 2010.[10] As a result, any disparities that these tools produce are harming a large population of women, a majority of whom have young children they are forced to leave behind. With these tools' widespread implementation, sentencing disparities created by miscalculating women's risk of recidivism have become increasingly problematic.[11]

One possible solution that criminology scholars have introduced is to create and implement different statistical risk assessment tools for men and women.[12] The tools for women would be based on research and theories of crime that consider women's unique risk factors to ensure that women are not sentenced unfairly due to risk factors that are not specific to them. While some criminologists previously suggested women-specific tools, the implications—legal and otherwise—of their actual deployment have not been fully addressed.

This Comment fills that gap, exploring the possible legal, theoretical, and policy implications of having different statistical risk assessment tools for men and women in sentencing. Part I offers a descriptive background, explaining the role of recidivism in sentencing, the risk assessment tools currently in use, and the ways in which these current tools based on men's criminal behavior misclassify women and contribute to inaccuracies in sentencing. Part II sets forth a proposal for a women-only tool as a solution

---

[7] Fact Sheet, *Incarcerated Women and Girls*, THE SENTENCING PROJECT (Nov. 2020), https://www.sentencingproject.org/publications/incarcerated-women-and-girls [https://perma.cc/YYK3-MFP2].

[8] *Id.* at 1.

[9] *Id.*

[10] Lauren E. Glaze & Laura M. Maruschak, *Parents in Prison and Their Minor Children*, BUREAU OF JUST. STAT. 2 (Aug. 2008), https://bjs.ojp.gov/content/pub/pdf/pptmc.pdf [https://perma.cc/3QA6-8UD5].

[11] I have strong reservations about the use of statistical risk assessment tools at sentencing. However, this paper's discussion is based on the assumption that they will continue to be used throughout the United States.

[12] Kelly N. Taylor & Kelley Blanchette, *The Women are Not Wrong: It is the Approach that is Debatable*, 8 CRIMINOLOGY & PUB. POL'Y 221, 227 (2009) (exploring the values of gender-informed risk assessment tools).

to these three issues. Part III provides an overview and analysis of potential legal, theoretical, and policy objections to using different tools for men and women in sentencing. I conclude (1) that gender-specific assessments could survive an equal protection challenge; (2) that such assessments for women should be implemented despite the need for further research and work on the conflation of sex and gender and the utilization of a gender binary in the United States criminal legal system; and (3) that they could be adapted for women defendants without opening the floodgates to a demand for assessments designed for every conceivable category of criminal defendant.

I. RECIDIVISM AND THE ISSUES WITH STATISTICAL RISK ASSESSMENT TOOLS

A. BACKGROUND

With the rise of evidence-based sentencing, predicting the risk that a sentenced defendant will reoffend, or their "recidivism" risk, has taken on an increasingly important role in the criminal legal system.[13] With laws addressing recidivism tracing back to 1695,[14] the assessment of an individual's future risk has been built into various stages of our nation's criminal legal system, from pre-trial detention to sentencing and prison placement to parole and probation revocation.[15] At sentencing in particular, courts use the risk of recidivism as a gauge of an individual's ability to conform to the law and attempt to tailor sentencing decisions to meet that risk.[16]

Throughout the majority of American history, this prediction of recidivism risk has been conducted through holistic clinical assessments based on personal observation of the subject.[17] However, these clinical assessments are being phased out as the criminal legal system transitions to

---

[13] *See generally* Dawinder Sidhu, *Moneyball Sentencing*, 56 B.C. L. REV. 671, 687 (2015) (explaining the historical and current role of recidivism risk in the U.S. criminal legal system).

[14] *Id.* at 685 (providing one example of a 1705 law in Virginia that addressed "the persistent problem of hog stealing by passing a statute that provided progressively more severe penalties for each subsequent offense") (citing LAWS OF VIRGINIA 276–78 (W. Hening ed., 1823)).

[15] *Id.* at 686 ("[P]rediction of future criminal conduct is an essential element in many of the decisions rendered throughout our criminal justice system, and a task performed countless times each day throughout the American system of criminal justice.") (quoting Jurek v. Texas, 428 U.S. 262, 275–76 (1975) (Stevens, J.) (plurality opinion) (internal quotation marks omitted).

[16] *Id.* at 685.

[17] *Id.* at 687.

statistical predictions of risk.[18] Statistical risk assessment tools use various characteristics that are supposedly correlated with risk of future criminality, known as criminogenic factors, and provide an estimated score of the individual's future risk of recidivism.[19] These criminogenic factors include criminal demographic variables (such as age, gender, and marital status), history variables (such as number of past convictions, past incarceration sentences, and number of violent or drug convictions), and socioeconomic variables (such as employment status and education).[20] The recidivism risk scores, and corresponding risk categories identified by the tools are then provided to the sentencing judges for consideration in determining custody level or community supervision.[21]

These statistical risk assessment tools are often presented as an unbiased way to address some of the major issues in the United States criminal legal system, such as mass incarceration.[22] Criminal legal actors, scholars, and progressive reform advocates have promoted these tools as a method to lower incarceration rates while ensuring community safety by identifying low-risk individuals and diverting them from incarceration or assigning them to supervised release while assigning high-risk individuals to custody.[23] Additionally, supporters of these tools argue that algorithm-based tools help to create a fairer system by replacing human decisions, which inherently have inconsistencies, implicit biases, and prejudices, with an objective

---

[18] Dawinder Sidhu, *Moneyball Sentencing*, 56 B.C. L. REV. 671, 688 (2015).

[19] *See, e.g.*, John Lightbourne, *Damned Lies & Criminal Sentencing Using Evidence-Based Tools*, 15 DUKE L. & TECH. REV. 327, 327 (2017); Nathan James, *Risk and Needs Assessment in Federal Prison System*, CONGRESSIONAL RESEARCH SERVICE 3 (July 2018), https://sgp.fas.org/crs/misc/R44087.pdf [https://perma.cc/E4LE-UATV]; PATRICIA VAN VOORHIS, EMILY SALISBURY, EMILY WRIGHT & ASHLEY BAUMAN, NAT'L INST. CORR., ACHIEVING ACCURATE PICTURES OF RISK AND IDENTIFYING GENDER RESPONSIVE NEEDS: TWO NEW ASSESSMENTS FOR WOMEN OFFENDERS 1 (2008) [hereinafter ACCURATE PICTURES OF RISK]; Michael Marcus, *MPC—The Root of the Problem: Just Desserts and Risk Assessment*, 61 FLA. L. REV. 751, 766 (2009); Emily Salisbury, Patricia Van Voorhis & Georgia V. Spiropoulos, *The Predictive Validity of a Gender-Responsive Needs Assessment: An Exploratory Study*, 55 CRIME & DELINQUENCY 550, 551 (2009) [hereinafter *Predictive Validity*].

[20] Sonja B. Starr, *Evidence-Based Sentencing and The Scientific Rationalization of Discrimination*, 66 STAN. L. REV. 803, 811 (2014).

[21] John Monahan & Jennifer L. Skeem, *Risk Assessment in Criminal Sentencing*, ANNUAL REVIEW OF CLINICAL PSYCHOLOGY 489, 495 (2016); *Predictive Validity*, *supra* note 19, at 551.

[22] Ben Green, *The False Promise of Risk Assessments: Epistemic Reform and the Limits of Fairness*, PROCEEDINGS OF THE 2020 CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY (FAT*) (2020).

[23] Starr, *supra* note 20, at 805; James, *supra* note 19, at 2; Monahan & Skeem, *supra* note 21, at 494–95.

assessment.[24] Based on these arguments, states throughout the country have increasingly adopted these tools in recent years.[25]

B. THE PROBLEM WITH STATISTICAL RISK ASSESSMENT TOOLS IN SENTENCING

Due to their outsized impact on sentencing, it is vital that these statistical risk assessment tools are fair and accurate assessments for all people subjected to them. However, many of the statistical risk assessment tools currently in use do not accurately consider the differences between men and women's pathways into and out of crime, the factors that affect their chances of recidivism, and how those factors impact their decisions.[26] The failure to take these differences into consideration can result in sentencing based on inaccurate risk scores.[27] Specifically, in the context of statistical risk assessment, most tools were designed for men and based on the behaviors and risk factors associated with men's criminality.[28] This is particularly concerning when one considers that studies demonstrate that women commit crime at much lower rates than men,[29] incarcerated women receive less

---

[24] Green, *supra* note 22, at 8; Sam Corbett-Davies, Sharad Goel & Sandra González-Bailón, *Even Imperfect Algorithms Can Improve the Criminal Justice System*, N.Y. TIMES (Dec. 20, 2017), https://www.nytimes.com/2017/12/20/upshot/algorithms-bail-criminal-justice-system.html; Cass R. Sunstein, *Algorithms, Correcting Biases* (Dec. 12, 2018) (unpublished manuscript) (on file with author) (arguing first that algorithms can overcome the harmful effects of cognitive biases; and second, that algorithms can be designed to avoid discrimination in its unlawful form).

[25] Foxwell, *supra* note 5, at 438–39 n.10 (listing at least 20 states: Arizona, Indiana, Kentucky, Michigan, Missouri, Ohio, Oklahoma, Pennsylvania, Utah, Virginia, Washington, West Virginia, Illinois, New Mexico, North Dakota, Maine, Minnesota, North Carolina, Texas, and Wisconsin as states who have started to use risk assessment in sentencing) (citing Starr, *supra* note 20, at 809 n.11).

[26] See subsection III.B for a discussion of the conflation of sex and gender and the use of a gender-binary in criminological literature and the U.S. criminal legal system.

[27] Criminological research and practice have historically focused on men, and women have been overlooked due to their low incarceration numbers and researchers' widespread reliance on gendered assumptions about women's criminal behavior. *E.g.*, PATRICIA VAN VOORHIS, JENNIFER PEILER, LOIS PRESSER, GEORGIA SPIROPOULIS & JENNIFER SUTHERLAND, NAT'L INST. CORR., CLASSIFICATION OF WOMEN OFFENDERS: A NATIONAL ASSESSMENT OF CURRENT PRACTICES AND THE EXPERIENCES OF THREE STATES 2 (2001) [hereinafter CURRENT PRACTICES]; *Predictive Validity*, *supra* note 19, at 551 (most people are classified by tools that were designed for men and based on the behaviors and risk factors associated with men's criminality); TIM BRENNAN & JAMES AUSTIN, NAT'L INST. CORR., WOMEN IN JAIL: CLASSIFICATION ISSUES 1 (1997).

[28] BRENNAN & AUSTIN, *supra* note 27, at 1.

[29] *2019 Crime in the United States*, UNIFORM CRIME REPORTING (last visited Feb. 27, 2022) (reporting that, in 2019, 72.5% of all arrestees were men and 78.9% of persons arrested

serious misconduct charges than men at the same custody level,[30] and women convicted of crimes reoffend, commit violent crimes, and recidivate at lower rates than men convicted of similar crimes.[31] When these differences between men and women's behaviors and violence risks are not accounted for, women are essentially being punished based on men's overall risk.[32] Experts have identified problems that stem from this failure to account for differences, such as issues with validity,[33] lack of gender-responsivity,[34] and over-classification.[35]

Individuals experience gendered pathways into and out of crime, which means that women's unique life experiences and offending contexts need to be considered when evaluating a woman's recidivism risk.[36] There is a new understanding that men and women (1) have risk factors that are unique to each gender, (2) have some risk factors that are the same but are differentially predictive, meaning that the factors change their risk scores in different directions or magnitudes, and (3) react differently to different situations and

---

for violent crimes were men) https://ucr.fbi.gov/crime-in-the-u.s/2019/crime-in-the-u.s.-2019 [https://perma.cc/GMK9-WWC2]; Örjan Falk, Märta Wallinius, Sebastian Lundström, Thomas Frisell, Henrik Anckarsäter & Nóra Kerekes, *The 1% of the Population Accountable for 63% of All Violent Crime Convictions*, 49 SOC. PSYCHIATRY PSYCHIATRIC EPIDEMIOLOGY 559, 559 (2013) (noting that the majority of violent crimes in Sweden are perpetrated by a small number of persistent offenders, typically men).

[30] Salisbury et al., *Predictive Validity*, *supra* note 19, at 556 (explaining that "the proportion of women in maximum custody who incurred serious prison misconducts [is similar to] the proportion of medium-custody men who committed" serious prison misconducts).

[31] Wainstein, *supra* note 6, at 30.

[32] Starr, *supra* note 20, at 825.

[33] Sarah M. Manchak, Jennifer L. Skeem, Kevin S. Douglas & Maro Siranosian, *Does Gender Moderate the Predictive Utility of the Level of Service Inventory-Revised (LSI-R) For Serious Violent Offenders?*, 36 CRIM. JUST. & BEHAV. 425, 428 (2009); Kelly Hannah-Moffat, *Actuarial Sentencing: An "Unsettled" Proposition*, 30 JUST. Q. 270, 279–82 (2012).

[34] Patricia Van Voorhis, Emily M. Wright, Emily J. Salisbury & Ashley Bauman, *Women's Risk Factors and Their Contributions to Existing Risk/Needs Assessment: The Current Status of a Gender-Responsive Supplement*, 37 CRIM. JUST. & BEHAV. 261, 263 (2010) [hereinafter *Women's Risk Factors*] (explaining that "gender-responsive literature suggests that female offenders are very different from male offenders, as evidenced by their unique paths into criminal behavior, the offenses in which they engage, and their decreased threat of violence across criminal justice settings"); Taylor & Blanchette, *supra* note 12, at 226.

[35] BRENNAN & AUSTIN, *supra* note 27, at 11; BARBARA BLOOM, BARBARA OWEN & STEPHANIE COVINGTON, NAT'L INST. CORR., GENDER-RESPONSIVE STRATEGIES: RESEARCH, PRACTICE, AND GUIDING PRINCIPLES FOR WOMEN OFFENDERS 19 (2003).

[36] Taylor & Blanchette, *supra* note 12, at 224.

risk factors.[37] Specifically, major risk factors for women that differ from men's risk factors include histories of abuse,[38] mental health needs,[39] employment,[40] and addiction.[41]

If risk factors impact men and women differently because of their unique experiences in our gendered society, then a tool that considers men's gendered reality—but not women's—will likely result in inaccurate and inequitable sentences. For example, statistical risk assessment tools often score full-time parents as "unemployed," a status that can increase their classification level, which corresponds to the type and length of a sentence.[42] This increased score misclassifies women because of their disproportionately high rate of assuming full-time parenting duties.[43] Additionally, most current classification models heavily weigh the seriousness of the current offense, which can harmfully impact victims of abuse—who are often women.[44] In particular, a murder conviction can place a woman in a high custody setting for killing her abuser even though women offenders and prison staff do not view the majority of women who commit a one-time offense in response to a sustained period of abuse as violent.[45]

---

[37] Manchak, Skeem, Douglas & Siranosian, *supra* note 33, at 426; *Women's Risk Factors*, *supra* note 34, at 263.

[38] MERRY MORASH, TIMOTHY S. BYNUM & BARBARA A. KOONS, NAT'L INST. JUST., WOMEN OFFENDERS: PROGRAMMING NEEDS AND PROMISING APPROACHES 1 (1998) (explaining that women in the criminal legal system are disproportionately victims of sexual or physical abuse, including childhood abuse); Manchak, Skeem, Douglas & Siranosian, *supra* note 33, at 426 (stating that 26% percent of women inmates have been exposed to sexual abuse during childhood compared to 5% of male inmates).

[39] *Predictive Validity*, *supra* note 19, at 559 (explaining that women are more likely to suffer from depression, anxiety, self-injurious behavior, as well as major mood disorders such as bipolar disorder, panic, post-traumatic stress, and eating disorders); Manchak, Skeem, Douglas & Siranosian, *supra* note 33, at 426 (discussing how mental disorders uniquely impact women's propensity for violence and that, women, generally at a lower risk for violence, are as likely to become involved in violence as men when they have a mental disorder).

[40] MORASH, BYNUM & KOONS, *supra* note 38, at 1 (explaining that women are more likely to be unemployed before incarceration).

[41] PATRICIA VAN VOORHIS, CTR. CRIM. JUST. RSCH. U. CINNCINATI, CLASSIFICATION OF WOMEN OFFENDERS: GENDER-RESPONSIVE APPROACHES TO RISK/NEEDS ASSESSMENT 4 (2004) (noting that women's pathways into substance abuse appear to be different than men's pathways, and a history of substance abuse has a slightly larger effect on women's recidivism than men's recidivism).

[42] CURRENT PRACTICES, *supra* note 27, at 12, 22.

[43] *Id.*

[44] *Id.* at 12–13.

[45] *Id.* at 9.

Not only do gendered differences lead to misclassifications for sentencing, but men and women in the same risk classifications also have differing probabilities of participating in future problematic behavior. For example, high-risk men have a higher probability of future problematic behavior than high-risk women, meaning they are more likely to reoffend.[46] This is critical to consider, given the studies demonstrating that incarcerated women incur less serious misconduct charges than men at the same custody level[47] and that maximum-custody women incur serious misconduct charges at roughly the same rate as medium-custody men. [48] As a result, a high-risk woman likely possesses a lower risk of reoffending or incurring misconduct charges and, therefore, should not be treated the same as a high-risk man. The information provided above illustrates how differences between men and women must be adequately considered to ensure accurate sentencing.

Skewed sentencing is particularly concerning when one considers the social impact of keeping women behind bars longer than necessary. As of 2019, 1.2 million women were under the supervision of the criminal legal system[49] and 58% of the women in state prisons had at least one child under the age of 18.[50] The number of children under 18 years old with a mother in prison has more than doubled between 1991 and 2010.[51] This is important because children who grow up with an incarcerated parent face many hurdles.[52] These hurdles include being "more likely to exhibit low self-esteem, depression, emotional withdrawal from friends and family, and inappropriate or disruptive behavior at home and in school" and, most saliently, having an increased risk of future delinquency or criminal behavior.[53] Additionally, the incarceration of a mother is especially disruptive for a child because women are most often the primary caregiver.[54] Therefore, sentencing tools more calibrated to women defendants will

---

[46] *See* CURRENT PRACTICES, *supra* note 27, at 5.

[47] *Predictive Validity*, *supra* note 19, at 556.

[48] *Id.* (explaining that the proportion of maximum-custody women who incurred serious prison misconducts is similar to the proportion of medium-custody men who committed serious prison misconducts).

[49] *Incarcerated Women and Girls*, *supra* note 7.

[50] *Id.*

[51] Glaze & Maruschak, *supra* note 10, at 2.

[52] LOIS M. DAVIS, MALCOLM V. WILLIAMS, KATHRYN PITKIN DEROSE, PAUL STEINBERG, NANCY NICOSIA, ADRIAN OVERTON, LISA MIYASHIRO, SUSAN TURNER, TERRY FAIN & EUGENE WILLIAMS, RAND CORP., UNDERSTANDING THE PUBLIC HEALTH IMPLICATIONS OF PRISONER REENTRY IN CALIFORNIA 117 (2011).

[53] *Id.*

[54] *Id.* at 119.

contribute to long-term social benefits by decreasing the future risk of delinquency and criminal behavior in the next generation.

II. A POSSIBLE SOLUTION—GENDER-SPECIFIC RISK ASSESSMENT TOOLS

The previous section outlined how statistical risk assessment tools are imposing inaccurate and inequitable sentences on women.[55] There are a few possible approaches that the United States criminal legal system can take to address the gender disparities that result from statistical risk assessment tools: (1) states can stop using statistical risk assessment tools entirely; (2) tools that are used for both men and women can consider gender-responsive factors and implement different risk score levels; or (3) separate statistical risk assessment tools can be used for men and women.[56]

The first two approaches are not promising. Beginning with the first approach, while statistical risk assessment tools have a wide array of issues that are beyond the scope of this Comment, they are not likely to be transitioned out of the United States criminal legal system anytime soon. "Numerous states use [statistical risk assessment] tools [in] sentencing, and a few [states] even require judges to consider them when making decisions."[57] Further, in the 2017 proposed draft, the Model Penal Code endorsed the use of statistical risk assessment tools to identify low-risk individuals in sentencing.[58]

As for the second approach, instead of simply utilizing tools that include women-specific or gender-responsive factors, feminist criminological researchers advocate for tools that are designed for women and based on women-focused research for multiple reasons. First, when widely used statistical risk assessment tools originally based on men are revalidated to include women, the revalidation studies occur after the tools were created, previously validated on men, and already in use without gender-responsive factors.[59] By the time the revalidation studies take place, the tools have to be

---

[55] Wainstein, *supra* note 6, at 54.

[56] *Id.* at 52.

[57] Lightbourne, *supra* note 19, at 332.

[58] MODEL PENAL CODE: SENTENCING § 6B.09(3) (AM. LAW INST., Proposed Final Draft 2017); moreover, eliminating actuarial tools entirely would not necessarily fix the problem if clinical or lay judge assessments of female criminality are also based on gendered assumptions not borne out by the data. *See, e.g.*, United Nations: Gender in the Criminal Justice System https://www.unodc.org/unodc/en/justice-and-prison-reform/cpcj-gender.html [https://perma.cc/6XXB-G97K].

[59] Patricia Van Voorhis, *On Behalf of Women Offenders: Women's Place in the Science of Evidence-Based Practice*, 11 CRIMINOLOGY & PUB. POL'Y 111, 113 (2012) (citing *Women's Risk Factors*, *supra* note 34, and Manchak, Skeem, Douglas & Siranosian, supra note 33).

pulled from use, reworked, and revalidated to include gender-responsive factors that are most relevant to women.[60] Therefore, even after being revalidated, the tools do not target many of the issues that are most predictive of women recidivism.[61] If researchers started with women-specific data, it is likely that the resulting tool would look very different than the current tools and would be more predictively accurate for women.[62] Third, statistical risk assessment tests that simply "add gender and stir" by controlling for gender assume that the influence of key factors is uniform between men and women which can hide important differences, such as their definition of employment or the impact of abuse on their propensity for violence.[63]

Assuming that statistical risk assessment tools are not going to vanish from sentencing in the near future, a solution incorporating these tools must be developed. This Comment focuses on one potential solution: the creation of separate statistical risk assessment tools for men and women.

While risk assessment tools that are specifically designed for women are rare, some criminologists already advocate for their use.[64] Criminology literature and research suggest that in order to accurately take gender-specific factors and pathways into consideration, risk assessment tools that assess women should be based on and designed for women.[65] In such a situation, one tool would be used to assess the recidivism risk for a man at sentencing based on criminological data derived from research conducted on men, and a separate tool would be used to assess the recidivism risk for a woman at sentencing based on criminological data derived from research conducted on women.

III. ANTICIPATING OBJECTIONS TO HAVING SEPARATE TOOLS FOR MEN AND WOMEN

The proposition of using different statistical risk assessment tools for men and women in sentencing is likely to elicit backlash. Assuming the United States criminal legal system will continue to incorporate statistical risk assessment tools in sentencing, it is vital that alternative solutions are found that can survive potential criticisms. This Section addresses some of the most likely legal, theoretical, and policy issues that could arise, offers

---

[60] *Id.*

[61] *Id.* (citing *Women's Risk Factors*, *supra* note 34).

[62] *Id.* (citing BLOOM ET AL., *supra* note 35).

[63] Kristy Holtfreter & Rhonda Cupp, *Gender and Risk Assessment: The Empirical Status of the LSI-R for Women*, 23 J. CONTEMP. CRIM. JUST. 366, 369 (2007).

[64] Starr, *supra* note 19, at 825; BRENNAN & AUSTIN, *supra* note 26, at 1.

[65] Wainstein, *supra* note 5, at 37.

potential solutions to those issues, and identifies areas where more research and discussion is needed.

The first subsection explores an Equal Protection Clause argument related to using separate statistical risk assessment sentencing tools for men and women. The following subsection explores potential theoretical concerns with designing and using different tools along binary (men–women) gender lines and the conflation of sex and gender, and suggests areas for future research and discussion. The final subsection analyzes the possible objection that using separate risk assessment tools based on gender would open the floodgates to having endless tools based on various characteristics.

A. WOULD THE USE OF SEPARATE STATISTICAL RISK ASSESSMENT TOOLS FOR MEN AND WOMEN IN SENTENCING VIOLATE THE EQUAL PROTECTION CLAUSE?

In the United States, formal equality, or the anti-classification approach, dominates the current legal doctrine.[66] Formal equality emphasizes equal treatment or process and is based on the idea that similar people should be treated similarly.[67] The anti-classification approach requires that statistical risk assessment tools ignore protected characteristics, including gender, unless their use passes constitutional scrutiny for equal protection.[68] Although there are potentially valid arguments that statistical risk assessment tools violate the Equal Protection Clause because they consider gender, the Supreme Court has never explicitly barred sentencing judges from considering such factors.[69] Rather, courts have generally ignored the Equal Protection issues surrounding statistical risk assessment tools (e.g., finding that constitutional limits on substantive considerations would interfere with the judiciary's ability to impose a proper sentence and would be inconsistent with historical practice).[70] Despite the courts' failure to rule on these equal protection issues so far, there is a colorable, though ultimately surmountable, equal protection issue with using gender-specific risk assessment tools.

---

[66] Green, *supra* note 22, at 2; Starr, *supra* note 20, at 827.

[67] Green, *supra* note 22, at 2.

[68] SHARAD GOEL, RAVI SHROFF, JENNIFER L. SKEEM, & CHRISTOPHER SLOBOGIN, RESEARCH HANDBOOK ON BIG DATA LAW, *The Accuracy, Equity, and Jurisprudence of Criminal Risk Assessment* 8–9 (Roland Vogl ed., 2019); Foxwell, s*upra* note 5, at 466.

[69] Lightbourne, *supra* note 19, at 337 n.62 (citing *Dodakian v. United States*, No. 14-cv-01188 (AJN)(SN), 2015 WL 11144511, at *12 (S.D.N.Y. Aug. 14, 2015)).

[70] Carissa Byrne Hessick & F. Andrew Hessick, *Recognizing Constitutional Rights at Sentencing*, 99 CALIF. L. REV. 47, 74, 83 (2011).

In *United States v. Virginia*, the Supreme Court announced that gender is a quasi-suspect class and that equal protection challenges involving classifications based on gender are subject to intermediate scrutiny.[71] Intermediate scrutiny means that the justification for using gender-based classifications must be "exceedingly persuasive."[72] Additionally, a state "must show that the [challenged] classification serves 'important governmental objectives and that the discriminatory means employed' are 'substantially related to the achievement of those objectives.'"[73] Despite the obstacles imposed by *United States v. Virginia*, the literature thus far tends to assume that the use of gender in risk assessment tools would easily pass intermediate scrutiny.[74]

With no binding precedent or clear answer on whether the consideration of gender in risk assessment tools is constitutionally permissible, this subsection explores the equal protection issues that arise from having two separate tools for men and women.

1. Equal Protection Objection

Critics of gender-specific risk assessment tools argue that using demographic characteristics in algorithms constitutes overt discrimination.[75] This criticism is in line with the idea that, under the Equal Protection Clause, statistical risk assessment tools should not consider protected characteristics, including gender. However, if the use of gender in these tools passes intermediate scrutiny, then the use is permissible.

Sonja Starr, a legal scholar who focuses on the use of predictive algorithms and disparities in sentencing, argues that presumptively unconstitutional discrimination under the Equal Protection Clause cannot be justified by statistical generalizations about a group regardless of "very

---

[71] United States v. Virginia, 518 U.S. 515, 531 (1996) ("Parties who seek to defend gender-based government action must demonstrate an 'exceedingly persuasive justification' for that action."); Foxwell, s*upra* note 5, at 452.

[72] Virginia, 518 U.S. at 532–33.

[73] *Id.* at 533 (citing Mississippi Univ. for Women v. Hogan, 458 U.S. 718, 724 (1982)).

[74] Foxwell, s*upra* note 5, at 452–53. The following Equal Protection Clause analysis and commentary is specific to gender because it is a quasi-suspect class that is subject to intermediate scrutiny. For that reason, the same analysis cannot be transferred to an inherently suspect class, such as race, that is subject to strict scrutiny (the hardest level of scrutiny to pass). Brett Snider, *Challenging Laws: 3 Levels of Scrutiny Explained*, Find Law (Jan. 27, 2014) https://www.findlaw.com/legalblogs/law-and-life/challenging-laws-3-levels-of-scrutiny-explained [https://perma.cc/NHG6-TNG6].

[75] Crystal S. Yang & Will Dobbie, *Equal Protection Under Algorithms: A New Statistical and Legal Framework*, 119 MICH. L. REV. 291, 295 (2020).

strong empirical support."[76] She points out that these risk assessment tools are statistical averages and that the tools' proponents "defend [them] on the basis that the averages are [correct]."[77] Starr argues that the justification that women pose a substantially lower recidivism risk and are being punished for men's recidivism risk "embraces a concept of 'actuarial fairness' and stands on unsound constitutional footing."[78] Specifically, she believes that this argument, which relies on statistical averages, would not constitute an "exceedingly persuasive justification" for a gender classification under the test laid out in *United States v. Virginia* because it amounts to "'overbroad generalizations' about the tendencies of men and women."[79]

Addressing Starr's argument, one possible path to constitutionality is that the evidence supporting differentiation based on gender for statistical risk assessment tools at sentencing is "exceedingly persuasive."[80] Starr relies on *Craig v. Boren*, a case in which the Supreme Court considered a challenge to a law that subjected men to a higher drinking age for some alcoholic beverages than women and was defended by the state with statistical evidence.[81] The Court noted that "prior cases have consistently rejected the use of sex as a decision-making factor even though the statutes in question certainly rested on far more predictive empirical relationships than this."[82] Starr argues that this language extended the prohibition on inferring individual tendencies from group statistics to these "hypothesized" gender differences.[83] However, nowhere in the *Craig v. Boren* holding did the Supreme Court prohibit inferences based on group statistics.[84] Rather, the Court held that the statistical evidence that the state relied on was too weak

---

[76] Starr, *supra* note 19, at 821; *Sonja B. Starr*, U. CHI. (last visited Dec. 15, 2021), https://www.law.uchicago.edu/faculty/starr [https://perma.cc/B5C5-XPED].

[77] *Id.* at 828.

[78] *Id.* at 825.

[79] *Id.* (citing United States v. Virginia, 518 U.S. 515, 533–34 (1996)).

[80] *Id.*

[81] *Id.*; *Craig v. Boren*, 429 U.S. 190, 192, 201 (1976).

[82] *Id.* at 202.

[83] Starr, *supra* note 20, at 825–26 (quoting United States v. Virginia, 518 U.S. 515, 533 (1996) ("The justification must be genuine, not hypothesized or invented *post hoc* in response to litigation."); J.E.B. v. Alabama *ex rel.* T.B., 511 U.S. 127, 135 (1994) (stating that "this Court consistently has subjected gender-based classifications to heightened scrutiny in recognition of the real danger that government policies that professedly are based on reasonable considerations in fact may be reflective of 'archaic and overbroad' generalizations about gender, or based on 'outdated misconceptions concerning the role of females in the home rather than in the 'marketplace and world of ideas.'"").

[84] Foxwell, s*upra* note 5, at 455; Craig v. Boren, 429 U.S. 190, 208–10 (1976).

to pass intermediate scrutiny.[85] This holding does not clearly indicate that the Supreme Court believes there is "a total ban on statistical evidence as a means of justifying gender-based classifications."[86] Since the literature on statistical risk assessment tools and their impact on gender disparities is more persuasive than the handful of studies presented in *Craig*, such research should be persuasive enough to pass intermediate scrutiny and to allow distinct tools to be used for men and women in sentencing.

Starr's reliance on *Frontiero v. Richardson* to argue that assumptions cannot be made based on group tendencies is similarly misplaced.[87] In *Frontiero*, the Supreme Court held that a military spousal benefit policy could not be based on the idea that most wives are financially dependent on their husbands.[88] While the Court acknowledged that "efficacious administration of governmental programs"—the reason cited for differentiating based on gender in the policy—has "some importance," it ultimately decided that the policy violated due process.[89] Even so, the decision does not prohibit *all* generalizations based on group data like Starr suggests. Rather, the Court's focus on the governmental purpose being too trivial to pass intermediate scrutiny leaves open the possibility that more important governmental reasons could provide the necessary justification for using group data to differentiate based on gender.[90]

Overall, Starr's argument that justifications which rely on statistical averages would not constitute an "exceedingly persuasive justification" is inapposite here. It is not clear from legal precedent that use of gendered statistical risk assessment tools would be rejected under intermediate scrutiny simply because the justification is based on group statistical averages. The important question seems to be whether using different statistical risk assessment tools for men and women serves an "important governmental objective," and whether the means employed—which would be using different tools—is "substantially related to the achievement of those objectives."

---

[85] Craig v. Boren, 429 U.S. 190, 201–02, 208–10 (1976).

[86] Foxwell, s*upra* note 4, at 455.

[87] Starr, *supra* note 20, at 826.

[88] Frontiero v. Richardson, 411 U.S. 677, 690–91 (1973).

[89] *Id.* at 690.

[90] Foxwell, s*upra* note 5, at 457.

2. Governmental Objectives Served by Gender-Based Classification

a. Deploying Separate Tools for Men and Women Would Increase Accuracy

Increased accuracy in sentencing is one important government interest that could satisfy the intermediate scrutiny analysis for gendered statistical risk assessment tools. Recently, courts emphasized the importance of accuracy in constructing statistical risk assessment tools.[91] While the cases that emphasized the importance of accuracy did not explicitly consider equal protection challenges, the courts noted personal characteristics, including gender, may need to be considered in forming risk predictions because promoting accuracy is an important goal that serves both the state and criminal defendants.[92] For example, the Wisconsin court in *State v. Loomis* identified the promotion of accuracy as a legitimate state objective when it noted that "if the inclusion of gender promotes accuracy, it serves the interests of institutions and defendants, rather than a discriminatory purpose."[93] The *Loomis* court further noted that discriminating along gender lines is necessary for the achievement of this objective because "any risk assessment tool which fails to differentiate between men and women will misclassify both genders."[94] Because promoting accuracy can be a way of achieving an important government interest, the degree to which gender improves the predictive accuracy of risk assessment tools may determine whether discrimination based on gender in statistical risk assessment tools survives intermediate scrutiny.[95] Based on the *Loomis* court noting that promoting accuracy is an important government interest for the purposes of an intermediate scrutiny evaluation, as well as other courts' recent emphasis on the importance of accuracy in constructing statistical risk assessment tools,[96] it would be unjust to sentence individuals based on tools that are not accurate for women because they fail to adequately consider gender differences.

---

[91] Yang & Dobbie, *supra* note 75, at 320.

[92] *Id.*; *see* Malenchik v. State, 928 N.E.2d 564, 571 (Ind. 2010); *see also* State v. Loomis, 881 N.W.2d 749, 766 (Wis. 2016).

[93] State v. Loomis, 881 N.W.2d 749, 766 (Wis. 2016) (citing Melissa Hamilton, *Risk–Needs Assessment: Constitutional and Ethical Challenges*, 52 AM. CRIM. L. REV. 231, 255 (Spring 2015).

[94] *Loomis*, 881 N.W.2d at 766.

[95] Yang & Dobbie, *supra* note 75, at 320.

[96] *Id.*; *see* Malenchik v. State, 928 N.E.2d 564, 571 (Ind. 2010); *see also* State v. Loomis, 881 N.W.2d 749, 766 (Wis. 2016).

Overall, promoting accuracy is one important governmental function that could allow statistical risk assessment tools that differentiate based on gender to pass intermediate scrutiny. In her survey of statistical risk assessment tools in sentencing and their impact on disparities between men and women, Wainstein's extensive literature review revealed that many gender-neutral tools or tools based on men contain issues with validity, over-classification, and gender-responsiveness.[97] Additionally, Wainstein identified the development and use of tools that were created specifically for women as the most promising path toward minimizing disparities in sentencing if statistical risk assessment tools continue to be in use throughout the United States.[98] As a result, assessing women with a tool that was based on and designed for them could serve the important governmental interest of promoting accuracy.

b. Appropriate Sentencing Better Protects the Community

A second potential objective is protecting the community through appropriate sentencing. The Supreme Court has noted "[t]he 'legitimate and compelling state interest' in protecting the community from crime cannot be doubted" and that "crime prevention is 'a weighty social objective.'"[99] Even Starr recognized that protecting the community from crime is "no doubt an important interest" and "even a 'compelling' one."[100] However, she does note that the Constitution requires that the connection between the differentiation on the basis of gender for statistical risk assessment tools and protecting the community be assessed.[101]

Despite Starr's warning, the relationship between gender and crime is quite strong. Gender is considered one of the most understood and well-documented correlates of crime, and the fact that women commit violent acts at a much lower rate than men is a "staple in criminology."[102] Overall, women are less dangerous than men, commit fewer violent crimes, and are rarely repeat offenders; when they do reoffend, their crimes seldom escalate in

---

[97] Wainstein, *supra* note 6, at 61.

[98] *Id.* at 61–62.

[99] Schall v. Martin, 467 U.S. 253, 264 (1984) (first quoting De Veau v. Braisted, 363 U.S. 144, 155 (1960); and then quoting Brown v. Texas, 443 U.S. 47, 52 (1979)).

[100] Starr, *supra* note 20, at 842.

[101] *Id.*

[102] Foxwell, s*upra* note 5, at 462 (quoting John Monahan, *The Inclusion of Biological Risk Factors in Violence Risk Assessments*, *in* BIOPREDICTION, BIOMARKERS, AND BAD BEHAVIOR 57, 67 (Ilina Singh, Walter P. Sinnott-Armstrong & Julian Savulescu eds., 2013)).

severity.[103] A national assessment that conducted interviews with corrections administrators and representatives from the fifty states and the Federal Bureau of Prisons found that "51 percent of respondents reported that either women pose less risk than men, or a much smaller proportion of women than men pose serious risks to institutional and public safety."[104] In 2018, only 18.3% of violent incidents were committed by women, while 77.0% were committed by men, and 4.7% were committed by men and women.[105] In addition, men are more likely to use a weapon while committing a violent offense, and serious injuries are more frequently associated with men.[106] Some scholars concluded that gender in risk assessment tools would pass a constitutional challenge under the Equal Protection Clause given the strength of the data on gender and crime and the governmental interest in preventing crime.[107]

Given this data, sentences can be decreased with the use of women-specific sentencing tools without a commensurate increase in crime. Beyond that, studies show that the diversion of women from custodial to non-custodial sentencing and the shortening of women's sentences would likely *reduce* crime, furthering the goal of protecting the community. It is well documented that imprisonment can lead to higher rates of recidivism compared to non-custodial sentences.[108] By diverting low-risk women from custodial sentences or assigning shorter custodial sentences, women will have less exposure to the criminogenic environments of prison and jail, thus lowering their likelihood of reoffending. Moreover, diverting women from prison and releasing them earlier could have downstream effects on crime otherwise committed by the next generation. This is important because children who have an incarcerated parent face many hurdles—being more likely to exhibit low self-esteem, depression, emotional withdrawal from friends and family, and inappropriate or disruptive behavior at home and in

---

[103] Hannah-Moffatt, *supra* note 33, at 283–84; David Olson, Arthur J. Lurigio & Megan Alderden, *Men are from Mars, Women are from Venus, but What Role Does Gender Play in Probation Recidivism?*, 5 JUST. RSCH. & POL'Y 33, 48 (2003) (noting that women are less likely to be rearrested).

[104] CURRENT PRACTICES, *supra* note 27, at 14.

[105] Rachel E. Morgan & Barbara A. Oudekerk, *Criminal Victimization, 2018*, BUREAU OF JUSTICE STATISTICS 12–13 (Sept. 2019), https://bjs.ojp.gov/content/pub/pdf/cv18.pdf [https://perma.cc/VT6D-7DLM].

[106] Lawrence A. Greenfeld, *Women Offenders*, BUREAU OF JUSTICE STATISTICS 3 (Dec. 1999), https://bjs.ojp.gov/content/pub/pdf/wo.pdf [https://perma.cc/VR8L-ZEDJ].

[107] Foxwell, s*upra* note 5, at 466; Monahan, *supra* note 102, at 68.

[108] José Cid, *Is Imprisonment Criminogenic?: A Comparative Study of Recidivism Rates Between Prison and Suspended Prison Sanctions*, 6 EUR. J. CRIMINOLOGY 459, 472 (2009).

school—and, most saliently, possess an increased risk of future delinquency and criminal behavior.[109] Notably, the incarceration of a mother is especially disruptive for a child because they are most commonly the primary caregiver.[110] As a result, diverting and releasing women sooner from custodial settings will both prevent a heightened recidivism risk and decrease their children's future risk of delinquency and criminal behavior, all of which will decrease crime. Given this clear connection to community protection, the use of gender-specific risk assessment tools strongly support a compelling government interest under intermediate scrutiny.

3. Tailoring to Important Governmental Objectives

In addition to identifying important governmental objectives, the state must also show that the means employed are "substantially related to the achievement of those objectives" to pass intermediate scrutiny.[111] In other words, the methods used to reach the important governmental objectives need to significantly contribute to the realization of those goals.

Some scholars have argued that gender-specific assessment tools fail this prong of the analysis because there are alternative ways to address these goals without using different tools for men and women.[112] A quick analysis demonstrates, however, that the three most commonly cited alternatives either fail to advance the important government objectives or do not further the objectives as well as gender-specific tools.

As explained in Section II, the first alternative, presenting a judge with a statistical risk assessment tool that omits gender produces inaccurate risk scores and classification levels. These inaccuracies then result in inappropriate sentences for many women, a scenario that would be counter to the two governmental objectives discussed above.

The second alternative is using a single tool for men and women that incorporates different cutoff scores to tie custody levels to similar types of people.[113] For example, such a tool could raise the cutoff score for women if infraction rates for men in maximum custody are greater than that for women.[114] While arguments could be made that this method relates to the governmental objectives by helping with some over-classification issues, this

---

[109] DAVIS ET AL., *supra* note 52, at 117.

[110] *Id.* at 119.

[111] United States v. Virginia, 518 U.S. 515, 558 (1996).

[112] Lightbourne, *supra* note 19, at 340.

[113] CURRENT PRACTICES, *supra* note 27, at 45.

[114] *Id.*

approach would not address the concerns that the tools do not adequately consider gender-responsive factors and that the tools are not always validated on a sufficient population of women. Without addressing all three concerns, the approach will not adequately achieve the potential governmental objectives of promoting accuracy and protecting the community through appropriate sentences.

The third alternative is expanding the variables used in the tools to accommodate the nature of women's offending.[115] For example, employment variables could be expanded, as they have already been in some states, to avoid classifying people who are homemakers or stay-at-home parents as unemployed.[116] While some men are stay-at-home parents, it occurs at a greater frequency among women.[117] As a result, this approach does acknowledge the need for gender-responsivity in criminogenic factors. However, it is not clear that the resulting tools would have been the same if it had started with and was designed for women.[118] In other words, tools that were built for and based on women would likely look very different than the tools originally designed for men. Additionally, including gender-responsive factors does not mean that the resulting tool was validated on women nor that it is better than a tool that started with women, both of which would contribute to over-classification. These alternative approaches fail to adequately address over-classification, lack of validity, and lack of gender responsivity, the three main ways in which statistical risk assessment tools are not accurate predictors of recidivism for women in sentencing.

Overall, these three alternative approaches fail to adequately address the potential governmental objectives. In contrast, creating risk assessment tools that are specifically designed for women and adequately consider their gendered pathways into and out of crime is "substantially related to the achievement of those objectives." This approach will address these issues by validating the tools on women populations, producing risk scores that accurately match women's behavior, and focusing on factors that are important indicators of women's recidivism risk. By addressing these issues, the separate tools for men and women will advance the important governmental objectives of accuracy and protecting the community through appropriate sentencing.

---

[115] *Id.* at 46.

[116] *Id.*

[117] Gretchen Livingston, *Stay-at-Home Moms and Dads Account for About One-in-Five U.S. Parents* (2018), https://www.pewresearch.org/fact-tank/2018/09/24/stay-at-home-moms-and-dads-account-for-about-one-in-five-u-s-parents [https://perma.cc/86TH-KS5A].

[118] *Women's Risk Factors*, *supra* note 34, at 263.

B. ARE THERE THEORETICAL OBJECTIONS TO IDENTIFYING INDIVIDUALS IN SENTENCING BY GENDER?

1. The Conflation of Sex and Gender

Since the criminological field, as well as the criminal legal system, tend to conflate sex and gender, most of the studies and literature use the words female or woman and male or man interchangeably.[119]

In writing this Comment, I faced two options on how to proceed. The first option was to wait to analyze gender disparities in sentencing, and the extent to which a more targeted tool might be a corrective measure, until there was adequate literature and research in existence that took the differences in sex and gender into account. The second option was to work with existing research and literature while acknowledging its flaws and shortcomings. This Comment chose the latter path. While this Comment draws on works that treat gender and sex as the same thing, it provides a starting point for the discussion on gender disparities, as well as calls attention to the overarching issue in the field of criminology and the criminal legal system with the hope of encouraging more inclusive works and policies.

2. Gender as a Spectrum

Beyond the conflation of sex and gender, most criminological works and many aspects of the criminal legal system also operate on a male-female binary without consideration of gender identities.[120] As a result, the greater gender spectrum is not acknowledged, and some people are forced into labels with which they do not identify. This Comment discussed having different statistical risk assessment tools for men and women. This proposal raises the question of how individuals who are not cisgender, as well as those who are gender nonbinary, fit into this framework. For example, what tool would individuals use in sentencing if they are transgender or nonbinary? Ultimately, gender being a spectrum does not change the fact that there is a

---

[119] Sex is a label (male or female) that is assigned by a doctor based on genitals and chromosomes at birth; Gender is a social or legal status that corresponds with a set of expectations from society about behaviors, characteristics, and thoughts; and Gender Identity is how one feels on the inside and expresses their gender through clothing, behavior, and personal appearance. *Sex and Gender Identity*, PLANNED PARENTHOOD (Feb. 26, 2022), https://www.plannedparenthood.org/learn/gender-identity/sex-gender-identity [https://perma.cc/U6EB-NARZ].

[120] Jace L. Valcore & Rebecca Pfeffer, *Systemic Error: Measuring Gender in Criminological Research*, 31 CRIM. JUST. STUD. 333, 333 (2018).

disparity between two groups—those in the data labeled as men, and those in the data labeled as women. However, it is an important question.

While this Comment does not claim to have a perfect answer, it encourages future work on this topic. The answer to this question starts with understanding why those labeled or identified as men and women in the data are different when it comes to crime. If the specific "thing" that makes men and women different—whether a certain testosterone level or a certain level of socialization—is identified, then tool assignments could be determined by that factor. However, there is no question that those who identify as "women" have long been ignored in criminology and attention to the role of gender in offending has only intensified over the last few decades.[121]

Identifying the underlying causes for the different criminal and recidivism tendencies of those labeled "men" and those labeled "women" would be a first step toward an informed policy with respect to risk assessment of transgender and gender nonbinary people. This research burden is not insurmountable; experts in an array of fields have already theorized explanations for the differences in criminal activity between men and women.

First, social theorists examined the "process[] through which the motiv[ation] to commit crime develop[s]."[122] For example, some social theorists suggest that most street crime stems from a need for fast cash, which is fostered by the "excitement and sensory stimulation" that accompanies street culture.[123] It is argued that this stimulation disproportionately attracts those identified as young men, which helps to explain why they commit more crimes than women.[124] While the attraction of young men to street crime might relate either to socialization (dominance or a need for performative masculinity), to some innate physical or hormone differences, or some combination of these, the fact remains that those who have been identified as "men" have, compared to those identified as "women," a higher rate of committing such crimes.

Second, evolutionary theorists argue that the physical aggression witnessed in some men "stems from an evolutionary need for survival and reproduction."[125] This argument suggests that men, in greater numbers than

---

[121] Wainstein, *supra* note 6, at 43.

[122] Foxwell, s*upra* note 5, at 464.

[123] Richard Wright & Volkan Topalli, *Choosing Street Crime*, *in* OXFORD HANDBOOK OF CRIMINOLOGICAL THEORY 461, 463 (Francis T. Cullen & Pamela Wilcox eds., 2013).

[124] *Id.*

[125] Foxwell, *supra* note 5, at 464; *see also* DAVID M. BUSS, EVOLUTIONARY PSYCHOLOGY: THE NEW SCIENCE OF THE MIND 298–303 (4th ed. 2011).

women, need to be aggressive to compete for access to resources and the opportunity to reproduce.[126] This theory assumes that the more aggressive tendencies we see (disproportionately, on average) in men are linked to their assigned sex at birth.

Third, neuroscience theories point to testosterone as a hormone that is linked to aggression, anti-social behavior, and violent crime.[127] People who are assigned as male at birth have several times the amount of testosterone as those assigned as female, and anti-social behavior is about four times as prevalent in people assigned male as those assigned female.[128] While this theory also focuses on one's sex at birth, it also raises interesting questions. If a higher risk for certain criminal conduct is linked to testosterone, what happens when people take hormone replacement therapy? Do people who are assigned female at birth with higher-than-average testosterone levels have similar recidivism risks to people who are assigned male at birth with lower-than-average testosterone levels?

Fourth, feminist theorists posit that gender is a social construct and define it as "the deeply entrenched institutionalization of sexual difference" that impacts how an individual reacts to different situations.[129] For example, Susan Okin, a feminist philosopher, argues that the experience of an individual is defined by their sex and that gender-neutral terms obscure this fact since we live in a gender-structured society.[130] Specifically, Okin argues that our gendered system has its roots in the traditional sex roles of a family.[131] She believes that people who are put into the traditionally female and male roles have different life experiences that affect their psychologies, modes of thinking, and patterns of moral development.[132] As a result, the unique psychologies of individuals who are perceived as male and female are key to understanding differences between men and women in all types of human activity, including crime. If this is the case, how do we assign people to specific statistical risk assessment tools? Based on the role they were put into as a child? On the expectations forced upon them by family or society? If a transgender person transitioned thirty years ago or one year ago, are they

---

[126] Foxwell, s*upra* note 5, at 464–65 (citing BUSS, *supra* note 123, at 298–303).

[127] Melissa Peskin, Yu Gao, Andrea L. Glenn, Anna Rudo-Hutt, Yaling Yang & Adrian Raine, *Biology and Crime*, *in* OXFORD HANDBOOK OF CRIMINOLOGICAL THEORY 22, 30 (Francis T. Cullen & Pamela Wilcox eds., 2012).

[128] *Id.*

[129] SUSAN MOLLER OKIN, JUSTICE, GENDER, AND THE FAMILY 6, (1989) (emphasis omitted).

[130] *Id.* at 11.

[131] *Id.* at 101.

[132] *Id.* at 106.

more or less likely to have the qualitatively distinct psychologies of someone who was raised with traditional male or female roles?

To the extent that these questions are a concern, it is with the tools themselves rather than with the overarching inequalities between the treatment of men and women. This paper encourages further discussion on these critical questions while urging the adoption of separate risk classification systems for 'men' and 'women' within today's binary criminal legal system.

C. POLICY CONCERNS WITH HAVING DIFFERENT TOOLS BASED ON GENDER

Lastly, another possible objection is that allowing the distinction between men and women would open the floodgates to having different tools for every other potential characteristic, such as race. While this Comment is focused on a single disparity between two groups, and not with the other potential disparities caused by these tools, it is a fair and critical question.

The first thing to note is that the government in other areas has successfully instituted different standards between men and women, such as education, without being hindered by the "what about other characteristics?" critique. For example, Title IX guarantees that women and men are treated in a like manner in all educational programs and applies to all educational institutions receiving federal funding in the United States.[133] To ensure that men and women are treated fairly, Title IX compares the entire men's and women's athletic programs at schools and withdraws federal funds if schools are not in compliance with its regulations.[134] Title IX demonstrates that it is possible to have men and women participating in separate activities with the goal of ensuring that women are not disadvantaged without an unmanageable volume of regulations coming through the floodgates.

Second, some of the arguments that are put forth in favor of separate tools for men and women would not apply to other characteristics. For example, commentators who focus on the impact of these tools on gender often highlight that women have a lower recidivism rate than men.[135] However, while the recidivism rate is lower for women than men, it is higher

---

[133] Title IX (Education Amendments Act of 1972, 2018); *Title IX*, WOMEN'S SPORTS FOUNDATION, https://www.womenssportsfoundation.org/advocacy_category/title-ix [https://perma.cc/F3YE-X6CM] (last visited Dec. 15, 2021).

[134] *Id*.

[135] Carissa Byrne Hessick, *Race and Gender as Explicit Sentencing Factors*, 13 J. GENDER, RACE & JUST. 127, 137 (2010).

for Black and Latino people than for white people.[136] So, the impact of using the gender effect argument in other circumstances, such as race, could produce more harm to racial minorities.[137] As a result, this argument is not likely to be used by minority groups aiming to have different tools for individuals based on race.

Similarly, if a white individual attempted to use this logic to argue that a tool that does not differentiate by race is unfair, they would have a much more difficult time making it through the equal protection analysis. First, the algorithm at most would have a disparate impact on the person; the algorithm itself does not treat white subjects differently. Thus, equal protection would have little to say about such an algorithm. If the question were instead whether a race-conscious algorithm would pass muster under the Equal Protection Clause, a litigant hoping for such an algorithm would have a difficult legal battle. That is because race is a suspect class subject to strict scrutiny under the Equal Protection Clause.[138] To pass strict scrutiny, a defendant must identify a government purpose compelling enough to discriminate along racial lines.[139] But the types of government interests that are already identified as compelling enough to pass strict scrutiny, such as national security,[140] diversity in education,[141] and remedying a history of past state discrimination,[142] would not support separate risk assessment tools for white individuals. Moreover, even after a compelling government interest that satisfies strict scrutiny is identified, the advocate must argue that the discrimination on the basis of race is *necessary* to accomplish the objective.[143] Here, it would be difficult to argue that there is no alternative. Rather, the fact is that differentiating by race in statistical risk assessment tools would exacerbate past state discrimination, harm minority groups, and potentially fail to produce a substantial change in leniency for the white individual in sentencing since white men are historically the norm in theory

---

[136] Id. at 129.

[137] *Id.*

[138] Regents of U. of Cal. v. Bakke, 438 U.S. 265, 291 (1978).

[139] Palmore v. Sidoti, 466 U.S. 429, 432–33 (1984).

[140] *See* Korematsu v. United States, 323 U.S. 214, 217–219 (1944); *see also* Hirabayashi v. United States, 320 U.S. 81, 100–02 (1943). While national security was identified as a government interest that is compelling enough to pass strict scrutiny, the example cases represent a racist history of Japanese American exclusion, incarceration, and internment. Korematsu, 323 U.S. at 217–19; Hirabayashi, 320 U.S. at 100–02.

[141] Grutter v. Bollinger, 539 U.S. 306, 308 (2003) (identifying student body diversity as a compelling state interest).

[142] Parents Involved in Cmty. Sch.s v. Seattle Sch. Dist. No. 1, 551 U.S. 701, 720 (2007) (citing Freeman v. Pitts, 503 U.S. 467, 494 (1992)).

[143] *Palmore*, 466 U.S. at 432–33.

and data. Overall, repurposing the gender-effect argument to race would not likely pass a strict scrutiny analysis.

Additionally, the proposition put forth in this paper would likely not result in a substantial increase in sentences or harm to men who are already being assessed by tools that are based on men's theories of crime and data. Sentencing is not a zero-sum game. Rather, there are neutral instruments that have a quantifiably higher error rate for women, which is resulting in women being put away for longer periods of time even though they offend at lower rates. By creating a separate tool for women, women would receive more-accurately tailored sentencing and men would not be substantially impacted. Rather, men will continue to be assessed by the same tools that have already been designed for them and based on their gender-specific factors and pathways into crime. Overall, the proposition put forth in this Comment would likely avoid opening the floodgates to having different tools for every other potential characteristic.

CONCLUSION

While statistical risk assessment tools were introduced as an unbiased method to address major issues in the United States criminal legal system, it is clear that current tools often contribute to inaccurate and inequitable sentences for women. Developing tools that are specifically created for women is one promising path forward to contribute to sentencing that is more-appropriately tailored to lowering recidivism. However, this approach requires men and women to be treated differently as they are assessed by different tools.

From a legal perspective, evidence supports that increasing accuracy and protecting the community could provide the necessary justification for gender-specific risk assessment tools. Further, the creation of tools that are specifically designed for women and adequately consider their gendered pathways into and out of crime is substantially related to the achievement of those objectives. Therefore, statistical risk assessment tools can ultimately surmount an equal protection challenge.

From an inclusion perspective, most of the criminological field and criminal legal system conflates sex and gender, and treats gender as a binary. By acknowledging these realities, this Comment hopes to shed light on some of the issues that need to be addressed as we work towards a more equitable and inclusive system. While these concerns are important, they should not derail attempts to address inequitable treatment. Rather, separate risk assessment tools for 'men' and 'women' should still be created and implemented to begin addressing issues that are currently harming members of our society.

From a policy perspective, it seems unlikely that having different tools for men and women would open the floodgates to have tools for every possible characteristic. Rather, gender-specific tools are a unique solution that causes no substantial harm to any of the groups involved and helps alleviate the disparities that women have suffered.