



National Institute of Justice

Stakeholder Statements Submitted in Response to the Department of Justice's Continued Review and Refinement of PATTERN

September 2022



Office for
Access to Justice
U.S. Department of Justice

NIJ | *National Institute
of Justice*

STRENGTHEN SCIENCE. ADVANCE JUSTICE.

U.S. Department of Justice
Office of Justice Programs
810 Seventh St. N.W.
Washington, DC 20531

Nancy La Vigne, Ph.D.

Director, National Institute of Justice

This and other publications and products of the National Institute of Justice can be found at:

National Institute of Justice

Strengthen Science • Advance Justice

NIJ.ojp.gov

Office of Justice Programs

Building Solutions • Supporting Communities • Advancing Justice

OJP.gov

The National Institute of Justice is the research, development, and evaluation agency of the U.S. Department of Justice. NIJ's mission is to advance scientific research, development, and evaluation to enhance the administration of justice and public safety.

The National Institute of Justice is a component of the Office of Justice Programs, which also includes the Bureau of Justice Assistance; the Bureau of Justice Statistics; the Office for Victims of Crime; the Office of Juvenile Justice and Delinquency Prevention; and the Office of Sex Offender Sentencing, Monitoring, Apprehending, Registering, and Tracking.

Opinions or conclusions expressed in this paper are those of the authors and do not necessarily reflect the official position or policies of the U.S. Department of Justice.



U.S. Department of Justice

Office of Justice Programs

National Institute of Justice

Office of the Director

Washington, D.C. 20531

Dear Colleagues,

On behalf of the U.S. Department of Justice, the National Institute of Justice (NIJ) and the Office for Access to Justice (ATJ) present *Stakeholder Statements Submitted in Response to the Department of Justice's Continued Review and Refinement of PATTERN*. The statements were offered in response to two stakeholder engagement sessions hosted collaboratively by NIJ and ATJ on Wednesday, September 28, and Thursday, September 29, 2022.

During the sessions, interested members of the public and criminal justice stakeholders provided comments and recommendations about further refinements to the Bureau of Prisons' risk assessment tool, the Prisoner Assessment Tool Targeting Estimated Risk and Needs (PATTERN). The comments and recommendations follow.

These sessions were convened as part of the Department's commitment to continuing the faithful implementation of the First Step Act (FSA), making ongoing improvements to PATTERN, and ensuring transparency in the research conducted in support of FSA implementation. NIJ remains dedicated to supporting the Department as this important work continues.

Sincerely,

A handwritten signature in black ink, appearing to read "Nancy La Vigne". The signature is fluid and cursive, with a long horizontal line extending to the right.

Nancy La Vigne, Ph.D.
Director
National Institute of Justice

Session 1
September 28, 2022

1. [American Civil Liberties Union](#) – Marissa Gerchick, Brandon Buskey
2. [Justice Action Network](#) – Cortland Broyles
3. [RTI International](#) – Matthew DeMichele
4. [The Sentencing Project](#) – Liz Komar

Session 2
September 29, 2022

5. [National Association of Criminal Defense Lawyers](#) – James Felman, Mary Price, Elizabeth Blackwood, Patricia Richman, Cortland Broyles
6. [Brennan Center for Justice](#) – Ames Grawert, Patricia Richman
7. [Council of State Governments Justice Center](#) – Sarah Desmarais, David D'Amora, Lahiz Tavárez
8. [Federal Public and Community Defenders](#) – Patricia Richman
9. [Melissa Hamilton, JD, Ph.D](#)



Formal Statement of the American Civil Liberties Union

For a Stakeholder Engagement Session on

First Step Act Implementation

Submitted to the Department of Justice's Office for Access to Justice

September 27, 2022

Authors

Marissa Gerchick

Data Scientist II, Algorithmic Justice Specialist, ACLU

Brandon Buskey

Director, ACLU Criminal Law Reform Project

On behalf of the American Civil Liberties Union (“ACLU”) and its more than 1.5 million members, supporters, and activists, and 53 nationwide affiliates, we thank you for the invitation to participate in this stakeholder engagement session regarding the PATTERN risk assessment built as a result of the First Step Act of 2018. The American Civil Liberties Union (ACLU) is dedicated to the principles of liberty, justice, and equality embodied in our nation’s Constitution and civil rights laws, and to protecting the civil liberties of all people in the United States. For years, the ACLU has been at the forefront of the fight against mass incarceration and its devastating impact on the people and their families who become ensnared in the criminal legal system, the failure to increase a proportional increase in public safety, and its disproportionate effect on communities of color.

The ACLU has serious concerns about PATTERN. We have repeatedly expressed these concerns over the last several years, along with other civil rights and legal organizations.¹ We believe PATTERN 1.3² continues to suffer from many of the systemic issues that affected previous versions of the tool. More than three years since PATTERN was first developed, the repeated failure to address these issues is unacceptable for many reasons, but most importantly because it unfairly prevents incarcerated people from benefitting fully from the earned time credit program. These continued problems underscore the inherent difficulties with using risk assessments to make consequential decisions about human liberty.

I. “Layers of Bias” in the PATTERN Risk Assessment Tool

The Department has asked for feedback on three important issues regarding the design and use of PATTERN. These topics are interrelated; thus, we apply a variant of the “layers of bias” framework for risk assessments,

¹ See, e.g., *Comment Letter to Department of Justice on PATTERN First Step Act*, available at <https://civilrights.org/resource/comment-letter-to-department-of-justice-on-pattern-first-step-act/> (Sept. 3, 2019); *Coalition Letter on the Use of PATTERN Risk Assessment in Prioritizing Release in Response to the COVID-19 Pandemic*, available at <https://www.aclu.org/letter/coalition-letter-use-pattern-risk-assessment-prioritizing-release-response-covid-19-pandemic> (April 3, 2020).

² See U.S. Department of Justice, *2021 Review and Revalidation of the First Step Act Risk Assessment Tool*, available at <https://nij.ojp.gov/library/publications/2021-review-and-revalidation-first-step-act-risk-assessment-tool> (December 2021).

where the concerns raised in one layer build on and exacerbate the concerns raised in previous layers.³

- In the first layer, we highlight continuing concerns about errors in the tool and the definition of recidivism that PATTERN employs.
- These problems compound issues in the next layer of bias, which is focused on concerns with the inputs used in PATTERN and PATTERN's scoring scheme.
- Finally, in the last layer, we discuss concerns related to algorithmic fairness. We emphasize that debates about whether PATTERN is “fair” cannot be resolved until the Department also addresses the concerns raised in the previous points.

II. First Layer of Bias: Errors with PATTERN and PATTERN's Recidivism Definition

In the first layer of analyzing a risk assessment like PATTERN, it is critical to understand whether the tool actually works as designed. Since it was developed, PATTERN has been plagued by repeated and serious errors with both the technical systems and human processes used to calculate risk scores.⁴ These errors have led to the misclassification of tens of thousands of people who are incarcerated – and those are just the errors we know about.⁵ As people who are incarcerated continue to report issues with the awarding of Earned Time Credits (ETCs) and release timelines,⁶ and as reports released by the

³ The “layers of bias” framework we use here to contextualize issues with PATTERN is a variation of the framing of Eckhouse et al. (2018) regarding layers of bias in quantitative risk assessments. See Eckhouse et al. (2018), *Layers of Bias: A Unified Approach for Understanding Problems with Risk Assessment*, available at <https://journals.sagepub.com/doi/abs/10.1177/0093854818811379>.

⁴ For a summary of the technical errors and human errors associated with PATTERN, see Grawert and Richman, *The First Step Act's Prison Reforms, Uneven Implementation and the Path Forward*, The Brennan Center for Justice, available at <https://www.brennancenter.org/our-work/research-reports/first-step-acts-prison-reforms> (Sept. 23, 2022).

⁵ See Carrie Johnson, “Flaws plague a tool meant to help low-risk federal prisoners win early release,” NPR, available at <https://www.npr.org/2022/01/26/1075509175/justice-department-algorithm-first-step-act> (Jan. 26, 2022).

⁶ See Erik Ortiz, “Thousands of Federal Inmates still await early release under Trump-era First Step Act,” NBC News, available at

Department every few months about PATTERN continue to identify new errors,⁷ we would urge the Department to seriously consider whether this tool works currently and how it and the public can be confident the tool will work in the future.

In this layer of bias, the Department must also evaluate the appropriate definition of recidivism it wants the tool to predict and whether the tool can actually predict that outcome.⁸ As we have stated in the past, PATTERN does not directly predict recidivism – PATTERN estimates the likelihood of rearrest or return to BOP custody following release.⁹ That distinction is critical. Overwhelming research has demonstrated that arrest is more reliably a measure of policing practices and priorities than actual crime, making arrest a racially-biased proxy for recidivism.¹⁰ For example, when it comes to traffic stops – the most common form of interaction between police and the public – study after study has demonstrated that police engage in persistent racial discrimination when making stops, frisks, searches and arrests.¹¹ Moreover, a

<https://www.nbcnews.com/news/us-news/thousands-federal-inmates-still-await-early-release-trump-era-first-st-rcna35162> (July 3, 2022).

⁷ See U.S. Department of Justice, *2021 Review and Revalidation of the First Step Act Risk Assessment Tool*, available at <https://nij.ojp.gov/library/publications/2021-review-and-revalidation-first-step-act-risk-assessment-tool> (December 2021).

⁸ PATTERN includes four separate tools. Two of these tools focus on “general recidivism,” which is defined as “as a return to BOP custody or a rearrest within three years of release from BOP custody, excluding all traffic offenses except driving under the influence and driving while intoxicated.” The other two tools focus on “violent recidivism,” defined as “a rearrest for a suspected act of violence within three years of release from BOP custody.” See pg. 10, U.S. Department of Justice, *2021 Review and Revalidation of the First Step Act Risk Assessment Tool*, available at <https://nij.ojp.gov/library/publications/2021-review-and-revalidation-first-step-act-risk-assessment-tool> (December 2021).

⁹ U.S. Department of Justice, *2021 Review and Revalidation of the First Step Act Risk Assessment Tool*, available at <https://nij.ojp.gov/library/publications/2021-review-and-revalidation-first-step-act-risk-assessment-tool> (December 2021).

¹⁰ See, e.g., American Civil Liberties Union, *A Tale of Two Countries: Racially Targeted Arrests in the Era of Marijuana Reform*, available at <https://www.aclu.org/report/tale-two-countries-racially-targeted-arrests-era-marijuana-reform> (April 2020); Lum & Isaac (2018), *To Predict and Serve*, available at <https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1740-9713.2016.00960.x>.

¹¹ See Baumgartner et al., (2016), *Targeting young men of color for search and arrest during traffic stops: evidence from North Carolina, 2002-2013*, available at <https://fbaum.unc.edu/articles/PGI-2016-Targeting.pdf>; Pierson et al., (2020), *A*

large percentage of arrests do not result in convictions.¹² Though focusing on arrests for suspected acts of violence is sometimes portrayed as a less biased or more accurate proxy, recent research has also called into question the reliance on arrests as a proxy for violent crime in risk assessment instruments.¹³ Taken together, this evidence suggests multiple, fundamental issues with using arrests as a proxy for recidivism.

These problems are compounded for people on post-release supervision, who are aggressively surveilled and often re-incarcerated for minor or technical violations of their conditions of release.¹⁴ In addition, judges are mandated to impose supervised release on the overwhelming majority of convicted individuals.¹⁵ This creates a severe risk that many on supervised release are “overconditioned.” Research consistently shows that placing more intrusive conditions on individuals than necessary increases their likelihood of recidivism.¹⁶

large-scale analysis of racial disparities in police stops across the United States, available at <https://www.nature.com/articles/s41562-020-0858-1>.

¹² For a discussion of the data from various jurisdictions about what percentage of arrests result in convictions, see Ames Grawert, *Brennan Center’s Public Comment on the First Step Act’s Risk and Needs Assessment Tool*, available at <https://www.brennancenter.org/our-work/research-reports/brennan-centers-public-comment-first-step-acts-risk-and-needs-assessment> (Sept. 3, 2019).

¹³ See, e.g., Fogliato et al. (2021), *On the validity of arrest as a proxy for offense: Race and the likelihood of arrest for violent crimes*. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 100-111), available at <https://dl.acm.org/doi/abs/10.1145/3461702.3462538>.

¹⁴ See, e.g., Jacob Schuman, *America’s Shadow Criminal Justice System*, The New Republic (2018), <https://newrepublic.com/article/148592/americas-shadow-criminal-justice-system>, providing an overview of supervised release system and how it often leads to revocation and re-incarceration, including for minor conduct and absent due-process protections.

¹⁵ U.S. Sentencing Comm’n, *Overview of Federal Criminal Cases: Fiscal Year 2019 at 10* (Apr. 2020), [https://www.ussc.gov/sites/default/files/pdf/research-and-publications/research-publications/2020/FY19 Overview Federal Criminal Cases.pdf](https://www.ussc.gov/sites/default/files/pdf/research-and-publications/research-publications/2020/FY19%20Overview%20Federal%20Criminal%20Cases.pdf).

¹⁶ See, e.g., Jennifer L. Doleac, “Study after Study Shows Ex-Prisoners Would be Better Off Without Intense Supervision,” Brookings Institution, available at <https://www.brookings.edu/blog/up-front/2018/07/02/study-after-study-shows-ex-prisoners-would-be-better-off-without-intense-supervision/> July 2018, for a literature review finding that more intense supervision for those on probation or parole results in worse outcomes.

In short, PATTERN’s definitions of recidivism more accurately predict systemic failures in our racially-biased systems of policing and post-release supervision than the risks posed by those seeking relief under the First Step Act. One critical but partial solution—repeatedly suggested by the ACLU and many other advocates and experts¹⁷—would be to limit PATTERN’s focus to convictions for new crimes, rather than arrests or returns to BOP custody for technical violations.

The Department’s stated reasoning in 2020 for refusing to adjust the definitions of recidivism in PATTERN fail to address these concerns. The Department has cited a lack of available data as one reason it cannot narrow PATTERN’s definition of recidivism to only focus on convictions.¹⁸ It also cited the fact that other federal agencies use a similar definition of recidivism, thus allowing for better comparisons of recidivism data.¹⁹ These are inadequate justifications for using a tool that is known to unfairly bar individuals from leaving prison early. We urge the Department to continue exploring ways to address data limitations. We also highlight that the Department should have access to data about *why* an individual is returned to BOP custody, and could limit the definition of general recidivism to exclude returns to custody for technical violations and minor offenses. If it has not already, the Department should immediately analyze this potential change and publicly share the results.

¹⁷ See, e.g., *Comment Letter to Department of Justice on PATTERN First Step Act*, available at <https://civilrights.org/resource/comment-letter-to-department-of-justice-on-pattern-first-step-act/> (Sept. 3, 2019), Ames Grawert, *Brennan Center’s Public Comment on the First Step Act’s Risk and Needs Assessment Tool*, available at <https://www.brennancenter.org/our-work/research-reports/brennan-centers-public-comment-first-step-acts-risk-and-needs-assessment> (Sept. 3, 2019).

¹⁸ See pg. 12-14, U.S. Department of Justice, *The First Step Act of 2018: Risks and Needs Assessment System – UPDATE*, available at <https://www.bop.gov/inmates/fsa/docs/the-first-step-act-of-2018-risk-and-needs-assessment-system-updated.pdf> (Jan. 2020).

¹⁹ See pg. 12-14, U.S. Department of Justice, *The First Step Act of 2018: Risks and Needs Assessment System – UPDATE*, available at <https://www.bop.gov/inmates/fsa/docs/the-first-step-act-of-2018-risk-and-needs-assessment-system-updated.pdf> (Jan. 2020).

III. Second Layer of Bias: Inputs to PATTERN and PATTERN's Scoring Scheme

In the second “layer” of bias, the inputs to PATTERN and the way the tool is scored add to the problems created by the previous layer. Here, the Department must recognize that every decision made in PATTERN's design process is a policy choice laden with values judgments, and must be analyzed as such.²⁰ It is a policy choice to set cut points for risk level categories.²¹ It is a policy choice with dangerous ramifications to use biased data – fueled by the systemic under-investment in and over-policing of marginalized communities – as inputs to the tool.²² It is even a policy choice to decide how and where to slice up specific data used as inputs to the tool.²³

The policy choices built in to PATTERN pose a high risk of discouraging incarcerated people from seeking rehabilitative programming, in direct opposition to the goals of the First Step Act. In its recent reports, the Department has highlighted the presence of dynamic variables as inputs to PATTERN.²⁴ Yet, despite the inclusion of some dynamic inputs, PATTERN is still largely dominated by static factors. For one stark example, *it appears to be mathematically impossible for men under the age of 26 to qualify as “minimum” risk – the lowest possible risk category – on both the general and*

²⁰ For further discussion of how PATTERN encodes policy choices, see Grawert and Richman, *The First Step Act's Prison Reforms, Uneven Implementation and the Path Forward*, The Brennan Center for Justice, available at <https://www.brennancenter.org/our-work/research-reports/first-step-acts-prison-reforms> (Sept. 23, 2022).

²¹ See pg. 12-13, U.S. Department of Justice, *First Step Act Annual Report, April 2022*, available at <https://www.ojp.gov/first-step-act-annual-report-april-2022> for information about the risk level categories and their cut points.

²² Information about the data used for PATTERN's input variables is described in <https://nij.ojp.gov/library/publications/2021-review-and-revalidation-first-step-act-risk-assessment-tool>.

²³ For a discussion of how methods for measuring data are not always value-neutral processes, see Harper et al. (2010), *Implicit Value Judgements in the Measurement of Health Inequalities*, available at <https://onlinelibrary.wiley.com/doi/10.1111/j.1468-0009.2010.00587.x>

²⁴ See, e.g., U.S. Department of Justice, *2021 Review and Revalidation of the First Step Act Risk Assessment Tool*, available at <https://nij.ojp.gov/library/publications/2021-review-and-revalidation-first-step-act-risk-assessment-tool> (December 2021).

*violent recidivism scales of PATTERN 1.3.*²⁵ Imagine a 25-year-old male, who scores the lowest number of points possible on every aspect of each recidivism tool, including every dynamic input to the tool. Even if this person had completed 10 programs, or 100 programs, or 1000 programs, they will still not be considered “minimum risk” by PATTERN – on the basis of their age alone.

The Department must continue to expand the pool of individuals who qualify as minimum or low risk, both by adjusting the risk level cut points and by placing greater relative emphasis on dynamic inputs that lower a risk classification. These *policy choices* would help shape PATTERN into a tool that is more consistent with the First Step Act’s goals of incentivizing rehabilitation, increasing release from BOP custody, and protecting public safety.

IV. Third Layer of Bias: Algorithmic Fairness Implications of PATTERN

Finally, in the last “layer” of bias, we address the Department’s request for information on improving the “fairness” of PATTERN. Even if the Department resolves all of the concerns we have already highlighted, there would still be fundamental issues with trying to make PATTERN “fair.”

To illustrate, there are several issues with the methodological choices the Department has made in measuring whether the tool is “fair,” as well as choices made in evaluating the “predictive validity” and the “racial and ethnic neutrality” of the tool.²⁶ For instance, grouping by race and computing the Area

²⁵ See U.S. Department of Justice, *2021 Review and Revalidation of the First Step Act Risk Assessment Tool*, available at <https://nij.ojp.gov/library/publications/2021-review-and-revalidation-first-step-act-risk-assessment-tool> (December 2021); also see U.S. Department of Justice, *First Step Annual Act Report, April 2022*, available at <https://www.ojp.gov/first-step-act-annual-report-april-2022>, for the most current cut points for the general recidivism tool. Using Tables 2.3 and 2.4 of the December 2021 report and Table 2a of the April 2022 report, we see that a male under 26 would receive 35 points based on their age alone under the general recidivism tool. If they received the minimum possible number of points on every other input, including the programs inputs, the lowest possible score they could receive is 13 – too high to be considered “minimum risk” on this scale of PATTERN. A similar logic follows with respect to the violent recidivism scale, where the lowest possible score a male under 26 could receive is 9, and scores of 7 or lower are required to be considered “minimum risk.”

²⁶ “Predictive validity” and “racial neutrality” are terms used throughout the Department’s reports about PATTERN, and relate to requirements of the First Step Act of 2018. See, e.g., U.S. Department of Justice, *2021 Review and Revalidation of*

Under the Curve (AUC) metric as the Department does presents an incomplete and potentially misleading picture of racial disparities.²⁷ In addition, the Department has measured and published separate results about the general recidivism and the violent recidivism tools – but in reality, risk level categories from the violent recidivism tool override those of the general recidivism tool. The tool’s evaluation by the Department should mirror its deployment in this regard.

More fundamentally, the Department should evaluate ways to immediately and equitably expand access to earned time credit. For example, the Department could consider classifying individuals as minimum or low risk based on their violent recidivism score, and only excluding individuals the general recidivism tool classifies as high risk. If properly restricted to convictions, as discussed above, this approach would use violent recidivism—the weightier public safety concern—as the primary risk consideration, rather than as an override.

V. Conclusion: The Department must act immediately to address these concerns.

Together, these layers of bias within PATTERN build upon and exacerbate each other. They must all be addressed to ensure the mandates of the First Step Act are realized. We urge the Department to act immediately to resolve these issues.

* * * * *

We thank the Office of Access to Justice for considering and including our comments on this critical subject. If you have any question, please contact ACLU Data Scientist Marissa Gerchick at mgerchick@aclu.org, or CLRP Director Brandon Buskey at bbuskey@aclu.org.

the First Step Act Risk Assessment Tool, available at <https://nij.ojp.gov/library/publications/2021-review-and-revalidation-first-step-act-risk-assessment-tool> (December 2021).

²⁷ For further discussion of this issue, see, e.g., Kallus & Zhou (2019), *The Fairness of Risk Scores Beyond Classification: Bipartite Ranking and the xAUC Metric*, available at <https://proceedings.neurips.cc/paper/2019/file/73e0f7487b8e5297182c5a711d20bf26-Paper.pdf>.

September 28, 2022

Nancy La Vigne
Director
National Institute of Justice
U.S. Department of Justice

Re: First Step Act Implementation Stakeholder Engagement Session

Dear Director La Vigne:

Thank you for convening this stakeholder engagement session on First Step Act implementation.

The Justice Action Network (JAN) is the country's largest bipartisan organization dedicated to advancing criminal justice reform at the federal and state levels. JAN brings together progressive and conservative partner groups, and collaborates with law enforcement, business leaders, and civil rights, victims' rights, and faith-based groups to advance reform. Our coalition works to pass life-changing legislation in Washington and state houses to make our criminal justice system fairer and more effective.

JAN and our partners worked together to pass the First Step Act in 2018 and have since been deeply committed to the swift, successful, and equitable implementation of the law's reforms.¹ Dating back to 2019, when the PATTERN tool was first unveiled, JAN provided feedback in a written comment to Attorney General Bill Barr, Bureau of Prisons Director Kathleen Hawk Sawyer, and National Institute of Justice Director David Muhlhausen on how the tool could be improved to be more accurate and equitable.² Since then, we have continued to work alongside groups including Arnold Ventures, FAMM, Federal Public and Community Defenders, National Association of Criminal Defense Lawyers, and other organizations on this issue.³

Ensuring that the law's second chance opportunities for accruing earned time credits for individuals who have successfully completed evidence-based recidivism reduction (EBRR) programming and other productive activities (PAs) are applied broadly is of paramount importance. In reviewing the Department's annual report published in April, JAN has three general non-exhaustive points of feedback. Specifically, we encourage the Justice Department to: (1) increase the number of dynamic factors that may help individuals reduce their PATTERN scores; (2) refine the cut points for risk level categories (RLCs) to allow the maximum number of individuals to qualify for earned time credits (TCs) while continuing to balance public safety requirements; and (3) redefine recidivism to exclude arrests and solely focus on convictions. Finally, we conclude this letter with remarks pertaining to the needs assessment that is outlined in the First Step Act, understanding that the purpose of today's session is to focus on the risk assessment tool.

First, PATTERN includes a number of static and dynamic risk factors that contribute to an individual's overall risk assessment score. While there are opportunities for individuals to reduce their general and violent risk assessment scores, mainly through participation in programs, treatment, and educational attainment, more factors should be identified and included in the tool.

One potential consideration could be to implement a decay factor in criminal history scores, making criminal history more dynamic. Under the current PATTERN tool, criminal history scores are static and do not change once submitted at the time of admission to prison. However, it is generally accepted that individuals age out of criminal behavior over time; recidivism rates substantially decrease by the age of 40.⁴

Second, we appreciate the effort made by the Department to adjust the cut points for general recidivism level categories, outlined in the April report. The readjustment should allow more individuals to qualify for TCs who may have previously been deemed ineligible. However, the violent recidivism score was not adjusted, which remains a concern. Considering that the those deemed to be minimum risk and low risk have incredibly low recidivism rates of 1.4 percent and 9.2 percent, respectively, and further considering that the violent recidivism score is used as an override tool over the general recidivism scores, JAN submits that the cut points needlessly limit the number of individuals eligible for TCs. The Department should work to adjust the cut points to allow more individuals to be eligible for TCs while continuing to account for public safety requirements.

Third, JAN welcomes the conversation on the definition of recidivism and would advocate for the definition to be limited to new convictions and exclude arrests. Basing the definition on re-arrests may be more expedient; it does, however, make the tool prone to exacerbating disparate racial impacts. There is evidence that policing practices often lead to more arrests of people of color than of their white counterparts.⁵ This in turn could lead to more racial disparities within the PATTERN tool and should therefore be addressed without delay.

Finally, recognizing that the Department has requested to specifically address the risk assessment tool for today's meeting, JAN has concerns about the needs assessment as stipulated by the First Step Act. We understand that the Bureau launched the Standardized Prisoner Assessment for Reduction in Criminality (SPARC-13) in January 2020, and we appreciate the level of detail included in this year's annual report. However, this needs assessment system has not been independently evaluated, and as the report outlines, work has only just begun this fall on a BOP partnership with NIJ to establish an external review committee. This independent evaluation is long overdue, and delay is particularly worrisome as individuals in federal custody have been prescribed treatment and programming under SPARC-13 for almost full two years.

In closing, we appreciate the Department for hosting today's stakeholder engagement session and for committing to improve the PATTERN tool to ensure it is accurate and equitable. The First Step Act has already improved the lives of over ten thousand people, and it is our hope that a renewed commitment from the Department will bring relief and second chances to thousands more deserving individuals. We urge that the recommendations and concerns outlined in this statement and as identified by our partners will be seriously considered and implemented in further iterations of PATTERN.

Sincerely,

Cortland Broyles
Federal Affairs Manager

¹ FIRST STEP ACT OF 2018, PUB. L. NO. 115-391.

² Justice Action Network public comment, *Comments in Response to the Department of Justice Report “The First Step Act of 2018 Risk and Needs Assessment System.”* September 3, 2019. <https://static1.squarespace.com/static/5d128ac9071d0b00017ea5e3/t/61e6ca836a1dc93c5bfbbac0/1642515078416/Justice+Action+Network+RNAS+Comments.pdf>.

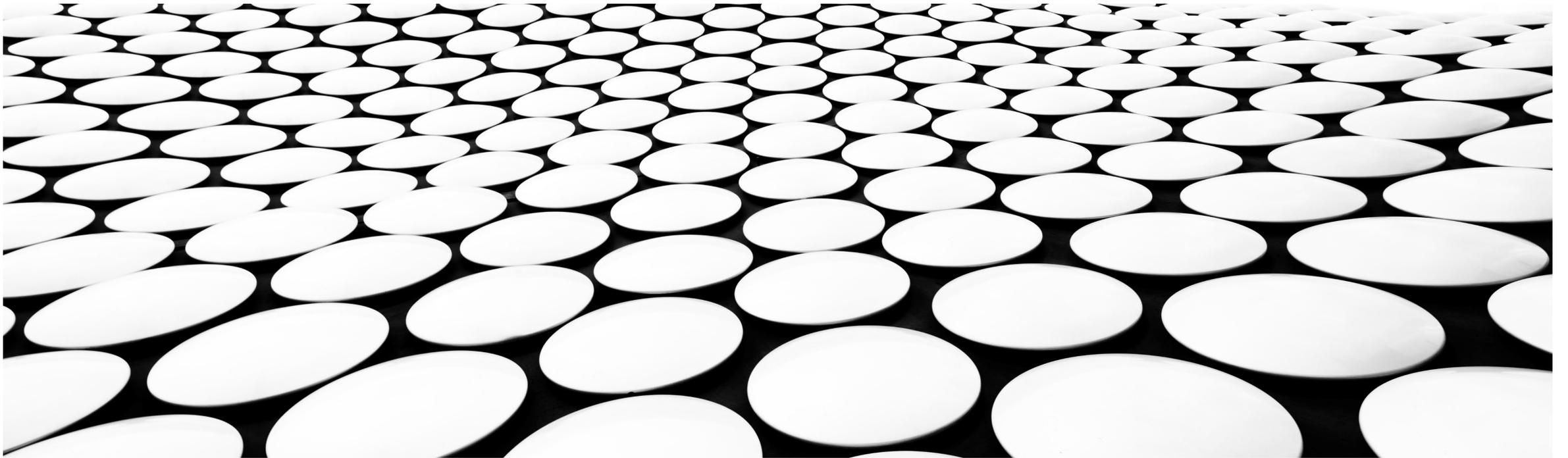
³ NACDL, FAIM and JAN Comment on Proposed Federal Bureau of Prisons FSA Earned Time Credits Rule. January 2021. <https://www.nacdl.org/getattachment/a39b698a-0b7d-4a15-b382-b0e9267058f2/nacdl-faim-jan-comments-to-bop-on-proposed-first-step-act-earned-time-credits-rule-january-2021.pdf>.

⁴ U.S. Sentencing Commission, *The Effects of Aging on Recidivism Among Federal Offenders*. https://www.ussc.gov/sites/default/files/pdf/research-and-publications/research-publications/2017/20171207_Recidivism-Age.pdf.

⁵ See, e.g., E. Davis et al, *Contacts Between Police and the Public, 2015*. Bureau of Justice Statistics, October 2018. <https://bjs.ojp.gov/content/pub/pdf/cpp15.pdf>.

RTI INTERNATIONAL

CENTER FOR COURTS AND CORRECTIONS RESEARCH



FIRST STEP ACT

Annual review

Error Correction

High Validity

Little evidence of prediction bias

QUESTION 1

Scoring and Inputs:

Is it all Age and Criminal History?

QUESTION 2

Scales:

**Do you need both General and
Violence?**

QUESTION 3

Race and Ethnicity:

**Can We Learn More about Native
American Trends (and Needs)?**

QUESTION 4

**Decision-making Aide:
Do Assessments Contribute to Better
Outcomes?**

QUESTION 5

**Evidence Based Programs:
Does PATTERN Link Individuals to
Risk Mitigating Programs?**

QUESTION 6

Outcomes:

**Can BOP Develop More Robust
Inclusion of Desistance?**

QUESTION 7

Standards:

Is FSA an Opportunity to Develop Standards?

QUESTION 8

Language Matters:

Is PATTERN the Right Name?

Dear Committee,

Hello.

I want to begin my remarks today by thanking the Office of Access to Justice for including RTI in this stakeholder engagement to comment on the PATTERN. I'm very excited to be included.

The First Step Act (FSA) (included in the Second Chance Reauthorization Act of 2018) was signed into law by President Trump on December 21, 2018, to improve criminal legal system outcomes, reduce the size of the federal prison population, and to maintain public safety.

The FSA allowed for sweeping reforms focused on reducing recidivism, creating incentives for success (e.g., earned time credit) in the Bureau of Prisons, reducing collateral consequences related to confinement (e.g., distance requirements, pilot programs for elderly), implementing sentencing reforms (e.g., change drug penalties, retroactive Fair Sentencing Act), and developing a system of oversight (e.g., Independent Review Committee, reports to Congress).

The FSA was implemented to reduce long prison sentences to federal prison and to improve the conditions in federal prisons. Nearly, upon the signing of the FSA, federal judges were able to impose shorter sentences than the statutorily defined mandatory minimums (e.g., sentencing safety valves). Likewise, people held in federal prison for crack cocaine offenses implemented prior to 2010, were now eligible to apply for a shorter prison sentence. The FSA implemented other practical improvements to the conditions of confinement including prohibiting the use of restraints on pregnant women and placing people close to their homes (within 500 miles). Additionally, the Bureau of Prisons (BOP) instituted an earned time credit program in which incarcerated individuals were able to reduce their sentences.

Risk Assessment System

The purpose of my remarks today is to comment on the Prisoner Assessment Tool Targeting Estimated Risk and Needs (PATTERN), a risk assessment system released by the Attorney General in 2019.

The PATTERN

The FSA requires the Department of Justice (DOJ) to review, validate, and release publicly its risk assessment tool (i.e., PATTERN) each year. The PATTERN is required to be assessed annually for (1) predictive validity, (2) dynamic validity, and (3) racial and ethnic neutrality (which refers to predictive bias).

The PATTERN achieves high predictive validity and minimal evidence of predictive bias. The Area Under the Curve (AUC) statistics are higher than what one finds with many assessments used in the criminal legal system – ranging between the low .70s to the low .80s. There is little evidence of significant predictive bias - black males and black females were underpredicted relative white individuals.

External Consultants: Fixed Major Errors

The DOJ/NIJ are to be commended for using independent consultants to conduct a quality check of earlier versions of the PATTERN. These consultants discovered what appears to be significant coding and analytic errors with the initial version of PATTERN. The external consultants reported in their revalidation:

“First, the coding and specification errors documented in the NIJ (2021) report have been corrected in the current report. Second, BOP has updated the data sources across several measures used in the current report to produce the most reliable and accurate data” (NIJ, December 2021).

Luckily, revisions were implemented to make progress on three goals:

(1) to increase opportunities for eligible individuals to earn time credits (and higher time credits for minimal and low risk) toward pre-release custody or supervised release

- (2) to mitigate the effects of various racial and ethnic disparities associated with previous risk classifications
- (3) to ensure that the PATTERN's risk level designations promote public safety and the reduction of recidivism.

These changes to the PATTERN with version 1.3 increase the Black/Hispanic male population estimated to score minimal and low risk by 36 and 26%, respectively. These are important findings given that people scoring in these lower categories are eligible to earn good time credits at a faster rate. To reiterate, the external consultants deserve an additional appreciation for discovering the coding and programming errors.

I believe the FSA provides an opportunity to identify and implement evidence-based programs that have the potential to reduce someone's likelihood of recidivism and (hopefully) increase someone's likelihood of being successful in their communities.

In the remainder of my comments, I pose 8 questions related to:

- Refinements to PATTERN's inputs and scoring scheme
- The inclusion of additional information in PATTERN and how that information may be used to improve prediction and fairness
- The current definition of recidivism used for PATTERN.

1. Scoring and Inputs: Is it all Age and Criminal History?

Now that I have highlighted that the PATTERN achieves high predictive validity and lacks significant differential prediction. I want to point to an area of potential improvement through empirical investigation.

The inputs to the PATTERN are rather basic. In fact, the instrument overfits on current age (age at assessment) and criminal history. Although there are more than one-dozen or so factors, and most of these looks familiar across other assessments, age and criminal history are upweighted to such an extent that I'm curious to know how much predictive validity the other factors contribute.

As someone that has studied, developed, and validated risk assessments for both pretrial and community supervision populations, I understand having to grapple with limited data and data of poor quality. With that said, however, it seems that the federal data system should have a more robust data collection system to truly capture needs and dynamic factors (this may come later with SPARC-13). The weights for age and criminal history make the PATTERN a bit more of static instrument – despite the inclusion of the low weighted items related to escapes, violence, and other items.

2. Scales: Do you need both General and Violence?

The PATTERN includes two scales – general and violent recidivism – for males and females. The violence scale does not appear too different from the general scale, and it suffers from the same issue of being upweighted by age and criminal history. I suspect the general scale predicts violent arrests about as well as the violence scale. It would seem the violent scale would be different from the general scale.

3. Race and Ethnicity: Can We Learn More about Native American Trends (and Needs)?

Although the PATTERN has few significant differences across racial and ethnic groups, there are issues with predictions and comparability for Native Americans. I would flag this as an opportunity for additional research to investigate the patterns of success and recidivism for Native peoples more specifically.

4. Decision-making Aide: Do Assessments Contribute to Better Outcomes?

Risk assessments – essentially – provide individual level estimates of an event occurring based on group data. That is all. Risk assessments, on their own, do not make decisions, they do not reduce recidivism, and they definitely do not fix racial/ethnic disparities.

Assessments are tools. I know it might seem rather rote to refer to assessments as tools as several scholars have pointed this out (Taxman, Monahan, amongst others). But I think we need to move away from reifying assessments and instead focus on what they can (and cannot) tell us.

Assessments can provide a systematic way to make decisions that provides transparency and (hopefully) reduce errors in human judgment.

What is important to figure out is if and how assessments (in this case the PATTERN) improve decisions. By improving decisions, this could include a lot of things, but mostly what the BOP appears to be doing is linking people up with the appropriate services. Do these services contribute to a reduction in recidivism? Do services improve the lives of released individuals (their families or communities)?

The challenge, therefore, for the PATTERN is less about predictive validity (as you have that) and more about ensuring that the PATTERN (or other RNS) includes treatment-relevant goals such that higher-risk folks get higher doses or more stringent programs or supervision. Similarly, lower-risk folks get lower doses or tailored programs for their needs. (Achieve the risk principle.)

The core objective for the PATTERN (and assessments more broadly) is to improve decisions that corrections staff make.

5. Evidence Based Programs: Does PATTERN Link Individuals to Risk Mitigating Programs?

A more difficult aspect for the BOP is to ensure that there are evidence-based programs suitable for individuals being released. Once the program evaluations are ready, then, individuals will be able to be linked by their risk level to the appropriate programs. That is, programs that mitigate or reduce the risk of one of the measured outcomes. Similarly, we should recognize that not everyone needs programs, but they may need assistance in general ways to restart their lives outside of prison.

6. Outcomes: Can BOP Develop More Robust Inclusion of Desistance?

This brings up the issue of outcomes. It seems that using rearrest or revocation to prison are logical/reasonable outcomes for the PATTERN to estimate risk. The DOJ should understand how their assessments are related to general public safety metrics. This is not to say that recidivism should be the only outcome of interest. Rather, as the recent *National Academies* report makes clear recidivism and desistance are different. Can the FSA an opportunity to track success by released individuals more broadly? Can the FSA be an opportunity to reduce barriers people face to be successful upon release?

7. Standards: Is FSA an Opportunity to Develop Standards?

Can the FSA and PATTERN development be an opportunity to think about and develop standards for assessment prediction, predictive bias, and adverse impact within criminal legal systems?

8. Language Matters: Is PATTERN the Right Name?

To conclude: I would offer a suggestion for the DOJ, which is to understand that language matters. Is the Prisoner Assessment Tool the correct name? We are learning that language can stigmatize and dehumanize, so I would encourage the DOJ to be introspective about their use of terminology.

Conclusion

Thank you again for the opportunity to comment on the PATTERN.

Kind regards,
Matthew DeMichele



Statement of

Liz Komar

Sentencing Reform Counsel
The Sentencing Project

First Step Act

Implementation

Stakeholder

Engagement Session

Office of Access to Justice
U.S. Department of Justice

September 28, 2022

Thank you for this opportunity to provide feedback on First Step Act implementation. Almost four years ago, Congress passed the First Step Act, a vital piece of legislation that gave many people hope. Congress recognized that people grow and change, and that it was in the interest of the American people and public safety to allow individuals to earn the ability to come home sooner by completing rehabilitative programs.

But today, the Bureau of Prisons has still failed to fully implement the First Step Act. The Bureau's delays in calculating and applying time credits, failure to provide sufficient access to rehabilitative programs, and PATTERN's overbroad limits on eligibility are deeply troubling. Individuals also must have access to the rehabilitative and productive programming that PATTERN incentivizes. The widespread practice of "augmentation," reassigning staff hired as teachers or to provide rehabilitative programming to correctional officer roles, has severely compromised the functioning and safety of federal prisons. The impact on programming has been significant: the waitlist for literacy education alone has grown to over 23,000 individuals.¹ And The Sentencing Project continues to hear from incarcerated individuals that Earned Time Credits are not being promptly counted and communicated to individuals, leading some to remain incarcerated past their expected earned release date.²

Since its development, the NIJ's usage of PATTERN has yielded well-documented racial disparities. Too many people are also excluded from relief by PATTERN's inadequate weighting of rehabilitative programming and productive activity and PATTERN's still overly punitive "cut points." We urge the NIJ to revise PATTERN by decreasing the weight of static factors and creating more opportunities for individuals to lower their recidivism scores.

Lower the weight accorded to criminal history and reward rehabilitation more.

When PATTERN was first released in July 2019, The Sentencing Project highlighted several major concerns with the tool: (1) Static factors (like age and type of criminal offense) comprised the bulk of an individual's risk score, while dynamic factors (like participation in rehabilitative programming) were underweighted; (2) There were limited opportunities for high- and medium-risk individuals to reduce their risk levels and thereby transition earlier to community corrections; and (3) The tool's weight for prior criminal history scores inherently produced a racially disparate effect.³ Our concerns remain largely unaddressed today.

¹ Department of Justice (2022), Federal Prison System FY2023 Performance Budget Congressional Submission, <https://www.justice.gov/file/1492946/download>, 33-34 & 43.

² For example, a memorandum regarding the Earned Time Credit automation tool appears to have been sent to many incarcerated individuals via Trulincs on September 8, 2022 stating that the tool had been launched. An individual at FCI Cumberland reported to The Sentencing Project, however, that as of September 26th, facility staff had not received any training or information on operating the automated credit system. He also reported that due to augmentation, a case manager is available at most once per week to address time credit related questions. As a result, by his calculation he was eligible for release to home confinement at least two weeks ago. Similarly, individuals at FCI Fort Dix and FCI Aliceville have reported that staff have received no guidance on how to roll out the tool, and others at FCI Butner and FCI El Reno report that caseworkers have stated that they do not have access to the tool yet.

³ Gotsch, K. (2021). Statement of Kara Gotsch, Deputy Director, The Sentencing Project, U.S. Department of Justice Stakeholder Listening Session on First Step Act Implementation. <https://www.sentencingproject.org/publications/statement-on-u-s-department-of-justice-stakeholder-listening-session-on-first-step-act-implementation/>.

Static factors currently play a heavy role in PATTERN scoring, and PATTERN scoring as a whole is heavily tilted toward adding points for negative behavior rather than providing opportunities to lower one's score. PATTERN should be rebalanced to lower the weight accorded criminal behavior and reward rehabilitation more.

PATTERN's heavy reliance on criminal history has troubling implications for fairness and equity. An individual's criminal history does not solely reflect past conduct, it also reflects the bias at every stage of the criminal legal system, from how an individual's neighborhood is policed, to the discretionary charging and plea bargaining practices of prosecutors, to the individual bias of judges, and the systemic bias of the criminal code. Any risk instrument that incorporates criminal history must be carefully constructed to avoid replicating such disparities – PATTERN falls short of that goal.

Given that PATTERN also relies on criminal history scores as calculated pursuant to the sentencing guidelines, it replicates the specific harms and flaws of the sentencing guidelines criminal history scoring system. As the Sentencing Commission's own draft simplification paper notes, the federal guidelines base the assignment of criminal history points on sentence length rather than on the past offense, which can "build[] in past discriminatory practices" regarding which individuals receive the lengthiest sentence. A Sentencing Commission analysis found Black men receive sentences 19% longer than similarly situated white men.⁴ And academic researchers found that amongst similarly situated defendants, Black men are also more likely to face charges carrying mandatory minimums.⁵ Those disparities are inherently replicated by the sentencing guidelines' criminal history scoring system.

Additionally, despite widespread agreement amongst sentencing experts that prior criminal convictions lose predictive value about future risk over time, PATTERN does not incorporate a "decay factor" that would discount the relevance of convictions as time passes. As Julian V. Roberts writes "[a]s the interval between episodes of offending increases, the probative value of criminal history as a predictor of future behavior declines... Sentencing theorists of all stripes, then, agree that the relevance of a previous conviction declines over time. At some point the prior offending should become extinct for the purposes of future sentencing."⁶ Yet PATTERN relies on the criminal history as scored at the moment the sentence was imposed.

To mitigate these flaws, criminal history's role in PATTERN scoring should be diminished by incorporating a decay factor, shifting focus away from sentence length, reducing the magnitude of overall impact of criminal history on the risk score, and by more heavily rewarding rehabilitative and productive activities.

Rehabilitative programming and productive activities should be rewarded more heavily both by increasing current point deduction values and by considering additional information. For example, the point deduction available for the completion of rehabilitative programming or a GED should be increased and individuals should receive additional point deductions for the

⁴ US Sentencing Commission (2017). Demographic Differences in Sentencing.

<https://www.ussc.gov/research/research-reports/demographic-differences-sentencing>

⁵ Rehavi, M. & Starr, S. (2014). Racial Disparity in Federal Criminal Sentences. *J. Pol. Econ.* 122, no. 6: 1320-54.

<https://repository.law.umich.edu/cgi/viewcontent.cgi?article=2413&context=articles> .

⁶ The Role of Criminal Record in the Sentencing Process. *Crime and Justice*, Vol. 22 (1997), University of Chicago Press, pp. 303-362, 335.

completion of higher education programs. Similarly, participation in work programs should result in greater point deductions. An individual who works in UNICOR for several years should receive more point deductions than an individual who participates for a single year. Additionally, institutional work assignments, such as working in food preparation or HVAC repair within a BOP facility, do not appear to be included within the current definition of work programs. The labor of incarcerated individuals is essential to the functioning of Bureau facilities and indisputably a form of work – the Bureau should recognize the value of this labor by including it within the “work program” calculation.

These modifications will help focus PATTERN more on an individual’s current risk and incentivize even more participation in rehabilitative programming and productive activity. While they will not eliminate PATTERN’s likelihood of perpetuating racial bias, they will decrease the extent to which PATTERN relies on factors particularly prone to racial disparities.

Less weight should be given to incident reports due to disparities.

Currently, PATTERN penalizes individuals for the same infractions three ways, adding points based on the existence of the incident, whether it was violent, and how recently the incident occurred. Of PATTERN’s 15 factors for men or 14 for women, three can potentially address the same disciplinary incident. That overemphasis is particularly concerning given the arbitrariness of discipline within federal facilities and well-documented disparities. Women are significantly more likely to be disciplined than men for minor infractions,⁷ and discipline rates vary by facility.⁸ For example, among maximum security facilities, at Thomson USP roughly 228 infractions were reported per 100 individuals, at Coleman I USP 80 per 100 individuals, and at Allenwood USP, 121 infractions were reported by 100 individuals. Given such significant variation, placing a heavy emphasis on incident reports injects arbitrariness into PATTERN scores and ultimately eligibility for relief. Limiting incident reports to fewer points and factors will improve fairness.

Recidivism should be measured by conviction, not arrest.

PATTERN is currently designed to predict whether someone will be rearrested for any offense – a metric riddled with racial disparities and not reflective of public safety. Racial bias and the over-policing of Black communities has led to significant disparities in arrest rates: “In 2016, Black Americans comprised 27% of all individuals arrested in the United States—double their share of the total population.”⁹ While some of those cases will ultimately go on to become convictions, many will be dismissed. As such, reliance on arrest histories, as opposed to convictions can perpetuate disparities, and also intrudes on the presumption of innocence. Researchers have criticized, for instance, plea bargaining guidelines for local prosecutors which

⁷ “In 2018, NPR and the Medill School of Journalism at Northwestern University collected data from prisons across the U.S. and found that women were disciplined two to three times more often than men for more minor violations of prison rules.” Shapiro, J. (2020). *Federal Report Says Women In Prison Receive Harsher Punishments Than Men*. NPR. <https://www.npr.org/2020/02/26/809269120/federal-report-says-women-in-prison-receive-harsher-punishments-than-men>.

⁸ Office of Justice Programs (Nov. 2021). *Federal Prisoner Statistics Collected under the First Step Act, 2021*. <https://bjs.ojp.gov/content/pub/pdf/fpscfsa21.pdf>.

⁹ *Report to the United Nations on Racial Disparities in the U.S. Criminal Justice System* (2018), The Sentencing Project. <https://www.sentencingproject.org/publications/un-report-on-racial-disparities/>.

rely on arrest histories rather than convictions.¹⁰ A tool which aims to predict merely which individuals are most likely to encounter the police and be rearrested – regardless of their innocence or guilt – will replicate the disparities in policing itself. Recidivism should be measured based on convictions, not arrests.

Create more realistic cut points.

NIJ should adopt more realistic cut points, particularly for risk of violent recidivism, which was unaddressed in NIJ's most recent revision. PATTERN outcomes do not lead to immediate release, they instead determine who is given the strongest incentives to engage in rehabilitative programming and who may be able to earn a change in the type of prerelease custody. The Department of Justice predicts that 80 percent of people classified as a medium risk of recidivism would not be rearrested for a violent offense.¹¹ Such a low rate of predicted recidivism is indicative of an overbroad category and contrary to public safety. The highest incentives to engage in rehabilitative activities should not be limited to individuals with next to no likelihood of reoffending – it is in the public interest to rehabilitate higher risk individuals.

Improve transparency and validity.

Finally, the NIJ and the Bureau should improve transparency regarding PATTERN scores. Given the significant liberty interests implicated by PATTERN scores, incarcerated individuals should have access to a clear and simple process to request their scores, including the inputs, and the ability to challenge their accuracy.

Thank you for this chance to provide feedback. The First Step Act is a critical opportunity to begin to build a more just and rehabilitative federal criminal legal system. Unfortunately, the potential has yet to be realized.

The NIJ, Bureau, and Department of Justice as a whole must prioritize fulfilling the First Step Act's promise. Beyond improvements to PATTERN, individuals must be able to access rehabilitative programming. Earned Time Credit accounting must be accurate and transparent. And finally, compassionate release should be significantly expanded. Thank you for your time.

¹⁰ Kutateladze, B. & Lawson, V. (2018). *Bad Arrests Lead to Bad Prosecution*. Cardozo Law Review. <http://cardozolawreview.com/wp-content/uploads/2018/08/KUTATELADZE.LAWSON.37.3.pdf>.

¹¹ U.S. Department of Justice (2022), First Step Act Annual Report, <https://www.ojp.gov/first-step-act-annual-report-april-2022>, 13.

Session 2
September 29, 2022

5. [National Association of Criminal Defense Lawyers](#) – James Felman, Mary Price, Elizabeth Blackwood, Patricia Richman, Cortland Broyles
6. [Brennan Center for Justice](#) – Ames Grawert, Patricia Richman
7. [Council of State Governments Justice Center](#) – Sarah Desmarais, David D'Amora, Lahiz Tavárez
8. [Federal Public and Community Defenders](#) – Patricia Richman
9. [Melissa Hamilton, JD, Ph.D](#)

MEMORANDUM

To: Nancy La Vigne, Ph.D, Director, National Institute of Justice

From: James Felman, Partner, Kynes Markman & Felman and Chair of NACDL's First Step Act Task Force; Mary Price, General Counsel, FAMM; Elizabeth Blackwood, Counsel and Director of NACDL's First Step Act Resource Center; Patricia Richman, National Sentencing Resource Counsel, Federal Public and Community Defenders; Cortland Broyles, Federal Affairs Manager, Justice Action Network

Re: PATTERN Improvements

Date: September 15, 2022

Thank you for the opportunity to meet with you and your colleagues to discuss our concerns with the PATTERN Risk Assessment Tool. We appreciated the open and frank exchange and look forward to future discussions with you. Following our recent conversation, you asked for a memorandum stating in short and simple terms our suggestions for improving the PATTERN Risk Assessment Tool. Our suggestions include the following:

1. **Greater transparency.** We believe the data relied on in developing the tool as well as the decisions, relying on that data, which led to the development of PATTERN, should be made public so that they are available for study and comment by academics and stakeholders. PATTERN is a high-stakes tool that controls how long a person is deprived of their liberty. It should not be a black box.
2. **More generous cut points.** While we applaud the recent changes to the cut points for the general recidivism scores, we believe further changes could be made to the cut points for both general and violent recidivism scores. We agree with the Department that setting cut points is a policy choice that should be guided by the First Step Act's goal of incentivizing as many people in federal prison as possible to engage in recidivism-reduction programming and productive activities. The recent changes are an important step towards this goal, but a more generous approach should be adopted. While we may not all agree on whether the cut points should be determined by reference to the relevant population being measured or absolute percentages of recidivism probability, we believe further ameliorative changes to the cut points are warranted.
3. **Recidivism should be measured by conviction, not arrest.** At present, the tool is designed to measure the probability that individuals in a group who share certain attributes will be re-arrested, not whether they have actually committed a new crime. Given disparate police practices in communities of color,

this definition of recidivism results in racially disparate impacts in the application of PATTERN.

4. **NIJ should incorporate a “decay factor” in criminal history scores.** Studies show significant decay in the predictive ability of a prior criminal event over time, but PATTERN locks in the criminal history as scored at the moment a person’s sentence was imposed. This makes PATTERN inaccurate for measuring recidivism potential because it fails to adjust its recidivism predictions to reflect the declining predictive effect of a prior record with age. And because criminal history is too often dictated by the color of one’s skin and one’s zip code, this practice embeds a racially disparate and unfair approach as criminal history is the most heavily weighted static factor included in PATTERN.
5. **More dynamic factors that result in the *subtraction* of points are needed.** PATTERN does not provide sufficient dynamic factors that would allow individuals the opportunity to lower their risk scores. We hope additional dynamic factors can be identified and added to the PATTERN tool and that the number of points associated with existing mechanisms to decrease a PATTERN score are increased.
6. **The recommendations of the IRC.** While we do not know what recommendations by the IRC have not yet been incorporated into PATTERN, we hope any such remaining recommendations will be considered and adopted. Additionally, the IRC’s recommendations and their implementation status should be made public.

**BRENNAN
CENTER**

FOR JUSTICE

September 27, 2022

Dear Director Rossi,

Thank you for the opportunity to speak at the upcoming Stakeholder Engagement Session on PATTERN, the First Step Act's risk assessment tool, convened by the Department of Justice's Office for Access to Justice.

I am writing today to share, as the Brennan Center's written submission, a new analysis we published last week on this very subject. Our report, *The First Step Act's Prison Reforms*, walks through areas of concern related to the Act's risk assessment tool, including its broad definition of recidivism and persistent racial biases in risk scoring.¹

I look forward to speaking on these issues during Thursday's session.

Very truly yours,

/s/

Ames C. Grawert
Senior Counsel & John L. Neu Justice Counsel, Justice Program

Brennan Center for Justice
at NYU School of Law
120 Broadway, Suite 1750
New York, New York 10271

¹ Ames Grawert and Patricia Richman, *The First Step Act's Prison Reforms*, Brennan Center for Justice, 2022, 3–5, <https://www.brennancenter.org/our-work/research-reports/first-step-acts-prison-reforms>.

The First Step Act's Prison Reforms

Uneven Implementation and the Path Forward

By **Ames C. Grawert and Patricia L. Richman**

PUBLISHED SEPTEMBER 23, 2022

Three years ago, Congress passed the First Step Act, the first major federal criminal justice reform legislation in nearly a decade.¹ The culmination of years of bipartisan advocacy, the law included both long-overdue changes to excessively punitive federal sentencing laws and reforms aimed at improving conditions in the federal prison system.

This brief examines the structure of the First Step Act's prison reforms, how they have been implemented, and what more Congress and the Department of Justice (DOJ) must do to realize their potential.

Background

In the 1980s, Congress enacted rigid mandatory penalties, which require judges to impose minimum terms of incarceration for certain federal crimes or when certain statutory criteria are satisfied.² Sometimes these penalties are triggered by specific conduct, such as possessing a firearm or possessing drugs above a specified threshold quantity.³ These laws significantly expanded the size of the federal prison system and led to an explosion in racial disparities in punishment, all without addressing drug use or improving public safety.⁴

The federal prison system strained under the effects of these penalties, with lawmakers describing a “state of

crisis” as the prison population climbed.⁵ Indeed, the federal prison population grew eightfold between the 1980s and the mid-2010s, outpacing growth in state-level incarceration, with weapon and drug offenses making up more than 60 percent of the growth in federal imprisonment.⁶ Mechanisms for checking excessive custodial sentences did not keep pace. Compassionate release, for example, which allows a federal court to reduce or end a prison sentence for “extraordinary and compelling circumstances,” was severely underused.⁷ People in prison had limited opportunities to earn early release through their conduct; “good time” credits, earned for good behavior while incarcerated, amounted to at most 47 days per year of incarceration.⁸

Congress began to rectify this situation in 2010 with the Fair Sentencing Act, which reduced the crack/powder cocaine sentencing ratio from 100:1 to 18:1 in drug trafficking cases and eliminated the five-year mandatory minimum for simple possession of crack.⁹ However, these changes applied only prospectively, meaning that people sentenced before the law went into effect remained in prison, serving the same wildly disproportionate sentences that Congress had just repudiated. Additionally, the act did not address the long-standing consensus that federal prisons were failing to provide meaningful programming and rehabilitation to incarcerated people.

The First Step Act attempted to address these shortcomings. Among other things, its sentencing reforms made the Fair Sentencing Act retroactive, permitting people sentenced under the old 100:1 crack/powder cocaine penalty scheme to apply for resentencing as if the Fair Sentencing Act's 18:1 ratio had been "in effect at the time the covered offense was committed."¹⁰ The act also revised other mandatory minimums and for the first time allowed people in federal prison to petition a federal court for compassionate release.¹¹ Additionally, and critically, it introduced a system for people to reduce their time spent in prison by participating in programming and activities.

The First Step Act's changes to federal sentencing laws had an immediate, significant impact. As of May 2021, roughly 3,700 people had benefited from a reduction in their sentence under the provision making the Fair Sentencing Act retroactive.¹² The average sentence was reduced by around three years, or roughly 25 percent.¹³ And, according to the Bureau of Prisons, the act has led to compassionate releases or sentence reductions in more than 4,200 other cases.¹⁴ The act's other sentencing reforms have begun to impact new cases, too. The restoration of judicial discretion in select drug cases benefited roughly 1,400 people in the first year of enactment.¹⁵

The FSA's Prison Reforms: Challenges and Successes

Unfortunately, errors and half-starts have marred the roll-out of the First Step Act's prison reforms.

The act aimed to create a system that would encourage people in prison to participate in programming designed to reduce recidivism — that is, the risk that they would come in contact with the criminal justice system, through arrest or otherwise, after release. This new system has several key components:

- **Incentives and rewards for program participation.** The act incentivizes engagement with "evidence-based recidivism reduction programming" (e.g., drug treatment or literacy programs) and "productive activities" (e.g., work or vocational training) by awarding time credits for participation that people can apply toward early transfer to supervised release, home confinement, or a residential reentry center (i.e., a halfway house).¹⁶ However, a laundry list of disqualifications excludes people convicted of certain offenses from participating.¹⁷
- **A risk assessment tool.** This tool is designed to determine the "recidivism risk of each prisoner" and the "type and amount of evidence-based recidivism reduction programming for each." Risk scores generated by the tool affect both the number of time credits individ-

uals can earn and how they can redeem them, making its design and implementation vital to the act's success.¹⁸

- **Expanded recidivism reduction programming opportunities.** The act also seeks to expand the availability of job training and other programming for all incarcerated individuals.¹⁹

Unlike the act's sentencing reforms, these changes to the prison system were phased in, to give the Bureau of Prisons (BOP) time to build out the relevant policies. The final rule governing the awarding of time credits did not go into effect until January 2022.

Along the way, the act's corrections reforms have hit a series of snags in implementation, requiring continued attention from policymakers to ensure that they succeed. What went wrong?

Rocky Start to Earned Time Credit Implementation

The core of the First Step Act's corrections reforms is a system that allows some people to receive earned time credits (ETCs) for participating in designated programming or activities behind bars, such as drug treatment or cognitive behavioral therapy. ETCs effectively shorten sentences, making them a powerful incentive for participating in programs. Implementation of the ETC system stalled, however, after the act's passage.²⁰ The BOP initially refused to award ETCs and issued draft regulations that would have sharply limited the reach of the program.²¹

That changed in January 2022 when the BOP, responding to broad and bipartisan criticism, issued a rule that dramatically increased the rate at which completed programs translate to ETCs. Simultaneously, the BOP calculated and awarded ETCs retroactively back to the date of the First Step Act's enactment.²² As a result, by April 5, 2022, more than 6,100 people had been transferred to supervised release and another 3,155 people had been transferred to prerelease custody — that is, a residential reentry center or home confinement.²³

Under the law and current regulations, eligible people can earn 10 to 15 days of ETCs for every 30-day period in which they successfully participate in programs or productive activities.²⁴ (People deemed minimum or low risk earn 15 days, while those classified at higher risk levels earn 10 days.) Once they have accrued time credits equal to the time remaining on their sentence, those credits can potentially be cashed in for either an early transfer to supervised release (by up to a year) or to prerelease custody.²⁵

Yet not all people are eligible to earn ETCs. The act excludes those convicted of many crimes.²⁶ These exclusions prevent nearly half of the federal prison population from benefiting from credits.²⁷ The exclusions also

appear to serve no policy purpose. According to the Independent Review Committee (IRC), the expert panel tasked with helping the BOP implement the law, there is “no significant difference in the collective recidivism-risk profiles of the BOP’s ETC-eligible and ETC-ineligible inmate populations.”²⁸

Additionally, despite the announcement of new rules improving the ETC system’s operation, it appears that the process of awarding ETCs continues to move slowly. People in federal prison have reported that they are not having their time credits applied or not being released as early as they should be, and that BOP staff have not received training or guidance in how to operate the new system. These missteps point to a need for continued oversight to ensure timely and accurate implementation.²⁹

Flawed Risk and Needs Assessment Tools

The First Step Act calls for an assessment system to evaluate each person’s risk of recidivism (defined as arrest or return to prison within three years of release) and criminogenic needs (factors that, unless addressed, may predict future contact with the criminal justice system).

The act’s corrections reforms rely on the system’s accuracy and fairness. People who score as minimum or low risk earn ETCs more quickly than those in higher risk categories. The BOP is also required to apply credits earned by people categorized as minimum or low risk toward prerelease custody or supervised release. In contrast, individuals classified as medium or high risk must meet additional criteria, including special approval by the warden, before their credits can be applied against their sentence. Additionally, those in the higher risk categories cannot earn early transfer to supervised release; they can apply their credits only to prerelease custody.³⁰

These high stakes make it particularly important that the risk and needs system be transparent, fair, and unbiased.³¹ Unfortunately, the part of the system focused on criminogenic needs was slow to be deployed. Worse, the system released by the DOJ for assessing risk, called PATTERN (Prisoner Assessment Tool Targeting Estimated Risks and Needs), remains flawed despite major revisions and renewed attention from the Biden administration.³²

Some elements of PATTERN have improved over time as the DOJ has responded to stakeholder criticism. But correcting all of its flaws would require fundamentally reevaluating what type of risk PATTERN measures and how it translates that risk to policy judgments — a top-to-bottom reconstruction that goes beyond the remedies the administration has proposed to date.

Background on PATTERN

Like any risk assessment tool, PATTERN was developed to predict the likelihood of a defined behavior based on

a series of inputs associated with that behavior. It seeks to predict a specific type of recidivism: “a return to BOP custody or a rearrest within three years of release from BOP custody, excluding all traffic offenses except driving under the influence and driving while intoxicated.”³⁸

The tool works by collecting information on a person and assigning them points based on factors in their background, which add up to a total risk score. Each factor carries a different weight; for example, completing a prison program will subtract one to three points from the total score. By contrast, being under the age of 26 can add a significant number of points. The BOP then determines how to translate these scores into a policy judgment — whether the person presents a minimum, low, medium, or high level of risk.³⁹

Different versions of the tool exist for men and for women, and for predicting general recidivism risk (defined as the likelihood of rearrest or return to BOP custody for any offense) and violent recidivism risk (the same likelihood but for an offense deemed violent).⁴⁰ These tools serve

Risk Assessment Tools in the Criminal Justice System

Risk assessment tools are designed to calculate the likelihood that someone will engage in certain behaviors based on characteristics they share with others in a group. Predictions tend to rely on a mix of dynamic factors, which people can change over time (such as participation in prison programming), and static factors, which people cannot change (such as age and criminal history). Before being put into practice, these actuarial tools are usually calibrated by comparing their predictions against data gathered from the real world.³³

Risk assessment isn’t math; it’s policy. Indeed, these models reflect policy choices at every phase, from judgments about risk tolerance to beliefs about criminal behavior. As a result, they remain both controversial and imperfect. They often classify individuals inaccurately and tend to misclassify people who are unlikely to pose a real threat as moderate or high risk.³⁴ An overreliance on static factors can also lead these tools to incorrectly classify Black people and others from overpoliced and disadvantaged communities as high risk.³⁵ Last, because tools are tested against historical data, they risk replicating biases in that data when predicting risk. For example, historical racial disparities in policing can lead an algorithm that was calibrated using arrest data to predict, incorrectly, that Black and white people reoffend at different rates — when in fact they are merely policed at different rates.³⁶ In the words of one expert, “No predictive tool is better than the data set from which it originates.”³⁷

different purposes, but — critically — when the risk tiers they produce diverge, the higher score governs. Practically, this means that someone’s potential ability to transfer to prerelease custody is limited until and unless they score into the low or minimum tiers on both assessments.⁴¹

Missteps in PATTERN’s Development (2019–Early 2022)

PATTERN’s rollout was characterized by implementation mistakes and policy missteps, raising concerns that the tool entrenched racial bias in the prison system, relied on an overly conservative definition of recidivism risk, and failed to account for people’s capacity for personal growth and change while incarcerated. These issues plagued the first years of implementation, even if (as discussed below) the DOJ has since taken steps to partially correct them.

- **Technical errors.** PATTERN required repeated technical corrections just to function as designed.⁴² Reviews conducted by the National Institute of Justice (NIJ), the DOJ agency tasked with reviewing the tool to guarantee its accuracy, revealed mistakes in how the model’s variables were defined and applied.⁴³ A January 2021 NIJ report identified scoring, coding, and weighting errors and recommended a revised version of PATTERN.⁴⁴ In the months that followed, the NIJ found even more errors, described in a December 2021 report.⁴⁵ These problems were compounded by human error by BOP staff in the scoring process. As a result, many people in BOP custody were assigned to the wrong risk category, even as the DOJ relied on PATTERN to make potentially life-and-death decisions about whether to transfer people to home confinement during the pandemic.⁴⁶
- **Racial bias.** In November 2021 the Bureau of Justice Statistics (BJS) reported that the vast majority of Black people in BOP custody — more than 70 percent — were classified by PATTERN as medium or high risk.⁴⁷ Further releases show that PATTERN continues to overpredict the likelihood that Black, Hispanic, and Asian people will commit new crimes or violate rules after leaving prisons, relative to white people in prison.⁴⁸

These biases stem in part from PATTERN’s focus on rearrest. Arrest is a poor proxy for criminal activity, as it may reflect policing decisions — where officers are deployed and their biases in making arrests — rather than actual criminality.⁴⁹ Designing a risk assessment tool based on arrest patterns also means relying on data tainted by decades of discriminatory policing, a concern raised by stakeholders years ago.⁵⁰

For that reason, the IRC, among others, has recommended a narrower definition of recidivism, one that “might better identify individuals likely to engage in serious criminal activity post-release” and potentially reduce the tool’s reliance on racially biased data.⁵¹ In

2020, however, the DOJ rejected calls to redesign PATTERN to include a narrower definition of recidivism, such as reconviction or re-incarceration, claiming such a definition would be unworkable because of data limitations.⁵²

- **Risk tolerance.** Translating a PATTERN score to a risk designation is a policy judgment. There is no objective quantifier of acceptable risk.⁵³ Instead, it is up to policymakers and the algorithm’s designers to define risk and decide how much of it their system should tolerate.⁵⁴ In making those judgments, the DOJ initially adopted a set of fairly conservative cut points — borders between risk categories — that were derived in part from the average predicted risk of recidivism of people released from the BOP.⁵⁵ But early data about recidivism among those scored by PATTERN showed that fewer than 2 percent of people scored as low risk were rearrested between July 2019 (when PATTERN was implemented) and September 2020. Further, only 4.5 percent of “high risk” individuals were rearrested in the study period.⁵⁶ That data led the Urban Institute to conclude that PATTERN overpredicted recidivism and that individuals with a higher PATTERN score should properly be classified in a lower risk category.⁵⁷
- **Variable weights.** Choosing which factors in someone’s background PATTERN should consider when scoring them is also a policy choice, as is the weighting of those factors — even if the latter is informed by mathematics.

Contrary to Congress’s intent, PATTERN has consistently overemphasized static factors like age and criminal history, which makes it difficult for people to change their assessment score over the course of their incarceration. Dynamic factors are weighted less heavily, and more than half of the dynamic factors in the model actually increase a person’s risk score.⁵⁸ Perhaps unsurprisingly, then, just 23–35 percent of people in one NIJ analysis had been able to reduce their assessed risk level “at the last assessment compared to the first.”⁵⁹ The stickiness of PATTERN assessments suggests that the tool does not yet appropriately account for personal growth — or program participation — during incarceration.⁶⁰

Revisions to PATTERN (April 2022–Present)

In April 2022 the DOJ finally announced a plan to address some of these issues. The attorney general approved a new version of PATTERN that attempts to fix the problems discussed above, and the BOP announced that it would use the new tool to rescore all individuals in its custody.⁶¹

Alongside this revision, the DOJ announced that it would revise the cut points for evaluating general recidivism risk to increase the percentage of people who qualify

TABLE 1

Share of Men Classified as Minimum or Low Risk

| | TOTAL | WHITE | BLACK | HISPANIC | NATIVE AMERICAN | ASIAN |
|---------------------|-------|-------|-------|----------|-----------------|-------|
| Original cut points | 44.0% | 55.7% | 31.7% | 51.0% | 21.7% | 70.0% |
| Revised cut points | 55.2% | 65.2% | 43.1% | 64.4% | 33.8% | 78.9% |

Source: U.S. Department of Justice, *First Step Act Annual Report*, 15–16.

for lower risk tiers. Under the new cut points, 55 percent of the BOP’s male population and 83 percent of its female population are projected to fall into the minimum- and low-risk categories under the general tool — up from 44 and 78 percent, respectively. Cut points for determining violent recidivism risk remain unchanged.⁶²

This change was specifically adopted to “help mitigate the effects of various racial and ethnic disparities associated with previous risk groupings” — that is, to reduce the tool’s racial bias — and will indeed result in more Black and Hispanic people being categorized as minimum or low risk, as shown in table 1.⁶³ However, the changes will not correct PATTERN’s tendency to overpredict the recidivism risk of nonwhite people, as the DOJ itself conceded.⁶⁴

Nor does this revision really address PATTERN’s overly conservative risk profile, because it leaves the cut points for evaluating risk of violent recidivism unchanged. As shown in table 2, those cut points erect a high bar for inclusion in the lower risk categories. As currently drafted, fully 80 percent of people classified as a medium risk of recidivism would go on to have no rearrest for a violent offense, to say nothing of reconviction or re-incarceration. And recall that risk assessment outcomes do not lead to immediate release; they instead determine “who is given more robust incentives to engage in rehabilitative

programming and who might earn a change in the type of prerelease custody.”⁶⁵

Changes to the general tool’s cut points will likely allow people to earn ETCs faster under the provision granting additional ETCs to people with low PATTERN scores.⁶⁶ But they may have limited impact on the ETC system’s broader functioning, as a high violent score will continue to override a low general score when determining someone’s eligibility to apply ETCs to prerelease custody or supervised release.⁶⁷ This limited impact also means racial disparities in access to prerelease custody will persist.

Taken together, the April 2022 revisions fall short of addressing every problem with PATTERN and the First Step Act’s use of risk assessment tools generally. But these remedial efforts represent progress, as does the greater level of transparency in the DOJ’s and NIJ’s more recent reports.⁶⁸ Furthermore, the DOJ’s April report pledged to “consider all legally permissible options for reducing the differential prediction based on race and ethnicity,” including potentially revisiting PATTERN’s focus on rearrest risk — an important, if overdue, break with the previous administration’s policy.⁶⁹

Delayed Criminogenic Needs Tool

In March 2022 the DOJ finally released a report on the last component of the act’s prison reforms: its assessment tool designed to identify and address each “prisoner’s specific criminogenic needs.”⁷⁰ The tool, SPARC-13, is intended to complement PATTERN by directing people in prison toward programming that meets their needs.

While it is too early to evaluate SPARC-13’s implementation, several details from the DOJ’s initial report stand out. For one, BOP staff report low levels of familiarity with the needs assessment system and how to administer it.⁷¹ The tool also appears to rely heavily on self-reporting by imprisoned people themselves, leading to a very high rate of refusal for some assessment areas. Nearly one-third of those assessed refused to be screened for trauma, for example.⁷² The BOP should work to reduce this refusal rate, including by making people aware that by refusing screening they may forgo opportunities to earn time credits.⁷³

TABLE 2

Predicted Likelihood of No Post-Release Arrest for a Violent Crime (Male)

| PATTERN RISK CATEGORY | PREDICTED LIKELIHOOD |
|-----------------------|----------------------|
| Minimum | 98.6% |
| Low | 90.8% |
| Medium | 79.8% |
| High | 63.9% |

Source: U.S. Department of Justice, *First Step Act Annual Report*, 13.

Lack of Adequate Prison Programming

For the ETC system to succeed, the BOP must significantly expand the availability of programming and productive activities. That is a heavy lift. Historically the BOP has failed to provide enough programming to satisfy the needs and wants of the people in its custody, even before accounting for the contemplated expansion.⁷⁴ Pandemic restrictions significantly hindered what programming was available; many programs were “highly impacted” by the virus, and some shut down entirely.⁷⁵ But the IRC warned in 2020 that “even a full return to pre-COVID-19 BOP programming levels will not be sufficient to make available evidence-based recidivism reduction programs and productive activities” for all eligible individuals in BOP custody by 2022, and it identified troubling demographic disparities in program participation among eligible individuals.⁷⁶

The final ETC system offers a partial response to this problem by allowing eligible people to earn credits if their program is interrupted through no fault of their own. But it is unclear whether this rule applies when a person is unable to even start a program. The rule’s explanatory text states that “inmates will not be penalized if specifically recommended [programs] are unavailable to them or at full enrollment at their facilities,” but the rule itself references only program “interruption.”⁷⁷ This lack of clarity is concerning, particularly in light of a recent BJS report showing that the programs providing the most hours of credit were, at least through the end of 2021, also the least available.⁷⁸ Some reports also indicate that BOP staff are not receiving clear guidance from DOJ leadership on how to implement new policies and regulations, and that programs may not be staffed or resourced to ensure a prompt and faithful rollout — further complicating an already byzantine system.

That said, recent information from the DOJ gives some reason for optimism. The DOJ’s latest publication describes a significant increase in the number of approved programs available to people in prison. During FY 2021, BOP staff also “recorded a marked increase in participation” in prison programming and productive activities. Furthermore, the BOP recently posted and filled a wide range of positions related to First Step Act implementation and contracted for evaluations of its programs. Last, the DOJ reported that more than \$362 million “in appropriated FSA funding” had been used “to expand reentry programs and their delivery.”⁷⁹ It is not clear how much of that funding was new, how much was distributed from elsewhere within the DOJ, and how much the BOP may still need to make up for programming shortfalls.⁸⁰

The Path Forward

For all its successes, the First Step Act continues to fall short of its promise. Yet its problems are fixable. Some solutions can be achieved simply through executive action; others call for congressional intervention, which may be (and certainly should be) achievable on a bipartisan basis.

The Department of Justice is equipped to make immediate policy changes that could drastically expand the reach of the First Step Act’s correctional reforms. The following steps would ensure that the department’s implementation efforts better align with congressional intent:

- **Prioritize transparency.** Despite long-standing requests from criminal justice reform advocates, among others, the DOJ has not released the data needed to fully assess PATTERN for accuracy and bias. Stakeholders will continue to regard PATTERN with skepticism and distrust until the department releases the information necessary to independently evaluate and validate PATTERN and adopts a practice of timely disclosures about progress and setbacks in its implementation.
- **Revise PATTERN.** Recent revisions to the risk assessment tool show that the DOJ understands the need to rebuild trust in this area of the act’s implementation, but much more needs to be done. The DOJ should start by reorienting PATTERN to focus on predicting a different type of recidivism — the risk of reconviction or re-incarceration, rather than rearrest. This change alone would reduce racial disparities in PATTERN’s risk predictions. Additionally, the DOJ should revisit its decision to leave unchanged PATTERN’s cut points for predicting violent recidivism.
- **Expand prison program offerings.** The list of programs that allow people to earn time credits has until recently been quite short, and it is unclear how much demand remains unmet since recent expansions. The BOP should continue to build out prison programming services and interpret rules expansively to ensure that people can earn credits where programs are inaccessible or booked. Additionally, the BOP should ensure that correctional staff are fully trained to implement new rules being promulgated by the DOJ, and provide transparency on how that training is conducted.

Congress should address structural problems in the First Step Act's design by taking these steps:

- **Broaden eligibility for earned time credits.** The percentage of people in prison eligible for ETCs is far too small: half of those incarcerated by the BOP are ineligible because of their offense of conviction. There is no public safety justification for these exclusions, and Congress should repeal them.
- **Decouple PATTERN from earned time credit eligibility.** The First Step Act's use of PATTERN to determine how much time individuals serve in prison is fundamentally flawed. It is also unusual: most state systems use risk assessments only for more limited purposes.⁸¹ Congress should amend the First Step Act

to limit PATTERN's role in its corrections reforms. Specifically, Congress should detach risk classification from ETC eligibility so that all people incarcerated by the BOP are incentivized to participate in programming and productive activities.

- **Increase funding for prison reforms.** Early in its implementation, the act suffered from funding shortfalls, with Congress even failing to appropriate new money for implementation during the law's first year. While recent reports indicate that the BOP's implementation efforts are now on better footing, policymakers should ensure, through oversight and sustained contact with BOP administrators, that the agency has the resources it needs to deploy high-quality prison programming to all people and in all facilities where it is needed.

Endnotes

- 1 First Step Act of 2018, Pub. L. 115-391, 132 Stat. 5194 (2018), <https://www.congress.gov/115/plaws/publ391/PLAW-115publ391.pdf> [<https://perma.cc/9TWB-U5HD>].
- 2 See, e.g., United States Sentencing Commission, *Report to the Congress: Cocaine and Federal Sentencing Policy*, 2002, 4–11, https://www.ussc.gov/sites/default/files/pdf/news/congressional-testimony-and-reports/drug-topics/200205-rtc-cocaine-sentencing-policy/200205_Cocaine_and_Federal_Sentencing_Policy.pdf [<https://perma.cc/36DL-K5NQ>] (providing background on the passage and structure of mandatory minimums in the Anti-Drug Abuse Acts of 1986 and 1988).
- 3 FAMI Foundation, *Federal Mandatory Minimums, Families Against Mandatory Minimums*, 2015, 1–16, <https://fami.org/wp-content/uploads/Chart-All-Fed-MMs.pdf> [<https://perma.cc/JW3E-963E>] (collecting mandatory minimums by offense, including by drug type and weight).
- 4 James Cullen, “Sentencing Laws and How They Contribute to Mass Incarceration,” Brennan Center for Justice, October 5, 2018, <https://www.brennancenter.org/our-work/analysis-opinion/sentencing-laws-and-how-they-contribute-mass-incarceration> [<https://perma.cc/8B98-S3ED>]; Ram Subramanian et al., *A Federal Agenda for Criminal Justice Reform*, Brennan Center for Justice, 2020, 19–20, <https://www.brennancenter.org/our-work/policy-solutions/federal-agenda-criminal-justice-reform> [<https://perma.cc/LQ72-DDSZ>]; and Paul J. Hofer, “Federal Sentencing After Booker,” *Crime & Justice* 48 (2019): 171–74, <https://www.journals.uchicago.edu/doi/abs/10.1086/701712>.
- 5 J. C. Watts Jr. et al., *Transforming Prisons, Restoring Lives: Final Recommendations of the Charles Colson Task Force on Federal Corrections*, Urban Institute, 2016, vii, <https://www.urban.org/sites/default/files/publication/77101/2000589-Transforming-Prisons-Restoring-Lives.pdf> [<https://perma.cc/U8QV-79T2>].
- 6 Watts et al., *Transforming Prisons, Restoring Lives*, 10.
- 7 28 U.S.C. § 3582(c)(1)(A). For the difficulty of securing compassionate release before the First Step Act’s passage, see *United States v. Brooker*, 976 F.3d 228, 231–32 (2d Cir. 2020); *United States v. McCoy*, 981 F.3d 271, 275–76 (4th Cir. 2020); *United States v. McGee*, 992 F.3d 1035, 1041–42 (10th Cir. 2021); *United States v. Maumau*, 993 F.3d 821, 830 (10th Cir. 2021) (discussing the history and under-use of compassionate release), *cert. denied*, 142 S. Ct. 583 (2021); U.S. Department of Justice, Office of the Inspector General, *The Federal Bureau of Prisons’ Compassionate Release Program*, 2013, 1, <https://www.oversight.gov/sites/default/files/oiq-reports/e1306.pdf> [<https://perma.cc/E5JD-9U5C>]; and Christie Thompson, “Old, Sick and Dying in Shackles,” *Marshall Project*, March 7, 2018, <https://www.themarshallproject.org/2018/03/07/old-sick-and-dying-in-shackles> [<https://perma.cc/T49J-JBR5>].
- 8 For background on this credit system — which was expanded by the First Step Act to 54 days per year, with the difference in credits awarded retroactively — see Good Conduct Time Credit Under the First Step Act, 84 Fed. Reg. 72,274, 72,274–76 (December 31, 2019) (codified at 28 C.F.R. § 523), <https://www.federalregister.gov/documents/2019/12/31/2019-27976/good-conduct-time-credit-under-the-first-step-act> [<https://perma.cc/E4LD-4Y6D>]. For the credit provision as it currently exists, see 18 U.S.C. § 3624(b).
- 9 Fair Sentencing Act of 2010, Pub. L. 111-220, 124 Stat. 2372 (2010). See also *Terry v. United States*, 593 U.S. ___, 141 S. Ct. 1858, 1861 (2021) (reviewing the history of cocaine sentencing laws).
- 10 First Step Act of 2018 § 404, 132 Stat. 5194, 5222 (2018).
- 11 First Step Act of 2018 § 603(b).
- 12 United States Sentencing Commission, *First Step Act of 2018 Resentencing Provisions Retroactivity Data Report*, 2021, 5, table 2 and n.1, <https://www.ussc.gov/sites/default/files/pdf/research-and-publications/retroactivity-analyses/first-step-act/20210519-First-Step-Act-Retro.pdf> [<https://perma.cc/W4B9-FWM5>].
- 13 United States Sentencing Commission, *First Step Act Resentencing Provisions Retroactivity Data Report*, 9, table 6. Note that due to data limitations, this analysis was based on a smaller sample — roughly 2,500 of the 3,700 cases the commission was aware of in total.
- 14 Federal Bureau of Prisons, “First Step Act: FSA Numbers,” accessed July 27, 2022, <https://www.bop.gov/inmates/fsa/> [<https://perma.cc/BV33-T8YE>].
- 15 Julie Zibulsky and Christine Kitchens, *The First Step Act of 2018: One Year of Implementation*, United States Sentencing Commission, 2021, 19–22, <https://www.ussc.gov/sites/default/files/pdf/research-and-publications/research-publications/2020/20200831-First-Step-Report.pdf> [<https://perma.cc/VK52-6JYH>].
- 16 First Step Act of 2018 § 101(a) (codified at 18 U.S.C. §§ 3632(d), 3633, 3635(3)).
- 17 First Step Act of 2018 § 101(a) (codified at 18 U.S.C. § 3632(d)(4)(D)).
- 18 See First Step Act of 2018 § 101(a) (codified at 18 U.S.C. § 3632(d)(4)(A) (providing higher amounts of earned time credit for some individuals deemed “minimum” or “low” risk) and (g)(1) (restricting eligibility to transfer to supervised release or prerelease custody to individuals in minimum- or low-risk categories, absent warden approval under specified circumstances).
- 19 See First Step Act of 2018 § 101(a) (codified at 18 U.S.C. §§ 3621(h)(4) & (6) (requiring the attorney general to “provide all prisoners with the opportunity to actively participate in evidence-based recidivism reduction programs, according to their specific criminogenic needs, throughout their entire term of incarceration”).
- 20 See, e.g., memorandum from Inspector General Michael E. Horowitz to Michael Carvajal, director, BOP, “Management Advisory Memorandum 22-007: Impact of the Failure to Conduct Formal Policy Negotiations on the Federal Bureau of Prisons’ Implementation of the FIRST STEP Act and Closure of Office of the Inspector General Recommendations,” November 15, 2021, <https://oig.justice.gov/reports/management-advisory-memorandum-impact-failure-conduct-formal-policy-negotiations-federal> [<https://perma.cc/YB5T-R6EW>].
- 21 FSA Time Credits, 85 Fed. Reg. 75,268 (proposed November 25, 2020) (superseded), <https://www.federalregister.gov/documents/2020/11/25/2020-25597/fsa-time-credits> [<https://perma.cc/P9K4-X47R>]. Under the proposed rule, an imprisoned person would have needed to complete eight hours of programming to accrue a “day” of “participation” in programming. FSA Time Credits, 85 Fed. Reg. at 75,269. That guidance would have interacted with the BOP’s sparse programmatic offerings — which in some cases meet for just a few hours a day — to make it very difficult to earn any meaningful amount of ETcs. See James M. Byrne, “The Effectiveness of Prison Programming: A Review of the Research Literature Examining the Impact of Federal, State, and Local Inmate Programming on Post-Release Recidivism,” *Federal Probation* 84, no. 1 (2020): 20, https://www.uscourts.gov/sites/default/files/84_1_1_0.pdf [<https://perma.cc/5S4B-B7EB>]. According to one estimate, “even if a person took every approved program the BOP offered . . . she would accrue only 8,143 hours of program time — over 600 hours short of what is necessary to obtain a year’s worth of time credits.” Federal Public and Community Defenders Legislative Committee letter to Federal Bureau of Prisons, Office of General Counsel, re: Comments on Docket No. BOP-1176P, RIN 1120-AB76, “FSA Time Credits,” January 22, 2021, 5, <https://perma.cc/RL3H-38QV>. An Urban Institute analysis found that “only 15 programs” in the BOP catalog “would provide credits of more

than three weeks.” Julie Samuels and Emily Tiry, *Implementation of the First Step Act: The Risk and Needs Assessment System*, Urban Institute, 2021, 4, 17, <https://www.urban.org/research/publication/implementation-first-step-act-risk-and-needs-assessment-system> [<https://perma.cc/ZK43-WVWZ>].

22 In the final rule, BOP acknowledged that certain aspects of the original proposal had been “inconsistent with the goals of the FSA” and corrected many of the concerns raised. The final rule aimed to create a “simpler . . . time credits program . . . that will more fully encourage and reward participation in evidence-based recidivism reduction programs and productive activities.” The final rule awards 10 days of ETCs for each 30-day period in which an eligible person participates in recommended programming, with an additional 5 days awarded to people who meet certain risk assessment criteria. FSA Time Credits, 87 Fed. Reg. 2705, 2706, 2708 (January 19, 2022) (to be codified at 28 C.F.R. pts. 523 & 541), <https://www.federalregister.gov/documents/2022/01/19/2022-00918/fsa-time-credits> [<https://perma.cc/95RC-DPHJ>]. U.S. Department of Justice, “Justice Department Announces New Rule Implementing Federal Time Credits Program Established by the First Step Act,” January 13, 2022, <https://www.justice.gov/opa/pr/justice-department-announces-new-rule-implementing-federal-time-credits-program-established> [<https://perma.cc/GK7C-FSLE>] (noting promulgation of revised regulation and retroactive application); and *Oversight of the Bureau of Prisons, Hearing Before the H. Comm. on the Judiciary*, 117th Cong. (2022) (statement of Michael D. Carvajal, director, Federal Bureau of Prisons), 4, <https://docs.house.gov/meetings/JU/JU08/20220203/114370/HHRG-117-JU08-Wstate-CarvajalM-20220203.pdf> [<https://perma.cc/TD7U-9Z2C>].

23 U.S. Department of Justice, *First Step Act Annual Report*, April 2022, 28–29, <https://www.ojp.gov/library/publications/first-step-act-annual-report-april-2022> [<https://perma.cc/A49Q-8LHN>].

24 FSA Time Credits (2022), 87 Fed. Reg. at 2706, 2708.

25 FSA Time Credits (2022), 87 Fed. Reg. at 2706–08 (implementing rule); 18 U.S.C. § 3632(d)(4) (describing credit eligibility); and 18 U.S.C. § 3624(g) (listing rules for usage of credits).

26 First Step Act of 2018 § 101(a) (codified at 18 U.S.C. § 3632(d)(4)(D)).

27 U.S. Department of Justice, *First Step Act Annual Report*, 28.

28 James M. Byrne et al., *Report of the Independent Review Committee Pursuant to the Requirements of Title I Section 107(g) of the First Step Act (FSA) of 2018 (P.L. 115-391)*, United States Senate Committee on the Judiciary, 2020, 2, <https://firststepact-irc.org/wp-content/uploads/2020/12/IRC-FSA-Title-I-Section-107g-Report-12-21-20.pdf> [<https://perma.cc/3SGR-QZTP>].

29 Erik Ortiz, “Thousands of Federal Inmates Still Await Early Release Under Trump-Era First Step Act,” NBC News, July 3, 2022, <https://www.nbcnews.com/news/us-news/thousands-federal-inmates-still-await-early-release-trump-era-first-st-rcna35162> [<https://perma.cc/Q86A-8M8W>].

30 See First Step Act of 2018, § 101(a) (codified at 18 U.S.C. § 3632(d)(4)(A)) (providing higher amounts of earned time credit for some individuals deemed “minimum” or “low” risk) and (g)(1) (restricting eligibility to transfer to supervised release or prerelease custody to individuals in minimum- or low-risk categories, absent warden approval under specified circumstances).

31 See, e.g., letter from David Patton, co-chair, Federal Public and Community Defenders Legislative Committee, to David B. Muhlhause, director, National Institute of Justice, re: DOJ First Step Act Listening Session on PATTERN, September 13, 2019 (available as Exhibit A to *The Federal Bureau of Prisons and Implementation of the First Step Act, Hearing Before the H. Judiciary Comm. Subcomm. on Crime, Terrorism, and Homeland Security*, 116th Cong. (2019) (statement of David E. Patton), 15, <https://www.congress.gov/116/meeting/house/110089/witnesses/HHRG-116-JU08-Wstate-PattonD-20191017.pdf> [<https://perma.cc/CYX2-6YMR>].

32 See U.S. Department of Justice, *The First Step Act of 2018: Risk and Needs Assessment System*, 2019, <https://nij.ojp.gov/first-step-act/the-first-step-act-of-2018-risk-and-needs-assessment-system.pdf> [<https://perma.cc/EC6L-5NRL>]. For revisions and administration support for further change, see “Revisions to PATTERN (April 2022–Present)” in this report.

33 Pamela M. Casey, Jennifer K. Elek, and Roger K. Warren, *Offender Risk & Needs Assessment Instruments: A Primer for Courts*, National Center for State Courts, 2014, 15, 20–22, https://www.ncsc.org/_data/assets/pdf_file/0018/26226/bja-rna-final-report-combined-files-8-22-14.pdf [<https://perma.cc/7R5V-VXEX>].

34 See letter from David E. Patton, co-chair, Federal Public and Community Defenders Legislative Committee, to Senator Richard Durbin, re: CORRECTIONS Act (S. 467), February 5, 2015, 8 (on file with the authors).

35 Watts et al., *Transforming Prisons, Restoring Lives*, 33–34; and Julia Angwin and Jeff Larson, “Bias in Criminal Risk Scores Is Mathematically Inevitable, Researchers Say,” ProPublica, December 30, 2016, <https://www.propublica.org/article/bias-in-criminal-risk-scores-is-mathematically-inevitable-researchers-say> [<https://perma.cc/NSS9-684G>].

36 See letter from Ames Grawert, senior counsel, Brennan Center for Justice, to David B. Muhlhause, director, National Institute of Justice, re: “Public Comment on ‘PATTERN,’ the First Step Act’s Risk and Needs Assessment Tool,” September 3, 2019, 4–5, <https://www.brennancenter.org/our-work/research-reports/brennan-centers-public-comment-first-step-acts-risk-and-needs-assessment> [<https://perma.cc/73CL-YXZC>].

37 Jessica M. Eaglin, “Constructing Recidivism Risk,” *Emory Law Journal* 67 (2017): 72, 94–99, <https://www.repository.law.indiana.edu/cgi/viewcontent.cgi?article=3647&context=facpub> [<https://perma.cc/6Q2X-YKXF>].

38 Federal Bureau of Prisons, “New FSA Resources Available,” October 26, 2020, https://www.bop.gov/resources/20201026_fsa_update.jsp [<https://perma.cc/NNE7-NWFA>].

39 National Institute of Justice, *2020 Review and Revalidation of the First Step Act Risk Assessment Tool*, 2021, 14, 17–19 and table 2.3, <https://www.ojp.gov/pdffiles1/nij/256084.pdf> [<https://perma.cc/P55N-SKRZ>].

40 While “general” and “violent” recidivism are technically different categories, both focus squarely on arrest risk. No conviction is necessary to qualify as a recidivating event in either case, and in fact the arrest need not even withstand the scrutiny of a grand jury. U.S. Department of Justice, *The First Step Act of 2018: Risk and Needs Assessment System — Update*, 2020, 12–14, <https://nij.ojp.gov/first-step-act/the-first-step-act-of-2018-risk-and-needs-assessment-system.pdf> [<https://perma.cc/BV9D-KL5B>].

41 U.S. Department of Justice, *First Step Act Annual Report*, 13–14.

42 Carrie Johnson, “Justice Department Works to Correct Bias in Deciding Who’s Released from Prison,” NPR, April 19, 2022, <https://www.npr.org/2022/04/19/1093538706/justice-department-works-to-curb-racial-bias-in-deciding-whos-released-from-pris> [<https://perma.cc/QX4F-UY65>]; and U.S. Department of Justice, *The First Step Act of 2018: Risk and Needs Assessment System*, 2019, <https://nij.ojp.gov/first-step-act/the-first-step-act-of-2018-risk-and-needs-assessment-system.pdf> [<https://perma.cc/EC6L-5NRL>].

43 U.S. Department of Justice, *The First Step Act of 2018: Risk and Needs Assessment System — Update*; U.S. Department of Justice, *First Step Act Implementation Fiscal Year 2020 90-Day Report*, 2020, <https://www.ojp.gov/pdffiles1/nij/254799.pdf> [<https://perma.cc/NF77-FV4E>]; U.S. Department of Justice, *The Attorney General’s First Step Act Section 3634 Annual Report*, 2020, <https://www.ojp.gov/Attorney-Generals-First-Step-Act-Section-3634-Annual-Report-December-2020> [<https://perma.cc/6BAN-AVMV>]; National Institute of Justice, *2020 Review and Revalidation*; and National Institute of Justice, *2021 Review and Revalidation of the First Step Act Risk*

- Assessment Tool, 2021, <https://www.ojp.gov/pdffiles1/nij/303859.pdf> [<https://perma.cc/3C6X-KC97>].
- 44** National Institute of Justice, *2020 Review and Revalidation*, 7.
- 45** National Institute of Justice, *2021 Review and Revalidation*, 7–8.
- 46** Ian MacDougall, “Bill Barr Promised to Release Prisoners Threatened by Coronavirus — Even as the Feds Secretly Made It Harder for Them to Get Out,” ProPublica, May 26, 2020, <https://www.propublica.org/article/bill-barr-promised-to-release-prisoners-threatened-by-coronavirus-even-as-the-feds-secretly-made-it-harder-for-them-to-get-out> [<https://perma.cc/Y2V5-SUBE>].
- 47** E. Ann Carson, *Federal Prisoner Statistics Collected Under the First Step Act, 2021*, U.S. Department of Justice, Bureau of Justice Statistics, 2021, 16, table 10, <https://bjs.ojp.gov/library/publications/federal-prisoner-statistics-collected-under-first-step-act-2021> [<https://perma.cc/3YFZ-8GW2>].
- 48** National Institute of Justice, *2021 Review and Revalidation*, 40.
- 49** Grawert, “Public Comment on ‘PATTERN,’” 3–4.
- 50** See Grawert, “Public Comment on ‘PATTERN,’” 2–4 (arguing that defining recidivism based on rearrest incorporates racial disparities in policing); and U.S. Department of Justice, *The First Step Act of 2018: Risk and Needs Assessment System — Update*, 13–14 (considering but rejecting a change to narrower definition of recidivism).
- 51** Byrne et al., *Report of the Independent Review Committee*, 3.
- 52** The DOJ also argued that rearrest is the definition of recidivism “most widely used in the federal system,” and that therefore retaining it for PATTERN “would make comparisons across risk assessment and criminal justice systems easier.” U.S. Department of Justice, *The First Step Act of 2018: Risk and Needs Assessment System — Update*, 2020, 14.
- 53** Samuels and Tiry, *Implementation of the First Step Act*, 9; and Amy Cyphert, “Reprogramming Recidivism: The First Step Act and Algorithmic Prediction of Risk,” *Seton Hall Law Review* 51, no. 2 (2020): 331–381, 351–353, <https://scholarship.shu.edu/cgi/viewcontent.cgi?article=1773&context=shlr> [<https://perma.cc/BB26-YW2C>].
- 54** Michael Tonry, “Legal and Ethical Issues in the Prediction of Recidivism,” *Federal Sentencing Reporter* 26, no. 3 (2014): 167–176, 167, https://scholarship.law.umn.edu/cgi/viewcontent.cgi?article=1535&context=faculty_articles [<https://perma.cc/R4KX-SKND>].
- 55** U.S. Department of Justice, *First Step Act Implementation Fiscal Year 2020*, 2. For criticisms, see Samuels and Tiry, *Implementation of the First Step Act*, 9.
- 56** U.S. Department of Justice, *The Attorney General’s First Step Act Section 3634 Annual Report*, 24.
- 57** Emily Tiry and Julie Samuels, “Three Ways to Increase the Impact of the First Step Act’s Earned Time Credits,” Urban Institute, April 30, 2021, <https://www.urban.org/urban-wire/three-ways-increase-impact-first-step-acts-earned-time-credits> [<https://perma.cc/7LES-WELQ>].
- 58** National Institute of Justice, *2020 Review and Revalidation*, 14.
- 59** National Institute of Justice, *2021 Review and Revalidation*, 28–29. See also Department of Justice, *First Step Act Annual Report*, 10.
- 60** For more on this subject, see, e.g., Grawert, “Public Comment on ‘PATTERN,’” 5–7.
- 61** U.S. Department of Justice, *First Step Act Annual Report*, 5, 9–12.
- 62** U.S. Department of Justice, *First Step Act Annual Report*, 12–18.
- 63** U.S. Department of Justice, *First Step Act Annual Report*, 14.
- 64** U.S. Department of Justice, *First Step Act Annual Report*, 12, 11 (“Thus, transitioning to PATTERN 1.3 will neither exacerbate nor solve these racial bias issues.”), and 20 (noting that changes to cut points “will not directly correct the racial differential prediction rates found to be associated with the PATTERN tool”).
- 65** *The First Step Act, the Pandemic, and Compassionate Release: What Are the Next Steps for the Federal Bureau of Prisons?*, Hearing Before the H. Judiciary Comm. Subcomm. on Crime, Terrorism, and Homeland Security, 117th Cong. (2022) (statement of Melissa Hamilton, professor of law and criminal justice, University of Surrey), 4, <https://docs.house.gov/meetings/JU/JU08/20220121/114349/HHRG-117-JU08-Wstate-HamiltonM-20220121.pdf> [<https://perma.cc/2Y8F-4TBZ>].
- 66** Department of Justice, *First Step Act Annual Report*, 20 (“The updated version of the tool and the revised cut points are anticipated to result in expanded opportunities for the earning of time credits under the FSA system, which will in turn incentivize a greater number of inmates to participate in evidence-based recidivism reduction programs and productive activities and reduce their predicted risk of recidivism even further.”).
- 67** Department of Justice, *First Step Act Annual Report*, 13–14.
- 68** National Institute of Justice, *Stakeholder Statements Submitted in Response to NIJ’s First Step Act Listening Sessions*, 2019, <https://www.ojp.gov/pdffiles1/nij/253115.pdf> [<https://perma.cc/25E8-DPDE>] (noting stakeholder requests for greater transparency).
- 69** Department of Justice, *First Step Act Annual Report*, 11–12.
- 70** 18 U.S.C. § 3632(a)(3).
- 71** Federal Bureau of Prisons, *First Step Act: Initial Review of the SPARC-13 Needs Assessment System*, U.S. Department of Justice, 2022, 14, https://www.bop.gov/inmates/fsa/docs/bop_fsa_needs_validation_report_2021.pdf [<https://perma.cc/XS45-GDX2>].
- 72** Federal Bureau of Prisons, *Initial Review of SPARC-13*, 12, table 4.
- 73** Federal Bureau of Prisons, *Initial Review of SPARC-13*, 10–13.
- 74** See *The Federal Bureau of Prisons and Implementation of the First Step Act* (statement of David E. Patton), 1, 5–7.
- 75** U.S. Department of Justice, *First Step Act Annual Report*, 33–35; and Byrne et al., *Report of the Independent Review Committee*, 1.
- 76** Byrne et al., *Report of the Independent Review Committee*, 2–3.
- 77** FSA Time Credits (2022), 87 Fed. Reg. at 2711.
- 78** Carson, *Federal Prisoner Statistics Collected*, 17–20 and table 11. This may have changed with the development of new programs and rollout of the January 2022 ETC rule. See U.S. Department of Justice, *First Step Act Annual Report*, 32–33 (“The passage of the Time Credits rule in January 2022 expanded the services which qualify as PAs.”).
- 79** Department of Justice, *First Step Act Annual Report*, 32–36, 45–48.
- 80** While the First Step Act authorized up to \$75 million annually for programming, Congress failed to appropriate that amount for the first year of implementation, forcing the DOJ to move money from elsewhere in its budget. Ames Grawert, “What Is the First Step Act — and What’s Happening With It?,” Brennan Center for Justice, June 23, 2020, <https://www.brennancenter.org/our-work/research-reports/what-first-step-act-and-whats-happening-it> [<https://perma.cc/7M3A-E9AU>].
- 81** See Amy Baron Evans et al., “Federal Defender Analysis of Corrections Act of 2015,” *Federal Sentencing Reporter* 28, no. 2 (2015): 151–153, 151 (“The system in the bill is novel and untested. State correctional systems . . . award time credits for participating in programs based on performance and/or disciplinary record, not risk assessment scores.”).

ABOUT THE AUTHORS

► **Ames Grawert** is senior counsel and John L. Neu Justice Counsel in the Brennan Center’s Justice Program. He leads quantitative and policy research focused on trends in crime and the collateral costs of mass incarceration. Additionally, he advocates for criminal justice reform at the state and federal levels. Previously Grawert served as an assistant district attorney in the Appeals Bureau of the Nassau County (New York) District Attorney’s Office, where he reviewed and litigated claims of actual innocence in addition to appellate work. Before entering public service, he was an associate at Mayer Brown LLP, where he represented people pro bono in post-conviction litigation.

► **Patricia L. Richman** is national sentencing resource counsel to the Federal Public and Community Defenders, where she leads legislative work. Previously Richman served as crime counsel to Sen. Dick Durbin on the Senate Judiciary Committee and as an assistant federal public defender in Maryland. Before that, Richman was an associate at Williams & Connolly LLP, a clerk for a federal district court judge in Philadelphia, and a special education teacher in Prince George’s County, Maryland.

ABOUT THE BRENNAN CENTER’S JUSTICE PROGRAM

The Brennan Center’s Justice Program seeks to secure our nation’s promise of equal justice for all by creating a rational, effective, and fair justice system. Its priority focus is to reduce mass incarceration. The program melds law, policy, and economics to produce new empirical analyses and innovative policy solutions to advance this critical goal.

ACKNOWLEDGMENTS

The Brennan Center gratefully acknowledges the Tow Foundation for its generous support of our work. This is an independent Brennan Center publication; the opinions expressed are those of the authors and do not necessarily reflect the views of our supporters.

The authors thank Shelley Fite for her input on federal appellate practice and recent developments in First Step Act case law, as well as Melissa Hamilton for her analysis of risk assessment tools and Sean Wilson for his insights into how new regulations have impacted prison staff and imprisoned people. The authors also thank Lauren-Brooke Eisen and Ram Subramanian for their strategic guidance of the report, as well as their careful revisions and insightful comments; Michael Waldman and John Kowal for their support of this work; and Calvin Chappell, Noah Kim, and Stephanie Wylie for their careful research and editing support. The authors also thank Maris Mapolski for her in-depth review of multiple drafts of this analysis. Last, the authors are grateful to Lisa Benenson, Matt Harwood, Zachary Laub, Alexandra Ringe, Janet Romero-Bahari, Derek Rosenfeld, and Alden Wallace for their review of this report, communications assistance, and work to ensure its successful publication.

August 2022

Advancing Fairness and Transparency: National Guidelines for Post-Conviction Risk and Needs Assessment

Sarah L. Desmarais, David A. D'Amora, Lahiz P. Tavárez



Justice Center
THE COUNCIL OF STATE GOVERNMENTS



BJA
Bureau of Justice Assistance
U.S. Department of Justice

Table of Contents

| | | | |
|--|-----------|--|-----------|
| Dedication | 3 | Guideline 7: Adopt agencywide strategies to minimize the potential that local implementation of a post-conviction risk and needs assessment instrument could promote disparities. | 25 |
| Acknowledgments | 3 | Principle III: Transparency | 28 |
| Introduction | 4 | Guideline 8: Provide system stakeholders with relevant information on the development, intended use, and validation of the post-conviction risk and needs assessment instrument. | 29 |
| Methodology | 5 | Guideline 9: Develop a written policy that guides the local use of the post-conviction risk and needs assessment instrument. | 32 |
| Usage | 5 | Guideline 10: Communicate the strengths and limitations of the post-conviction risk and needs assessment instrument to the general public. | 34 |
| Timing | 5 | Principle IV: Effective Communication and Use | 37 |
| Purpose | 6 | Guideline 11: Anchor communication of post-conviction risk and needs assessment results in the RNR principles. | 38 |
| Intended Audience | 6 | Guideline 12: Contextualize the results of the post-conviction risk and needs assessment instrument. | 40 |
| Principle I: Accuracy | 7 | Guideline 13: Develop a template for communicating the individual results of the post-conviction risk and needs assessment instrument to all relevant stakeholders, including the person being assessed. | 42 |
| Guideline 1: Conduct a local evaluation of the post-conviction risk and needs assessment instrument to ensure that the instrument is suitable for the agency’s population. | 8 | Glossary | 45 |
| Guideline 2: Meet the minimum performance thresholds of post-conviction risk and needs assessments completed in the field according to statistical standards. | 11 | Bibliography | 51 |
| Guideline 3: Use a continuous quality improvement process to ensure successful implementation of the post-conviction risk and needs assessment instrument. | 13 | | |
| Guideline 4: Use a multistep approach to assess risk and needs over time. | 17 | | |
| Principle II: Fairness | 19 | | |
| Guideline 5: Examine the results of the post-conviction risk and needs assessment instrument for predictive bias and disparate impact across groups. | 20 | | |
| Guideline 6: Apply post-conviction risk and needs assessment instrument results to individual cases in keeping with the Risk-Need-Responsivity (RNR) principles. | 23 | | |

Dedication

We are pleased to dedicate this project to Dr. Ed Latessa. Dr. Latessa spent his entire career working to improve the criminal justice system, especially in the areas of risk and needs assessment and correctional programming. He was an invaluable advisor to The Council of State Governments (CSG) Justice Center and the Bureau of Justice Assistance within the U.S. Department of Justice's Office of Justice Programs for well over a decade. His advice and counsel improved our knowledge and strengthened the impact of our work in the field. Dr. Latessa transformed the fields of community supervision and correctional habilitation and treatment in the United States, and for that, we are eternally grateful.

Acknowledgments

This project was accomplished through the collaborative efforts of researchers, risk and needs assessment instrument developers, practitioners, and leaders in the field who gave generously of their time and expertise over the course of 2.5 years. It draws on an extensive review of literature and related research, observations from the field, feedback from national experts, several multidisciplinary forums and advisory group discussions, and a rigorous review process.

Thank you to this publication's group of expert advisors whose input greatly informed the development of these guidelines. The following collaborators provided insights and resources that were truly invaluable:

Project Credits

Writing: **Sarah L. Desmarais**, Policy Research Associates; **David A. D'Amora**, CSG Justice Center; **Lahiz P. Tavárez**, CSG Justice Center

Advising: **Elizabeth Lyon**, **Marshall Clement**, **Megan Quattlebaum**, CSG Justice Center

Editing: **Katy Albis**, **Leslie Griffin**, **Emily Morgan**, CSG Justice Center

Public Affairs: **Aisha Jamil**, CSG Justice Center

Design: **Michael Bierman**

Web Development: **Eleventy Group**

Advisory Group

Ashlee Barnes, Virginia Commonwealth University

Mark Bergstrom, Pennsylvania Commission on Sentencing

Guy Bourgon, Public Safety Canada (ret.), Clinical and Consulting Psychology Practice

Kristofer "Bret" Bucklen, Pennsylvania Department of Corrections

Christina Campbell, University of Cincinnati

Bree Derrick, Idaho Department of Correction

Kevin Douglas, Simon Fraser University

Grant Duwe, Minnesota Department of Corrections

Joel Dvoskin, University of Arizona College of Medicine

Carey Green, Criminal Justice Assistance Division, Texas Department of Criminal Justice

Stephen Hart, Protect International Risk and Safety Services

Kirk Heilbrun, Drexel University

Natalie Jennifer Jones, Evaluation Division of the Social Sciences and Humanities Research Council of Canada

KiDeuk Kim, Urban Institute

Daryl Kroner, Southern Illinois University

Edward Latessa, University of Cincinnati

Brian Lovins, Justice System Partners

Eyitayo Onifade, Clark Atlanta University

Osonde Osoba, Center for Scalable Computing and Analysis, Pardee RAND Graduate School

Sarah Picard, Center for Court Innovation

Charles Robinson, Administrative Office of the United States Courts

Ralph Serin, Criminal Justice Decision Making Laboratory, Carleton University

Christopher Slobogin, Vanderbilt University Law School

Raymond "Chip" Tafrate, Central Connecticut State University

Faye Taxman, George Mason University

Gina Vincent, University of Massachusetts Medical School, Center for Mental Health Services Research

This project was supported by Grant No. 2019-ZB-BX-K002 awarded by the Bureau of Justice Assistance. The Bureau of Justice Assistance is a component of the Department of Justice's Office of Justice Programs, which also includes the Bureau of Justice Statistics, the National Institute of Justice, the Office of Juvenile Justice and Delinquency Prevention, the Office for Victims of Crime, and the Office of Sex Offender Sentencing, Monitoring, Apprehending, Registering, and Tracking. Points of view or opinions in this document are those of the author and do not necessarily represent the official position or policies of the U.S. Department of Justice.

Introduction

Risk and needs assessment instruments are widely used in criminal justice settings across the United States to inform decisionmaking at various stages in case processing. These guidelines were developed to promote accuracy, fairness, transparency, and effective communication and use of risk and needs assessment instruments to inform decisionmaking following conviction.

Methodology

The guidelines are drawn from empirical research and reflect the perspectives of an advisory group comprising nearly 30 researchers, practitioners, and policymakers. Advisory group members participated in several discussions on key concerns in the development, validation, and implementation of post-conviction risk and needs assessment instruments. These guidelines also consider existing statements and guidance on the use of risk and needs assessment instruments written by diverse groups and for diverse contexts.

Usage

These guidelines pertain to the use of post-conviction risk and needs assessment instruments to inform decisions and case planning that occur after court disposition—specifically, after conviction and sentencing. They may also be used in the application of assessment results to inform decisionmaking and case planning in the context of alternative forms of criminal justice processing, such as after a decision has been made to offer a diversion program.

Timing

Many of the guidelines describe processes or steps that should ideally be taken prior to implementation of a post-conviction risk and needs assessment instrument. But many corrections and community supervision agencies are already using a post-conviction risk and needs assessment instrument. For those agencies, we recommend:

- (1) Conducting an informal review of the extent to which current policies and practices adhere to the action items under each guideline.
- (2) Setting specific goals and identifying priorities for implementing the remaining guidelines over a realistic time period.

Whatever the setting, the guidelines presume that the intended use of post-conviction risk and needs assessment instruments is to support accurate, fair, and transparent decisions regarding a person's risk of recidivism. These guidelines also presume that the purpose of post-conviction risk and needs assessment instruments is, ultimately, to promote public safety and positive outcomes for people in the criminal justice system through the least restrictive means possible.

Purpose

The purpose of this publication is to provide the information that criminal justice agencies need to ensure that the implementation of post-conviction risk and needs assessment instruments promotes accuracy, fairness, transparency, and effective communication and use. The sections of the publication are organized by these four principles. Each section provides the rationale for the related guideline, recommends actions that should be taken or requirements that must be met to follow the guideline, and reviews practical considerations for planning, implementation, and continuous quality improvement (CQI).

Intended Audience

The intended audience for this publication includes people who support agency administrators, supervisors, and other stakeholders involved in selecting or implementing post-conviction risk and needs assessment instruments, the development of related policy, and decisions regarding their ongoing use. These individuals may include trainers, quality assurance personnel, research partners, or other consultants. The content is also relevant for developers of post-conviction risk and needs assessment instruments and researchers or evaluators who may evaluate the performance of assessment results in studies or in practice. Assessors and their supervisors additionally may find utility in the information presented here to support training and CQI-related efforts.

Other stakeholders may find that some of the additional information provided herein supports a deeper understanding of the guidelines. For instance, system actors (e.g., judges, attorneys, service providers, or probation/parole officers) or people in contact with the criminal justice system may find that this publication helps them understand what informs the application of assessment results in individual case processing.

Principle I: Accuracy

The first four guidelines speak to strategies that agencies can use to promote accuracy in the use of post-conviction risk and needs assessment instruments. Accuracy refers to the degree to which the assessment results predict the recidivism outcomes they were designed to predict, as measured in relation to the observed rate and severity of criminal behavior. Promoting accuracy also involves considering whether the post-conviction risk and needs assessment instruments are completed and used as intended to inform case decisions and planning within facilities and in the community.

“

Criminal justice is a domain where it is imperative to exercise maximal caution and humility in the deployment of statistical tools. (Partnership on AI, 2019, 33)

”

We recommend the following guidelines to promote accuracy of post conviction risk and needs assessment instruments:

1. Conduct a local evaluation of the post conviction risk and needs assessment instrument to ensure that the instrument is suitable for the agency’s population.
2. Meet the minimum performance thresholds of post conviction risk and needs assessments completed in the field according to statistical standards.
3. Use a continuous quality improvement (CQI) process to ensure successful implementation of the post conviction risk and needs assessment instrument.
4. Use a multistep approach to assess risk and needs over time.

Guideline 1: Conduct a local evaluation of the post-conviction risk and needs assessment instrument to ensure that the instrument is suitable for the agency's population.

Overview

This guideline establishes that the post-conviction risk and needs assessment instrument can be completed reliably and with acceptable levels of accuracy in predicting the outcome(s) of interest in practice. Many different factors can affect the reliability and validity of the assessment results, including assessors, information used to complete the assessment, resources, policies, and practices. A local evaluation may produce information that not only bears on the reliability and validity of the assessments but also elucidates potential issues or concerns in local practices and policies that should be addressed.

Action Items

Establish inter-rater reliability prior to using the instrument.

Inter-rater reliability is relevant to any assessment instrument to ensure that the assessment is completed consistently and accurately.¹ Indeed, establishing that assessments can be completed with consistency across independent assessors is a necessary criterion for establishing validity.² Even if the assessment does not involve an interview and most—or even all—of the items are completed using official records, there may be differences in how that information is extracted and interpreted by assessors or errors may occur in the coding process. To that end, all individuals who will be conducting the assessments should complete a minimum of three practice cases to consensus after training and prior to using the instrument in the field.³

There are different ways to ensure that these practice cases occur. For instance, the trainers hired to conduct the pre-service training (see [Guideline 3](#)) may provide additional case materials and ratings that agencies may use for these practice cases.⁴ If not, agency representatives may ask the expert trainer or other qualified professional to help them develop practice cases and ratings. Alternatively, agencies may develop case studies that experienced in-house assessors have coded and use them for consensus ratings as new assessors are trained. Another strategy may be to have new assessors review case materials and complete the assessment for an individual who is currently being assessed—or has recently been assessed—by a more experienced assessor in the agency. Assessors should have these opportunities to practice *after training* but *before use in practice* to increase their understanding of the assessment process, get feedback on their ratings, and gain experience. Information on specific metrics for determining whether inter-rater reliability is acceptable is provided later on (see [Guideline 2](#)).

1. Douglas et al., "Clinical Forensic Psychology."

2. Douglas, Skeem, and Nicholson, "Research Methods in Violence Risk Assessment," 325–46; Gottfredson and Moriarty, "Statistical Risk Assessment," 178–200.

3. Vincent, Guy, and Grisso, *Risk Assessment in Juvenile Justice*.

4. Ibid.

Complete a local validation, ideally prior to using the instrument, to confirm that the assessment results are predicting recidivism using local data and in the context of current and local practices.

Agencies should complete a local validation to demonstrate [predictive validity](#). Predictive validity is not a property of the post-conviction risk and needs assessment instrument itself but rather a property of the assessment results.⁵ So, while we can be confident that an instrument already validated in large research studies or in other jurisdictions will produce reliable and valid assessment results, we cannot assume that the same level of reliability and predictive validity will be achieved locally. Factors that can affect the validity and reliability of assessment results include local record keeping practices; assessor attitudes, training, and knowledge; and variations in penal codes and base rates of recidivism.⁶ Indeed, absent the necessary information and time, implementation with fidelity—and, consequently, reliability and validity—is not possible, even with highly motivated, knowledgeable, and well-trained staff.

Agencies should conduct a local validation study to establish performance in relation to jurisdiction-specific rates of recidivism, ideally prior to using the instrument.⁷ This may be possible to achieve through a retrospective study design, for example, if assessors can extract the information needed to complete the assessment and document recidivism from existing records (e.g., jail/prison records, court records, etc.).

If the assessment requires information that is not available in local records or requires an interview, then an alternative study design would be to conduct a pilot implementation with a subset of cases and test the validity of these assessments prospectively (i.e., looking forward) prior to a full-scale implementation. A sample size of 500 people would likely be sufficient for the local validation.⁸ Further discussion of the research methods for conducting local validations are available elsewhere (see, for example, the [Public Safety Risk Assessment Clearinghouse](#)). We provide information on specific metrics for determining whether predictive validity is acceptable later on (see [Guideline 2](#)).

Revalidate assessment results at least every 5 years—or sooner if there are major policy or population changes—to verify that the assessment results continue to meet minimum performance thresholds.

Regular revalidation will ensure that the assessment instruments continue to be completed as intended and that the results continue to demonstrate acceptable reliability and validity. There may be changes over time in the reliability and validity of assessment results for both expected and unexpected reasons. In particular, there may be meaningful changes in the makeup of the criminal justice population over time because of reforms in policing, charging, or prosecution, for example. Such changes will call for a re-examination of the assessment results to ensure that they continue to meet minimum performance thresholds (see [Guideline 2](#)).

5. American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, *Standards for Educational and Psychological Testing*.

6. Mills, Jones, and Kroner, "An Examination of the Generalizability of the LSI-R and VRAG Probability Bins," 565–85; de Vogel and de Vries Robbé, "Adapting Risk Assessment Tools to New Jurisdictions," 26–39.

7. Vincent, Guy, and Grisso, *Risk Assessment in Juvenile Justice*.

8. For more information on this sample size estimation, see Hanson et al., *A Five-Level Risk and Needs System*.

Even in the absence of policy and practice changes, we recommend revalidation of assessment results at routine intervals. There will inevitably be drift from coding and administration protocols as time since training elapses or with staffing changes and turnover. As part of the planning process—ideally, prior to using the instrument—agencies should establish a timeline for revalidation and identify and allocate resources and staffing to support the revalidation. We recommend that revalidation occur at least every 5 years to balance system demands and resources and to allow sufficient time from implementation to evaluation for recidivism outcomes to be observed and documented.

Consult with experts such as university partners or other experienced evaluators, as needed, to ensure that local evaluations adhere as much as possible to best practices in risk and needs assessment research and standards in test validation.

There are many factors related to the design and methods of an evaluation that affect the reliability and validity of its findings.⁹ To that end, there are established standards that should be applied, to the extent possible, to ensure that local validations are conducted in a sufficiently rigorous manner. These standards are found in psychological and educational testing,¹⁰ accepted practices in risk assessment research methods,¹¹ and guidelines for reporting risk assessment research methods and findings.¹² Applying such standards will also promote the likelihood that the evaluation’s findings are an accurate reflection of the performance of the post-conviction risk and needs assessment results and cannot be attributed to other factors.

Agencies may not have the in-house expertise, resources, and knowledge to design and field a validation study that would stand up to public and peer review. We recommend that agencies consult with experts such as university partners or other experienced evaluators to inform the methods of their local evaluation efforts.

9. For instance, validation studies should account for time at risk and length of follow-up. Time at risk refers to the amount of time for which an individual may actually be able to engage in criminal behavior and length of follow-up refers to the period from assessment to the end of the follow-up. Time at risk and follow-up periods are critical for understanding the base rates of criminal behavior. For some individuals in the study, these values may be the same; for others they may be different. For example, if someone is assessed at the point of admission to a prison, incarcerated for 2 years, and then followed for another 2 years in the community, the follow-up period would be 4 years, but actual time at risk is only 2 years.

10. American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, *Standards for Educational and Psychological Testing*.

11. Douglas, Skeem, and Nicholson, “Research Methods in Violence Risk Assessment,” 325–46.

12. Singh et al., “Reporting Guidance for Violence Risk Assessment Predictive Validity Studies,” 15–22.

Guideline 2: Meet the minimum performance thresholds of post-conviction risk and needs assessments completed in the field according to statistical standards.

Overview

Establishing that post-conviction risk and needs assessments completed in the field meet the minimum performance thresholds according to statistical standards is critical to evaluating accuracy. The performance metrics provided reflect well-established statistical standards for measuring the strength or degree of agreement not only among assessors but also between the assessment results and recidivism. To be clear, these are minimum performance thresholds. Agencies may elect to require more—but not less—stringent performance thresholds than the minimums provided here as a matter of policy for all cases or for specific contexts.

Action Items

Demonstrate good agreement or better among assessors for post-conviction risk and needs assessments completed in the field.

Agreement among assessors may be evaluated using different statistical approaches, depending on the rating or scoring of interest, and each approach may have advantages and disadvantages. For categorical ratings such as yes/no or low/moderate/high, the level of observed agreement is the most straightforward and easiest to calculate (number of agreements/number of agreements and disagreements), but it does not account for expected agreement. [Kappa](#) considers both observed and expected agreement, but it can produce erroneous results when there is little to no response variability.¹³ [Intra-class coefficient \(ICC\)](#) looks at whether assessors rate individuals similarly on a continuous scale as opposed to across categories.

The guidelines for interpreting the strength or practical significance of assessor agreement reflect well-established standards in social and epidemiological sciences.¹⁴ Based upon these standards, [good agreement](#) is indicated by:

- Observed agreement among assessors of 80 percent or greater
- Kappa = .60–.74
- ICC = .60–.74

For agencies wishing to adopt more stringent criteria, [excellent agreement](#) is indicated by:

- Observed agreement among assessors of 90 percent or greater
- Kappa = .75–1.00
- ICC = .75–1.00

13. Feinstein and Cicchetti, "High Agreement but Low Kappa," 543–49.

14. Cicchetti, "The Precision of Reliability and Validity Estimates Re-Visited," 695–700.

Demonstrate good validity or better in predicting the likelihood of recidivism with post-conviction risk and needs assessments completed in the field.

To evaluate the performance of post-conviction risk and needs assessments in predicting the likelihood of recidivism, consider two different metrics:¹⁵ (1) the observed rates of criminal behavior at each risk level and (2) an overall index of predictive validity.

The observed rate of criminal behavior at each risk level represents a simple calculation examining the proportion of individuals who went on to recidivate within each risk level. To demonstrate, if 75 people were classified as Risk Level 3 and 5 of them recidivated, then the observed rate of criminal behavior at Risk Level 3 is 6.7 percent (i.e., $5/75 \times 100$). There are no set performance standards or established benchmarks for what would be “good or better” for this metric. However, as risk levels increase, so too should the observed rates of criminal behavior. In other words, at higher levels, we would expect to see higher rates of criminal behavior than at lower levels. Additionally, observed rates of criminal behavior should increase at each subsequent risk level. For example, if the observed rate of criminal behavior at Risk Level 3 is 6.7 percent, then the observed rate of criminal behavior at Risk Level 2 should be less than 6.7 percent, and at Risk Level 4 greater than 6.7 percent.¹⁶

Because the observed rate of criminal behavior at each risk level is a purely descriptive metric,¹⁷ we recommend using the areas under receiver operating characteristic curves (AUCs) to provide an overall index of predictive validity. AUCs represent the likelihood that a randomly selected individual who recidivated during the follow-up period received a higher risk score than a randomly selected individual who did not recidivate during the follow-up period. AUC is a preferred metric because its values are not affected by rates of recidivism to the same degree that other metrics are influenced by rates of recidivism.¹⁸ The guidelines for interpreting the strength or practical significance of AUCs, again, reflect well-established research standards.¹⁹ Based upon these standards, good validity will be indicated by AUC values of .65–.70. For agencies wishing to adopt more stringent criteria, excellent validity will be indicated by AUC values of .71–1.00.²⁰

15. While there are additional metrics that agencies and others may wish to examine, such as the false positive and false negative rates or the positive predictive value (PPV) and negative predictive value (NPV), we do not recommend their use. We highlight two issues of particular importance here. First, post-conviction risk and needs assessment instruments do not make binary predictions about future criminal behavior (e.g., yes or no), nor do they make binary decisions about an individual (e.g., detain or release). Rather, they estimate the likelihood of recidivism using multiple categories or levels to provide decisionmakers with the information necessary to make such decisions. Consequently, these metrics do not reflect how the instruments are designed or intended to be used. Second, these metrics are dependent upon sample size and recidivism rates. As a result, values that may be interpreted as reflecting poor validity could instead represent errors in a small number of cases or successful mitigation of recidivism. More specifically, PPV and NPV are based upon a single threshold or cutoff, but there are no post-conviction risk and needs assessment instruments that use a single threshold or cutoff. Instead, they typically use at least three risk levels or categories. To calculate the PPV and NPV, then, requires selecting a threshold; this may include the use of a single numerical score or risk level as the threshold, or the PPV and NPV may be calculated for each risk level. The former is what is done most frequently; however, this does not reflect how the assessment results are used in practice. The latter is more akin to how the instruments are used in practice, but the calculated values will be affected by the relatively small number of cases at each level. Specifically, even a small number of “errors” may dramatically affect the observed PPV or NPV. Because fewer individuals are typically assessed at higher relative to lower risk levels, this means that even within a single validation study, the estimates of PPV and NPV will be less stable for higher than lower risk levels. Further, the base rate of recidivism in a given jurisdiction puts boundaries on the possible range of values: PPV will increase with increases in the prevalence of recidivism, while NPV will decrease with increases in recidivism. This means that in jurisdictions with relatively low rates of recidivism, it is not possible to observe high PPVs. Only with higher rates of recidivism will higher PPVs be observed. The converse is true for NPV.
16. Ideally, the increase in the observed rate of criminal behavior would be statistically significant from one level to the next; however, this may not be realistic if there are small numbers of people assessed at each level and low base rates of criminal behavior. Consequently, a substantive increase in the observed rate of criminal behavior from one level to the next is sufficient.
17. Singh, “Predictive Validity Performance Indicators in Violence Risk Assessment,” 8–22; Helmus and Babchishin, “Primer on Risk Assessment,” 8–25.
18. Smith, “The Effects of Base Rate and Cutoff Point Choice,” 83–11; Rice and Harris, “Comparing Effect Sizes in Follow-up Studies,” 615–20.
19. Cohen’s *d* is a measure of the difference between the averages of two groups. It is the most commonly used measure against which to interpret the strength of association in the social and epidemiological sciences.
20. An AUC value of 1.00 indicates perfect discrimination between those who went on to recidivate from those who did not recidivate during follow-up period; .50 indicates discrimination at chance levels; and 0.00 indicates completely incorrect discrimination (i.e., all those who did not recidivate were identified as higher risk for recidivism while those who did recidivate were identified as lower risk and vice versa).

Guideline 3: Use a CQI process to ensure successful implementation of the post-conviction risk and needs assessment instrument.

Overview

Even the most well-established, vetted, and validated post-conviction risk and needs assessment instrument may fail to produce the desired results if not implemented with fidelity. Successful implementation will require significant planning and resources at the outset, as well as the establishment and deployment of strategies to monitor the implementation and assessment processes over time. Deliberate, pre-planned CQI efforts will allow for prompt identification of issues that may interfere with the effectiveness of post-conviction risk and needs assessment instruments and enable the deployment of strategies to address those issues, thereby promoting the accuracy of assessment results.

Action Items

Document the protocols for applying the post-conviction risk and needs assessment instrument.

Protocols for administration should document how, when, and for whom and by whom assessments will be completed. Agencies should develop and document these administration protocols as part of the planning process, ideally before using the instrument. For agencies that have already implemented a post-conviction risk and needs assessment instrument, documentation of current administration protocols should be prioritized and completed over a short, but feasible, timeframe. Doing so will not only help promote accuracy in the assessments, ensuring that they are completed as intended, but also will provide clarity and transparency on the appropriate—and inappropriate—use of the post-conviction risk and needs assessment instrument.

The protocols should describe the required and recommended sources of information to use to complete the ratings. They also should describe which individuals should be assessed, when assessments should be completed for them, and what decisions and processes the results should inform.²¹ Additionally, protocols should describe when re-assessments should be conducted; this may include specification of the timeframe for routine re-assessment (e.g., every 6 months)²² or certain conditions that would prompt re-assessment (e.g., change in relationships, employment, housing, health, legal status, etc.).²³ For more on re-assessment, see [Guideline 4](#).

21. For example, post-conviction risk and needs assessments may be required for individuals in specific programs or charged with certain offenses. In terms of timing and decisions, there may be a requirement to complete an initial assessment within 2 weeks of intake to a new program or agency to inform case planning or within 4 weeks of release from a program or setting to inform release planning.

22. Barnes et al., "Validity of Initial, Exit, and Dynamic Juvenile Risk Assessment," 21–38; Lloyd et al., "Reassessment Improves Prediction of Criminal Recidivism," 568–81.

23. Buchanan et al., "Psychiatric Violence Risk Assessment," 340.

In addition to documenting the administration protocols, we recommend that agencies document protocols for how, when, and by whom CQI will be conducted. These CQI protocols should ideally be developed and documented in consultation with diverse stakeholders—including instrument developers or other experts, staff, supervisors, and administrators—before using the instrument. However, they should also be revisited periodically as there may be emergent issues that call for changes in the frequency or focal points of CQI-related efforts. Thinking through and planning for CQI before using the instrument will ensure that the necessary data, resources, and staffing are available to support CQI over time.

Prior to their use of the instrument in practice, provide all assessors with training on the rating procedures and protocols for applying assessment results to inform case plans.

Assessors should complete all required training prior to using a post-conviction risk and needs assessment instrument in practice. The minimum training requirements for administration may be specified by the instrument developers, but at the very least should include training on the strategies to gather and interpret information, procedures for rating items, and how to interpret results. Typically, trainings on these fundamentals are provided by experts, including the instrument developers or others who are well trained and qualified in using the instrument. They may be offered live in person, online, or available on demand via a licensed provider or organization. These training options and modalities will be determined largely as a function of the instrument selected, as well as agency resources, needs, and practical considerations (e.g., staff schedules, onboarding processes, certification or credentialing requirements, etc.). Assessors additionally should complete one (or more) practice cases. Either during or after training, most trainers will provide the practice case materials and the experts' "gold standard" ratings against which to compare trainees' assessments.²⁴ In total, assessors should complete and receive feedback on at least four to five practice cases before they begin to use the instrument in the field (i.e., one or two in the context of the training and three after completing the training; see [Guideline 1](#)).

Assessors also should receive training on the site-specific policies and protocols for applying assessment results to inform case plans before they begin using the instrument in the field. This training can be completed in conjunction with or after the training on the administration of the instrument. It should cover local policies regarding for whom and when post-conviction risk and needs assessments will be conducted and for what purpose(s), ensuring that these uses match the tasks for which the instrument was developed. This training should be provided by local experts, supervisors, or other administrators involved in developing and overseeing the implementation of the post-conviction risk and needs assessment instrument, as appropriate.

Supervisors and others who will be involved in or affected by the implementation should receive some level of training on the procedures and protocols for the use of the post-conviction risk and needs assessment instrument. This may include a short, overview presentation on the basic approach and use of the post-conviction risk and

24. Vincent, Guy, and Grisso, *Risk Assessment in Juvenile Justice*.

needs assessment instrument for other agency staff, decisionmakers, or community representatives. Supervisors will need sufficient knowledge of how to conduct the assessment and use the assessment results in decisions and case planning to be able to conduct CQI, including [case reviews](#) and booster sessions, as described below. It may not be necessary for them to attend all the trainings or to complete all the practice cases. However, the more supervisors and other local leaders demonstrate commitment to the use of the post-conviction risk and needs assessment instrument, the more buy-in there may be from staff and other stakeholders.

Agencies that already use an instrument should develop a strategy to provide this training within a 6-month period.

Complete case reviews at least twice yearly during implementation to identify problems to correct through individual coaching or booster training.

At least twice per year, supervisors should conduct case reviews that examine:

- Fidelity to the rating and scoring guidelines.
- Adherence to the implementation protocols.
- Concordance among assessment results and case decisions, resource allocation, and service provision.

As part of the planning process, ideally before using the instrument, we recommend that agencies develop a case review checklist that includes observable indicators of fidelity issues or errors in the assessments themselves (e.g., missing ratings, inconsistencies between item scores and risk levels) as well as the steps in the assessment process (e.g., collecting information from records and interviews, if appropriate). The checklist should also include items for documenting the population(s) that should be assessed, if appropriate, and the timing of the assessments as detailed in the local administration protocols. Finally, the checklist should include items that speak to whether case decisions, resource allocation, and service provision are in line with the assessment results. We recommend the application of the Risk-Need-Responsivity model²⁵ as a framework for examining concordance between assessment results and practices; see [Guideline 6](#) for more on this.

The **Risk Assessment Quality Improvement (RAQI) protocol**²⁶ provides a starting point for structuring the case review checklist and CQI process. However, agency leaders may wish to consult with the developers of the specific instrument they use or other experts to ensure that they have adequately captured the relevant issues in their local case review checklist. When case reviews reveal issues that need to be addressed, supervisors may wish to address them at the individual or group level, as they come up, or during the annual booster training sessions. However, it is imperative that supervisors and other stakeholders consider whether the identified issues stem uniquely from assessors themselves (e.g., knowledge or motivational concerns) or whether they rise to the broader system or interagency level (e.g., lack of specificity in the protocols, unavailability of required documents, etc.).

25. Bonta and Andrews, *Risk-Need-Responsivity Model*.

Conduct booster training at least annually for all assessors during implementation.

At least once per year, assessors using the instrument should complete a booster training session to prevent drift and promote assessment accuracy. These booster sessions should:

- Review rating procedures and protocols for using assessment results.
- Require completion of at least one—but ideally two—practice cases to good agreement or better with a “gold standard” or expert rating.
- Address any other issues identified in the case reviews.

Booster sessions may be conducted by experienced in-house assessors or by outside experts (e.g., the instrument developers or other experts qualified to train on the instrument).²⁶ As with the initial training sessions, these sessions could be held for all assessors in a group—whether in person or online—or through an on-demand format that could be accessed by individual assessors as needed. There is no one best or recommended approach. The booster session training format and modality may be informed by the instrument that is selected for implementation, the needs and resources of the agency, and other practical considerations (e.g., the number of staff to be trained, specific training needs identified in the CQI reviews).

Practice cases can be completed during or after the booster session in various ways, such as having a staff member present a case to the group for assessment and review or using practice cases developed in collaboration with the expert trainer or other qualified professional. The goal of these practice cases is to offer real-time feedback on the accuracy of the assessment ratings and the connection between assessment results and case planning. Finally, while booster sessions are an opportunity to discuss and address issues identified via case reviews, it may be necessary to provide individual assessors or teams with more specific and targeted feedback through team meetings or one-on-one supervision, as appropriate.²⁷

26. Vincent, Guy, and Grisso, *Risk Assessment in Juvenile Justice*.

27. *Ibid.*

Guideline 4: Use a multistep approach to assess risk and needs over time.

Overview

While not required, a multistep approach to reassess risk and needs over time may contribute to greater accuracy and efficiency in the post-conviction risk and needs assessment process. Agencies may wish to implement a multistep approach for various reasons. In particular, the use of a risk screening instrument as an initial step in a comprehensive post-conviction risk and needs assessment process may help expedite initial decisionmaking and case processing. Additionally, the routine re-administration of post-conviction risk and needs assessment instruments that include dynamic factors and needs will afford the detection of changes in risk and needs that can be used to amend risk management strategies and case plans.

Action Items

Follow post-conviction risk screening instruments, if used, with a comprehensive risk and needs assessment only for those identified as being at potentially heightened risk of recidivism.

Although the terms are often used interchangeably, screening and assessment refer to two different, but related, processes. In the post-conviction context, screening refers to the universal implementation of a short, easily administered set of items to quickly identify individuals who are potentially at heightened risk of recidivism and should receive a more in-depth, comprehensive risk and needs assessment. In other words, screening instruments can be used as a first step to identify and “screen out” individuals who pose limited risk of recidivism and, thus, do not need to be evaluated further and to identify and “screen in” those who are at *potentially* heightened risk of recidivism and therefore warrant further, more comprehensive evaluation of their risk and needs.

Screening, by definition, is not a standalone process. Instead, the addition of screening to a comprehensive post-conviction risk and needs assessment process, while not necessary, may prove useful for agencies with large caseloads that are seeking to prioritize resources. Screening instruments are designed to cast a wide net; they are calibrated during the development and validation process to over- (as opposed to under-) estimate risk. That is, they are intentionally designed to reduce the likelihood of false negatives—individuals who are misidentified as low risk for recidivism. However, in doing so, the number of false positives—individuals who are misidentified as being at heightened risk of recidivism—will be high. If used, screening instruments must be followed by a comprehensive risk and needs assessment for those “screened in.”

To be clear, post-conviction risk and needs assessment instruments can be implemented in the absence of risk screening instruments. Agencies that are seeking to adopt an evidence-based assessment approach do not need to implement a universal risk screening protocol. However, the opposite is not true. Do not use risk screening instruments in lieu of comprehensive post-conviction risk and needs assessment instruments. This is a misapplication of screening instruments and will overestimate risk, which, in turn, will contribute to unnecessary

individual, assessor, and system costs and can potentially contribute to increases in recidivism. For these reasons, using risk screening instruments in lieu of comprehensive post-conviction risk and needs assessment instruments also will threaten fairness.

Re-administer post-conviction risk and needs assessment instruments that include dynamic factors and needs at routine intervals to monitor individual progress and inform amendments to case planning, as needed.

Re-administering post-conviction risk and needs assessment instruments that include dynamic factors and needs will improve assessment accuracy.²⁸ Dynamic risk factors and needs, by definition, are capable of change. For that reason, post-conviction risk and needs assessment instruments that include dynamic factors and needs require re-administration over time. Doing so will provide not only a measure of change—if any—in an individual's risk level overall but also an opportunity to review the appropriateness—and effectiveness—of the current risk management and intervention strategies. An overall decrease in risk level across repeated assessments may prompt consideration of a reduction in the level of supervision and services. An overall increase in risk level may suggest the need for greater supervision and services. Alternatively, a lack of change in risk level may prompt consideration of whether the appropriate factors are being targeted at the appropriate level via the intervention and, if so, in such a way as to promote individual responsibility.²⁹

The timeframe for re-administering the post-conviction risk and needs assessment may depend on the instrument selected but also should account for the assessment's purpose, population, context, and local resources.³⁰ It may be useful to consult with the instrument developers or other experts to ascertain a timeframe for re-administration that balances resources with utility.

The emphasis here has been on the re-administration of post-conviction risk and needs assessment instruments that include dynamic factors and needs. We anticipate that static risk factors will change little, if at all, by definition. Change in static factors, if any, will typically be in the direction of increased rather than decreased risk (e.g., new charges or convictions that contribute to a higher criminal history rating). However, we may also see some reductions in risk level over time—even on risk factors thought to be static in nature—if items specify behaviors in a certain timeframe (e.g., convictions in the prior 2 years). So, agencies are encouraged to document circumstances in which re-administration of any post-conviction risk and needs assessment instruments, even those composed of static risk factors, may be needed to promote assessment accuracy.

28. Barnes et al., "Validity of Initial, Exit, and Dynamic Juvenile Risk Assessment," 21–38; Lloyd et al., "Reassessment Improves Prediction of Criminal Recidivism," 568–81.

29. See Bonta and Andrews, *Risk-Need-Responsivity Model*.

30. For example, some instruments may include more acute dynamic factors in which we might expect more frequent or rapid change, while others may include more stable dynamic factors that might change more slowly over months or years, if at all. As another example, some settings may confer more stability, expose individuals to fewer changes in their environment, or afford fewer opportunities for intervention, resulting in relatively limited change in functioning and risk. Alternatively, some agencies may implement post-conviction risk and needs assessments to support periods of transition, whether in or out of a particular setting or program. We may anticipate periods of transition to be times during which there will be considerable fluctuation in risk and needs. As a final example, some populations may show more or less change; we may anticipate greater change in risk and needs among some people convicted for first-time offenses or younger people in the criminal justice system, but less change in risk and needs among those who have had longer or more chronic justice system involvement.

Principle II: Fairness

The next three guidelines speak to strategies that agencies can use to promote fairness in the use of post-conviction risk and needs assessment instruments. Broadly speaking, fairness refers to the equitable use of the results of post-conviction risk and needs assessment instruments to inform case decisions, resource allocation, and services overall. However, fairness as it relates to disparities in racial, ethnic, gender, or other characteristics such as mental illness in post-conviction processes should consider, more specifically, the degree to which assessment results have the same meanings and applications across groups defined by these characteristics. Fairness should be considered in the development, validation, and implementation of post-conviction risk and needs assessment instruments.

“

[O]ne cannot expect any risk assessment tool to reverse centuries of racial injustice or gender inequality. That bar is far too high. But, one can hope to do better.

(Berk, Heidari, Jabbari, Kearns, and Roth, 2017, 35)

”

We recommend the following guidelines to promote fairness in the use of post conviction risk and needs assessment instruments:

5. Examine the results of the post conviction risk and needs assessment instrument for predictive bias and disparate impact across groups.
6. Apply post conviction risk and needs assessment instrument results to individual cases in keeping with the Risk Need Responsivity (RNR) principles.
7. Adopt agencywide strategies to minimize the potential that local implementation of a post conviction risk and needs assessment instrument could promote disparities.

Guideline 5: Examine the results of the post-conviction risk and needs assessment instrument for predictive bias and disparate impact across groups.

Overview

A post-conviction risk and needs assessment instrument is not necessarily biased or unfair simply because one group of people is rated higher or lower, on average, compared to another group of people.³¹ Instead, consider (a) how assessment results relate to recidivism across groups and (b) how assessment results are used to inform decisions across groups. These two considerations speak to predictive bias and disparate impact, respectively. Predictive bias is present when assessment results demonstrate different levels of predictive validity across groups, whereas disparate impact is present when the assessment results are applied inequitably across groups.

These two concepts are related but are not dependent upon each other. Predictive bias relates to assessment accuracy across groups but does not necessarily lead to disparate impact. Assessment results can show some differences in predictive accuracy between groups but still demonstrate equitable, positive impacts on case decisions and outcomes.³² For example, assessment results might demonstrate slightly better predictive accuracy for White than Black people but still contribute to less restrictive placements for both White and Black people. Further, disparate impact does not require the presence of predictive bias. Even if assessment results have similar levels of predictive accuracy across groups, they still may be used in different ways to inform case decisions and outcomes for different groups. For example, assessment results may demonstrate similar predictive accuracy for Black and White people, but judges and other decisionmakers may be more likely to deviate from assessment results in an upward direction (i.e., increase estimated risk) and impose more restrictive conditions for Black people than for White people.³³

The Standards for Educational and Psychological Testing³⁴ specify that test bias exists when scores function differently for different groups of people, which implies an adverse impact on one group compared to another.³⁵ For these reasons, we recommend that agencies focus on whether there is evidence of disparate impact in considerations of fairness.

31. Vincent and Viljoen, "Racist Algorithms or Systemic Problems?" 1576–84; Reynolds and Suzuki, "Bias in Psychological Assessment," 82–113.

32. See, for example, Lowder, Grommon, and Ray, *Improving the Accuracy and Fairness of Pretrial Release Decisions*; Lowenkamp, DeMichele, and Klein Warren, "Replication and Extension of the Lucas County PSA Project."

33. Orton, Hogan, and Wormith, "An Examination of the Professional Override," 0093854820942270; Marlowe et al., "Employing Standardized Risk Assessment in Pretrial Release Decisions," 361–76.

34. American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, *Standards for Educational and Psychological Testing*.

35. Meade and Fetzer, "Test Bias, Differential Prediction, and a Revised Approach," 738–61.

Action Items

Establish whether the likelihood of recidivism increases in similar ways across risk levels for members of groups defined by race, ethnicity, and gender.

We recommend asking the developers of post-conviction risk and needs assessment instruments for evidence that the likelihood of recidivism increases in comparable ways across risk levels from group to group. If this information cannot be provided or the instrument was developed locally, agencies can examine the performance indicators described in [Guideline 2](#) across groups.

Two key questions should be answered. First, do risk levels relate to the rates of recidivism as expected within groups defined by race, ethnicity, and gender (i.e., do higher observed rates of criminal behavior correspond to higher risk levels rather than lower risk levels)? It is possible—even likely—that recidivism rates will differ within a given risk level from one group to the next. What matters is whether the recidivism rates increase across risk levels *within* groups in the anticipated way. Second, do the performance indicators meet the minimum thresholds described in [Guideline 2](#)? Again, there may be some differences among groups, but what matters is that the performance indicators still meet statistical standards for predictive accuracy from group to group.³⁶

Test whether assessment results identify individual risk levels and needs and predict recidivism in the same way from group to group.

A critical step in evaluating the fairness of a post-conviction risk and needs assessment instrument is determining whether the assessment results predict recidivism in the same way, regardless of group membership. Said another way, we need to test statistically whether the strength (and direction) of the relationship between assessment results and recidivism differs systematically as a function of race, ethnicity, or gender.

Agencies can use various statistical methods to find out whether the average risk rating relates to the average recidivism rate in the same way for each group. We recommend using the methods that represent the state of the art and have been applied in peer-reviewed publications that test for racial, ethnic, and gender biases in risk and needs assessment.³⁷ Because these methods are complex, we recommend consulting with a researcher or evaluator with specific expertise in regression analysis or other statistical methods if that expertise is not available in house.

36. While there has been considerable emphasis on false positives and false negatives as metrics for understanding fairness, we do not recommend their use for both pragmatic and statistical reasons, two of which we highlight here. First, calculating false positives and false negatives requires assessment results to be used to categorize people into two groups based on whether they will or will not recidivate. However, post-conviction risk and needs assessment instruments do not produce such binary classifications; instead, they place people within risk levels or categories. As a result, a single threshold or cut-off must be chosen, above which someone is designated as testing “positive” for recidivism and below which they are designated as testing “negative.” This is not typically how instruments are used in practice, limiting the external validity—or practical relevance—of such metrics. Second, the false positive and false negative rates will differ dramatically as a function of the threshold selected, as well as the rate of recidivism during follow-up. Consequently, the generalizability of the results across jurisdictions and even within jurisdictions over time is very limited. For further discussion regarding the limitations of false positives and false negatives as metrics of fairness in the context of risk assessment, see Helmus and Babchishin, “Primer on Risk Assessment,” 8–25; Freeman, Hu, and Jannetta, *Racial Equity*.

37. We recommend testing a moderation model, which involves conducting multiple regression analysis in which the assessment results, grouping variable (e.g., gender or race), and their interaction term are entered as predictors of recidivism. Only if the interaction term is a statistically significant predictor of recidivism is there evidence of predictive bias. See, for example, Skeem and Lowenkamp, “Risk, Race, and Recidivism,” 680–712; Lowder et al., “Racial Bias and LSI-R Assessments,” 210–33; Cohen and Lowenkamp, “Revalidation of the Federal PTRR,” 234–60.

Compare how assessment results relate to case decisions, resource allocation, and service provision across groups.

At the core of concerns regarding the use of post-conviction risk and needs assessment is whether assessment results are applied in different ways for different groups and more specifically, whether the use of assessment results leads to more punitive and restrictive responses for marginalized groups.³⁸ The question that agencies must answer is whether there is evidence that the way assessment results are used to inform case decisions, resource allocation, and service provision contributes to greater racial, ethnic, or gender disparity than the strategies through which these processes are conducted otherwise.

To answer this question, we recommend, at a minimum, examining case decisions, resource allocation, and service provision across groups as part of a CQI strategy—for example, as part of routine data monitoring or case reviews every 6 months. Specifically, every 6 months, agencies should have a plan to examine the following metrics within groups defined by race, ethnicity, and gender:

- Percentage of each type of case decision.
- Assigned levels of classification, supervision, or condition.
- Average number of services provided overall.
- Percentage of each type of service.

However, to fully answer this question would require an evaluation design that allows for a systematic comparison of (1) case decisions, resource allocation, and service provision made using the results of a post-conviction risk and needs assessment instrument to (2) case decisions, resource allocation, and service provision made without the assessment results. A [randomized controlled trial](#) is the most rigorous design but challenging to do in the context of real-world practice. Alternative evaluation designs that may be more feasible include a [quasi-experimental, between-groups design](#), or a [pre-post test design](#). While these evaluation designs are limited in the degree to which findings speak to disparate impact that can be attributed to the assessment results (as opposed to other factors), they can still help agencies identify where there are systematic differences in case outcomes to address.

38. Vincent and Viljoen, "Racist Algorithms or Systemic Problems?" 1576–84.

Guideline 6: Apply post-conviction risk and needs assessment instrument results to individual cases in keeping with the RNR principles.

Overview

Applying the RNR model can promote fairness by providing a structure for guiding and, specifically, limiting the scope of the use of post-conviction risk and needs assessment instruments. The RNR model³⁹ is widely recognized as an evidence-based framework for promoting positive public safety and case outcomes through the practical application of the results from post-conviction risk and needs assessment instruments. Briefly:

- The **Risk principle** entails matching the level of supervision, resources, and services with the individual's assessed level of risk of recidivism.
- The **Need principle** specifies that interventions should target dynamic factors and needs that increase the risk of recidivism for that individual.
- The **Responsivity principle** involves tailoring risk management strategies and services to a person's specific abilities, motivations, and strengths as part of the case planning process.

Together, these three principles emphasize an individualized approach that is informed by assessment results and limited in scope.

Action Items

Use assessment results to inform the appropriate level of intervention needed to manage the assessed risk of recidivism.

Consistent with the Risk principle, assessment results should inform the least restrictive level of intervention needed to manage a person's risk of recidivism. The greater the estimated level of risk, the greater the supervision, resources, and services that should be allocated and vice versa. The objective is to use the post-conviction risk and needs assessment results to help identify the minimum level of intervention, if any, that is necessary to manage a person's potential risk to public safety. Assessment results should not be used to justify a higher level of intervention than appropriate for the offense(s) of conviction.

Some post-conviction risk and needs assessment instruments provide case management recommendations regarding the type and number of hours of supervision and services required at a given level of recidivism risk. There also have been efforts to develop recommendations regarding the frequency and intensity of intervention that are not instrument specific such as the **five-level risk and needs system** [↗](#). Briefly, the five-level system seeks to provide a common language to communicate information about risk and needs, and it recommends the appropriate intensity and type of risk management and intervention strategies indicated by a given risk and needs level. Other criteria for specific domains of intervention and treatment also may be relevant, such as The ASAM Criteria for the level and intensity of treatment services for people with addictions and co-occurring conditions.

39. Bonta and Andrews, *Risk-Need-Responsivity Model*.

In practice, we recommend that agencies develop local guidelines regarding the frequency and intensity of supervision and services vis-à-vis the assessment results prior to using the post-conviction risk and needs assessment instrument. For agencies that have already implemented a post-conviction risk and needs assessment instrument, development of local guidelines should be prioritized and completed over a short, but feasible, timeframe. Agencies may need to revisit these guidelines and adapt them over time as the population, availability of services, or other local resources change.

Identify the dynamic factors and needs to be addressed through intervention.

Consistent with the Need principle, interventions should target the dynamic (i.e., changeable) factors and needs that contribute to risk of recidivism for that individual. The reasons and motivations that lead to criminal behavior can differ dramatically from person to person, even among those who have the same factors present in their social environment.⁴⁰ Post-conviction risk and needs assessment instruments that include at least some dynamic risk and needs factors provide critical, person-specific information regarding behaviors, beliefs, or other factors to be targeted in intervention. Instruments that predominantly—or entirely—comprise static risk factors are more limited in their utility with respect to guiding the tailored interventions we recommend here.

Further, targeting dynamic risk and needs factors for intervention will de-emphasize historical factors that cannot be changed such as age at first arrest. In doing so, we can move away from factors that are known sources of bias and act as proxies for race, ethnicity, or gender.

Although applying the Need principle can be challenging, it is not “all or nothing.” As adherence to the Need principle increases—or with better “treatment match”—the likelihood of positive case outcomes, including public safety, increases.⁴¹ Because dynamic factors and needs can change over time, the re-administration of the post-conviction risk and needs assessment instrument at routine intervals to inform any needed amendments to case plans will promote the likelihood of success. (See [Guideline 4](#) for more on dynamic factors and re-administration).

Maximize reductions in recidivism by tailoring the interventions to individual motivations, strengths, and abilities.

Consistent with the Responsivity principle, reductions in recidivism will be maximized by tailoring interventions to case-specific barriers or facilitators to successful habilitation, including individual motivations, abilities, and strengths. In the development and amendment of case plans, assessors may consider two types of responsivity: [general responsivity](#) and [specific responsivity](#).

40. For example, one person may be heavily influenced by antisocial peers and have few prosocial contacts to buffer against these influences. Another person may also have antisocial peers, but their risk of recidivism is driven by problems related to substance use rather than the antisocial influence of these peers. An intervention focused on positive peer support, then, may mitigate risk in the former example, while a substance use intervention may have greater effectiveness in the latter. Similarly, if there is no indication of substance use as a factor, then a substance use-focused intervention may do more harm than good.

41. Singh et al., “From Risk Assessment to Risk Management,” 1–9; Vieira, Skilling, and Peterson-Badali, “Matching Court-Ordered Services with Treatment Needs,” 385–401; Nelson and Vincent, “Matching Services to Criminogenic Needs,” 1136–53.

General responsivity refers to the use of interventions that have demonstrated effectiveness in addressing criminogenic risk factors and needs, particularly approaches that use social learning or cognitive behavioral methods.⁴² The most effective interventions may use diverse, evidence-based strategies such as prosocial modeling, positive reinforcement, or problem-solving skill development that meet an individual where they are. General responsivity also emphasizes the importance of establishing a warm, respectful, trusting, and collaborative working alliance to promote positive treatment outcomes. Strategies that reflect cultural humility and a multicultural orientation, for example, may help facilitate strong working alliances and foster more just and equitable practices.⁴³

Specific responsivity refers to the tailoring of services to address individual and environmental factors that may affect treatment outcomes. This may include the use of specialized interventions such as culturally tailored interventions, trauma-informed approaches, or gender-specific services. Specific responsivity also should include consideration of environmental factors such as the institutional culture, staff skills or attitudes, and barriers to service access and use. Specific responsivity represents an opportunity not only to promote positive treatment outcomes in an individual case but also to address factors that may be contributing to racial, ethnic, and gender disparities more broadly.

Guideline 7: Adopt agencywide strategies to minimize the potential that local implementation of a post-conviction risk and needs assessment instrument could promote disparities.

Overview

Ultimately, it is how a post-conviction risk and needs assessment instrument is used in practice that will determine whether it contributes to the unfair treatment of people across groups defined by race, ethnicity, and gender. Instruments differ in their contents, methods, and purposes.⁴⁴ The information used to complete the assessments—a potential source of **systemic bias**—also differs as a function of local policies and practices as well as record keeping.⁴⁵ As a result, the performance, meaning, and application of assessment results may differ from setting to setting and population to population.

It is unlikely that any one strategy, including the use of post-conviction risk and needs assessment instruments, will *eliminate* racial, ethnic, or gender inequities in the criminal justice system. However, strategies employed in the system should not exacerbate these inequities either. Consequently, it is imperative that agencies take the steps necessary to minimize the potential that the use of post-conviction risk and needs assessment instruments promotes disparities in their local setting, context, or jurisdiction.

42. Dowden and Andrews, "The Importance of Staff Practice in Delivering Effective Correctional Treatment," 203–14.

43. Mosher et al., "Cultural Humility," 221–33.

44. Vincent and Viljoen, "Racist Algorithms or Systemic Problems?" 1576–84.

45. Mayson, "Bias in, Bias Out," 2122–473.

Action Items

Select and implement post-conviction risk and needs assessment instruments based on their performance, content, and context.

There is no one instrument that is “fairest.” Instead, the following information should be considered to support the selection and implementation of any post-conviction risk and needs assessment instrument to ensure that it does not perpetuate inequities:

- Predictive accuracy metrics across groups, as described above.
- The implications of factors that are known sources of bias or may act as proxies for race, ethnicity, and gender.
- The context(s) in which assessment results will be used.

In the process of selecting a post-conviction risk and needs assessment instrument, agencies should consider the degree to which the instrument includes factors that are known sources of bias or may serve as proxies for race, ethnicity, or gender (e.g., criminal history, gang affiliation, employment, education level, debt, or housing stability). The information captured in these items may reflect bias or marginalization resulting from systemic and structural inequities, and, consequently, their inclusion may contribute to disparities.⁴⁶ For example, information on criminal history may reflect biases in local policing, prosecutorial, and judicial practices. There are many post-conviction risk and needs assessment instruments that include a wide range of static and dynamic factors.⁴⁷ As such, we recommend that agencies avoid instruments that emphasize factors that are known sources of bias or may serve as proxies for race, ethnicity, or gender.⁴⁸

That said, the inclusion of such factors does not necessarily mean that the instrument will produce biased assessment results, nor does the exclusion of such factors mean that an instrument will be free from bias.⁴⁹ For these reasons, consideration of evidence regarding the performance of assessment results across groups, as described in [Guideline 5](#), will provide information that is essential to the selection process. For example, if two or more post-conviction risk and needs assessment instruments are roughly comparable, the instrument that minimizes differences in predictive accuracy among groups should be selected.⁵⁰

Finally, it is important to consider the context in which assessment results will be used. Certain instruments may be more appropriate for some decisions or applications than others. For instance, if the task at hand is one of classification, then a well-validated instrument that comprises primarily static factors may be acceptable (assuming there is limited evidence of group differences in assessment results). If the context also requires the development of case plans, an instrument that additionally includes dynamic risk and needs factors would be more appropriate.

46. Starr, “Evidence-Based Sentencing,” 803–72; Harcourt, “Risk as a Proxy for Race,” 237–43; Eckhouse et al., “A Unified Approach,” 185–209.

47. Desmarais, Johnson, and Singh, “Performance of Recidivism Risk Assessment Instruments,” 206–22.

48. Hart, “Culture and Violence Risk Assessment,” 76–96.

49. Mayson, “Bias in, Bias Out,” 2122–473; Skeem and Lowenkamp, “Using Algorithms,” 259–78.

50. Skeem and Lowenkamp, “Using Algorithms,” 259–78; Vincent and Viljoen, “Racist Algorithms or Systemic Problems?” 1576–84; Shepherd and Lewis-Fernandez, “Forensic Risk Assessment and Cultural Diversity,” 427–38.

Develop and implement strategies to support equitable and safe case decisions, resource allocation, and service provision.

Agencies can—and should—develop and institute strategies to support equitable and safe case decisions, resource allocation, and service provision. The use of assessment results should be clearly articulated in local administration protocols and policies governing the use of post-conviction risk and needs assessment instruments, as described in [Guidelines 2](#) and [9](#), respectively. Clear guidance on when, for whom, and how post-conviction risk and needs assessment instruments will be completed and applied to inform decisionmaking will reduce the potential that the use of assessment instruments is biased. When case decisions, resource allocation, or service provision deviate from assessment results, people managing these decisions should provide justification to explain why such deviations are appropriate. Ultimately, assessment results are just one source of information that agencies should consider during the case planning process.⁵¹ There may be case-related issues (e.g., specific offenses for which there are blanket policies such as sex offenses) or other considerations that inform individual decisions (e.g., current caseload size, availability of placements or programming, and/or limited staff resources).

Ongoing CQI, as described in [Guideline 3](#), will provide the opportunity to monitor the implementation and use of post-conviction risk and needs assessment results to ensure that their application supports equitable and safe case decisions, resource allocation, and service provision. If there is evidence of predictive bias or disparate impact across groups, various strategies can be implemented to increase the fairness of the process. Such strategies range from changes to the prediction model⁵² to clear and direct policies for usage. We strongly advise against professionals relying on their intuition rather than the results of post-conviction risk and needs assessment instruments due to abundant evidence showing that unaided human judgments are less accurate and more biased overall.⁵³ The clearly documented evidence of systemic bias in the criminal justice system, in particular, requires checks and balances on personal judgment in decisionmaking. A better assessment instrument option might be one that relies less on factors that are known sources of bias or may act as proxies for race, ethnicity, and gender and instead focuses on a person's current behavior and functioning.

51. Vincent and Viljoen, "Racist Algorithms or Systemic Problems?" 1576–84.

52. See, for example, Berk and Kuchibhotla, "Improving Fairness."

53. Grove et al., "Clinical Versus Mechanical Prediction," 19–30; Jung et al., "Simple Rules," 771–800; Lin et al., "The Limits of Human Predictions," eaz0652.

Principle III: Transparency

The third set of guidelines speak to strategies that agencies can utilize to promote transparency in the use of post-conviction risk and needs assessment instruments. Transparency refers to how and what information about the content, structure, and application of these instruments is disseminated to stakeholders. Transparency is relevant in both the development and implementation of risk and needs assessment instruments and requires a proactive communication strategy.

“

Transparency is a necessary step to accountability. (Eaglin, 2017, 111)

”

We recommend the following guidelines to promote transparency of post conviction risk and needs assessment instruments:

- 8.** Provide system stakeholders with relevant information on the development, intended use, and validation of the post conviction risk and needs assessment instrument.
- 9.** Develop a written policy that guides the local use of the post conviction risk and needs assessment instrument.
- 10.** Communicate the strengths and limitations of the post conviction risk and needs assessment instrument to the general public.

Guideline 8: Provide system stakeholders with relevant information on the development, intended use, and validation of the post-conviction risk and needs assessment instrument.

Overview

All system stakeholders should have the information they need to understand the assessment process and be able to use this information to determine for themselves whether the process is fair and the results are accurate. This means that the information must be both available and understandable. Yet it is neither realistic nor necessary for the entirety or specifics of the process to be understood by everyone. For example, a defense attorney, defendant, or community member does not necessarily need to know the specifics of a technology such as the mechanism of machine learning algorithms;⁵⁴ however, they should have enough information to be able to question the assessment content and results, and how they are being used.⁵⁵ By informing system stakeholders about the development, intended use, and validation of post-conviction risk and needs assessment instruments, we can achieve greater transparency (and accuracy) than is possible through assessments of risk and needs based on human judgments alone.

Ideally, instrument developers and researchers will make the information described below available from the outset. Indeed, the availability and accessibility of this information to the public should be key considerations when selecting post-conviction risk and needs assessment instruments.

Action Items

Articulate the purpose for which the post-conviction risk and needs assessment instrument was developed, including the intended settings, populations, and outcomes.

Post-conviction risk and needs assessment instruments differ in their intended purpose, setting, population, and outcome. Some instruments were designed with a primary focus on estimating the likelihood of recidivism, while others were designed to also inform case planning, including supervision and intervention. Some instruments were designed for specific settings (e.g., jail, prison, reentry, community-based supervision) or populations (e.g., people in detention, on parole or probation, etc.), while others were designed for more general application. Many were designed to estimate general risk of recidivism, including committing a new crime or violating conditions of probation or parole. Some are focused specifically on assessing risk of violence, and others estimate risk of specific forms of violence such as sexual violence and domestic violence. Moreover, the timeframe over which instruments estimate risk may differ from days to weeks to months to years. Given this wide variation across instruments, it is important to clearly state the purpose, setting, population, and outcomes for which the post-conviction risk and needs assessment instrument was developed to ensure that it is applied as intended.

54. Chiao, "Fairness, Accountability and Transparency," 126–39.

55. Carlson, "The Need for Transparency," 303–29.

Explain the set of factors the assessment considers—including their definitions, scoring, and weighting—in a manner that can be understood by different audiences, particularly those who will be using the results and those who will be affected by them.

Agencies should describe the assessment’s administration approach and data sources in sufficient detail so stakeholders can understand the process and any issues that may arise regarding the veracity of the information gathered (e.g., misrepresentation of circumstances or events, incomplete data, data entry errors). Although there is considerable overlap, not all post-conviction risk and needs assessment instruments include the same risk and needs factors, nor are these factors defined, scored, and weighted similarly across instruments. For example, depending on the instrument, “criminal history” may include age at first arrest, number of prior arrests, or number of convictions in the past 10 years, among many other definitions. One instrument might define “substance use” as any current drug use, while the other might define it as any lifetime alcohol or drug addiction. Because definitions, measurement, and weighting of these factors will affect risk estimates and have different implications for different groups, the general definitions, rating guidelines, and weighting must be described in plain language. This language should appear not just in technical manuals but also in other easily accessible outlets such as agency websites.

Additionally, the manner through which information is gathered to inform item ratings and the sources of this information also differ among post-conviction risk and needs assessment instruments. Some post-conviction risk and needs assessment instruments exclusively use information from official records, whereas others incorporate self-reporting. Others, still, require structured interviews with the individual being assessed or people with whom they interact (e.g., family members) and observations of behavior and functioning. Some post-conviction risk and needs assessment instruments are computerized and automated, while others are completed on paper. Again, sufficient description of the methods for information gathering should be provided for system stakeholders to consider the sources and potential issues with the information needed to complete the assessments.

Describe how risk levels are assigned.

The manner in which item ratings are combined to produce risk levels representing an estimated likelihood of recidivism differs across instruments. Some instruments use a simple [checklist approach](#) that involves adding item ratings to arrive at a total score.⁵⁶ Other instruments use an [algorithmic approach](#) that combines and weights item ratings using more advanced statistical models. These total scores are cross-referenced (by hand or via a computer program) with actuarial tables that describe probabilities or rates of recidivism seen in development, norming, or validation samples. Other instruments use a structured professional judgment approach in which assessors rate the items for their presence, severity, and relevance and use them to estimate the risk level based on their

56. Burgess, “Factors Determining Success or Failure on Parole,” 221–34.

professional judgment (rather than computed scores). Finally, other instruments may use a checklist or statistical approach to produce an initial risk level that can be adjusted for individual case circumstances or considerations; in other words, an [adjusted actuarial method](#).⁵⁷

Agencies should identify and clearly describe the method of assigning risk levels for both lay and technical audiences. For the lay audience, a simple description of the general approach for how item ratings and risk scores relate to risk levels and how these risk levels, in turn, relate to recidivism may suffice. This information is typically included in instrument manuals, but agencies should also provide it in other easily accessible outlets (e.g., websites, information repositories, printed documents). For the technical audience, links or contacts for further detailed information on the mathematical models and training data should be provided.

Outline the training requirements for people administering the instrument, including CQI elements described previously.

As described in [Guideline 3](#), all assessors must complete *all* required training before they complete a post-conviction risk and needs assessment in practice, including:

- Training on the strategies to gather and interpret information, procedures for rating items, how to interpret results, and how to apply results to inform practices.
- Completion of four to five practice cases.

All assessors should complete booster trainings, at least annually, after initial training. While instrument manuals typically contain general training requirements, agencies should also make these requirements available to local stakeholders who will be involved in, or affected by, the use of the post-conviction risk and needs assessment instrument as part of the implementation process. A summary of the local plans for implementing the initial training, case reviews, and booster training should also be outlined in a written policy and made available by the agency for review and input from stakeholders. See [Guideline 10](#) for more on specific strategies to support stakeholder involvement.

Publish the findings of validation studies examining the post-conviction risk and needs assessment instrument in a manner that is accessible to a variety of audiences.

The traditional approach of publishing the findings of validation studies examining post-conviction risk and needs assessment instruments in scholarly journal articles or agency reports is not sufficient. The findings must be made available to system [stakeholders](#) in forms that are readily accessible, understandable, and useful to them. This can be achieved in many different ways and formats; for example, in a high level, short overview of the study findings or a more detailed research brief summarizing the study purpose, methods, and findings. If such summaries or research briefs do not already exist, they should be developed through collaboration between researchers and

57. Picard et al., *Beyond the Algorithm*.

system stakeholders to ensure the accuracy and comprehensibility of the content. It may be helpful to consult with experts in science communication to ensure that the study findings are written in a manner that is not just accessible but also understandable to a variety of audiences. When these products are complete, agencies should make them available through posting/linking on websites, social media, information repositories, or other outlets that can be accessed by the public and are not behind a paywall.

Guideline 9: Develop a written policy that guides the local use of the post-conviction risk and needs assessment instrument.

Overview

Developing a written policy, ideally before using the instrument, will not only guide local practices but also help system stakeholders understand how the use of the post-conviction risk and needs assessment instrument may affect people in the criminal justice system. The written policy should describe the sources of information used to complete the assessments, including potential pitfalls that may exist in these sources, and the contexts in which and how the assessment results will be used. Doing so will promote greater transparency—and accountability—in the use of post-conviction risk and needs assessment instruments to inform case decisions, resource allocation, and service provision in the agency. For these reasons, developing a written policy and amending it as necessary is essential to the successful implementation of a post-conviction risk and needs assessment instrument.

Action Items

Describe the source(s) of information that will be used to complete the post-conviction risk and needs assessments locally and identify potential pitfalls, such as data quality or biases, that may exist in these sources.

The written policy should include the protocols that were developed to guide the local use of the post-conviction risk and needs assessment instrument, as discussed in [Guideline 3](#). The protocols should describe the sources of information that will be used to complete the ratings, how and by whom that information will be gathered (e.g., record review, interviews, self-report questionnaires, etc.), and what potential concerns there may be with the data quality or potential biases that may exist in the data. For instance, there may be known issues as they relate to local record keeping for certain types of information, or there may be concerns that stem from the nature of the information source more generally. That said, it is important to balance concern regarding potential biases with the actual veracity of the information.⁵⁸ It may be valuable to conduct informal reviews of sources for accuracy of

58. For example, while there are reasons to question the veracity of self-reported information, we often find that self-report of criminal behavior is more—not less—accurate than official records. Similarly, collateral informants, such as family members, are often used to corroborate information; however, there may be cases in which there has been limited contact between family members and the individual being assessed and, consequently, family members may not provide accurate information on current behaviors, functioning, and circumstances.

information at both the individual and system levels. As previously noted, it is not necessary to detail all possible pitfalls in the data but, rather, to sufficiently describe the information sources for stakeholders to consider and question the veracity of data used to complete the assessments.

Define the contexts in which and how the results of the post-conviction risk and needs assessment instrument will be used to inform case decisions, resource allocation, and service provision.

Because many post-conviction risk and needs assessments have been developed for various settings and populations, it is important to define how and in which specific context(s) the results of the post-conviction risk and needs assessment instruments will be used locally. As noted above, these issues will have been addressed in the development of the local administration protocols (see [Guideline 3](#)). The task is now to ensure that these protocols are adequately described in the broader written policy. Specifically, it should include a clear description of which individuals should be assessed, when initial and repeat assessments should be completed (e.g., within 2 weeks of intake and every 6 months after), and how results will inform decisions and processes.

Create the opportunity for input on the written policy from stakeholders.

Seeking input from stakeholders contributes to transparency by creating an opportunity for individuals, various groups, and members of the public to understand and influence decisions that may affect them—directly or indirectly. To do so in a meaningful way, agencies should seek input at various points in the policy development process and on the specific issues where the input has a real potential to help inform the policy. Sometimes the opportunity for shaping the policy will be limited; at other times, there may be greater flexibility and opportunity for influencing the policy. Inviting input from stakeholders does not mean that agencies must necessarily change policy in response to the feedback gathered. Rather, it provides a forum for considering and responding to a wide range of views and concerns, as possible and appropriate. It is also an opportunity to foster trust, gain buy-in, and improve interagency and agency-community relations.

Different strategies can be used to gather input on the written policy from stakeholders. These could include individual interviews, focus groups, community cafés, study circles, written response requests (via email or other format), mail or online surveys, electronic polling, or public meetings, hearings, or workshops. In selecting strategies, agencies should consider what information stakeholders may need to make informed contributions and whether stakeholders may benefit from hearing from each other. Agencies should also determine if there are specific groups that may need additional outreach to ensure that their opinions are heard, whether there is a need to have comments on public record, and the timeframe for review and input. Regardless of the strategy, agencies should gather input from diverse stakeholders to ensure a wide range of views and concerns are considered and to promote meaningful involvement and inclusion with respect to race, ethnicity, gender, mental illness, and other

characteristics. Once agencies have gathered the input, it is their responsibility to balance and interpret it, decide whether to change the policy to address concerns or views that were shared, and report to stakeholders how their input was considered and used.

Creating the opportunity for input from stakeholders can be challenging, but the benefits are significant. It can help support better outcomes by facilitating implementation of policy that is better understood by stakeholders and reflects their interests and values. Further, gathering stakeholder input can develop system capacity to solve and manage issues that may stem from differing views and misunderstandings regarding post-conviction risk and needs assessment instruments.

Establish a process and timeline to review and update the written policy, as necessary.

Because it may be challenging to do on an ad hoc basis, we recommend that agencies establish a process and timeline to review and update the written policy, ideally during the planning period prior to using the instrument. As described in [Guideline 1](#) in relation to revalidation efforts, we recommend that agencies identify and allocate resources and staffing to support the policy review and update during this planning process. Specifically, we recommend that agencies conduct the policy review and update following the instrument revalidation at least every 5 years. Doing so will ensure that the policy review and update can account for the findings of the revalidation in addition to other changes in the agency (e.g., staffing and resources) and local criminal justice practices, policies, and populations, as relevant. Note that agencies may also need to review and update the policy between revalidations to account for major circumstantial changes.

Guideline 10: Communicate the strengths and limitations of the post-conviction risk and needs assessment instrument to the general public.

Overview

Agencies may use a variety of strategies to communicate to the general public the strengths and limitations of the post-conviction risk and needs assessment instrument selected for local implementation. Some strategies such as public meetings, briefings, or telephone contact require person-to-person communication, whereas others can be accomplished remotely such as through printed information (e.g., fact sheets, newsletters, bulletins), websites, information repositories, press, and social media. There is no one best or most appropriate approach for communicating this information to the general public. Instead, agencies should consider a range of factors, including the current level of knowledge and understanding of criminal justice processes, public preferences for receiving information, and forms of communication that may be more or less effective across groups as a function of accessibility, language, literacy, and trust. The key is to ensure that the information is available, understandable, and accessible across groups. At a minimum, we recommend that the information be included on agency websites.

Ideally, this would occur prior to implementation. For agencies already using a post-conviction risk and needs assessment instrument, however, this should be completed as soon as possible.

It can reasonably be expected that, without such efforts, community members will have limited knowledge and understanding of post-conviction risk and needs assessments. In addition to promoting the principle of transparency, there are benefits to ensuring that community members have the information necessary to evaluate and develop an informed opinion on the post-conviction risk and needs assessment instrument. Indeed, public opinion can have a substantial impact on the success or failure of policy implementation as it relates to post-conviction risk and needs assessment instruments or otherwise.⁵⁹

Action Items

Make information on the instrument’s purpose, content, and validation available for easy access by the general public.

Building upon the actions necessary to meet the requirements of [Guideline 8](#), agencies should describe in lay terms the purpose, content, and validation of the post-conviction risk and needs assessment instrument selected for use. Specifically, agencies should make the following information available for easy access by the general public:

- What the post-conviction risk and needs assessment instrument is and is not designed to do.
- How item selection and weighting minimize racial, ethnic, and gender disparities in assessment results while promoting accuracy.
- How the post-conviction risk and needs assessment instrument has been evaluated, including studies of predictive bias and disparate impact, as well as any limitations or gaps in research that remain to be addressed.

Whether this requires additional communication strategies beyond those implemented under [Guideline 8](#) will need to be determined on a case-by-case basis. To that end, agencies should evaluate whether the general public is aware of and has access to the materials, outlets, websites, etc. through which this information is currently made available. They should also assess whether the information presented is likely to be understood by a wide audience. Additionally, agencies should implement a plan for how to raise community awareness of where this information is located (e.g., through press release, social media).

Describe the process through which the post-conviction risk and needs assessment instrument was selected for implementation.

A critical aspect of transparency is outlining the process through which the post-conviction risk and needs assessment instrument was selected for implementation. Agencies should describe the issues and evidence that were considered as well as the stakeholders—individuals or groups, as appropriate—who participated in the selection process. In addition to the considerations outlined in these guidelines, there are various resources available to support agencies in instrument selection such as the Public Safety Risk Assessment Clearinghouse’s

59. Burstein, “The Impact of Public Opinion on Public Policy,” 29–40.

Tool Selector [↗](#) or the guidance provided in the report on **Risk Assessment Instruments Validated and Implemented in Correctional Settings in the United States** [↗](#). Whatever the process, agencies should ensure that they are documenting each step, including decisions made along the way. The goal is to describe the process in sufficient detail so that community members will understand how and why a certain instrument was selected from the many instruments available for use. Documenting this process also may benefit agencies themselves by establishing institutional knowledge that may be lost over time as a result of staffing changes or turnover.

Clearly state how the post-conviction risk and needs assessment instrument will be used locally.

Again, the requirement is not to duplicate efforts described in [Guideline 9](#) but to make the policy available to the general public in a clear and concise manner once it is finalized and whenever it is revised. Community members need not know the details of the training and administration protocols but, rather, more generally, how and when assessments will be completed, for whom, and to inform what types of decisions. This communication may naturally flow from efforts under [Guideline 9](#) to create opportunities for input on the written policy from stakeholders. However, if community members are not included in that process, agencies must identify the strategy(ies) that will be used to ensure that a clear statement of use is released publicly before using the instrument.

Explain how the accuracy and impact of the post-conviction risk and needs assessment instrument on case outcomes will be monitored overall and across groups.

Drawing from the plan derived to meet the requirements of [Guidelines 2](#) and [5](#), prepare a short description of the methods that will be implemented to examine the performance and consequences of assessment results as they are locally used. Agencies should write this description in lay language and provide information on the general approach, key indicators of performance and impact, and, importantly, efforts that will be implemented should these efforts highlight any issues of concern. This should include a brief summary of actions that may be taken at the [individual](#), [group](#), or [system](#) levels.

Principle IV: Effective Communication and Use

The final three guidelines speak to strategies that agencies can utilize to promote effective communication and the use of post-conviction risk and needs assessment instrument results. The manner in which assessors communicate individual assessment results can greatly affect their impact on decisionmaking and, consequently, their effectiveness. It is only through effective communication of assessment results that they can appropriately inform case decisions, resource allocation, and service provision.⁶⁰ Improper communication of individual assessment results can undermine efforts to promote accuracy, fairness, and transparency in the use of post-conviction risk and needs assessment instruments. Communication, then, must be a central consideration in planning, training, and implementation.

“

Improper risk communication can render a risk assessment that was otherwise well-conducted completely useless—or even worse than useless, if it gives consumers the wrong impression. (Heilbrun, Dvoskin, Hart, & McNiel, 1999, 94)

”

We recommend the following guidelines to promote the effective communication and use of post conviction risk and needs assessment instruments:

- 11.** Anchor communication of post conviction risk and needs assessment results in the RNR principles.
- 12.** Contextualize the results of the post conviction risk and needs assessment instruments.
- 13.** Develop a template for communicating the individual results of the post conviction risk and needs assessment instrument to all relevant stakeholders, including the person being assessed.

60. Heilbrun et al., “Violence Risk Communication,” 91–105.

Guideline 11: Anchor communication of post-conviction risk and needs assessment results in the RNR principles.

Overview

In addition to supporting fairness, as described in [Guideline 6](#), the application of the RNR model to the post-conviction risk and needs assessment can promote effective communication and use of the results. Specifically, RNR provides a framework for helping assessors identify what information should be communicated about the assessment results and the recommended intervention to different stakeholder groups.⁶¹ As discussed further in [Guideline 13](#), effective communication does not mean sharing all information derived during the assessment process but, rather, focusing on what information is necessary to support decisionmaking. Indeed, when presented with too much information, decisionmakers will rely on prior experiences and personal biases, including stereotypes, to discern the relative importance and weight of the various pieces of information.⁶²

Action Items

Describe assessment results as placing an individual in a particular risk level that informs the minimum level of intervention needed to mitigate their risk of recidivism rather than assigning a specific probability or likelihood of recidivism to the individual.

The results of most post-conviction risk and needs assessment instruments provide information on how the individual was rated, scored, and ranked in terms of their risk of recidivism in relation to a group of people who were assessed using the post-conviction risk and needs instrument.⁶³ It would be a mistake—and potentially misleading—to assign specific probability or likelihood of recidivism to the individual. Instead, assessment results should be described as placing an individual in a particular risk level that informs the minimum level of intervention needed to mitigate their assessed risk of recidivism. In keeping with the Risk principle, intervention, including supervision and services, should be commensurate with the assessed level of risk. That is, individuals at the lowest risk level should receive the least intensive intervention and those at the highest risk level, the most intensive intervention. However, the most intensive intervention should still represent the least restrictive conditions within which the risk can be managed. As such, the most intensive intervention could still be community placement and services.⁶⁴

Judges and other decisionmakers often desire a combination of categorical and numerical information on risk.⁶⁵ Consequently, the rate of recidivism observed among those who were placed in that risk level in the norming or validation samples can be shared but with the clear specification that this is not to be understood as the individual's absolute probability or likelihood of recidivism.

61. Heilbrun, Newsham, and Pietruszka, "Risk Communication." 2016

62. Tversky and Kahneman, "Judgment under Uncertainty," 1124.

63. Harris, Lowenkamp, and Hilton, "Evidence for Risk Estimate Precision," 111–27.

64. Desmarais and Lowder, "Principles and Practices of Risk Assessment," 593–603.

65. Kwartner, Lyons, and Boccaccini, "Judges' Risk Communication Preferences," 185–94; Evans and Salekin, "Violence Risk Communication," 143–64.

Identify the presence of risk and protective factors that contribute to the assessment results, emphasizing the dynamic factors and needs that should be addressed through intervention.

In keeping with the Need principle, communication of assessment results should emphasize the dynamic factors and needs that contribute to recidivism risk for that individual and that, consequently, should be addressed in intervention. This communication should not just name the risk and protective factor in the abstract. Instead, it should provide a very brief operational definition of the factors as specified by the post-conviction risk and needs assessment instrument as well as a description of the specific behavior, attitude, or circumstance as it presents in the person who was assessed. These definitions and descriptions should be very short—just a few words will do in many cases—but sufficient to convey the issue that needs to be addressed.⁶⁶

While static risk factors may be relevant to individual risk, they are not modifiable. As such, they should be communicated briefly, if at all, in relation to the initial intensity of intervention recommended, unless subsequent behavior (e.g., supervision failure, new crime) results in a higher static risk score. Unless static risk factors can be translated into some modifiable form, they should not be integrated into case planning. Further, we recommend that ratings for items that were not deemed to be relevant to individual risk (whether present or not) are excluded from communication.⁶⁷ Including these ratings may inadvertently—and mistakenly—convey that the items should be addressed through intervention.

Explain case-specific barriers or facilitators to successful habilitation, above and beyond those described previously.

Facilitate the application of the Responsivity principle by articulating any case-specific issues that may undermine or otherwise detract from the effectiveness of intervention. As discussed in [Guideline 6](#), this may include communicating case-specific barriers or facilitators to successful habilitation that should be considered in the development and tailoring of case plans (i.e., specific responsivity), including those that relate to the individual and their environment. This may also include identifying and recommending interventions with demonstrated effectiveness in addressing the dynamic factors and needs identified during the assessment process (i.e., general responsivity).⁶⁸

66. Heilbrun et al., "Assessing Normative Approaches to Communicating Violence Risk," 187–96.

67. Storey, Watt, and Hart, "An Examination of Violence Risk Communication," 39–55.

68. Dowden and Andrews, "The Importance of Staff Practice in Delivering Effective Correctional Treatment," 203–14.

Guideline 12: Contextualize the results of the post-conviction risk and needs assessment instrument.

Overview

The ways in which assessment results are communicated to stakeholders will determine how they are used.⁶⁹ Consequently, communicating information about the context surrounding the assessment process and its results is necessary for balancing concerns of public safety with the promotion of individual rights and habilitation in subsequent decisionmaking and intervention.⁷⁰ Risk of recidivism is not an individual trait. Rather, it will depend upon the complex interaction of a person's characteristics with their social and physical environments. To that end, it is critical that the recipient of information about assessment results understands the circumstances surrounding the assessment and its results, the situations in which risk of recidivism would be elevated, and what can be done to prevent it.

Action Items

State the likelihood and, when possible, the type(s) of criminal behavior anticipated in the absence of interventions over the timeframe(s) specified by the instrument.

Communication of assessment results should include a clear statement on the likelihood of recidivism anticipated in the absence of intervention, as estimated using the post-conviction risk and needs assessment instrument. This may include a simple statement of the assessment results reporting placing an individual in a particular risk level, as described in [Guideline 11](#). To the extent possible, however, the type(s) of criminal behavior anticipated in the absence of interventions, over the timeframe(s) specified by the instrument, should be clearly described. Recidivism is not one type of behavior. Instead, the behavior that would constitute "recidivism" varies in nature, frequency, and severity. For this reason, it may not be sufficient to make a general statement regarding risk level. Instead, the nature (e.g., nonviolent, violent, sexually violent) and severity of the anticipated behavior(s) as well as the potential harm to victim(s) should be specified to the extent possible. Further, communication should specify over what timeframe(s) and in what setting(s) the assessment results are intended to estimate risk of recidivism, if specified by the instrument. Because risk of recidivism is time and context dependent, a statement regarding how the level of risk might change over time or across settings can help inform case decisions and prioritize resource allocation and intervention. For example, such a statement could help identify outcomes that need more immediate intervention for victims and public safety compared to those that may help support successful habilitation in the long term.

69. Heilbrun et al., "Violence Risk Communication," 91-105.

70. Ignelzi et al., "Best Practices," 452-54.

The degree and specificity to which such information is known to assessors may differ. Some post-conviction risk and needs assessment instruments, for example, may produce different risk estimates for different types of recidivism (e.g., any criminal behavior, violent behavior, technical violations/infractions) and over different timeframes (e.g., 1 month, 6 months, 2 years, 5 years), while others may only speak to risk of recidivism in a general or aggregate way.

Define the parameters of the assessment results.

In addition to addressing the context of the assessed risk, the context of the assessment itself should be communicated. This information will help support evaluations of the accuracy and fairness of the assessment results, promote transparency, and, ultimately, provide for the appropriate application of assessment results. To that end, the purpose of the assessment should be clearly stated—not only the conditions that prompted the assessment (e.g., intake to a new facility) but also the decision(s) and processes the assessment results are intended to inform (e.g., custody level, case planning, program placement, etc.). There also should be a clear but brief description of the administration protocols, including the sources of information used and any concerns or limitations regarding that information (see [Guideline 3](#)). As previously noted, data sources can reflect bias and affect the accuracy and fairness of the assessment results. As such, a cautionary statement regarding confidence in the accuracy of the current assessment results may be warranted and, if so, factors that affected confidence such as mixed or inconsistent information in the data sources that could not be resolved. Finally, conditions that would prompt re-assessment should be specified (see [Guideline 4](#) for more on re-assessment).

Identify the type and approximate intensity of interventions that are likely to reduce the anticipated risk of recidivism and support successful case outcomes.

The estimated likelihood of recidivism reflects the absence of intervention. Consequently, it is necessary to specify the minimum level of intervention needed to manage that risk of recidivism (Risk principle), the interventions that are likely to be successful for that individual (Need principle), and any case-specific considerations for promoting the effectiveness of the intervention (Responsivity principle; see [Guideline 6](#)). It is also important to clearly communicate how assessment results inform those interventions. Again, assessment results will only improve outcomes if they are used to inform decisionmaking, resource allocation, and service provision in meaningful ways. For these reasons, identifying the type and intensity of interventions that are likely to be effective is a critical step in effective risk communication.

Few post-conviction risk and needs assessment instruments will produce specific recommendations for risk management and intervention. Instead, and as discussed in [Guideline 6](#), agencies should develop local guidelines describing the frequency, type, and intensity of supervision and services that can be quickly referenced to inform this communication before using the instrument.

Guideline 13: Develop a template for communicating the individual results of the post-conviction risk and needs assessment instrument to all relevant stakeholders, including the person being assessed.

Overview

A template for communication helps outline and structure what information assessors will share with different groups of stakeholders about the assessment process and results. The development and implementation of a standard template for the written communication of individual assessment results will improve comprehension and use of the results and can reduce assessor effort and time. Specifically, having a standard template for written communication may help overcome barriers to effective communication and use of assessment results by reducing chances for factual error, misrepresentation of assessment information, or presentation of misleading or irrelevant information. Using a standard template for written communication can also streamline the presentation of assessment results for ease of understanding and increase the predictability of information that will be communicated.

While the focus here is on the development of a standard template for written communication, we suggest that this template also be used as the foundation for oral communication of the findings such as in a courtroom or in meetings.

Action Items

Provide a structure and format for presenting the assessment results in a manner that is clear, concise, predictable, and consistent across assessors and cases.

The exact structure and format of the written communication template may vary from agency to agency. However, the following information should be included:

- A brief statement on the instrument that was used and the sources of information used to complete the assessment.
- The estimated risk levels.
- The identified risk and protective factors, needs, and case-specific barriers or facilitators to successful community reintegration.
- The recommendations for intervention (if appropriate).

Also, we recommend against selectively reporting individual item ratings. Doing so may overburden the audience and unintentionally emphasize individual factors in ways that are inconsistent with their contributions to the overall risk estimates.

Importantly, assessors should receive training on the communication template and how to use it as part of the pre-service training process (see [Guideline 3](#)) to maximize its use and effectiveness. Other stakeholders also should receive a brief training on the template to increase the predictability of information communicated about the assessment and to support their comprehension of assessment results. If possible, integrate the template into existing electronic reporting tools or as a fillable form to promote implementation of the template with fidelity.

Use communication strategies that promote comprehension and reduce the impact of potentially problematic information.

Assessment information can be complex and difficult to understand. Agencies must make efforts to promote comprehension through the use of evidence-based communication strategies. In particular, people tend to comprehend more and make more informed decisions when the important information is easy to evaluate and understand. To that end, assessment results should be presented in accordance with cognitive expectations (e.g., higher numbers mean greater risk).⁷¹ Further, less is indeed more when it comes to the communication of assessment results.⁷² Consequently, only the most relevant information about risk and needs should be communicated such as the risk factors and needs that were present and relevant rather than those that were not. Efforts also should be made to avoid technical jargon and information that requires inferences or interpretation;⁷³ for example, it is better to provide the estimated likelihood of recidivism as a percentage rather than as a number out of 100 (or some other denominator).⁷⁴

Additionally, the typical presentation of historical information first—followed by information about the current case and present functioning—may unintentionally emphasize and anchor decisionmaking in what has occurred in the past (i.e., criminal history) rather than the present circumstances and current functioning of the individual.⁷⁵ For this reason, we recommend structuring the template to follow the Situation, Background, Assessment, and Recommendation (SBAR) communication strategy⁷⁶ such that information on the current case and circumstances is presented first, followed by the background or historical case information, then the assessment results (i.e., estimated risk level), and last, recommendations for supervision and services. Finally, communication of assessment information should avoid language that may be biased and inadvertently perpetuate prejudicial beliefs. Guidelines for [bias-free language](#) should be consulted in the development of the communication template.⁷⁷

71. Peters et al., "Numeracy Skill," 741–48.

72. Peters et al., "Less Is More," 169–90.

73. Heilbrun et al., "Violence Risk Communication," 91–105.

74. Monahan et al., "Communicating Violence Risk," 121–26.

75. Dror, "A Hierarchy of Expert Performance," 121–27.

76. Thomas, Bertram, and Johnson, "The SBAR Communication Technique," 176–80.

77. See, for example, "Bias-Free Language," APA Style, accessed March 25, 2021, <https://apastyle.apa.org/style-grammar-guidelines/bias-free-language/>.

Tailor communication to the target audience, with the potential for different templates for different stakeholders, but avoid sharing assessment results beyond relevant stakeholders.

For communication to be effective, it must be tailored to the target audience. This may mean developing different templates or modifying the standard template for different stakeholders. For example, how information on the assessment process and results is shared with a judge may differ from how this information is shared with the individual who was assessed or a service provider to whom the individual may be referred. Some audiences may need detailed information, while others may only need a high level summary. Tailoring communication requires knowing how the audience prefers to receive information, what information is relevant to them and their decisionmaking, and their level of knowledge about post-conviction risk and needs assessment, generally, and the instrument used, specifically.⁷⁸ It also is important to consider factors that may affect communication accessibility such as literacy, preferred language(s), or abilities.

There may be gaps in knowledge regarding these various communication preferences, needs, and barriers. Agencies can most easily gather the information needed to help tailor communication to the target audience by involving diverse stakeholders in the template development process.

Share the template with stakeholders for review and feedback prior to finalizing it.

Sharing the template with stakeholders prior to using the instrument will afford them the opportunity for input that can be used to promote the appropriateness, acceptability, and effectiveness of the template across audiences. For agencies that have already implemented a post-conviction risk and needs assessment instrument, developing and sharing a communication template with stakeholders for review and feedback should be prioritized and completed over a short, but feasible, timeframe. As discussed in relation to the written policy ([Guideline 9](#)), seeking input from stakeholders can also promote transparency in the use of post-conviction risk and needs assessment instruments, foster trust, and improve relations. Again, inviting input from stakeholders does not mean that agencies must necessarily change the template in response to the feedback, but it does provide the opportunity for ensuring that the assessment information is being communicated as intended; the organization of information makes sense to the audience; the desired content is included; and the language and format are appropriate. With such input, the template may be better received and given greater consideration by stakeholders in their decisionmaking.

78. Heilbrun, "Prediction Versus Management Models Relevant to Risk Assessment," 347; Schopp, "Communicating Risk Assessments," 939.

Glossary

Acceptable levels of accuracy: The accuracy with which the results of post-conviction risk and needs assessment instruments predict the outcome they were intended to predict (e.g., recidivism) indicated by area under the curve (AUC) values of .64–.71. (See [good validity](#).)

Accuracy: The degree to which results of post-conviction risk and needs assessment instruments predict the recidivism outcomes they were designed to predict.

Adjusted actuarial method: An actuarial approach to post-conviction risk and needs assessment in which the statistically derived risk estimate can be adjusted for individual case circumstances or considerations through the use of professional judgment (i.e., professional or clinical override) to increase or decrease the risk estimate.

Algorithmic (or actuarial) approach: An approach to post-conviction risk and needs assessment that combines and weights item ratings using statistical models that produce risk levels representing an estimated likelihood of recidivism. The total scores are cross-referenced (by hand or via computer program) with actuarial tables that describe probabilities or rates of recidivism seen in development, norming, or validation samples.

Area under the curve (AUC): In this context, a predictive validity performance indicator measuring the probability that a randomly selected person who recidivated during follow-up would have received a higher risk score or level using a given risk assessment approach than a randomly selected person who did not recidivate during follow-up.

Between-groups design: An evaluation design in which one compares outcomes between two or more groups that receive different interventions to measure the effectiveness of an intervention; for example, comparing placement decisions of one group of people who were assessed using a post-conviction risk and needs assessment instrument (i.e., intervention group) with another group that was not (i.e., comparison group).

Bias-free language: Language that demonstrates inclusive treatment of people and sensitivity with respect to race, ethnicity, gender, age, and other categories or identities. It involves avoiding terminology that may be hurtful, offensive, or perpetuate prejudicial beliefs.

Case review: Part of the continuous quality improvement (CQI) process, case reviews examine fidelity to the rating and scoring guidelines, adherence to the implementation protocols, and concordance between assessment results and case decisions, resource allocation, and service provision.

Checklist approach: An approach to post-conviction risk and needs assessment that involves simply adding item ratings to arrive at a total score of the number of items endorsed as present, where lower scores reflect lower risk of recidivism and higher scores reflect higher risk of recidivism.

Continuous quality improvement (CQI): A structured process that expands upon basic quality assurance methods, examining aggregate data on processes, practices, and outcomes to identify areas for improvement at the organizational or system level and to implement needed improvements.

Disparate impact: When results of post-conviction risk and needs assessment instruments are applied inequitably across groups, leading to adverse agency or system-level responses to one group of people, such as a group defined by race, ethnicity, or gender, as compared to another group.

Dynamic risk factors: Factors that contribute to risk but can change over time (e.g., social networks, thinking patterns, housing, substance use, finances, etc.), also called criminogenic needs. Dynamic risk factors not only add to the predictive ability of an assessment instrument, they represent those areas that can be changed through programming and interventions.

Effective communication and use: When the results of the post-conviction risk and needs assessment instruments are shared, discussed, and applied with strategies that promote understanding, accuracy, transparency, and positive case outcomes.

Evaluation: The systematic investigation of the results of a post-conviction risk and needs assessment instrument to determine its performance and effect on case decisions, resource allocation, and service provision.

Excellent agreement: Concordance among assessors who administer a post-conviction risk and needs assessment instrument indicated by (1) observed agreement of 90 percent or greater, (2) Kappa of .75–1.00, or (3) intra-class coefficient (ICC) of .75–1.00.

Excellent validity: The accuracy with which the results of post-conviction risk and needs assessment instruments predict recidivism indicated by area under the curve (AUC) values of .71–1.00.

Fairness: The equitable use of results from post-conviction risk and needs assessment instruments to inform case decisions, resource allocation, and service provision overall. This principle considers the degree to which assessment results have similar meanings and applications across groups, as it relates to racial, ethnic, and gender disparities in post-conviction processes.

Fidelity: The degree to which a post-conviction risk and needs assessment instrument is used as intended, including adherence to scoring guidelines, administration protocols, and local policies for use in practice.

General responsivity: Subprinciple of the Responsivity principle positing that the use of cognitive social learning methods will be most effective at reducing recidivism.

Good agreement: Concordance among assessors who administer a post-conviction risk and needs assessment instrument indicated by (1) observed agreement of 80 percent or greater, (2) [Kappa](#) of .60–.74, or (3) intra-class coefficient ([ICC](#)) of .60–.74.

Good validity: Accuracy with which the results of post-conviction risk and needs assessment instruments predict recidivism indicated by area under the curve ([AUC](#)) values of .64–.71. (See [acceptable levels of accuracy](#).)

“Gold standard”: An assessment completed by an instrument developer or other expert that serves as the criterion against which to compare the accuracy of ratings completed by an assessor in the context of training or use in practice.

Group level: Characteristics, knowledge, attitudes, skills, behaviors, or other attributes of multiple people together.

Individual level: Characteristics, knowledge, attitudes, skills, behaviors, or other attributes of a single person.

Inter-rater reliability: The degree to which assessors who administer a post-conviction risk and needs assessment instrument achieve the same results when assessing the same person. This is a property of the assessment results rather than of the post-conviction risk and needs assessment instrument itself.

Intra-class coefficient (ICC): The measure of inter-rater reliability representing the strength of agreement among multiple assessors on continuous variables (e.g., total scores), statistically corrected for chance.

Item: Component of a post-conviction risk and needs assessment instrument that is used to document the presence and/or severity of a risk or needs factor.

Kappa: Measure of inter-rater reliability representing the percentage of categorizations (e.g., low, moderate, or high risk) upon which multiple assessors agreed, statistically corrected for chance.

Minimum level of intervention: The lowest amount and intensity of supervision, resources, and services that is necessary to manage an identified level of recidivism risk.

Need principle: The principle of the Risk-Need-Responsivity model positing that treatment and case management should target the identified dynamic risk factors and criminogenic needs that can be positively impacted through services, supervision, and supports to reduce recidivism. The greater the number of dynamic risk factors and criminogenic needs are addressed through interventions, the greater positive impact those interventions will have on reducing recidivism.

Norming: In the development of an actuarial post-conviction risk and needs assessment instrument, the process through which population-based recidivism rates for each risk level or category are established. Individual assessment results are then compared against these risk levels or categories.

Observed agreement: The measure of inter-rater reliability representing the percentage of categorizations (e.g., low, moderate, or high risk) upon which multiple assessors agreed.

Observed rates of criminal behavior: The proportion of people within each risk level who went on to recidivate divided by the total number of people who were rated at that risk level.

Outcomes: In the context of post-conviction risk and needs assessment validation, the specific form(s) of recidivism that is being forecasted (e.g., general offending, violent crime, sexual violence).

Performance thresholds: Well-established scientific standards for measuring the strength or degree of agreement among assessors (i.e., inter-rater reliability) or between the assessment results and recidivism (i.e., predictive validity).

Population: The specific group(s) of people in the criminal justice system (e.g., people in detention, on parole or probation, etc.) for which a risk and needs assessment instrument is intended and validated for use.

Predictive bias: When the results of a post-conviction risk and needs assessment instrument consistently demonstrate different levels of predictive validity across groups (e.g., race, ethnicity, gender).

Predictive validity: The accuracy with which results of the post-conviction risk and needs assessment instrument forecast the outcomes they were intended to predict (e.g., recidivism). This is a property of the assessment results rather than of the assessment instrument itself.

Pre-post test design: An evaluation design in which the outcome of interest is assessed at least two times (i.e., pre-test and post-test) in order to measure the effectiveness of a new treatment or intervention; for example, recidivism rates or detention rates are examined before and after the implementation of a post-conviction risk and needs assessment instrument.

Protective factors: Characteristics of a person (e.g., attitudes, substance use), their environment (e.g., neighborhood, family, peers), or situation (e.g., housing, employment) that is associated with a decrease in the likelihood of recidivism.

Protocols for administration: Written documentation that describes for whom post-conviction risk and needs assessments will be completed and by whom, the sources of information that should be used to complete the assessments, what decisions and processes they inform, and when re-assessments should be conducted.

Purpose: The primary goal of implementing a risk and needs assessment instrument (e.g., predicting the likelihood of recidivism, informing case planning, etc.).

Quasi-experimental design: A type of between-groups evaluation design in which one compares outcomes between two or more groups to measure the effectiveness of a given intervention. In this evaluation design, there is no random assignment; rather, participants are assigned to groups based on other criteria. For example,

such an evaluation might involve comparing placement decisions in a jurisdiction where a post-conviction risk and needs assessment instrument has been implemented (i.e., the intervention group) to placement decisions in another jurisdiction that has not implemented a post-conviction risk and needs assessment instrument (i.e., the comparison group).

Randomized controlled trial (RCT): A type of between-groups evaluation design in which one compares outcomes between two or more groups of participants who are randomly assigned to receive different interventions to measure the effectiveness of an intervention. For example, such an evaluation might involve comparing placement decisions for participants who were randomly assigned to be assessed using a post-conviction risk and needs assessment instrument (i.e., the intervention group) to placement decisions for participants who were randomly assigned not to be assessed using a post-conviction risk and needs assessment instrument (i.e., the control group) within one jurisdiction.

Responsivity principle: The principle of the Risk-Need-Responsivity model positing that individual and system-level efforts to provide cognitive behavioral treatment and reduce barriers to positive learning outcomes (e.g., tailoring to reading ability, motivation, strengths) will promote the effectiveness of interventions in reducing recidivism.

Risk and needs assessment: The process of estimating the likelihood of future criminal behavior and identifying the dynamic risk and needs factors that may serve as treatment targets in the development of risk management and treatment plans.

Risk and needs assessment instrument: An instrument—composed of empirically or theoretically based risk (and in some tools also protective) factors—used to estimate the likelihood of future criminal behavior and to inform decisionmaking following convictions.

Risk-Need-Responsivity (RNR) principles: The RNR principles are a set of research-based guiding principles that, when implemented correctly, can help reduce reoffending and violations of conditions of probation and parole and help policymakers, administrators, and practitioners determine how to allocate resources, deliver services, and provide the right people with the right supports and services to have the greatest impact on recidivism and public safety.

Risk principle: The principle of the Risk-Need-Responsivity model dictating that the level and intensity of supervision, treatment, and other services should be proportionate to a person’s assessed level of risk of recidivism.

Risk screening instrument: A short, easily administered set of items to quickly identify (1) individuals who are at potentially heightened risk of recidivism and who should, therefore, receive a more in-depth, comprehensive risk and needs assessment (i.e., screened “in”) versus (2) individuals who pose limited risk of recidivism and, thus, do not need to be evaluated further (i.e., screened “out”).

Setting: The specific location or stage of criminal justice processing (e.g., jail, prison, reentry, community-based supervision, etc.) in which a risk and needs assessment instrument is intended and validated for use.

Specific responsivity: The subprinciple of the Responsivity principle emphasizing the importance of considering and addressing individual and environmental characteristics that may act as barriers to intervention effectiveness; for example, building relevant staff skills, addressing prejudicial beliefs among staff, or “fine-tuning” services or interventions such as modifying cognitive behavioral treatment to account for a cognitive impairment associated with mental illness.

Stakeholders: An individual or group with a vested interest in a criminal justice agency’s work, including professionals who work within or with the criminal justice system, such as judges, attorneys, service providers, and probation/parole officers, as well as people in the criminal justice system and their families.

Static risk factors: Factors that are unchanging or that cannot be changed through deliberate intervention (e.g., age, prior offenses). Static risk factors contrast with dynamic risk factors (or criminogenic needs), which can be used to inform the targets of supervision and human service interventions.

Structured professional judgment: An approach to post-conviction risk and needs assessment in which assessors estimate risk by considering a set number of factors that are empirically and theoretically associated with the outcome of interest. Total scores are not used to make the final judgments of risk; instead, assessors consider the relevance of each item to the person being assessed as well as whether there are any case-specific factors not explicitly included in the list.

System level: Organizations, policies, laws, practices, and structures that comprise a system such as the criminal justice system.

Systemic bias: Disparities in criminal justice system responses to one group of people with a protected characteristic (e.g., race, ethnicity, and gender) compared to another group, stemming from both current and historical discriminatory policies and practices. An example of systemic bias is higher rates of conviction among Black people compared to White people despite similar rates of criminal behavior. Although we have chosen to use the term “systemic bias” here, it is often interchangeable with “structural bias.”

Transparency: The degree to which information about the content, structure, and application of post-conviction risk and needs assessment instruments is disseminated to stakeholders in an understandable manner.

Validation: An empirical evaluation used to determine the predictive validity of the results of a post-conviction risk and needs assessment instrument. (See [predictive validity](#).)

Written communication template: A template that outlines and structures what information assessors will share with stakeholders about the assessment process and results in written communications (e.g., reports).

Bibliography

- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 2014.
- Barnes, Ashlee R., Nordia A. Campbell, Valerie R. Anderson, Christina A. Campbell, Eytayo Onifade, and William S. Davidson. "Validity of Initial, Exit, and Dynamic Juvenile Risk Assessment: An Examination across Gender and Race/Ethnicity." *Journal of Offender Rehabilitation* 55, no. 1 (2016): 21–38. <https://doi.org/10.1080/10509674.2015.1107004>.
- Berk, Richard, Hoda Heidari, Shahin Jabbari, Michael Kearns, and Aaron Roth. "Fairness in Criminal Justice Risk Assessments: The State of the Art." *arXiv* (2017). arXiv:1703.09207.
- Berk, Richard A., and Arun Kumar Kuchibhotla. "Improving Fairness in Criminal Justice Algorithmic Risk Assessments Using Conformal Prediction Sets." *arXiv: Applications* (2020). <https://arxiv.org/abs/2008.11664v1>.
- Bonta, James, and Donald A. Andrews. *Risk-Need-Responsivity Model for Offender Assessment and Rehabilitation*. Ottawa, Canada: Public Safety Canada, 2007. <https://www.publicsafety.gc.ca/cnt/rsrscs/pblctns/rsk-nd-rspnsvty/rsk-nd-rspnsvty-eng.pdf>.
- Breaux, Ariel, Yenys Castillo, Lara Guzmán-Hosta, Antoinette Kavanaugh, Rebecca Rivas, Danielle Rynczak, and Mark Worthen. "Race in Forensic Evaluations: The AP-LS Practice Committee Offers Practical Considerations." *American Psychology-Law Society Newsletter* (May 2021). https://ap-ls.wildapricot.org/resources/EmailTemplates/2021_05%20May%20AP-LS%20Newsletter/ClinicalForensicPractice5_21.pdf.
- Buchanan, Alec, Renee Binder, Michael Norko, and Marvin Swartz. "Psychiatric Violence Risk Assessment." *American Journal of Psychiatry* 169, no. 3 (2012): 340. <https://doi.org/10.1176/appi.ajp.2012.169.3.340>.
- Burgess, Ernest W., "Factors Determining Success or Failure on Parole." In *The Workings of the Indeterminate Sentence Law and the Parole System in Illinois*, edited by A. W. Bruce, E.W. Burgess, J. Landesco, and A.J. Harno: 221–34. Springfield: Illinois Board of Parole, 1928.
- Burstein, Paul. "The Impact of Public Opinion on Public Policy: A Review and an Agenda." *Political Research Quarterly* 56, no. 1 (2003): 29–40. <https://doi.org/10.1177/02F106591290305600103>.
- Carlson, Alyssa M. "The Need for Transparency in the Age of Predictive Sentencing Algorithms." *Iowa Law Review* 103 (2017): 303–29.
- Chiao, Vincent. "Fairness, Accountability and Transparency: Notes on Algorithmic Decision-Making in Criminal Justice." [In English], *International Journal of Law in Context* 15, no. 2 (2019): 126–39. <https://doi.org/10.1017/S1744552319000077>.
- Cicchetti, Domenic V. "The Precision of Reliability and Validity Estimates Re-Visited: Distinguishing between Clinical and Statistical Significance of Sample Size Requirements." *Journal of Clinical and Experimental Neuropsychology* 23, no. 5 (2001): 695–700. <https://doi.org/10.1076/jcen.23.5.695.1249>.
- Cohen, Thomas H., and Christopher T. Lowenkamp. "Revalidation of the Federal PTR: Testing the PTR for Predictive Biases." *Criminal Justice and Behavior* 46, no. 2 (2019): 234–60. <https://psycnet.apa.org/doi/10.1177/0093854818810315>.
- de Vogel, Vivienne, and Michiel de Vries Robbé. "Adapting Risk Assessment Tools to New Jurisdictions." In *International Perspectives on Violence Risk Assessment*, edited by Jay P. Singh, Stål Bjørkly, and Seena Fazel. American Psychology-Law Society Series, 26–39. New York: Oxford University Press, 2016.
- Desmarais, Sarah L., and Evan M. Lowder. "Principles and Practices of Risk Assessment in Mental Health Jail Diversion Programs." *CNS Spectrums* 25, no. 5 (Oct 2020): 593–603. <https://doi.org/10.1017/S1092852919001652>.
- Desmarais, Sarah L., Kiersten L. Johnson, and Jay P. Singh. "Performance of Recidivism Risk Assessment Instruments in U.S. Correctional Settings." *Psychological Services* 13, no. 3 (Aug 2016): 206–22. <https://doi.org/10.1037/ser0000075>.
- Douglas, Kevin S., Randy K. Otto, Sarah L. Desmarais, and Randy Borum. "Clinical Forensic Psychology." *Handbook of Psychology, Second Edition 2* (2012). <https://doi.org/10.1002/9781118133880.hop202008>.
- Douglas, Kevin S., Jennifer L. Skeem, and Elizabeth Nicholson. "Research Methods in Violence Risk Assessment." In *Research Methods in Forensic Psychology*, edited by Barry Rosenfeld and Stephen D. Penrod, 325–46. John Wiley & Sons, Inc., 2011.

Dowden, Carig, and D. A. Andrews. "The Importance of Staff Practice in Delivering Effective Correctional Treatment: A Meta-Analytic Review of Core Correctional Practice." [In English], *International Journal of Offender Therapy and Comparative Criminology* 48, no. 2 (April 2004): 203–14. <https://doi.org/10.1177/0306624x03257765>.

Dror, Itiel E. "A Hierarchy of Expert Performance." *Journal of Applied Research in Memory and Cognition* 5, no. 2 (2016): 121–27. <https://doi.org/10.1016/j.jarmac.2016.03.001>.

Eckhouse, Laurel, Kristian Lum, Cynthia Conti-Cook, and Julie Ciccolini. "Layers of Bias: A Unified Approach for Understanding Problems with Risk Assessment." *Criminal Justice and Behavior* 46, no. 2 (2018): 185–209. <https://doi.org/10.1177/0093854818811379>.

Evans, Stephanie A., and Karen L. Salekin. "Violence Risk Communication: What Do Judges and Forensic Clinicians Prefer and Understand?" *Journal of Threat Assessment and Management* 3, no. 3-4 (2016): 143–64. <https://doi.org/10.1037/tam0000062>.

Feinstein, Alvan R., and Domenic V. Cicchetti. "High Agreement but Low Kappa: I. The Problems of Two Paradoxes." *Journal of Clinical Epidemiology* 43, no. 6 (1990): 543–49.

Freeman, Kelly Roberts, Cathy Hu, and Jesse Jannetta. *Racial Equity and Criminal Justice Risk Assessment*. Washington, DC: Urban Institute, 2021. <https://www.urban.org/research/publication/racial-equity-and-criminal-justice-risk-assessment>.

Gottfredson, Stephen D., and Laura J. Moriarty. "Statistical Risk Assessment: Old Problems and New Applications." *Crime & Delinquency* 52, no. 1 (2006): 178–200. <https://journals.sagepub.com/doi/10.1177/001128705281748>.

Grove, William M., David H. Zald, Boyd S. Lebow, Beth E. Snitz, and Chad Nelson. "Clinical Versus Mechanical Prediction: A Meta-Analysis." *Psychological Assessment* 12, no. 1 (2000): 19–30.

Hamilton, Melissa. *Risk Assessment Tools in the Criminal Legal System – Theory and Practice: A Resource Guide*. Washington, DC: National Association of Criminal Defense Lawyers, 2020. <https://www.nacdl.org/Document/RiskAssessmentReport>.

Hanson, R. Karl, Guy Bourgon, Robert J. McGrath, Daryl G. Kroner, David A. D'Amora, Shenique S. Thomas, and Lahiz P. Tavarez. *A Five-Level Risk and Needs System: Maximizing Assessment Results in Corrections through the Development of a Common Language*. New York: The Council of State Governments Justice Center, 2017.

Harcourt, Bernard E., "Risk as a Proxy for Race: The Dangers of Risk Assessment." *Federal Sentencing Reporter* 27, no. 4 (2015): 237–43.

Harris, Grant T., Christopher T. Lowenkamp, and N. Zoe Hilton. "Evidence for Risk Estimate Precision: Implications for Individual Risk Communication." *Behavioral Sciences & the Law* 33, no. 1 (February 2015): 111–27. <https://doi.org/10.1002/bsl.2158>.

Hart, Stephen D., "Culture and Violence Risk Assessment: The Case of Ewert V. Canada." *Journal of Threat Assessment and Management* 3, no. 2 (2016): 76–96. <https://doi.org/10.1037/tam0000068>.

Heilbrun, Kirk. "Prediction Versus Management Models Relevant to Risk Assessment: The Importance of Legal Decision-Making Context." *Law and Human Behavior* 21, no. 4 (1997): 347–59. <https://doi.org/10.1023/A:1024851017947>.

Heilbrun, Kirk, Joel Dvoskin, Stephen Hart, and Dale McNeil. "Violence Risk Communication: Implications for Research, Policy, and Practice." *Health, Risk & Society* 1, no. 1 (1999): 91–105. <https://doi.org/10.1080/13698579908407009>.

Heilbrun, Kirk, Rebecca Newsham, and Victoria Pietruszka. "Risk Communication: An International Update." In *International Perspectives on Violence Risk Assessment*. New York: Oxford University Press, 2016.

Heilbrun, Kirk, Melanie L. O'Neill, Tomika N. Stevens, Lisa K. Strohman, Quinten Bowman, and Yi-Wen Lo. "Assessing Normative Approaches to Communicating Violence Risk: A National Survey of Psychologists." *Behavioral Sciences & the Law* 22, no. 2 (March 2004): 187–96. <https://doi.org/10.1002/bsl.570>.

Helmus, L. Maaik, and Kelly M. Babchishin. "Primer on Risk Assessment and the Statistics Used to Evaluate Its Accuracy." *Criminal Justice and Behavior* 44, no. 1 (2017): 8–25. <https://doi.org/10.1177/0093854816678898>.

Ignelzi, James, Bob Stinson, James Raia, Thomas Osinowo, Larry Ostrowski, and Jennifer Schwirian. "Best Practices: Utilizing Risk-of-Violence Findings for Continuity of Care." *Psychiatric Services* 58, no. 4 (2007): 452–54. <https://doi.org/10.1176/ps.2007.58.4.452>.

Jung, Jongbin, Connor Concannon, Ravi Shroff, Sharad Goel, and Daniel G. Goldstein. "Simple Rules to Guide Expert Classifications." *Statistics in Society* 183, no. 3 (2020): 771–800. <https://rss.onlinelibrary.wiley.com/doi/10.1111/rssa.12576>.

- Kwartner, Phylissa P., Phillip M. Lyons, and Marcus T. Boccaccini. "Judges' Risk Communication Preferences in Risk for Future Violence Cases." *International Journal of Forensic Mental Health* 5, no. 2 (2006): 185–94. <https://doi.org/10.1080/14999013.2006.10471242>.
- Lin, Zhiyuan J., Jongbin Jung, Sharad Goel, and Jennifer Skeem. "The Limits of Human Predictions of Recidivism." *Science Advances* 6, no. 7 (Feb 2020): eaaz0652. <https://doi.org/10.1126/sciadv.aaz0652>.
- Lloyd, Caleb D., R. Karl Hanson, Dylan K. Richards, and Ralph C. Serin. "Reassessment Improves Prediction of Criminal Recidivism: A Prospective Study of 3,421 Individuals in New Zealand." [In English], *Psychological Assessment* 32, no. 6 (2020): 568–81. <https://doi.org/10.1037/pas0000813>.
- Lowder, Evan M., Megan M. Morrison, Daryl G. Kroner, and Sarah L. Desmarais. "Racial Bias and LSI-R Assessments in Probation Sentencing and Outcomes." [In English], *Criminal Justice and Behavior* 46, no. 2 (2019): 210–33. <https://doi.org/10.1177/0093854818789977>.
- Lowder, Evan M., Eric Grommon, and Bradley R. Ray. *Improving the Accuracy and Fairness of Pretrial Release Decisions: A Multi-Site Study of Risk Assessments Implemented in Four Counties*. Washington, DC: Bureau of Justice Programs, U.S. Department of Justice, 2020.
- Lowenkamp, Christopher T., Matthew DeMichele, and Lauren Klein Warren. "Replication and Extension of the Lucas County PSA Project." *Advancing Pretrial Policy and Research*. 2020.
- Marlowe, Douglas B., Timothy Ho, Shannon M. Carey, and Carly D. Chadick. "Employing Standardized Risk Assessment in Pretrial Release Decisions: Association with Criminal Justice Outcomes and Racial Equity." *Law and Human Behavior* 44, no. 5 (Oct 2020): 361–76. <https://doi.org/10.1037/lhb0000413>.
- Mayson, Sandra G., "Bias in, Bias Out." *The Yale Law Journal* 128, no. 8 (2019): 2122–473.
- Meade, Adam W., and Michael Fetzer. "Test Bias, Differential Prediction, and a Revised Approach for Determining the Suitability of a Predictor in a Selection Context." *Organizational Research Methods* 12, no. 4 (2009/10/01 2009): 738–61. <https://doi.org/10.1177/1094428109331487>.
- Mills, Jeremy F., Michael N. Jones, and Daryl G. Kroner. "An Examination of the Generalizability of the LSI-R and VRAG Probability Bins." *Criminal Justice and Behavior* 32, no. 5 (2005): 565–85. <https://doi.org/10.1177/0093854805278417>.
- Monahan, John, Kirk Heilbrum, Eric Silver, Erik Nabors, Jonathan Bone, and Paul Slovic. "Communicating Violence Risk: Frequency Formats, Vivid Outcomes, and Forensic Settings." *The International Journal of Forensic Mental Health* 1, no. 2 (2002): 121–26. <https://doi.org/10.1080/14999013.2002.10471167>.
- Mosher, David K., Joshua N. Hook, Laura E. Captari, Don E. Davis, Cirleen DeBlaere, and Jesse Owen. "Cultural Humility: A Therapeutic Framework for Engaging Diverse Clients." *Practice Innovations* 2, no. 4 (2017): 221–33. <https://doi.org/10.1037/pri0000055>.
- Nelson, Rebecca J., and Gina M. Vincent. "Matching Services to Criminogenic Needs Following Comprehensive Risk Assessment Implementation in Juvenile Probation." *Criminal Justice and Behavior* 45, no. 8 (2018): 1136–53. <https://doi.org/10.1177/0093854818780923>.
- Orton, Laura C., Neil R. Hogan, and J. Stephen Wormith. "An Examination of the Professional Override of the Level of Service Inventory–Ontario Revision." *Criminal Justice and Behavior* (2020): 0093854820942270. <https://doi.org/10.1177/0093854820942270>.
- Partnership on AI. *Report on Algorithmic Risk Assessment Tools in the U.S. Criminal Justice System* (San Francisco: Partnership on AI, 2019). <https://partnershiponai.org/wp-content/uploads/2021/08/Report-on-Algorithmic-Risk-Assessment-Tools.pdf>.
- Peters, Ellen, Nathan Dieckmann, Anna Dixon, Judith H Hibbard, and CK Mertz. "Less Is More in Presenting Quality Information to Consumers." *Medical Care Research and Review* 64, no. 2 (2007): 169–90. <https://doi.org/10.1177/10775587070640020301>.
- Peters, Ellen, Judith Hibbard, Paul Slovic, and Nathan Dieckmann. "Numeracy Skill and the Communication, Comprehension, and Use of Risk-Benefit Information." *Health Affairs* 26, no. 3 (May/June 2007): 741–48. <https://doi.org/10.1377/hlthaff.26.3.741>.
- Picard, Sarah, Matt Watkins, Michael Rempel, and Ashmini G. Kerodal. *Beyond the Algorithm: Pretrial Reform, Risk Assessment, and Racial Fairness*. New York: Center for Court Innovation, 2019. https://www.courtinnovation.org/sites/default/files/media/document/2019/Beyond_The_Algorithm.pdf.
- Reynolds, Cecil R., and Lisa A. Suzuki. "Bias in Psychological Assessment: An Empirical Review and Recommendations." In *Handbook of Psychology: Assessment Psychology*, Vol. 10, 2nd Ed., 82–113. (Hoboken, NJ: John Wiley & Sons, Inc., 2013). <https://doi.org/10.1002/0471264385.wei1004>.

- Rice, Marnie E., and Grant T. Harris. "Comparing Effect Sizes in Follow-up Studies: Roc Area, Cohen's D, and R." *Law and Human Behavior* 29, no. 5 (2005): 615–20. <https://doi.apa.org/doi/10.1007/s10979-005-6832-7>.
- Schopp, Robert F. "Communicating Risk Assessments: Accuracy, Efficacy, and Responsibility." *American Psychologist* 51, no. 9 (1996): 939–44. <https://doi.org/10.1037/0003-066X.51.9.939>.
- Shepherd, Stephane M., and Roberto Lewis-Fernandez. "Forensic Risk Assessment and Cultural Diversity: Contemporary Challenges and Future Directions." *Psychology, Public Policy, and Law* 22, no. 4 (2016): 427–38. <https://doi.org/10.1037/law0000102>.
- Singh, Jay P., "Predictive Validity Performance Indicators in Violence Risk Assessment: A Methodological Primer." [In English], *Behavioral Sciences and the Law* 31, no. 1 (January/February 2013): 8–22. <https://doi.org/10.1002/bsl.2052>.
- Singh, Jay P., Sarah L. Desmarais, Brian G. Sellers, Tatiana Hylton, Melissa Tirotti, and Richard A. Van Dorn. "From Risk Assessment to Risk Management: Matching Interventions to Adolescent Offenders' Strengths and Vulnerabilities." [In English], *Children and Youth Services Review* 47 (December 2014): 1–9. <https://doi.org/10.1016/j.childyouth.2013.09.015>.
- Singh, Jay P., Suzanne Yang, Edward P. Mulvey, and Ragee Group. "Reporting Guidance for Violence Risk Assessment Predictive Validity Studies: The Ragee Statement." *Law and Human Behavior* 39, no. 1 (February 2015): 15–22. <https://doi.org/10.1037/lhb0000090>.
- Skeem, Jennifer L., and Christopher T. Lowenkamp. "Risk, Race, and Recidivism: Predictive Bias and Disparate Impact." *Criminology: An Interdisciplinary Journal* 54, no. 4 (2016): 680–712. <https://doi.org/10.1111/1745-9125.12123>.
- Skeem, Jennifer, and Christopher Lowenkamp. "Using Algorithms to Address Trade-Offs Inherent in Predicting Recidivism." *Behavioral Sciences & the Law* 38, no. 3 (2020): 259–78. <https://onlinelibrary.wiley.com/doi/10.1002/bsl.2465>.
- Slobogin, Christopher. "Primer on Risk Assessment for Legal Decision-Makers." *Vanderbilt Criminal Justice Program* (2020). <https://scholarship.law.vanderbilt.edu/faculty-publications/1182/>.
- Smith, William R. "The Effects of Base Rate and Cutoff Point Choice on Commonly Used Measures of Association and Accuracy in Recidivism Research." *Journal of Quantitative Criminology* 12, no. 1 (1996/03/01 1996): 83–111. <https://doi.org/10.1007/BF02354472>.
- Starr, Sonja B. "Evidence-Based Sentencing and the Scientific Rationalization of Discrimination." *Stanford Law Review* 66, no. 4 (2014): 803–72.
- Storey, Jennifer E., Kelly A. Watt, and Stephen D. Hart. "An Examination of Violence Risk Communication in Practice Using a Structured Professional Judgment Framework." *Behavioral Sciences & the Law* 33, no. 1 (January 2015): 39–55. <https://doi.org/10.1002/bsl.2156>.
- Thomas, Cynthia M., Evelyn Bertram, and Doreen Johnson. "The SBAR Communication Technique: Teaching Nursing Students Professional Communication Skills." *Nurse Educator* 34, no. 4 (2009): 176–80. <https://doi.org/10.1097/NNE.0b013e3181aaba54>.
- Tversky, Amos, and Daniel Kahneman. "Judgment under Uncertainty: Heuristics and Biases." *Science* 185, no. 4157 (1974): 1124. <https://doi.org/10.1126/science.185.4157.1124>.
- Vieira, Tracey A., Tracey A. Skilling, and Michele Peterson-Badali. "Matching Court-Ordered Services with Treatment Needs: Predicting Treatment Success with Young Offenders." [In English], *Criminal Justice and Behavior* 36, no. 4 (April 2009): 385–401. <https://doi.org/10.1177/0093854808331249>.
- Vincent, Gina M., Laura S. Guy, and Thomas Grisso. *Risk Assessment in Juvenile Justice: A Guidebook for Implementation*. Chicago: John D. and Catherine T. MacArthur Foundation, 2012.
- Vincent, Gina M., and Jodi L Viljoen. "Racist Algorithms or Systemic Problems? Risk Assessments and Racial Disparities." *Criminal Justice and Behavior* 47, no. 12 (2020): 1576–84. <https://doi.org/10.1177%2F0093854820954501>.
- Zottola, Samantha A., Sarah L. Desmarais, Evan M. Lowder, and Sarah E. Duhart Clarke. "Evaluating Fairness of Algorithmic Risk Assessment Instruments: The Problem with Forcing Dichotomies." *Criminal Justice and Behavior* (August 2021). <https://doi.org/10.1177%2F00938548211040544>.

**Statement of Patricia L. Richman
National Sentencing Resource Counsel
Federal Public and Community Defenders
First Step Act Implementation: Stakeholder Engagement Sessions
September 29, 2022**

Thank you for inviting the views of the Federal Public and Community Defenders on the First Step Act's (FSA) risk assessment tool, PATTERN ("Prisoner Assessment Tool Targeting Estimated Risks and Needs"). My name is Patricia Richman, and I am National Sentencing Resource Counsel for the Federal Public and Community Defenders.¹ Federal Public and Community Defenders, along with other appointed counsel under the Criminal Justice Act, represent 80 to 90 percent of individuals charged with federal crimes.

Since the FSA's enactment, we have submitted detailed comments and testimony to BOP, DOJ and Congress about the implementation of the FSA's correctional reforms, which I incorporate here.² Most recently, I worked with the

¹ For the past four years, I have had the opportunity to closely observe implementation efforts by the Department of Justice (DOJ) and Bureau of Prisons (BOP). In 2019, I served as detailee counsel to Senator Dick Durbin, where I assisted his oversight of DOJ's development and release of its risk assessment tool, PATTERN. After leaving the Senate Judiciary Committee, I returned to the Defenders, where I train attorneys around the country about the FSA's back-end reforms, track FSA litigation against the BOP, and receive reports about implementation successes and failures on a near-daily basis.

² See Letter from David Patton & Jon Sands, Co-Chairs, Federal Public & Community Defenders Legislative Committee, and Lisa Freeland, Chair Defender Services Advisory Group, to Hons. Dick Durbin & Chuck Grassley (May 4, 2021), <https://bit.ly/2UK2gSf>; Letter from David Patton and Jon Sands, Co-Chairs, Federal Public & Community Defenders Legislative Committee to Office of General Counsel, Federal Bureau of Prisons,

Statement of Patricia L. Richman
September 29, 2022

Brennan Center to produce a report analyzing the status of the FSA's prison reforms and to make several recommendations, which I have attached as an exhibit and also incorporate here.³

Before turning to substance, I would like to thank DOJ for the steps it has taken over the past year to “honor the [FSA’s] promise” of “a path to an early return home for eligible incarcerated people who invest their time and energy in programs that reduce recidivism.”⁴ The final Earned Time Credit Rule (ETC Rule) that DOJ promulgated in January, 2022, was transformative; we have been so glad to see thousands return home as a direct result of its adoption.

Re: Comments on Docket No. BOP 1176P, RIN 1120-AB76, “FSA Time Credits” (Jan 22, 2021) (FD ETC Comment), <https://bit.ly/3keZSgV>; *Hearing on the Fed. Bureau of Prisons and Implementation of the First Step Act Before the Comm. on the Judiciary of the H. Subcomm. on Crime, Terrorism, and Homeland Sec., 116th Cong.* (Statement of David Patton at 5), (Oct. 17, 2019), <https://bit.ly/3kfSAJF> pdf; Letter from David Patton and Jon Sands, Co-Chairs, Federal Public & Community Defenders Legislative Committee, to David B. Muhlhausen, *Re: DOJ First Step Act Listening Session on PATTERN* (Sept. 13, 2019) (2019 PATTERN Letter) <https://bit.ly/36C7vGd>.

³ Ames C. Grawert & Patricia L. Richman, *The First Step Act’s Prison Reforms: Uneven Implementation and the Path Forward*, Brennan Center for Justice (Sept. 23, 2022), <https://bit.ly/3DSJmwX>.

⁴ Press Release, *Justice Department Announces New Rule Implementing Federal Time Credits Program Established by the First Step Act* (Jan. 13, 2022) (quoting Attorney General Merrick B. Garland), <https://bit.ly/3DSJGvF>.

I would also like to thank DOJ for increased transparency⁵ in its recent reports.⁶ In those reports, DOJ has confirmed long-standing stakeholder fears that the FSA’s use of risk assessment tools would “exacerbate racial and socioeconomic disparities.”⁷ In November 2021, the Bureau of Justice Statistics reported that the vast majority of Black people in BOP custody—more than 70 percent—were classified by PATTERN as medium or high risk.⁸ Then, in December 2021, NIJ reported that PATTERN overpredicts⁹ the likelihood that Black, Hispanic, and

⁵ Although transparency has increased, there is room for improvement. Dr. Melissa Hamilton, an expert in risk assessments, recently explained to Congress that although recent reports are “helpful in providing a host of various statistics to provide outsiders a better understanding of PATTERN. . . .this is not a fully acceptable alternative to making publicly available an anonymized version of the dataset(s) for independent researchers. I have discovered various statistical and textual errors in the NIJ Report itself. As with the revelation of problems in the initial PATTERN development by new consultants, verification of the work of these consultants might well be better confirmed by others.” See *Hearing on the ‘The First Step Act, The Pandemic, and Compassionate Release: What Are the Next Steps for the Federal Bureau of Prisons? Before the United States House of Representatives Committee on the Judiciary Subcommittee on Crime, Terrorism, and Homeland Security Oversight*, 117th Cong. 4 (Jan. 21, 2022) (statement of Dr. Melissa Hamilton), <https://bit.ly/3xT5nI7>.

⁶ See Off. of Att’y Gen. *U.S. Dep’t of Justice, First Step Act Annual Report* (Apr. 2022), <https://perma.cc/A49Q-8LHN> (April 2022 DOJ Report); Nat’l Inst. Justice, *2021 Review and Revalidation of the First Step Act Risk Assessment Tool* (Dec. 2021), <https://perma.cc/3C6X-KC97> (December 2021 NIJ Report); E. Ann Carson, Bureau of Justice Statistics, *Federal Prisoner Statistics Collected under the First Step Act* (Nov. 2021) (BJS Report), <https://perma.cc/3YFZ-8GW2>; Nat’l Inst. Justice, *2020 Review and Revalidation of the First Step Act Risk Assessment Tool* (Jan. 2021), <https://bit.ly/3cNrm93>.

⁷ See, e.g., Statement for the Record of the ACLU, Justice Roundtable, and The Leadership Conference in Response to DOJ April 3 and 5 Listening Sessions (Apr. 12, 2019), <https://bit.ly/3dK5z5D>; see also Nat’l Inst. Justice, *Stakeholder Statements Submitted in Response to NIJ’s First Step Act Listening Sessions* (Jul. 2020), <https://bit.ly/3kjGm2E>.

⁸ Carson, BJS report.

⁹ According to NIJ, “overprediction means that the instrument assigns a higher recidivism risk score for the race group relative to white group.” See NIJ, *Predicting Recidivism: Continuing to Improve the Bureau of Prisons’ Risk Assessment Tool, PATTERN* (Apr. 19, 2022), <https://bit.ly/3BNfqjd>

Statement of Patricia L. Richman
September 29, 2022

Asian people will be arrested or returned to custody for rules violations after leaving prisons.¹⁰

DOJ has taken some steps to respond to these disparities. In April 2022, it announced that it had revised the cut points for the general risk scales in order to, *inter alia*, “help mitigate the effects of various racial and ethnic disparities associated with the previous risk groupings.”¹¹ Even with these changes, significant disparities persist. In April, DOJ estimated that 56.9 percent of Black men would remain in the higher risk categories, in contrast to only 34.8 percent of white men.¹²

DOJ must move faster and more aggressively to remediate these problems. Although we are glad that DOJ is soliciting the views of outside voices, we are concerned to find the DOJ still in “listening mode.” The time for meaningful collaboration and a two-way dialogue between DOJ and external stakeholders is long past due. Since before the FSA’s enactment, Defenders—along with many others—have submitted detailed comments and testimony proposing reforms and specific changes to mitigate bias in risk assessment.¹³ Despite this, as I speak today, a racially biased tool continues to dictate who benefits from the FSA’s reforms. We urge DOJ to accelerate its work by immediately publishing a proposal to correct PATTERN’s deficiencies that includes specific deliverables, due dates, and a continuing commitment to greater transparency.

¹⁰ December 2021 NIJ Report at 4.

¹¹ April 2022 DOJ Report at 14.

¹² April 2022 DOJ Report at 14-15.

¹³ *See supra*, note 2; *see also* Stakeholder Statements, *supra* note 7.

Now, turning to today’s agenda. For the reasons explained below, and in prior submissions incorporated here, the Department should rework PATTERN’s inputs and scoring scheme to reduce embedded bias and improve people’s ability to reduce their risk score. In addition, the Department should redefine “recidivism” to include only convictions.

I. Refinements to PATTERN’s inputs and scoring scheme

A. *Criminal History.*

DOJ should revise PATTERN to correct its overemphasis on criminal history.¹⁴ Under the currently approved version of PATTERN (1.3), a person can be assigned from 0 to 40 points for criminal history; and 40 points alone can trigger designation as “medium risk” on the male and female general risk scales. This is inconsistent with Congress’ requirement that “prisoners at each risk level have a meaningful opportunity to reduce their classification during the period of incarceration.”¹⁵ Not only that, but the effect of criminal history is multiplied across PATTERN through other inputs that also count criminal history.¹⁶

¹⁴ See Jennifer L. Skeem & Christopher T. Lowenkamp, *Risk, Race, and Recidivism: Predictive Bias and Disparate Impact*, 54 CRIMINOLOGY 680, 705 (2016) (raising the possibility that “efforts could be undertaken in the risk assessment domain to rely less heavily on criminal history”).

¹⁵ 18 U.S.C. § 3632(a)(5); *see also* 18 U.S.C. § 3632(a)(4) (requiring the method for reassessing recidivism risk be “based on factors . . . that are dynamic and that can reasonably be expected to change while in prison”).

¹⁶ These factors include: sex offender, history of escape, history of violence, all incident reports, serious incident reports, and time since last serious incident report.

Experts have long warned that an overemphasis on static factors like criminal history in risk assessment can embed racial disparities and lead to unfairness. This is because criminal history is not just the product of participation in crime, but also of biased practices throughout the criminal legal system (for example, overpolicing underprivileged neighborhoods, disparate victim reporting, prosecutorial decisions).¹⁷

One proposal to mitigate these problems would be to reduce the heavy weight criminal history carries and incorporate a “decay factor” in its calculation.¹⁸ “[A]ny presumption that the vast majority of offenders pose a constant and lifelong risk is not supported by empirical evidence,” and studies show a “significant decay in the predictive ability of a prior criminal event.”¹⁹ “All of this empirical knowledge strongly calls for recidivism premiums and risk assessment tools to curtail the use of criminal history from a temporal perspective.”²⁰

PATTERN does not follow this approach. Instead, a person’s criminal history upon entry to the BOP (usually, as calculated at the time of sentencing by the

¹⁷ See Amy Baron-Evans & David Patton, *A Response to Judge Pryor’s Proposal to “Fix” the Guidelines: A Cure Worse than the Disease*, 29 FED. SENT’G. REP. 104, 112 (Dec. 1, 2016-Feb. 1, 2017) (collecting studies); see also Radley Balko, Op-Ed., *There’s Overwhelming Evidence that the Criminal-Justice System is Racist. Here’s the Proof*. WASH. POST (Updated Jun. 10, 2020), <https://wapo.st/3xTx6bF>.

¹⁸ See Melissa Hamilton, *Back to the Future: The Influence of Criminal History on Risk Assessment*, 20 BERKELEY J. CRIM L. 75, 123 (2015) (compiling studies).

¹⁹ *Id.* at 124.

²⁰ *Id.* at 125.

district court) is set in stone throughout their term of imprisonment.²¹ So, even if twelve years have passed since a person was sentenced to sixty days' imprisonment for petty theft, PATTERN will continue to count it.²² This differs from how the U.S. Sentencing Guidelines instructs courts to calculate criminal history. For example, short sentences of at least sixty days and less than a year and one month are not counted towards criminal history if they were "imposed more than ten years prior to the defendant's commencement of the instant offense."²³ DOJ should incorporate the periodic recalculation of criminal history scores into PATTERN to reflect its diminishing predictive value over time.

B. *Disciplinary Infractions.*

PATTERN also embeds bias and unfairness by giving disciplinary infractions heavy dynamic weight. A single guilty disciplinary infraction for a man in BOP custody could conceivably trigger the addition of points under several categories of the general recidivism scale: "all incident reports" (+1); "serious incident reports" (+1); "time since last incident report" (less than 3 months, +3). Further, it appears that a single disciplinary event could also receive points if it is deemed to be

²¹ Criminal history is defined as "[t]he number of criminal history points taken from the most recent BRAVO available." December 2021 NIJ Report. This is typically "derived from the US Sentencing Guidelines Criminal History Points, as reflected in the final judgment and the [Statement of Reasons]. If not found in either the Judgment or [Statement of Reasons], use the points assessed by the [U.S. Probation Office] in the [presentence report]." See Bureau of Prisons Program Statement 5100.08, CN-1 at 8 (Sept. 14, 2019), https://www.bop.gov/policy/progstat/5100_008cn.pdf

²² *Id.*

²³ See U.S.S.G. §4A1.1 App. Note 2.

Statement of Patricia L. Richman
September 29, 2022

“violent” or an “escape” by triggering the variables of “history of violence” (+7) or history of escape (+9).”²⁴

The piling on of points for disciplinary infractions is problematic for several reasons. First, the prison disciplinary process comes with weak procedural protections, raising questions of factual accuracy and reliability. In addition, the disciplinary regime affords prison officials broad discretion, potentially incorporating their biases.²⁵ “Social scientists have demonstrated that there is a relationship between race and disciplinary infractions in prison,” although many have “failed to consider the role that racial bias play.”²⁶

Second, the acts that can qualify for discipline are often relatively minor. For example, BOP’s inmate discipline program counts the following as infractions: malingering, feigning illness, obscene language, being unsanitary or untidy, or being absent from an assignment.²⁷ Disciplinary infractions can also vary widely across prisons, meaning that an infraction often says more about the place of a person’s incarceration and its culture than that person’s conduct. A November 2021

²⁴ In a 2020 example of scoring people under PATTERN, DOJ gave several examples where individuals who committed a serious violent incident in prison were penalized through infractions enhancements, and by adding 7 points under “history of violence.” See Dep’t of Just., *The First Step Act of 2018: Risk and Needs Assessment System – UPDATE*, 32 (Jan. 2020), <https://bit.ly/3Sy9qS4>.

²⁵ See Andrea C. Armstrong, *Race, Prison Discipline, and the Law*, 5 UC IRVINE L. REV. 759, 761, 762 (2015) (exploring the “potential implications of implicit bias research in the prison disciplinary context”).

²⁶ See Katie M. Becker, *Race and Prison Discipline: A Study of North Carolina Prisons*, 43 N. CAROLINA CENT. L. REV. 2, 3-4 (gathering studies).

²⁷ Bureau of Prisons, Program Statement 5100.08, Inmate Discipline Program (2011).

BJS report confirmed that there are wide disparities in how disciplinary programs are administered across BOP. A random selection of five institutions demonstrates the point:

| | Population (#) ²⁸ | Total Infractions (#) | Rate of Infractions by Person on Average | Total at Greatest Severity (#) | Rate of Infractions at Greatest Severity (%) |
|-------------------------------|------------------------------|-----------------------|--|--------------------------------|--|
| Big Sandy, USP | 1249 | 1196 | 0.95 | 366 | 30% |
| Allenwood, USP | 550 | 760 | 1.4 | 143 | 18% |
| Atlanta USP (medium security) | 1767 | 1192 | 0.67 | 576 | 48% |
| Marianna-FCI | 503 | 86 | 0.17 | 49 | 57% |
| Terminal Island-FCI | 705 | 95 | 0.13 | 26 | 27% |

As this sample shows, both the rate and severity of infractions vary widely across institutions.

Finally, the heavy weight of disciplinary infractions in PATTERN also lends itself to abuse by staff. Reports of retaliation and abuse of individuals who are incarcerated in BOP are common.²⁹ An abusive correctional officer could strategically use disciplinary infractions to threaten incarcerated people with an

²⁸ Note that the population numbers in this Table were from December 31, 2020, as published on BOP's website. The data included may not be exact matches to BJS. See Bureau of Prisons, *Population Statistics* (Dec. 31, 2020), available at <https://bit.ly/3C1R4P7>.

²⁹ See, e.g., Letter from Hon. Ron Wyden, & Hon. Jeffery A. Merkeley, United States Senators (D-Or.) to Collette Peters, Dir. Of Fed. Bureau of Prisons (Aug. 18, 2022) (requesting an update about press reports of "retaliation and violence against inmates at FCI Sheridan for speaking about their experiences regarding unaddressed medical needs; small cell confinement; and limited access to family and lawyers as a result of the pandemic"), <https://bit.ly/3Cexggt>.

increased risk score, and a corresponding loss of the ability to apply earned time credits.

C. *Number of programs completed.*

DOJ should clarify which programs are counted under the “number of programs completed” variable and consider amending it to require “participation” rather than “completion.” PATTERN provides people the opportunity to reduce their risk score based on the the number of programs they have “successfully completed.” NIJ has defined this variable as “The number of ACE, Brave, Challenge, Drug Education, Life Connections, Parenting, Skills, Sex Offender Residential Treatment, STAGES, and Step-Down courses successfully completed during the current incarceration.”³⁰ Programs are combined into groupings, with the maximum (11+) triggering a reduction of 8 to 12 points, depending on the scale. It is unclear what programs this variable includes: All programs? Only the ten specifically listed? Those included in the approved program guide?³¹

The “completion” requirement is also problematic. In the context of earned time credits, BOP agreed with comments that “by focusing on completion, BOP diminishes the value of participation and weakens the incentive structure Congress enacted,” and that “there are myriad situations where people would successfully

³⁰ December 2021 NIJ Report at 12; *see also 2020 Review and Revalidation of the First Step Act Risk Assessment Tool*, App’x A (Jan. 2021) (Operationalization of PATTERN variables), <https://bit.ly/3DVQvwj> (listing “Adult Continuing Education (ACE), Parenting, Life Connections Program, Brave, Challenge, Skills, Sex Offender (residential or nonresidential), Stages, and Step-Down Courses.”)

³¹ *See* Bureau of Prisons, *First Step Act Approved Programs Guide* (Aug. 2022), <https://bit.ly/3LJUCxo>.

participate in an approved program and—through no fault of their own—be prevented from, or delayed in, completing it.”³² In addition, BOP agreed that “[t]ransfers, program resource and staffing limitations, and facility movement restrictions all impact program completion, as do length of sentence, program availability and waitlists.” BOP should clarify what programs are counted by this variable and consider eliminating the “completion” requirement.

D. *Work programs completed.*

PATTERN limits the number of points that can be awarded for “work programs completed” to a maximum of 2 or 4 points, depending on the scale. This variable is too narrowly defined and accorded too little weight.

First, the variable does not appear to include non-UNICOR work assignments. It is defined as “the number of technical and vocational courses completed during the current incarceration,” including federal industry employment UNICOR.³³ In other words, it seems that a person could spend 10 years diligently working in a non-UNICOR job as a janitor while in prison but see no reduction in their risk profile as a result.

Second, even if other types of work *do* count under this variable, only “completed” programs result in the subtraction of points. NIJ has specified that for UNICOR employment, working at least one day constitutes completion, but it does

³² See FSA Time Credits, 87 Fed. Reg. 2705, 2711 (codified at 28 C.F.R. 523 & 541).

³³ December 2021 NIJ Report at 12.

not define the meaning of “completion” for any other work programs included in the variable. In the context of the earned time credit rule, however, BOP has agreed that a “prison job is not a ‘program to complete,’” because it has no set duration, and its success is based on continued employment and supervisor evaluations.”³⁴ That logic applies with equal force here: a person’s continuing and successful employment should be counted in their favor as reducing their risk of recidivism. It should not count just once but be credited repeatedly over the period of a person’s successful employment. For example, a person could earn a reduction in their risk score for every year of successful employment.

In sum, DOJ should broaden the definition of this variable to include ongoing work assignments, remove the “completion” requirement, and give this variable greater dynamic weight in reducing a person’s risk score.

II. Recidivism

DOJ should reorient PATTERN to focus on predicting the risk of reconviction, rather than mere rearrest or return to custody (which can encompass technical violations of supervised release). As we explain in our Brennan Center report:

Arrest is a poor proxy for criminal activity, as it may reflect policing decisions—where officers are deployed and their biases in making arrests—rather than actual criminality. Designing a risk assessment tool based on arrest patterns also means relying on data

³⁴ FSA Time Credits at 2710 (agreeing with comments).

tainted by decades of discriminatory policing, a concern raised by stakeholders years ago.³⁵

For this and other reasons, DOJ should redefine recidivism to include only convictions,³⁶ and also consider narrowing the *type* of convictions to exclude petty offenses and misdemeanors.³⁷ Although DOJ has in the past rejected such calls on the ground that a narrower definition would be unworkable because of data limitations,³⁸ there is evidence to the contrary. For example, the United States Sentencing Commission has repeatedly managed to compile reconviction information about large groups of released people for its recidivism reports.³⁹ Further, U.S. Probation creates thousands of presentencing reports each year that count only convictions, suggesting that the federal government has access to this information.

³⁵ Grawert & Richman, *First Step* at 4.

³⁶ *See id.* at 6.

³⁷ The Independent Review Committee similarly recommended that NIJ and its consultants explore the feasibility of employing a “felony arrest only” outcome measure. While Defenders agree that the definition should be narrowed, we are concerned that the use of arrest as a proxy for criminal conduct embeds inequity and racial disparities. *See* Independent Review Committee, *Report of the Independent Review Committee Report Pursuant to the Requirements of Title I Section 107(g) of the First Step Act (FSA) of 2018* (P.L. 115-391) 3 (December 21, 2020).

³⁸ *See* U.S. Dep’t of Just., *The First Step Act of 2018: Risk and Needs Assessment System—Update* at 13-14 (Jan. 2020), <https://bit.ly/3SfiGut>.

³⁹ *See, e.g.*, USSC, *Recidivism Among Federal Offenders* 4 (2019), <https://bit.ly/3r9WnKR> (including data regarding reconviction); USSC, *Recidivism Among Federal Offenders: A Comprehensive Overview* 15 (2016), <https://bit.ly/2NCafYW> (using three different measures of recidivism, including rearrest, reconviction & reincarceration); USSC, *Recidivism Among Federal Drug Trafficking Offenders* 4 (2017) <https://bit.ly/3dLDZ7W> (same).

Statement of Patricia L. Richman
September 29, 2022

III. Conclusion

We are grateful for today's listening session, and hope that there will be additional opportunities for feedback and discussion moving forward. The Federal Public Community Defenders are eager to partner with the Department of Justice to expand the reach and improve the fairness of the FSA's prison reforms.

EXHIBIT A

The First Step Act's Prison Reforms

Uneven Implementation and the Path Forward

By **Ames C. Grawert and Patricia L. Richman**

PUBLISHED SEPTEMBER 23, 2022

Three years ago, Congress passed the First Step Act, the first major federal criminal justice reform legislation in nearly a decade.¹ The culmination of years of bipartisan advocacy, the law included both long-overdue changes to excessively punitive federal sentencing laws and reforms aimed at improving conditions in the federal prison system.

This brief examines the structure of the First Step Act's prison reforms, how they have been implemented, and what more Congress and the Department of Justice (DOJ) must do to realize their potential.

Background

In the 1980s, Congress enacted rigid mandatory penalties, which require judges to impose minimum terms of incarceration for certain federal crimes or when certain statutory criteria are satisfied.² Sometimes these penalties are triggered by specific conduct, such as possessing a firearm or possessing drugs above a specified threshold quantity.³ These laws significantly expanded the size of the federal prison system and led to an explosion in racial disparities in punishment, all without addressing drug use or improving public safety.⁴

The federal prison system strained under the effects of these penalties, with lawmakers describing a “state of

crisis” as the prison population climbed.⁵ Indeed, the federal prison population grew eightfold between the 1980s and the mid-2010s, outpacing growth in state-level incarceration, with weapon and drug offenses making up more than 60 percent of the growth in federal imprisonment.⁶ Mechanisms for checking excessive custodial sentences did not keep pace. Compassionate release, for example, which allows a federal court to reduce or end a prison sentence for “extraordinary and compelling circumstances,” was severely underused.⁷ People in prison had limited opportunities to earn early release through their conduct; “good time” credits, earned for good behavior while incarcerated, amounted to at most 47 days per year of incarceration.⁸

Congress began to rectify this situation in 2010 with the Fair Sentencing Act, which reduced the crack/powder cocaine sentencing ratio from 100:1 to 18:1 in drug trafficking cases and eliminated the five-year mandatory minimum for simple possession of crack.⁹ However, these changes applied only prospectively, meaning that people sentenced before the law went into effect remained in prison, serving the same wildly disproportionate sentences that Congress had just repudiated. Additionally, the act did not address the long-standing consensus that federal prisons were failing to provide meaningful programming and rehabilitation to incarcerated people.

The First Step Act attempted to address these shortcomings. Among other things, its sentencing reforms made the Fair Sentencing Act retroactive, permitting people sentenced under the old 100:1 crack/powder cocaine penalty scheme to apply for resentencing as if the Fair Sentencing Act's 18:1 ratio had been "in effect at the time the covered offense was committed."¹⁰ The act also revised other mandatory minimums and for the first time allowed people in federal prison to petition a federal court for compassionate release.¹¹ Additionally, and critically, it introduced a system for people to reduce their time spent in prison by participating in programming and activities.

The First Step Act's changes to federal sentencing laws had an immediate, significant impact. As of May 2021, roughly 3,700 people had benefited from a reduction in their sentence under the provision making the Fair Sentencing Act retroactive.¹² The average sentence was reduced by around three years, or roughly 25 percent.¹³ And, according to the Bureau of Prisons, the act has led to compassionate releases or sentence reductions in more than 4,200 other cases.¹⁴ The act's other sentencing reforms have begun to impact new cases, too. The restoration of judicial discretion in select drug cases benefited roughly 1,400 people in the first year of enactment.¹⁵

The FSA's Prison Reforms: Challenges and Successes

Unfortunately, errors and half-starts have marred the roll-out of the First Step Act's prison reforms.

The act aimed to create a system that would encourage people in prison to participate in programming designed to reduce recidivism — that is, the risk that they would come in contact with the criminal justice system, through arrest or otherwise, after release. This new system has several key components:

- **Incentives and rewards for program participation.** The act incentivizes engagement with "evidence-based recidivism reduction programming" (e.g., drug treatment or literacy programs) and "productive activities" (e.g., work or vocational training) by awarding time credits for participation that people can apply toward early transfer to supervised release, home confinement, or a residential reentry center (i.e., a halfway house).¹⁶ However, a laundry list of disqualifications excludes people convicted of certain offenses from participating.¹⁷
- **A risk assessment tool.** This tool is designed to determine the "recidivism risk of each prisoner" and the "type and amount of evidence-based recidivism reduction programming for each." Risk scores generated by the tool affect both the number of time credits individ-

uals can earn and how they can redeem them, making its design and implementation vital to the act's success.¹⁸

- **Expanded recidivism reduction programming opportunities.** The act also seeks to expand the availability of job training and other programming for all incarcerated individuals.¹⁹

Unlike the act's sentencing reforms, these changes to the prison system were phased in, to give the Bureau of Prisons (BOP) time to build out the relevant policies. The final rule governing the awarding of time credits did not go into effect until January 2022.

Along the way, the act's corrections reforms have hit a series of snags in implementation, requiring continued attention from policymakers to ensure that they succeed. What went wrong?

Rocky Start to Earned Time Credit Implementation

The core of the First Step Act's corrections reforms is a system that allows some people to receive earned time credits (ETCs) for participating in designated programming or activities behind bars, such as drug treatment or cognitive behavioral therapy. ETCs effectively shorten sentences, making them a powerful incentive for participating in programs. Implementation of the ETC system stalled, however, after the act's passage.²⁰ The BOP initially refused to award ETCs and issued draft regulations that would have sharply limited the reach of the program.²¹

That changed in January 2022 when the BOP, responding to broad and bipartisan criticism, issued a rule that dramatically increased the rate at which completed programs translate to ETCs. Simultaneously, the BOP calculated and awarded ETCs retroactively back to the date of the First Step Act's enactment.²² As a result, by April 5, 2022, more than 6,100 people had been transferred to supervised release and another 3,155 people had been transferred to prerelease custody — that is, a residential reentry center or home confinement.²³

Under the law and current regulations, eligible people can earn 10 to 15 days of ETCs for every 30-day period in which they successfully participate in programs or productive activities.²⁴ (People deemed minimum or low risk earn 15 days, while those classified at higher risk levels earn 10 days.) Once they have accrued time credits equal to the time remaining on their sentence, those credits can potentially be cashed in for either an early transfer to supervised release (by up to a year) or to prerelease custody.²⁵

Yet not all people are eligible to earn ETCs. The act excludes those convicted of many crimes.²⁶ These exclusions prevent nearly half of the federal prison population from benefiting from credits.²⁷ The exclusions also

appear to serve no policy purpose. According to the Independent Review Committee (IRC), the expert panel tasked with helping the BOP implement the law, there is “no significant difference in the collective recidivism-risk profiles of the BOP’s ETC-eligible and ETC-ineligible inmate populations.”²⁸

Additionally, despite the announcement of new rules improving the ETC system’s operation, it appears that the process of awarding ETCs continues to move slowly. People in federal prison have reported that they are not having their time credits applied or not being released as early as they should be, and that BOP staff have not received training or guidance in how to operate the new system. These missteps point to a need for continued oversight to ensure timely and accurate implementation.²⁹

Flawed Risk and Needs Assessment Tools

The First Step Act calls for an assessment system to evaluate each person’s risk of recidivism (defined as arrest or return to prison within three years of release) and criminogenic needs (factors that, unless addressed, may predict future contact with the criminal justice system).

The act’s corrections reforms rely on the system’s accuracy and fairness. People who score as minimum or low risk earn ETCs more quickly than those in higher risk categories. The BOP is also required to apply credits earned by people categorized as minimum or low risk toward prerelease custody or supervised release. In contrast, individuals classified as medium or high risk must meet additional criteria, including special approval by the warden, before their credits can be applied against their sentence. Additionally, those in the higher risk categories cannot earn early transfer to supervised release; they can apply their credits only to prerelease custody.³⁰

These high stakes make it particularly important that the risk and needs system be transparent, fair, and unbiased.³¹ Unfortunately, the part of the system focused on criminogenic needs was slow to be deployed. Worse, the system released by the DOJ for assessing risk, called PATTERN (Prisoner Assessment Tool Targeting Estimated Risks and Needs), remains flawed despite major revisions and renewed attention from the Biden administration.³²

Some elements of PATTERN have improved over time as the DOJ has responded to stakeholder criticism. But correcting all of its flaws would require fundamentally reevaluating what type of risk PATTERN measures and how it translates that risk to policy judgments — a top-to-bottom reconstruction that goes beyond the remedies the administration has proposed to date.

Background on PATTERN

Like any risk assessment tool, PATTERN was developed to predict the likelihood of a defined behavior based on

a series of inputs associated with that behavior. It seeks to predict a specific type of recidivism: “a return to BOP custody or a rearrest within three years of release from BOP custody, excluding all traffic offenses except driving under the influence and driving while intoxicated.”³⁸

The tool works by collecting information on a person and assigning them points based on factors in their background, which add up to a total risk score. Each factor carries a different weight; for example, completing a prison program will subtract one to three points from the total score. By contrast, being under the age of 26 can add a significant number of points. The BOP then determines how to translate these scores into a policy judgment — whether the person presents a minimum, low, medium, or high level of risk.³⁹

Different versions of the tool exist for men and for women, and for predicting general recidivism risk (defined as the likelihood of rearrest or return to BOP custody for any offense) and violent recidivism risk (the same likelihood but for an offense deemed violent).⁴⁰ These tools serve

Risk Assessment Tools in the Criminal Justice System

Risk assessment tools are designed to calculate the likelihood that someone will engage in certain behaviors based on characteristics they share with others in a group. Predictions tend to rely on a mix of dynamic factors, which people can change over time (such as participation in prison programming), and static factors, which people cannot change (such as age and criminal history). Before being put into practice, these actuarial tools are usually calibrated by comparing their predictions against data gathered from the real world.³³

Risk assessment isn’t math; it’s policy. Indeed, these models reflect policy choices at every phase, from judgments about risk tolerance to beliefs about criminal behavior. As a result, they remain both controversial and imperfect. They often classify individuals inaccurately and tend to misclassify people who are unlikely to pose a real threat as moderate or high risk.³⁴ An overreliance on static factors can also lead these tools to incorrectly classify Black people and others from overpoliced and disadvantaged communities as high risk.³⁵ Last, because tools are tested against historical data, they risk replicating biases in that data when predicting risk. For example, historical racial disparities in policing can lead an algorithm that was calibrated using arrest data to predict, incorrectly, that Black and white people reoffend at different rates — when in fact they are merely policed at different rates.³⁶ In the words of one expert, “No predictive tool is better than the data set from which it originates.”³⁷

different purposes, but — critically — when the risk tiers they produce diverge, the higher score governs. Practically, this means that someone’s potential ability to transfer to prerelease custody is limited until and unless they score into the low or minimum tiers on both assessments.⁴¹

Missteps in PATTERN’s Development (2019–Early 2022)

PATTERN’s rollout was characterized by implementation mistakes and policy missteps, raising concerns that the tool entrenched racial bias in the prison system, relied on an overly conservative definition of recidivism risk, and failed to account for people’s capacity for personal growth and change while incarcerated. These issues plagued the first years of implementation, even if (as discussed below) the DOJ has since taken steps to partially correct them.

- **Technical errors.** PATTERN required repeated technical corrections just to function as designed.⁴² Reviews conducted by the National Institute of Justice (NIJ), the DOJ agency tasked with reviewing the tool to guarantee its accuracy, revealed mistakes in how the model’s variables were defined and applied.⁴³ A January 2021 NIJ report identified scoring, coding, and weighting errors and recommended a revised version of PATTERN.⁴⁴ In the months that followed, the NIJ found even more errors, described in a December 2021 report.⁴⁵ These problems were compounded by human error by BOP staff in the scoring process. As a result, many people in BOP custody were assigned to the wrong risk category, even as the DOJ relied on PATTERN to make potentially life-and-death decisions about whether to transfer people to home confinement during the pandemic.⁴⁶
- **Racial bias.** In November 2021 the Bureau of Justice Statistics (BJS) reported that the vast majority of Black people in BOP custody — more than 70 percent — were classified by PATTERN as medium or high risk.⁴⁷ Further releases show that PATTERN continues to overpredict the likelihood that Black, Hispanic, and Asian people will commit new crimes or violate rules after leaving prisons, relative to white people in prison.⁴⁸

These biases stem in part from PATTERN’s focus on rearrest. Arrest is a poor proxy for criminal activity, as it may reflect policing decisions — where officers are deployed and their biases in making arrests — rather than actual criminality.⁴⁹ Designing a risk assessment tool based on arrest patterns also means relying on data tainted by decades of discriminatory policing, a concern raised by stakeholders years ago.⁵⁰

For that reason, the IRC, among others, has recommended a narrower definition of recidivism, one that “might better identify individuals likely to engage in serious criminal activity post-release” and potentially reduce the tool’s reliance on racially biased data.⁵¹ In

2020, however, the DOJ rejected calls to redesign PATTERN to include a narrower definition of recidivism, such as reconviction or re-incarceration, claiming such a definition would be unworkable because of data limitations.⁵²

- **Risk tolerance.** Translating a PATTERN score to a risk designation is a policy judgment. There is no objective quantifier of acceptable risk.⁵³ Instead, it is up to policymakers and the algorithm’s designers to define risk and decide how much of it their system should tolerate.⁵⁴ In making those judgments, the DOJ initially adopted a set of fairly conservative cut points — borders between risk categories — that were derived in part from the average predicted risk of recidivism of people released from the BOP.⁵⁵ But early data about recidivism among those scored by PATTERN showed that fewer than 2 percent of people scored as low risk were rearrested between July 2019 (when PATTERN was implemented) and September 2020. Further, only 4.5 percent of “high risk” individuals were rearrested in the study period.⁵⁶ That data led the Urban Institute to conclude that PATTERN overpredicted recidivism and that individuals with a higher PATTERN score should properly be classified in a lower risk category.⁵⁷
- **Variable weights.** Choosing which factors in someone’s background PATTERN should consider when scoring them is also a policy choice, as is the weighting of those factors — even if the latter is informed by mathematics.

Contrary to Congress’s intent, PATTERN has consistently overemphasized static factors like age and criminal history, which makes it difficult for people to change their assessment score over the course of their incarceration. Dynamic factors are weighted less heavily, and more than half of the dynamic factors in the model actually increase a person’s risk score.⁵⁸ Perhaps unsurprisingly, then, just 23–35 percent of people in one NIJ analysis had been able to reduce their assessed risk level “at the last assessment compared to the first.”⁵⁹ The stickiness of PATTERN assessments suggests that the tool does not yet appropriately account for personal growth — or program participation — during incarceration.⁶⁰

Revisions to PATTERN (April 2022–Present)

In April 2022 the DOJ finally announced a plan to address some of these issues. The attorney general approved a new version of PATTERN that attempts to fix the problems discussed above, and the BOP announced that it would use the new tool to rescore all individuals in its custody.⁶¹

Alongside this revision, the DOJ announced that it would revise the cut points for evaluating general recidivism risk to increase the percentage of people who qualify

TABLE 1

Share of Men Classified as Minimum or Low Risk

| | TOTAL | WHITE | BLACK | HISPANIC | NATIVE AMERICAN | ASIAN |
|---------------------|-------|-------|-------|----------|-----------------|-------|
| Original cut points | 44.0% | 55.7% | 31.7% | 51.0% | 21.7% | 70.0% |
| Revised cut points | 55.2% | 65.2% | 43.1% | 64.4% | 33.8% | 78.9% |

Source: U.S. Department of Justice, *First Step Act Annual Report*, 15–16.

for lower risk tiers. Under the new cut points, 55 percent of the BOP’s male population and 83 percent of its female population are projected to fall into the minimum- and low-risk categories under the general tool — up from 44 and 78 percent, respectively. Cut points for determining violent recidivism risk remain unchanged.⁶²

This change was specifically adopted to “help mitigate the effects of various racial and ethnic disparities associated with previous risk groupings” — that is, to reduce the tool’s racial bias — and will indeed result in more Black and Hispanic people being categorized as minimum or low risk, as shown in table 1.⁶³ However, the changes will not correct PATTERN’s tendency to overpredict the recidivism risk of nonwhite people, as the DOJ itself conceded.⁶⁴

Nor does this revision really address PATTERN’s overly conservative risk profile, because it leaves the cut points for evaluating risk of violent recidivism unchanged. As shown in table 2, those cut points erect a high bar for inclusion in the lower risk categories. As currently drafted, fully 80 percent of people classified as a medium risk of recidivism would go on to have no rearrest for a violent offense, to say nothing of reconviction or re-incarceration. And recall that risk assessment outcomes do not lead to immediate release; they instead determine “who is given more robust incentives to engage in rehabilitative

programming and who might earn a change in the type of prerelease custody.”⁶⁵

Changes to the general tool’s cut points will likely allow people to earn ETCs faster under the provision granting additional ETCs to people with low PATTERN scores.⁶⁶ But they may have limited impact on the ETC system’s broader functioning, as a high violent score will continue to override a low general score when determining someone’s eligibility to apply ETCs to prerelease custody or supervised release.⁶⁷ This limited impact also means racial disparities in access to prerelease custody will persist.

Taken together, the April 2022 revisions fall short of addressing every problem with PATTERN and the First Step Act’s use of risk assessment tools generally. But these remedial efforts represent progress, as does the greater level of transparency in the DOJ’s and NIJ’s more recent reports.⁶⁸ Furthermore, the DOJ’s April report pledged to “consider all legally permissible options for reducing the differential prediction based on race and ethnicity,” including potentially revisiting PATTERN’s focus on rearrest risk — an important, if overdue, break with the previous administration’s policy.⁶⁹

Delayed Criminogenic Needs Tool

In March 2022 the DOJ finally released a report on the last component of the act’s prison reforms: its assessment tool designed to identify and address each “prisoner’s specific criminogenic needs.”⁷⁰ The tool, SPARC-13, is intended to complement PATTERN by directing people in prison toward programming that meets their needs.

While it is too early to evaluate SPARC-13’s implementation, several details from the DOJ’s initial report stand out. For one, BOP staff report low levels of familiarity with the needs assessment system and how to administer it.⁷¹ The tool also appears to rely heavily on self-reporting by imprisoned people themselves, leading to a very high rate of refusal for some assessment areas. Nearly one-third of those assessed refused to be screened for trauma, for example.⁷² The BOP should work to reduce this refusal rate, including by making people aware that by refusing screening they may forgo opportunities to earn time credits.⁷³

TABLE 2

Predicted Likelihood of No Post-Release Arrest for a Violent Crime (Male)

| PATTERN RISK CATEGORY | PREDICTED LIKELIHOOD |
|-----------------------|----------------------|
| Minimum | 98.6% |
| Low | 90.8% |
| Medium | 79.8% |
| High | 63.9% |

Source: U.S. Department of Justice, *First Step Act Annual Report*, 13.

Lack of Adequate Prison Programming

For the ETC system to succeed, the BOP must significantly expand the availability of programming and productive activities. That is a heavy lift. Historically the BOP has failed to provide enough programming to satisfy the needs and wants of the people in its custody, even before accounting for the contemplated expansion.⁷⁴ Pandemic restrictions significantly hindered what programming was available; many programs were “highly impacted” by the virus, and some shut down entirely.⁷⁵ But the IRC warned in 2020 that “even a full return to pre-COVID-19 BOP programming levels will not be sufficient to make available evidence-based recidivism reduction programs and productive activities” for all eligible individuals in BOP custody by 2022, and it identified troubling demographic disparities in program participation among eligible individuals.⁷⁶

The final ETC system offers a partial response to this problem by allowing eligible people to earn credits if their program is interrupted through no fault of their own. But it is unclear whether this rule applies when a person is unable to even start a program. The rule’s explanatory text states that “inmates will not be penalized if specifically recommended [programs] are unavailable to them or at full enrollment at their facilities,” but the rule itself references only program “interruption.”⁷⁷ This lack of clarity is concerning, particularly in light of a recent BJS report showing that the programs providing the most hours of credit were, at least through the end of 2021, also the least available.⁷⁸ Some reports also indicate that BOP staff are not receiving clear guidance from DOJ leadership on how to implement new policies and regulations, and that programs may not be staffed or resourced to ensure a prompt and faithful rollout — further complicating an already byzantine system.

That said, recent information from the DOJ gives some reason for optimism. The DOJ’s latest publication describes a significant increase in the number of approved programs available to people in prison. During FY 2021, BOP staff also “recorded a marked increase in participation” in prison programming and productive activities. Furthermore, the BOP recently posted and filled a wide range of positions related to First Step Act implementation and contracted for evaluations of its programs. Last, the DOJ reported that more than \$362 million “in appropriated FSA funding” had been used “to expand reentry programs and their delivery.”⁷⁹ It is not clear how much of that funding was new, how much was distributed from elsewhere within the DOJ, and how much the BOP may still need to make up for programming shortfalls.⁸⁰

The Path Forward

For all its successes, the First Step Act continues to fall short of its promise. Yet its problems are fixable. Some solutions can be achieved simply through executive action; others call for congressional intervention, which may be (and certainly should be) achievable on a bipartisan basis.

The Department of Justice is equipped to make immediate policy changes that could drastically expand the reach of the First Step Act’s correctional reforms. The following steps would ensure that the department’s implementation efforts better align with congressional intent:

- **Prioritize transparency.** Despite long-standing requests from criminal justice reform advocates, among others, the DOJ has not released the data needed to fully assess PATTERN for accuracy and bias. Stakeholders will continue to regard PATTERN with skepticism and distrust until the department releases the information necessary to independently evaluate and validate PATTERN and adopts a practice of timely disclosures about progress and setbacks in its implementation.
- **Revise PATTERN.** Recent revisions to the risk assessment tool show that the DOJ understands the need to rebuild trust in this area of the act’s implementation, but much more needs to be done. The DOJ should start by reorienting PATTERN to focus on predicting a different type of recidivism — the risk of reconviction or re-incarceration, rather than rearrest. This change alone would reduce racial disparities in PATTERN’s risk predictions. Additionally, the DOJ should revisit its decision to leave unchanged PATTERN’s cut points for predicting violent recidivism.
- **Expand prison program offerings.** The list of programs that allow people to earn time credits has until recently been quite short, and it is unclear how much demand remains unmet since recent expansions. The BOP should continue to build out prison programming services and interpret rules expansively to ensure that people can earn credits where programs are inaccessible or booked. Additionally, the BOP should ensure that correctional staff are fully trained to implement new rules being promulgated by the DOJ, and provide transparency on how that training is conducted.

Congress should address structural problems in the First Step Act's design by taking these steps:

- **Broaden eligibility for earned time credits.** The percentage of people in prison eligible for ETCs is far too small: half of those incarcerated by the BOP are ineligible because of their offense of conviction. There is no public safety justification for these exclusions, and Congress should repeal them.
- **Decouple PATTERN from earned time credit eligibility.** The First Step Act's use of PATTERN to determine how much time individuals serve in prison is fundamentally flawed. It is also unusual: most state systems use risk assessments only for more limited purposes.⁸¹ Congress should amend the First Step Act

to limit PATTERN's role in its corrections reforms. Specifically, Congress should detach risk classification from ETC eligibility so that all people incarcerated by the BOP are incentivized to participate in programming and productive activities.

- **Increase funding for prison reforms.** Early in its implementation, the act suffered from funding shortfalls, with Congress even failing to appropriate new money for implementation during the law's first year. While recent reports indicate that the BOP's implementation efforts are now on better footing, policymakers should ensure, through oversight and sustained contact with BOP administrators, that the agency has the resources it needs to deploy high-quality prison programming to all people and in all facilities where it is needed.

Endnotes

- 1 First Step Act of 2018, Pub. L. 115-391, 132 Stat. 5194 (2018), <https://www.congress.gov/115/plaws/publ391/PLAW-115publ391.pdf> [<https://perma.cc/9TWB-U5HD>].
- 2 See, e.g., United States Sentencing Commission, *Report to the Congress: Cocaine and Federal Sentencing Policy*, 2002, 4–11, https://www.ussc.gov/sites/default/files/pdf/news/congressional-testimony-and-reports/drug-topics/200205-rtc-cocaine-sentencing-policy/200205_Cocaine_and_Federal_Sentencing_Policy.pdf [<https://perma.cc/36DL-K5NQ>] (providing background on the passage and structure of mandatory minimums in the Anti-Drug Abuse Acts of 1986 and 1988).
- 3 FAMI Foundation, *Federal Mandatory Minimums, Families Against Mandatory Minimums*, 2015, 1–16, <https://fami.org/wp-content/uploads/Chart-All-Fed-MMs.pdf> [<https://perma.cc/JW3E-963E>] (collecting mandatory minimums by offense, including by drug type and weight).
- 4 James Cullen, “Sentencing Laws and How They Contribute to Mass Incarceration,” Brennan Center for Justice, October 5, 2018, <https://www.brennancenter.org/our-work/analysis-opinion/sentencing-laws-and-how-they-contribute-mass-incarceration> [<https://perma.cc/8B98-S3ED>]; Ram Subramanian et al., *A Federal Agenda for Criminal Justice Reform*, Brennan Center for Justice, 2020, 19–20, <https://www.brennancenter.org/our-work/policy-solutions/federal-agenda-criminal-justice-reform> [<https://perma.cc/LQ72-DDSZ>]; and Paul J. Hofer, “Federal Sentencing After Booker,” *Crime & Justice* 48 (2019): 171–74, <https://www.journals.uchicago.edu/doi/abs/10.1086/701712>.
- 5 J. C. Watts Jr. et al., *Transforming Prisons, Restoring Lives: Final Recommendations of the Charles Colson Task Force on Federal Corrections*, Urban Institute, 2016, vii, <https://www.urban.org/sites/default/files/publication/77101/2000589-Transforming-Prisons-Restoring-Lives.pdf> [<https://perma.cc/U8QV-79T2>].
- 6 Watts et al., *Transforming Prisons, Restoring Lives*, 10.
- 7 28 U.S.C. § 3582(c)(1)(A). For the difficulty of securing compassionate release before the First Step Act’s passage, see *United States v. Brooker*, 976 F.3d 228, 231–32 (2d Cir. 2020); *United States v. McCoy*, 981 F.3d 271, 275–76 (4th Cir. 2020); *United States v. McGee*, 992 F.3d 1035, 1041–42 (10th Cir. 2021); *United States v. Maumau*, 993 F.3d 821, 830 (10th Cir. 2021) (discussing the history and under-use of compassionate release), *cert. denied*, 142 S. Ct. 583 (2021); U.S. Department of Justice, Office of the Inspector General, *The Federal Bureau of Prisons’ Compassionate Release Program*, 2013, 1, <https://www.oversight.gov/sites/default/files/oiq-reports/e1306.pdf> [<https://perma.cc/E5JD-9U5C>]; and Christie Thompson, “Old, Sick and Dying in Shackles,” *Marshall Project*, March 7, 2018, <https://www.themarshallproject.org/2018/03/07/old-sick-and-dying-in-shackles> [<https://perma.cc/T49J-JBR5>].
- 8 For background on this credit system — which was expanded by the First Step Act to 54 days per year, with the difference in credits awarded retroactively — see Good Conduct Time Credit Under the First Step Act, 84 Fed. Reg. 72,274, 72,274–76 (December 31, 2019) (codified at 28 C.F.R. § 523), <https://www.federalregister.gov/documents/2019/12/31/2019-27976/good-conduct-time-credit-under-the-first-step-act> [<https://perma.cc/E4LD-4Y6D>]. For the credit provision as it currently exists, see 18 U.S.C. § 3624(b).
- 9 Fair Sentencing Act of 2010, Pub. L. 111-220, 124 Stat. 2372 (2010). See also *Terry v. United States*, 593 U.S. ___, 141 S. Ct. 1858, 1861 (2021) (reviewing the history of cocaine sentencing laws).
- 10 First Step Act of 2018 § 404, 132 Stat. 5194, 5222 (2018).
- 11 First Step Act of 2018 § 603(b).
- 12 United States Sentencing Commission, *First Step Act of 2018 Resentencing Provisions Retroactivity Data Report*, 2021, 5, table 2 and n.1, <https://www.ussc.gov/sites/default/files/pdf/research-and-publications/retroactivity-analyses/first-step-act/20210519-First-Step-Act-Retro.pdf> [<https://perma.cc/W4B9-FWM5>].
- 13 United States Sentencing Commission, *First Step Act Resentencing Provisions Retroactivity Data Report*, 9, table 6. Note that due to data limitations, this analysis was based on a smaller sample — roughly 2,500 of the 3,700 cases the commission was aware of in total.
- 14 Federal Bureau of Prisons, “First Step Act: FSA Numbers,” accessed July 27, 2022, <https://www.bop.gov/inmates/fsa/> [<https://perma.cc/BV33-T8YE>].
- 15 Julie Zibulsky and Christine Kitchens, *The First Step Act of 2018: One Year of Implementation*, United States Sentencing Commission, 2021, 19–22, <https://www.ussc.gov/sites/default/files/pdf/research-and-publications/research-publications/2020/20200831-First-Step-Report.pdf> [<https://perma.cc/VK52-6JYH>].
- 16 First Step Act of 2018 § 101(a) (codified at 18 U.S.C. §§ 3632(d), 3633, 3635(3)).
- 17 First Step Act of 2018 § 101(a) (codified at 18 U.S.C. § 3632(d) (4)(D)).
- 18 See First Step Act of 2018 § 101(a) (codified at 18 U.S.C. § 3632(d)(4)(A) (providing higher amounts of earned time credit for some individuals deemed “minimum” or “low” risk) and (g)(1) (restricting eligibility to transfer to supervised release or prerelease custody to individuals in minimum- or low-risk categories, absent warden approval under specified circumstances).
- 19 See First Step Act of 2018 § 101(a) (codified at 18 U.S.C. §§ 3621(h)(4) & (6) (requiring the attorney general to “provide all prisoners with the opportunity to actively participate in evidence-based recidivism reduction programs, according to their specific criminogenic needs, throughout their entire term of incarceration”).
- 20 See, e.g., memorandum from Inspector General Michael E. Horowitz to Michael Carvajal, director, BOP, “Management Advisory Memorandum 22-007: Impact of the Failure to Conduct Formal Policy Negotiations on the Federal Bureau of Prisons’ Implementation of the FIRST STEP Act and Closure of Office of the Inspector General Recommendations,” November 15, 2021, <https://oig.justice.gov/reports/management-advisory-memorandum-impact-failure-conduct-formal-policy-negotiations-federal> [<https://perma.cc/YB5T-R6EW>].
- 21 FSA Time Credits, 85 Fed. Reg. 75,268 (proposed November 25, 2020) (superseded), <https://www.federalregister.gov/documents/2020/11/25/2020-25597/fsa-time-credits> [<https://perma.cc/P9K4-X47R>]. Under the proposed rule, an imprisoned person would have needed to complete eight hours of programming to accrue a “day” of “participation” in programming. FSA Time Credits, 85 Fed. Reg. at 75,269. That guidance would have interacted with the BOP’s sparse programmatic offerings — which in some cases meet for just a few hours a day — to make it very difficult to earn any meaningful amount of ETcs. See James M. Byrne, “The Effectiveness of Prison Programming: A Review of the Research Literature Examining the Impact of Federal, State, and Local Inmate Programming on Post-Release Recidivism,” *Federal Probation* 84, no. 1 (2020): 20, https://www.uscourts.gov/sites/default/files/84_1_1_0.pdf [<https://perma.cc/5S4B-B7EB>]. According to one estimate, “even if a person took every approved program the BOP offered . . . she would accrue only 8,143 hours of program time — over 600 hours short of what is necessary to obtain a year’s worth of time credits.” Federal Public and Community Defenders Legislative Committee letter to Federal Bureau of Prisons, Office of General Counsel, re: Comments on Docket No. BOP-1176P, RIN 1120-AB76, “FSA Time Credits,” January 22, 2021, 5, <https://perma.cc/RL3H-38QV>. An Urban Institute analysis found that “only 15 programs” in the BOP catalog “would provide credits of more

than three weeks.” Julie Samuels and Emily Tiry, *Implementation of the First Step Act: The Risk and Needs Assessment System*, Urban Institute, 2021, 4, 17, <https://www.urban.org/research/publication/implementation-first-step-act-risk-and-needs-assessment-system> [<https://perma.cc/ZK43-WVWZ>].

22 In the final rule, BOP acknowledged that certain aspects of the original proposal had been “inconsistent with the goals of the FSA” and corrected many of the concerns raised. The final rule aimed to create a “simpler . . . time credits program . . . that will more fully encourage and reward participation in evidence-based recidivism reduction programs and productive activities.” The final rule awards 10 days of ETCs for each 30-day period in which an eligible person participates in recommended programming, with an additional 5 days awarded to people who meet certain risk assessment criteria. FSA Time Credits, 87 Fed. Reg. 2705, 2706, 2708 (January 19, 2022) (to be codified at 28 C.F.R. pts. 523 & 541), <https://www.federalregister.gov/documents/2022/01/19/2022-00918/fsa-time-credits> [<https://perma.cc/95RC-DPHJ>]. U.S. Department of Justice, “Justice Department Announces New Rule Implementing Federal Time Credits Program Established by the First Step Act,” January 13, 2022, <https://www.justice.gov/opa/pr/justice-department-announces-new-rule-implementing-federal-time-credits-program-established> [<https://perma.cc/GK7C-FSLE>] (noting promulgation of revised regulation and retroactive application); and *Oversight of the Bureau of Prisons, Hearing Before the H. Comm. on the Judiciary*, 117th Cong. (2022) (statement of Michael D. Carvajal, director, Federal Bureau of Prisons), 4, <https://docs.house.gov/meetings/JU/JU08/20220203/114370/HHRG-117-JU08-Wstate-CarvajalM-20220203.pdf> [<https://perma.cc/TD7U-9Z2C>].

23 U.S. Department of Justice, *First Step Act Annual Report*, April 2022, 28–29, <https://www.ojp.gov/library/publications/first-step-act-annual-report-april-2022> [<https://perma.cc/A49Q-8LHN>].

24 FSA Time Credits (2022), 87 Fed. Reg. at 2706, 2708.

25 FSA Time Credits (2022), 87 Fed. Reg. at 2706–08 (implementing rule); 18 U.S.C. § 3632(d)(4) (describing credit eligibility); and 18 U.S.C. § 3624(g) (listing rules for usage of credits).

26 First Step Act of 2018 § 101(a) (codified at 18 U.S.C. § 3632(d)(4)(D)).

27 U.S. Department of Justice, *First Step Act Annual Report*, 28.

28 James M. Byrne et al., *Report of the Independent Review Committee Pursuant to the Requirements of Title I Section 107(g) of the First Step Act (FSA) of 2018 (P.L. 115-391)*, United States Senate Committee on the Judiciary, 2020, 2, <https://firststepact-irc.org/wp-content/uploads/2020/12/IRC-FSA-Title-I-Section-107g-Report-12-21-20.pdf> [<https://perma.cc/3SGR-QZTP>].

29 Erik Ortiz, “Thousands of Federal Inmates Still Await Early Release Under Trump-Era First Step Act,” NBC News, July 3, 2022, <https://www.nbcnews.com/news/us-news/thousands-federal-inmates-still-await-early-release-trump-era-first-st-rcna35162> [<https://perma.cc/Q86A-8M8W>].

30 See First Step Act of 2018, § 101(a) (codified at 18 U.S.C. § 3632(d)(4)(A)) (providing higher amounts of earned time credit for some individuals deemed “minimum” or “low” risk) and (g)(1) (restricting eligibility to transfer to supervised release or prerelease custody to individuals in minimum- or low-risk categories, absent warden approval under specified circumstances).

31 See, e.g., letter from David Patton, co-chair, Federal Public and Community Defenders Legislative Committee, to David B. Muhlhause, director, National Institute of Justice, re: DOJ First Step Act Listening Session on PATTERN, September 13, 2019 (available as Exhibit A to *The Federal Bureau of Prisons and Implementation of the First Step Act, Hearing Before the H. Judiciary Comm. Subcomm. on Crime, Terrorism, and Homeland Security*, 116th Cong. (2019) (statement of David E. Patton), 15, <https://www.congress.gov/116/meeting/house/110089/witnesses/HHRG-116-JU08-Wstate-PattonD-20191017.pdf> [<https://perma.cc/CYX2-6YMR>].

32 See U.S. Department of Justice, *The First Step Act of 2018: Risk and Needs Assessment System*, 2019, <https://nij.ojp.gov/first-step-act/the-first-step-act-of-2018-risk-and-needs-assessment-system.pdf> [<https://perma.cc/EC6L-5NRL>]. For revisions and administration support for further change, see “Revisions to PATTERN (April 2022–Present)” in this report.

33 Pamela M. Casey, Jennifer K. Elek, and Roger K. Warren, *Offender Risk & Needs Assessment Instruments: A Primer for Courts*, National Center for State Courts, 2014, 15, 20–22, https://www.ncsc.org/_data/assets/pdf_file/0018/26226/bja-rna-final-report-combined-files-8-22-14.pdf [<https://perma.cc/7R5V-VXEX>].

34 See letter from David E. Patton, co-chair, Federal Public and Community Defenders Legislative Committee, to Senator Richard Durbin, re: CORRECTIONS Act (S. 467), February 5, 2015, 8 (on file with the authors).

35 Watts et al., *Transforming Prisons, Restoring Lives*, 33–34; and Julia Angwin and Jeff Larson, “Bias in Criminal Risk Scores Is Mathematically Inevitable, Researchers Say,” ProPublica, December 30, 2016, <https://www.propublica.org/article/bias-in-criminal-risk-scores-is-mathematically-inevitable-researchers-say> [<https://perma.cc/NSS9-684G>].

36 See letter from Ames Grawert, senior counsel, Brennan Center for Justice, to David B. Muhlhause, director, National Institute of Justice, re: “Public Comment on ‘PATTERN,’ the First Step Act’s Risk and Needs Assessment Tool,” September 3, 2019, 4–5, <https://www.brennancenter.org/our-work/research-reports/brennan-centers-public-comment-first-step-acts-risk-and-needs-assessment> [<https://perma.cc/73CL-YXZC>].

37 Jessica M. Eaglin, “Constructing Recidivism Risk,” *Emory Law Journal* 67 (2017): 72, 94–99, <https://www.repository.law.indiana.edu/cgi/viewcontent.cgi?article=3647&context=facpub> [<https://perma.cc/6Q2X-YKXF>].

38 Federal Bureau of Prisons, “New FSA Resources Available,” October 26, 2020, https://www.bop.gov/resources/20201026_fsa_update.jsp [<https://perma.cc/NNE7-NWFA>].

39 National Institute of Justice, *2020 Review and Revalidation of the First Step Act Risk Assessment Tool*, 2021, 14, 17–19 and table 2.3, <https://www.ojp.gov/pdffiles1/nij/256084.pdf> [<https://perma.cc/P55N-SKRZ>].

40 While “general” and “violent” recidivism are technically different categories, both focus squarely on arrest risk. No conviction is necessary to qualify as a recidivating event in either case, and in fact the arrest need not even withstand the scrutiny of a grand jury. U.S. Department of Justice, *The First Step Act of 2018: Risk and Needs Assessment System — Update*, 2020, 12–14, <https://nij.ojp.gov/first-step-act/the-first-step-act-of-2018-risk-and-needs-assessment-system.pdf> [<https://perma.cc/BV9D-KL5B>].

41 U.S. Department of Justice, *First Step Act Annual Report*, 13–14.

42 Carrie Johnson, “Justice Department Works to Correct Bias in Deciding Who’s Released from Prison,” NPR, April 19, 2022, <https://www.npr.org/2022/04/19/1093538706/justice-department-works-to-curb-racial-bias-in-deciding-whos-released-from-pris> [<https://perma.cc/QX4F-UY65>]; and U.S. Department of Justice, *The First Step Act of 2018: Risk and Needs Assessment System*, 2019, <https://nij.ojp.gov/first-step-act/the-first-step-act-of-2018-risk-and-needs-assessment-system.pdf> [<https://perma.cc/EC6L-5NRL>].

43 U.S. Department of Justice, *The First Step Act of 2018: Risk and Needs Assessment System — Update*; U.S. Department of Justice, *First Step Act Implementation Fiscal Year 2020 90-Day Report*, 2020, <https://www.ojp.gov/pdffiles1/nij/254799.pdf> [<https://perma.cc/NF77-FV4E>]; U.S. Department of Justice, *The Attorney General’s First Step Act Section 3634 Annual Report*, 2020, <https://www.ojp.gov/Attorney-Generals-First-Step-Act-Section-3634-Annual-Report-December-2020> [<https://perma.cc/6BAN-AVMV>]; National Institute of Justice, *2020 Review and Revalidation*; and National Institute of Justice, *2021 Review and Revalidation of the First Step Act Risk*

- Assessment Tool, 2021, <https://www.ojp.gov/pdffiles1/nij/303859.pdf> [<https://perma.cc/3C6X-KC97>].
- 44** National Institute of Justice, *2020 Review and Revalidation*, 7.
- 45** National Institute of Justice, *2021 Review and Revalidation*, 7–8.
- 46** Ian MacDougall, “Bill Barr Promised to Release Prisoners Threatened by Coronavirus — Even as the Feds Secretly Made It Harder for Them to Get Out,” ProPublica, May 26, 2020, <https://www.propublica.org/article/bill-barr-promised-to-release-prisoners-threatened-by-coronavirus-even-as-the-feds-secretly-made-it-harder-for-them-to-get-out> [<https://perma.cc/Y2V5-SUBE>].
- 47** E. Ann Carson, *Federal Prisoner Statistics Collected Under the First Step Act*, 2021, U.S. Department of Justice, Bureau of Justice Statistics, 2021, 16, table 10, <https://bjs.ojp.gov/library/publications/federal-prisoner-statistics-collected-under-first-step-act-2021> [<https://perma.cc/3YFZ-8GW2>].
- 48** National Institute of Justice, *2021 Review and Revalidation*, 40.
- 49** Grawert, “Public Comment on ‘PATTERN,’” 3–4.
- 50** See Grawert, “Public Comment on ‘PATTERN,’” 2–4 (arguing that defining recidivism based on rearrest incorporates racial disparities in policing); and U.S. Department of Justice, *The First Step Act of 2018: Risk and Needs Assessment System — Update*, 13–14 (considering but rejecting a change to narrower definition of recidivism).
- 51** Byrne et al., *Report of the Independent Review Committee*, 3.
- 52** The DOJ also argued that rearrest is the definition of recidivism “most widely used in the federal system,” and that therefore retaining it for PATTERN “would make comparisons across risk assessment and criminal justice systems easier.” U.S. Department of Justice, *The First Step Act of 2018: Risk and Needs Assessment System — Update*, 2020, 14.
- 53** Samuels and Tiry, *Implementation of the First Step Act*, 9; and Amy Cyphert, “Reprogramming Recidivism: The First Step Act and Algorithmic Prediction of Risk,” *Seton Hall Law Review* 51, no. 2 (2020): 331–381, 351–353, <https://scholarship.shu.edu/cgi/viewcontent.cgi?article=1773&context=shlr> [<https://perma.cc/BB26-YW2C>].
- 54** Michael Tonry, “Legal and Ethical Issues in the Prediction of Recidivism,” *Federal Sentencing Reporter* 26, no. 3 (2014): 167–176, 167, https://scholarship.law.umn.edu/cgi/viewcontent.cgi?article=1535&context=faculty_articles [<https://perma.cc/R4KX-SKND>].
- 55** U.S. Department of Justice, *First Step Act Implementation Fiscal Year 2020*, 2. For criticisms, see Samuels and Tiry, *Implementation of the First Step Act*, 9.
- 56** U.S. Department of Justice, *The Attorney General’s First Step Act Section 3634 Annual Report*, 24.
- 57** Emily Tiry and Julie Samuels, “Three Ways to Increase the Impact of the First Step Act’s Earned Time Credits,” Urban Institute, April 30, 2021, <https://www.urban.org/urban-wire/three-ways-increase-impact-first-step-acts-earned-time-credits> [<https://perma.cc/7LES-WELQ>].
- 58** National Institute of Justice, *2020 Review and Revalidation*, 14.
- 59** National Institute of Justice, *2021 Review and Revalidation*, 28–29. See also Department of Justice, *First Step Act Annual Report*, 10.
- 60** For more on this subject, see, e.g., Grawert, “Public Comment on ‘PATTERN,’” 5–7.
- 61** U.S. Department of Justice, *First Step Act Annual Report*, 5, 9–12.
- 62** U.S. Department of Justice, *First Step Act Annual Report*, 12–18.
- 63** U.S. Department of Justice, *First Step Act Annual Report*, 14.
- 64** U.S. Department of Justice, *First Step Act Annual Report*, 12, 11 (“Thus, transitioning to PATTERN 1.3 will neither exacerbate nor solve these racial bias issues.”), and 20 (noting that changes to cut points “will not directly correct the racial differential prediction rates found to be associated with the PATTERN tool”).
- 65** *The First Step Act, the Pandemic, and Compassionate Release: What Are the Next Steps for the Federal Bureau of Prisons?*, Hearing Before the H. Judiciary Comm. Subcomm. on Crime, Terrorism, and Homeland Security, 117th Cong. (2022) (statement of Melissa Hamilton, professor of law and criminal justice, University of Surrey), 4, <https://docs.house.gov/meetings/JU/JU08/20220121/114349/HHRG-117-JU08-Wstate-HamiltonM-20220121.pdf> [<https://perma.cc/2Y8F-4TBZ>].
- 66** Department of Justice, *First Step Act Annual Report*, 20 (“The updated version of the tool and the revised cut points are anticipated to result in expanded opportunities for the earning of time credits under the FSA system, which will in turn incentivize a greater number of inmates to participate in evidence-based recidivism reduction programs and productive activities and reduce their predicted risk of recidivism even further.”).
- 67** Department of Justice, *First Step Act Annual Report*, 13–14.
- 68** National Institute of Justice, *Stakeholder Statements Submitted in Response to NIJ’s First Step Act Listening Sessions*, 2019, <https://www.ojp.gov/pdffiles1/nij/253115.pdf> [<https://perma.cc/25E8-DPDE>] (noting stakeholder requests for greater transparency).
- 69** Department of Justice, *First Step Act Annual Report*, 11–12.
- 70** 18 U.S.C. § 3632(a)(3).
- 71** Federal Bureau of Prisons, *First Step Act: Initial Review of the SPARC-13 Needs Assessment System*, U.S. Department of Justice, 2022, 14, https://www.bop.gov/inmates/fsa/docs/bop_fsa_needs_validation_report_2021.pdf [<https://perma.cc/XS45-GDX2>].
- 72** Federal Bureau of Prisons, *Initial Review of SPARC-13*, 12, table 4.
- 73** Federal Bureau of Prisons, *Initial Review of SPARC-13*, 10–13.
- 74** See *The Federal Bureau of Prisons and Implementation of the First Step Act* (statement of David E. Patton), 1, 5–7.
- 75** U.S. Department of Justice, *First Step Act Annual Report*, 33–35; and Byrne et al., *Report of the Independent Review Committee*, 1.
- 76** Byrne et al., *Report of the Independent Review Committee*, 2–3.
- 77** FSA Time Credits (2022), 87 Fed. Reg. at 2711.
- 78** Carson, *Federal Prisoner Statistics Collected*, 17–20 and table 11. This may have changed with the development of new programs and rollout of the January 2022 ETC rule. See U.S. Department of Justice, *First Step Act Annual Report*, 32–33 (“The passage of the Time Credits rule in January 2022 expanded the services which qualify as PAs.”).
- 79** Department of Justice, *First Step Act Annual Report*, 32–36, 45–48.
- 80** While the First Step Act authorized up to \$75 million annually for programming, Congress failed to appropriate that amount for the first year of implementation, forcing the DOJ to move money from elsewhere in its budget. Ames Grawert, “What Is the First Step Act — and What’s Happening With It?,” Brennan Center for Justice, June 23, 2020, <https://www.brennancenter.org/our-work/research-reports/what-first-step-act-and-whats-happening-it> [<https://perma.cc/7M3A-E9AU>].
- 81** See Amy Baron Evans et al., “Federal Defender Analysis of Corrections Act of 2015,” *Federal Sentencing Reporter* 28, no. 2 (2015): 151–153, 151 (“The system in the bill is novel and untested. State correctional systems . . . award time credits for participating in programs based on performance and/or disciplinary record, not risk assessment scores.”).

ABOUT THE AUTHORS

► **Ames Grawert** is senior counsel and John L. Neu Justice Counsel in the Brennan Center’s Justice Program. He leads quantitative and policy research focused on trends in crime and the collateral costs of mass incarceration. Additionally, he advocates for criminal justice reform at the state and federal levels. Previously Grawert served as an assistant district attorney in the Appeals Bureau of the Nassau County (New York) District Attorney’s Office, where he reviewed and litigated claims of actual innocence in addition to appellate work. Before entering public service, he was an associate at Mayer Brown LLP, where he represented people pro bono in post-conviction litigation.

► **Patricia L. Richman** is national sentencing resource counsel to the Federal Public and Community Defenders, where she leads legislative work. Previously Richman served as crime counsel to Sen. Dick Durbin on the Senate Judiciary Committee and as an assistant federal public defender in Maryland. Before that, Richman was an associate at Williams & Connolly LLP, a clerk for a federal district court judge in Philadelphia, and a special education teacher in Prince George’s County, Maryland.

ABOUT THE BRENNAN CENTER’S JUSTICE PROGRAM

The Brennan Center’s Justice Program seeks to secure our nation’s promise of equal justice for all by creating a rational, effective, and fair justice system. Its priority focus is to reduce mass incarceration. The program melds law, policy, and economics to produce new empirical analyses and innovative policy solutions to advance this critical goal.

ACKNOWLEDGMENTS

The Brennan Center gratefully acknowledges the Tow Foundation for its generous support of our work. This is an independent Brennan Center publication; the opinions expressed are those of the authors and do not necessarily reflect the views of our supporters.

The authors thank Shelley Fite for her input on federal appellate practice and recent developments in First Step Act case law, as well as Melissa Hamilton for her analysis of risk assessment tools and Sean Wilson for his insights into how new regulations have impacted prison staff and imprisoned people. The authors also thank Lauren-Brooke Eisen and Ram Subramanian for their strategic guidance of the report, as well as their careful revisions and insightful comments; Michael Waldman and John Kowal for their support of this work; and Calvin Chappell, Noah Kim, and Stephanie Wylie for their careful research and editing support. The authors also thank Maris Mapolski for her in-depth review of multiple drafts of this analysis. Last, the authors are grateful to Lisa Benenson, Matt Harwood, Zachary Laub, Alexandra Ringe, Janet Romero-Bahari, Derek Rosenfeld, and Alden Wallace for their review of this report, communications assistance, and work to ensure its successful publication.

Statement of Melissa Hamilton
First Step Act Implementation: Stakeholder Engagement Session
September 29, 2022

Thank-you for the invitation to participate in this timely and important event. I am a law professor, with a law degree and a doctorate in criminology. I have been working in the area of algorithmic risk assessment for years and have been avidly following the development and implementation of PATTERN and its various iterations. Herein I highlight some important issues for consideration.

I. Definition of Recidivism

Various arguments have been made before about the problematic issues of using any arrest or technical violation as the definition of recidivism (e.g., bias, errors). I also promote the use of convictions as an alternative. The additional point here is that regardless of which measure used, only serious offenses should be considered. PATTERN risk scores are not directly tied to public safety in the sense of the person being immediately released based on a low-risk prediction. PATTERN is more simply about identifying which individuals have the opportunity to complete rehabilitative programs in order to earn a limited number of days' early release at some point in the (often distant) future. It is not evident that the risk upon some future release of committing (or being arrested for) a technical violation or a minor offense should prevent an individual from benefiting after successfully completing programs. Is the likelihood, even certainty for argument's sake, of the person violating a no drinking supervision term or committing a minor theft really a type of risk that should entirely preclude them from the benefits of the First Step Act?

Instead, it is more likely that the real concern is about the types of risk that do pose a safety risk to the community, albeit even in the distant future. Hence, more attention should be paid to limiting the recidivist event in some way to serious crimes. This could be to focus on serious (e.g., felony level) and violent offending. Care could also be taken on the definition of violence as it often gets expanded beyond what really matters. In the risk assessment world, it is too common to lump together into a violence category even minor levels (low-level threats, misdemeanor assault), crimes that do not necessarily have a violent element (burglary), and non-contact offending (online harassment).

II. Preference for False Positives

PATTERN has been intentionally calibrated to prefer false positives over false negatives. I define these terms for purposes here by dichotomizing the PATTERN predictions into two groups. The lower risk lumps together the PATTERN outcomes of minimum and low. The higher risk combines the PATTERN outcomes of median and high. The reason is due to the FSA threshold for the earned time credits at the division between low and medium. In any event, a false positive here would mean a prediction of higher risk when the person was successful upon release, meaning not rearrested. A false negative is the lower risk person who was rearrested.

Table 1 provides the false positive and false negative rates for the four PATTERN scales.¹

Table 1. False Positive and False Negative Rates

| Scale | False positive rate | False negative rate |
|----------------|---------------------|---------------------|
| Male general | 33% | 23% |
| Male violent | 70% | 7% |
| Female general | 38% | 23% |
| Female violent | 78% | 4% |

In all scales, the false positive rates are significantly higher than the false negative rates. This means that there has been a judgement call that when the algorithm errs, the error is far more likely to err on the side of inaccurately classifying individuals as higher risk. This is a human choice, not a scientifically-based result. This means that another choice can be made toward the other direction in order to reduce the false positive rate by increasing the threshold settings for the risk groupings.

III. Calibration Inequalities

The 2021 Validation Report provides some metrics that indicate that PATTERN does not operate equivalently based on race/ethnicity in terms of over- or underpredicting for different groups, depending on the scale. The measurements used and the discussion may not have been understandable to all audiences. I provide here another way to see the data. In this form, the statistics indicate that PATTERN is not calibrated equally for all. In other words, a low-risk or a high-risk prediction does not mean the same thing in terms of the likelihood of rearrest depending on race/ethnicity. Appendix A contains the tables with statistical information.²

The most equivalent performance is found in the male general scale between white and black individuals. The recidivism rates are roughly equal (when rounding) at each risk bin for these two groups in the male general scale. Still, the other racial groups varied significantly. For example, in the low-risk group, whereas white and black individuals had a 31% recidivism rate, the rate for Native-Americans was far greater at 52% and the rate for Asians lesser at 21%.

The equivalence of recidivism rates as between White and Blacks does not follow through to the other three scales. In the male violent scale, recidivism rates varied: low risk (White 8% versus Black 12%), medium (White 19% versus Black 24%), high (White 30% versus Black 39%). The female general recidivism scale did not observe equal rates in risk groups across White and Black (e.g., low risk with 40% White and 30% Black) or other groups.

Calculating additional statistics for gender (combining racial/ethnic groups), the results for the general scale are in Table 2.

¹ National Institute of Justice, 2021 Review and Revalidation of the First Step Act Risk Assessment Tool (Dec. 2021). The numbers are derived from the FY14-15 (fiscal years 2015-2015) validation sample. The numbers are not significantly different for the FY16 or FY17 samples.

² *Id.* The data to create the tables derives from the FY 2014-2015 validation sample in Table A1.

Table 2. Rearrest Rates by Gender on the General Scales

| | Minimum | Low | Medium | High |
|--------|---------|-----|--------|------|
| Male | 11% | 30% | 53% | 74% |
| Female | 10% | 35% | 57% | 77% |

Table 2 indicates that males are less likely to be rearrested than females in each of the low, medium, and high-risk categories. The results for gender for the violent scale are in Table 3.

Table 3. Rearrest Rates by Gender on the Violent Scales

| | Minimum | Low | Medium | High |
|--------|---------|-----|--------|------|
| Male | 1% | 9% | 21% | 35% |
| Female | 2% | 6% | 20% | 36% |

Results for the violent scale are closer, but disparities are evident.

The point is that the same risk categorization (i.e., minimum, low, medium, or high) does not mean the same likelihood of recidivism by race/ethnicity or by gender.

IV. General Versus Violent Risk Scales

PATTERN scores by gender and on each of two scales, one predicting any rearrest and the other any violent rearrest. But when the individual's final risk category is assigned, the higher of the two scales is used. The problem is that, in the end, the final assignment is conflating very different things. The first is mixing up the risk of any rearrest (i.e., technical violation) and a rearrest for a violent crime. The other problem is that the proportional likelihood of rearrest is very different. For instance, suggest two individuals both of whom were assigned the high-risk group for purposes of PATTERN and the opportunity for benefits under the FSA. One was high risk based on the general recidivism scale (but was lower on the violent scale) while the other was high risk associated with their score on the violent recidivism scale. For two males, the first one is associated with a 74% recidivism rate (for any technical violation) while the other is associated with a 35% violence recidivism rate. These two share the same outcome but pose very different types of danger.

The suggestion here, then, is to revisit the two different scales and the rule placing the higher score as the final category.

V. Cutpoints

The choice of where to place the cutpoints to distinguish between the four categories of risk is just that: a choice. There are no standards or norms in the field about what any of the terms (e.g., low, medium, or high) actually mean other than presenting as relative to each other. Thus, other than the ordinal categorization tending to link to a higher expected recidivism rate (however defined by the particular tool), there is no commonality of what rate or range of rates qualify for any group. The point here, then, is that if policymakers want to maximize the incentives under the FSA to participate in rehabilitative programming, the selection of higher cutpoints does not violate the scientific principles of the risk assessment field.

VI. Disciplinary Infractions

The significant reliance upon the disciplinary system is concerning from perspectives of fairness, transparency, and accuracy. The disciplinary process has a low due process threshold. Plus, disciplinary practices vary considerably across BOP institutions. Thus, it could simply be the luck of the draw as to which prison the person is assigned. If sent to a prison whose culture is to overpolice in terms of citing often and for the most negligible of acts, the person's likelihood of an increased risk score based on that practice rises. It is simply that some cultures more or less rely upon the formal disciplinary process as a means of maintaining security. Some institutions simply as a matter of course use other tools to improve the health and security of their environments.

Disciplinary infractions also carry the imprint of biases, which are then baked into PATTERN, likely yielding biased results.

VII. Protective Factors

PATTERN could incorporate protective factors. While it does not appear that current assessments would include them, the system could be modified so that the needs assessment protocol could also evaluate for relevant protective factors (e.g., self-control, attending counseling, prosocial relationships with others, positive recreation hours).

VIII. Fairness

A stringent process should be implemented to provide individuals with their scored PATTERN documents and a right to challenge perceived errors. There is a vast literature that shows prisoners respond to procedural fairness. The validation studies on PATTERN have shown numerous errors in the scoring across thousands of prisoners. Thus, this clearly is not an error-free program.

Likewise, there should be regular audits of PATTERN scoring and checks on which databases are being used. While the Bureau of Prisons has evidently automated the scoring of PATTERN, this does not mean there are not errors being made. Plus, with an automation process, this could mean that the errors are systematic. There is already experience with PATTERN that it was not being scored or using the definitions for the factors that were intended by the original development team. In other words, there is already a history of confirming with a later audit that significant inaccuracies resulted. Thus, sufficient reasons exist for an audit process around all aspects of PATTERN.

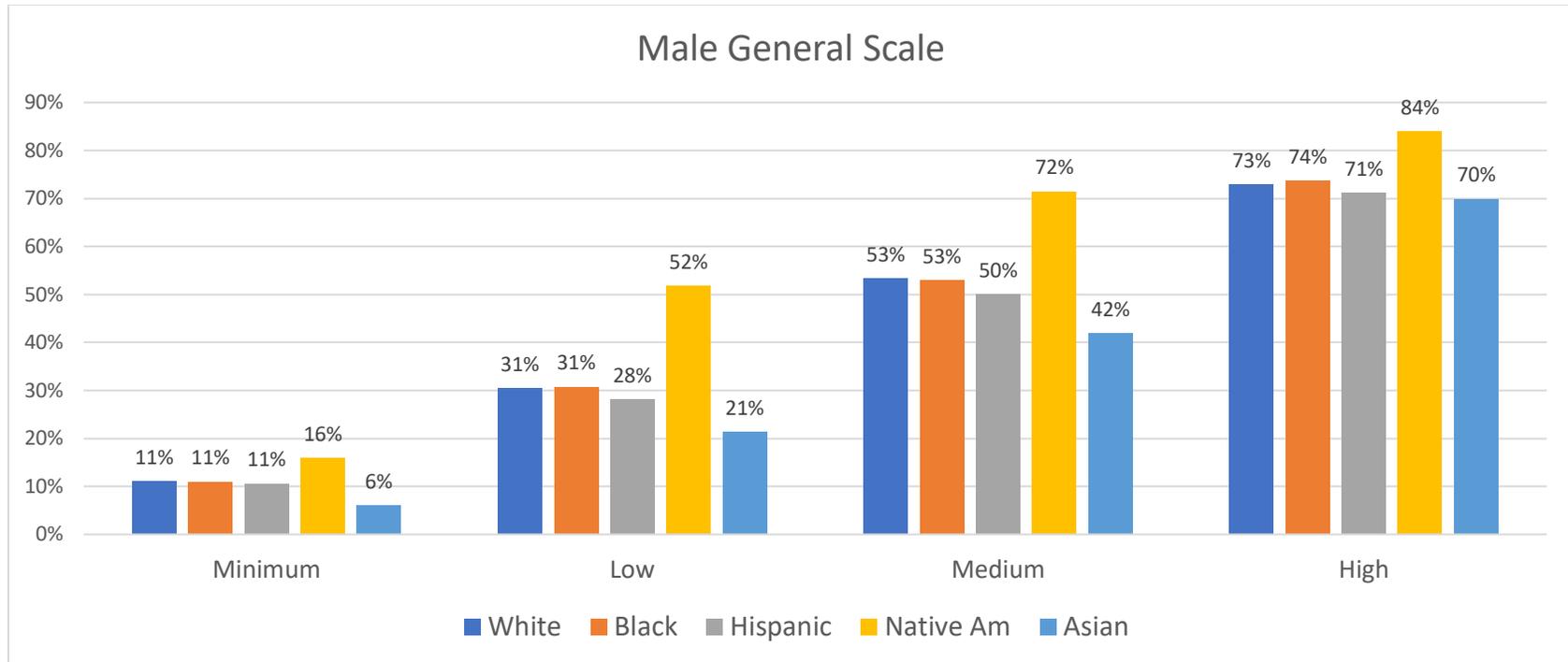
IX. Transparency

I lend my voice to the others in calling for greater transparency in making available the datasets for independent researchers to assess the performance and fairness of PATTERN. The consultants have done an admirable job in providing a number of statistics. Yet there are certainly statistical metrics that I, as an experienced risk assessment empirical scholar, would want to have that have not been provided.

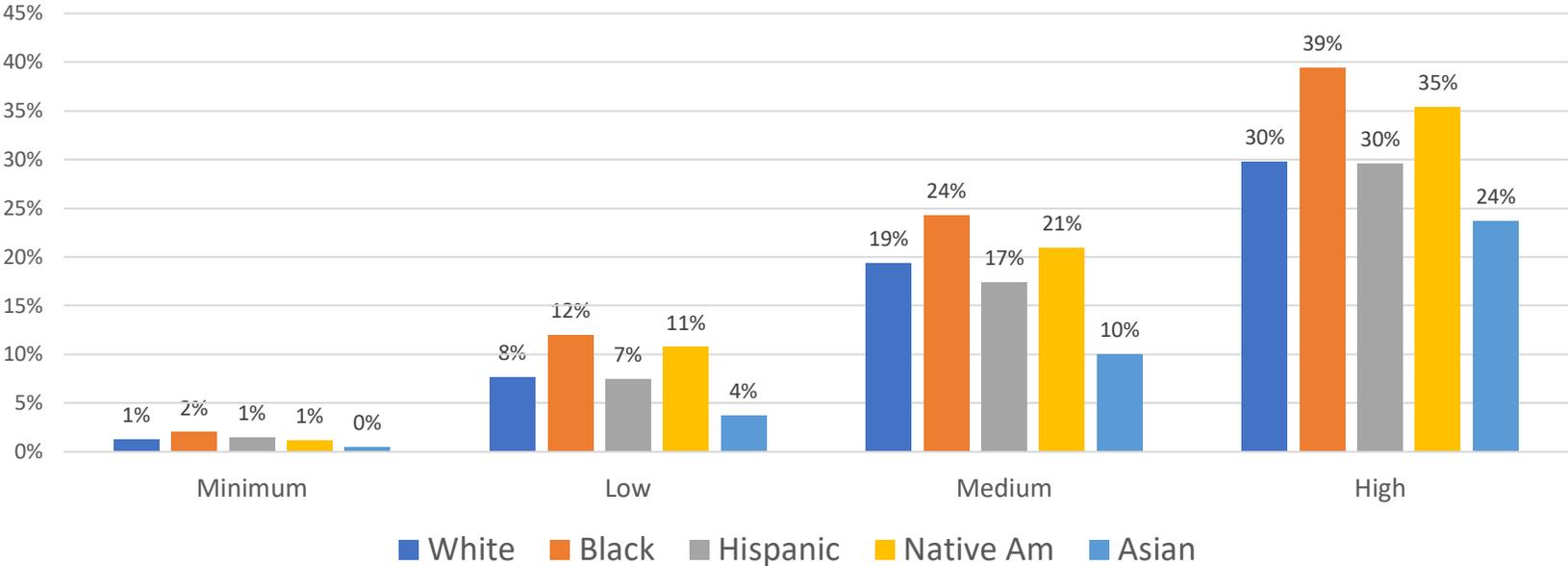
Melissa Hamilton, JD, Ph.D
Professor of Law & Criminal Justice
University of Surrey School of Law

Appendix A

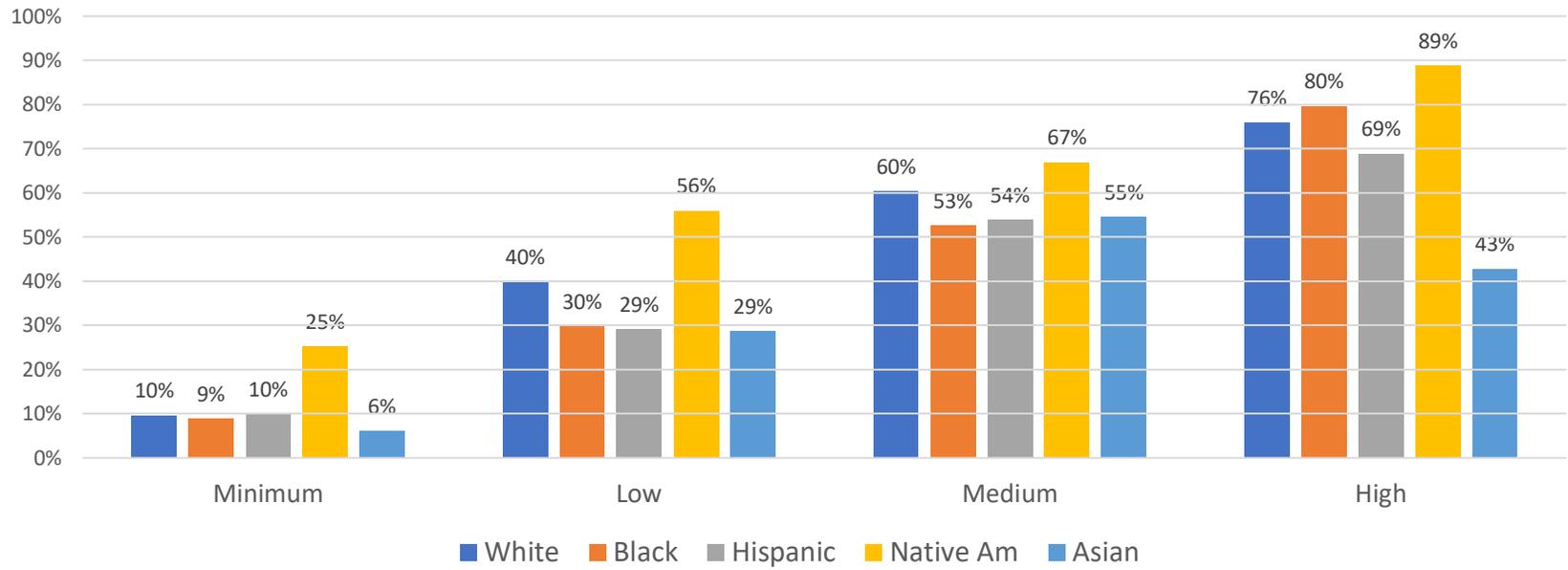
Tables of Recidivism Rates by Scale and Race/Ethnicity



Male Violent Scale



Female General Scale



Female Violent Scale

