



September 3, 2019

David B. Muhlhausen, Ph.D.
Director of the National Institute of Justice
810 7th St NW
Washington, DC 20531

Dear Dr. Muhlhausen,

The Leadership Conference on Civil and Human Rights, The Leadership Conference Education Fund, the American Civil Liberties Union, the Center on Race, Inequality, and the Law at NYU Law, The Justice Roundtable, Media Mobilizing Project, and Upturn respectfully submit the following comments in response to the July 19, 2019 release of the FIRST STEP Act of 2018: Risk and Needs Assessment System report.

The Leadership Conference is a coalition charged by its diverse membership of more than 200 national organizations to promote and protect the civil rights of all persons in the United States. Creating constructive pathways to redemption and rehabilitation in the criminal legal system has long been a top priority of The Leadership Conference and its members. The pervasive, unequal treatment of people of color and people who are low-income undermines our democracy's founding promise of equality under the law. We write to urge the National Institute of Justice ("NIJ"), the Federal Bureau of Prisons ("BOP"), and the Department of Justice ("DOJ") to acknowledge and correct the racial and gender biases in the Prisoner Assessment Tool Targeting Estimated Risk and Needs ("PATTERN") assessment tool algorithm. Ultimately, we request adjustments to the risk score system and validation of this tool by independent experts before it is used to assess anyone in federal prison. Therefore, we ask the DOJ to suspend the use of PATTERN until it adequately addresses these concerns.

On May 8, 2018, The Leadership Conference urged the House Judiciary Committee to vote "No" on the FIRST STEP Act¹ because we feared the Act's lack of transformative "front end" reform would stall our justice system in the broken status quo.² Further, we criticized the bill for "using risk assessment tools in an unconventional manner [because they] are unreliable and exacerbate racial and socioeconomic disparities."³ We predicted that the law's risk assessment provisions would result in incarcerated people being unable to decrease their risk category in order to earn credits toward early release to a residential reentry center or a halfway house.⁴ After members of Congress made key changes to move the bill toward meaningful reform, we ultimately supported the legislation while continuing to articulate concerns regarding the use of a "risk and needs assessment tool." We submitted additional materials in April⁵ and June⁶ of 2019, detailing our outstanding concerns with the development and implementation of the new "risk and needs assessment" system as required by law. Now, nine months after the passage of the Act, we worry that these initial fears are substantiated. The development and implementation of the risk and needs assessment tool set forth in the July 19 report represents DOJ's failure to adhere to the statute's requirements



and threatens to significantly undermine the spirit of the law by indefinitely codifying our criminal legal system’s existing racial and gender disparities.

At municipal, state, and federal levels, civil and human rights organizations, community-based advocates, computer scientists, attorneys, and policy experts have all warned against the use of risk and needs assessments and other algorithmic tools in the criminal legal system because they carry the potential to entrench existing racial bias and injustice.² Our stance is nuanced. Though we discourage the use of these tools, we acknowledge their increasing presence in our legal system and advocate for appropriate safeguards to ensure their accuracy and fairness. Now, we reiterate our standing concerns about the BOP’s use of a risk and needs assessment to evaluate incarcerated persons for early release. The BOP should rectify the inequities in PATTERN’s racial outcomes and cease using different algorithms for different genders. The DOJ risk scores system has never been independently validated by non-IRC data scientists.

I. Background regarding the FIRST STEP Act’s proposed risk and needs assessment system.

The DOJ has advanced PATTERN as a new gender-specific risk and needs assessment tool that fulfills the FIRST STEP Act’s statutory requirement to assign a “recidivism score” to each incarcerated person that predicts their risk of committing a new crime within three years after release.⁷ The new tool purports to make use of both static risk factors and dynamic risk factors traditionally associated with either an increase or reduction in risk. While the FIRST STEP Act requires both risk and needs systems,⁸ to date, the DOJ has only released the risk system and asserted that the needs system is forthcoming. On July 19, 2019, NIJ, a research arm of the DOJ’s Office of Justice Programs, published a report to accompany Attorney General William P. Barr’s release of PATTERN.⁹ The report touts the tool’s power to maximize the number of incarcerated people eligible to earn credits toward early release, identify those qualified to participate in rehabilitative programming, and to ensure public safety.¹⁰ Specifically, NIJ claims that PATTERN: (1) “achieves a high level of predictive performance and surpasses” other tools used to evaluate inmate risk; (2) “more appropriately aligns with the goals of the FIRST STEP ACT [by] mak[ing] greater use of dynamic factors; (3) exhibits that “predictive performance is unbiased across racial and ethnic classifications.” Each of these claims is misleading, for reasons we detail at length below. In particular:

- The lack of transparency surrounding the design and development of PATTERN undermines accountability and frustrates the ability of outside researchers, academics, and advocates to effectively test tools and advocate for incarcerated persons.
- The claim that PATTERN achieves a high level of predictive performance is belied by the fact that algorithms based on historical trends do not, and cannot, accurately predict crime.
- The assertion that PATTERN makes greater use of dynamic factors is misleading at best, and several factors employed by PATTERN deserve more scrutiny.
- The claim that PATTERN’s predictive performance is unbiased across racial and ethnic classifications ignores historical and enduring patterns of racial bias and discrimination that infect the data upon which PATTERN relies.



- The separation of risk calculations on the basis of gender raises serious constitutional concerns.

Academics and practitioners alike have serious doubts about the effectiveness, fairness, and ability of tools like PATTERN to assess risk.¹¹ The American Bar Association (“ABA”) has identified areas requiring consideration or clarification to help ensure that PATTERN will produce accurate, individualized outcomes that remain racially, ethnically, and gender neutral.¹² We share the ABA’s concerns and doubts raised by academics and practitioners. Below, we set forth a set of additional concerns about whether PATTERN fulfills any of the claims made by the NIJ.¹³

II. The lack of transparency associated with the design and development of PATTERN undermines accountability.

The NIJ asserts that the system’s risk assessment tool was built upon and replaces the Bureau Risk and Verification Observation - Recidivism tool (“BRAVO-R”). However, remarkably little information is publicly available about the development, effectiveness, or accuracy of BRAVO-R. The absence of information about the tool on which PATTERN was based makes it incredibly difficult to substantiate the claims that PATTERN is an acceptable replacement for BRAVO-R.¹⁴ For example, the only public-facing documentation on the BOP website that references BRAVO-R is the Legal Resource Guide to the Federal Bureau of Prisons, from March 2019, which does not even refer to the tool by name:

“Placement in the SOTP-R is reserved for high risk sexual offenders, based on the extent and seriousness of the inmate’s offending history, as determined by a formal risk assessment conducted by BOP staff.”¹⁵

In short, a full evaluation of PATTERN and the process by which it was developed demands far greater transparency regarding the tools that led to it. If BRAVO-R served as the foundation for PATTERN, the DOJ must provide a full and clear description of BRAVO-R, including the data, methods, and decision-making process used to develop, validate, and test it independently. Given the important role that BRAVO-R’s initial scores play in PATTERN, it is essential for more information about BRAVO-R to be made publicly available. The absence of such information frustrates efforts to properly evaluate PATTERN, raising significant questions about its utility.¹⁶ More can and must be done to make PATTERN’s design, architecture, and training data open to independent research, review, testing, validation, and criticism.¹⁷

III. NIJ’s assertion that “PATTERN achieves a high level of predictive performance” is belied by the fact that algorithms based on historical trends do not, and cannot, accurately predict crime.

The claim that PATTERN achieves a “high level of predictive performance” in forecasting recidivism over three years suffers from several distinct, but overlapping, problems.

First, it is critical to unpack the outcome being measured. PATTERN attempts to forecast recidivism, but it



does so by measuring “any arrest or return to BOP custody following release.”¹⁸ However, the notion that a risk assessment tool can accurately predict *recidivism*¹⁹ misunderstands the power of the tool and the underlying measurement problem. At best, because PATTERN is reliant upon rearrest data, it can only truly provide a baseline probability that an individual may encounter, and be arrested by, law enforcement.²⁰ Indeed, the use of “arrest as a measure of criminality fundamentally assumes that people who do the same things are arrested at the same rates.”²¹ In other words, risk assessment tools like PATTERN ask us to assume that arrest is itself not a biased measure of criminality. Of course, a substantial body of evidence indicates that this is not the case.

Second, the main evidence for the claim that “PATTERN achieves a high level of predictive performance” appears to be the Area Under the Curve (AUC) scores.²² While the AUC is a relevant measurement of predictive validity, it is not the only measurement — nor should it be. As one scholar argues, “the AUC provides an incomplete portrayal of predictive validity” because, among other things, “it does not capture how well a risk assessment’s tool’s predictions of risk agree with actual observed risk.”²³ Only reporting the AUC does not, in and of itself, offer sufficient evidence of predictive utility. NIJ should more comprehensively test PATTERN on a range of predictive measurements, and publicly release those results.

IV. NIJ’s assertion that “PATTERN. . . makes greater use of dynamic factors” is misleading at best, and several factors deserve more scrutiny.

The SA mandates the BOP’s use of a risk and needs assessment tool which uses both static and dynamic factors²⁴ and other history information to assign a risk category of high, medium, low, or minimum probability of violent or nonviolent recidivism for incarcerated persons in federal detention.²⁵ The NIJ created PATTERN to fulfill that mandate.²⁶ Unfortunately, the NIJ’s assertion that PATTERN makes greater use of dynamic factors rings hollow for several reasons. It is a claim undermined by the fact that static factors considered by PATTERN consistently outweigh the dynamic factors that inform one’s PATTERN score. Moreover, to the extent an individual seeks to engage in behavior that would improve their risk category, the finite weight given to dynamic factors and the limited BOP programming available to do so frustrate the goals of the FIRST STEP ACT. Finally, several concerns remain with the selection of particular dynamic factors included in the various PATTERN models.

A specific example shows how PATTERN undervalues dynamic factors.

A review of PATTERN’s static and dynamic factors, along with an example of how those factors and the BOP’s rehabilitative offerings interact in an individual case, is instructive. PATTERN’s *static factors* include: age at first conviction, current incarceration reason, sex offender status, age at time of PATTERN assessment, criminal history score, history of violence, history of escape, voluntary surrender, and education score.²⁷ PATTERN’S *dynamic factors* include: general infraction convictions, serious and violent infraction convictions, technical/vocational courses, beneficial programs, participation in the federal industry employment UNICOR program, drug education, and fiscal responsibility non-compliance.



With these factors in mind, assuming a general recidivism model, take the example of a male deemed medium risk who was convicted of robbery in his teenage years (12 points) where he panicked, fought (1 point), and attempted to escape (2 points). Twenty years after that first conviction, this person has been apprehended for a minor federal drug offense (30 points). While serving time for the second offense, this person participates in programs that reduce his recidivism score like drug education classes (-1 point) and completes drug treatment (-2 points).²⁸ Despite his best efforts to lower his recidivism score and reintegrate into society, the static factors will consistently outweigh any subsequent progress. The heavy weight of static factors discourages incarcerated persons from participating in rehabilitation programs because their efforts ultimately will have no impact on their eligibility for release.

These concerns are compounded when one considers the limited availability of rehabilitative programming in the BOP. The BOP boasts 17 programs: 11 psychology services programs, five education and vocational programs, and one chaplaincy services program.²⁹ However, in the 102 federal detention facilities in the U.S., no facility offers all 17 programs.³⁰ Waitlists for rehabilitation programs are unfathomably long, and the programming varies in quality.³¹ For example, the waiting list for the BOP's literacy program is 16,000 people.³² The process of awarding program placement to people in prison is opaque, infused with too much discretion, and hampered by limitations on program availability to make participation a reliable measure of one's fitness for release.³³

Despite proclaiming the “greater use of dynamic factors,” NIJ concedes that static factors are weighted more than dynamic factors, making it nearly impossible for many people labeled as medium risk to benefit from their participation in rehabilitative programming.³⁴ This is inconsistent with the spirit of the statute.

Moreover, two specific dynamic factors deserve more scrutiny.

The first dynamic variable is “non-compliance with financial responsibility,” which the report states is “an offender’s willingness to use income earned during incarceration for payment toward victim restitution and dependents.” This dynamic factor is only included for predicting general recidivism in PATTERN’s female model. As an initial matter, it is unclear how this variable will be measured given that it is defined, by its own terms, as “willingness.”³⁵ To the extent this willingness is measured through historical data — which, again, is unclear as the report is vague on the matter — there are significant problems.

Incarcerated people’s ability to pay fines or fees associated with a conviction should never serve as the basis to deny them freedom or the ability to meaningfully reduce their risk score. Individuals in prison are impoverished and do not make a living wage. Even if given the most lucrative jobs for a whole year, imprisoned persons cannot make more than \$307.05 per year.³⁶ On such a modest salary, incarcerated people can barely afford their life in prison, let alone pay conviction-associated fees and fines.³⁷ One egregious example of the penal system punishing poverty is the cost of feminine hygiene products. Traditionally, menstrual hygiene products at prison commissaries have been extraordinarily expensive. In response to the public outrage, BOP changed its policy to offer these products to women for free in August 2017.³⁸ Further, Section 511 of the FIRST STEP Act now directs the Director of BOP to offer tampons and



sanitary napkins for free.³⁹ To the extent that PATTERN’s “non-compliance with financial responsibility” dynamic variable was developed on and based on historical data — when those incarcerated were required to pay exorbitant rates for basic menstrual hygiene products — PATTERN risks penalizing those who may have otherwise directed earned income toward victim restitution and dependents, but instead choose necessary hygiene products. This is not a productive result. NIJ must more clearly describe how this variable was and will be measured.

The second dynamic variable is number of technical or vocational courses “created as a count metric.” This dynamic factor is only included for predicting general recidivism – and not violent recidivism – for PATTERN’s male and female models. As a threshold issue, it is unclear from the report what, exactly, this dynamic variable is actually measuring. Given that, NIJ must provide more clarity. To the extent that this variable is measuring how many technical or vocational courses someone completes, there appear to be three options for scoring this variable: either someone has not taken a technical or vocational course (0), is currently in progress or has only finished one ($>0, \leq 1$), or has taken more than 1 (>1). According to PATTERN’s methodology, those incarcerated would receive 0 points for completing more than one course but receive between -2 and -4 for no technical or vocational courses. This logic makes no sense, given the available evidence on how vocational and technical courses help reduce future incarcerations.⁴⁰ Therefore, we encourage NIJ to eliminate the factor of ability to pay when deciding incarcerated people’s eligibility for release altogether.

V. NIJ’s claim that PATTERN’s “predictive performance is unbiased across racial and ethnic classifications” ignores historical and enduring patterns of racial bias and discrimination that infect the data upon which PATTERN relies.

A fundamental, widespread criticism of risk assessment tools is their reliance on data that correlates with race to produce forecasts about an individual’s behavior.⁴¹ Doing so raises a concern that the forecasts made by the tool will reflect, reproduce, and exacerbate racial bias. The NIJ is well aware of this concern, stating that “nearly all indicators likely to be used within a risk assessment model have the potential to be correlated with socio-economic status (SES), race, and/or ethnicity.”⁴² Accordingly, PATTERN’s developers attempted to measure the likelihood of non-white vs. white individuals being scored as minimum or low risk through the “relative rate index” (RRI) to “assess the magnitude of disparity across risk categories.”⁴³ The RRI, which, to our knowledge, has traditionally been used to assess disproportionate minority contact within the juvenile justice system, is but one way to assess certain disparities.

In truth, NIJ’s claim that PATTERN is “unbiased across racial and ethnic classifications” is dependent on a constrained definition of racial bias as a statistical matter. The report defines a “racially unbiased” tool as one in which a tool is “correctly calibrated or standardized within racial and ethnic groups.”⁴⁴ Further, though the report cites critical academic findings regarding metrics of racial fairness and risk assessment systems, it does not fully grapple with them. This research has shown that, given differing base rates of arrest amongst two groups, even if risk scores are well-calibrated, a system cannot satisfy other metrics, like equal false-positive rates.⁴⁵ Of course, given that the underlying base rate here—rearrests—diverge



along racial lines, it is “mathematically impossible to develop a model that will be fair in the sense of having equal predictive value across groups, and fair in the sense of treating members of groups similarly in retrospect.”⁴⁶

However, that impossibility does not mean we should just give up and simply define whether a risk assessment system is racially biased or not through that single metric. Instead, it calls for a deeper analysis. NIJ must more expansively test PATTERN for racial bias, using a variety of available statistical measures including false positives and false negatives.⁴⁷ We encourage NIJ to engage the many researchers in the field of computer science and data science working on fairness, accountability, and transparency, to conduct this kind of analysis.⁴⁸

VI. PATTERN’S separate risk calculations on the basis of gender raises serious constitutional concerns

One of PATTERN’s defining features is its use of gender-responsive modeling to build a tool that forecasts recidivism risk at different levels for men and women. Such modeling is based on the proposition that men and women have differing recidivism risks based, in large part, on their gender.⁴⁹ The inclusion of gender — and, in turn, the creation of a separate tool for men and women — raises serious constitutional concerns, such concerns deriving from the Supreme Court’s rejection of gender classifications that are grounded in statistical generalizations about groups.⁵⁰ To the extent PATTERN’s differential scores for men and women are grounded in efforts to infer individual tendencies from group statistics, those efforts are constitutionally suspect.⁵¹ The NIJ should abandon this model altogether.

Conclusion

We appreciate the efforts undertaken by the DOJ to fulfill the mandate of the FIRST STEP ACT and are grateful for the opportunity to offer comments regarding the development of the risk and needs assessment to be used for that purpose. Unfortunately, however, our review of the NIJ’s report regarding the development and implementation of PATTERN leads to several troubling conclusions. First, the assertion that PATTERN “achieves a high level of predictive performance” is belied by the fact that risk assessment systems based on historical trends do not, and cannot, accurately predict recidivism.⁵² Second, the claim that PATTERN makes “greater use of dynamic factors” is misleading and several existing dynamic factors are problematic on their face.⁵³ Third, for a variety of reasons, PATTERN’s predictive performance cannot be understood as unbiased across racial and ethnic classifications.⁵⁴ Finally, using separate risk calculations for men and women raises serious constitutional questions.

Therefore, we urge the National Institute of Justice, the Federal Bureau of Prisons, and the Department of Justice to address our fundamental concerns about PATTERN. While we acknowledge the release of 3,100 incarcerated persons because of the sentencing and good conduct provisions that our organizations advocated for inclusion in the FIRST STEP ACT, we remain concerned about the racial, ethnic and gender biases that the “risk assessment system,” PATTERN, exhibits — which could be holding back thousands



more from the freedom they deserve.⁵⁵

We look forward to hearing from you soon. If you have any questions about the issues raised in this comment letter, please contact Sakira Cook, Director, Justice Reform at cook@civilrights.org or Antoine Prince Albert III, Technology Fellow at apalbert@civilrights.org.

Respectfully submitted,

American Civil Liberties Union
 Center on Race, Inequality, and the Law at NYU Law
 The Justice Roundtable
 The Leadership Conference Education Fund
 The Leadership Conference on Civil and Human Rights
 Media Mobilizing Project
 Upturn

Enclosure:
 American Bar Association (ABA) Comment Letter

¹ Formerly Incarcerated Reenter Society Transformed Safely Transitioning Every Person Act, P.L. 115-391, 115th Congress, Dec 21, 2018.
² The Leadership Conference on Civil & Human Rights, Vote “No” on The FIRST STEP Act, May 8, 2018, <https://civilrights.org/resource/vote-no-first-step-act/>.
³ *Id.*
⁴ *Id.*
⁵ Statement for the Record of The ACLU, Justice Roundtable, and The Leadership Conference in Response to Department of Justice (DOJ) April 3 and 5 Listening Sessions, Apr. 12, 2019, <https://civilrights.org/resource/statement-for-the-record-of-the-aclu-justice-roundtable-and-the-leadership-conference-in-response-to-department-of-justice-doj-april-3-and-5-listening-sessions/>.
⁶ Response of the ACLU, Justice Roundtable, and The Leadership Conference to Hudson Institute’s Request for Supplemental Information, Jun. 14, 2019, <https://civilrights.org/resource/response-of-the-aclu-justice-roundtable-and-the-leadership-conference-to-hudson-institutes-request-for-supplemental-information/>.
⁷ The National Institute of Justice (associated with the U.S. Dep’t of Justice’s Office of Justice Programs), The First Step Act of 2018: Risk and Needs Assessment System (Jul. 2019), https://nij.ojp.gov/sites/g/files/xyckuh171/files/media/document/the-first-step-act-of-2018-risk-and-needs-assessment-system_1.pdf [herein after NIJ Report].
⁸ P. L. 115-391 (18 U.S.C.) § 3632(a)(3) and (5) (explaining that “the Attorney General, in consultation with the Independent Review Committee . . . shall develop and release publicly . . . a risk and needs assessment system . . . which shall be used to-- . . . (3) determine the type and amount of evidence-based recidivism reduction programming that is appropriate for each prisoner and assign each prisoner to such programming accordingly, and based on the prisoner’s specific criminogenic needs, and in accordance with subsection (b) . . . and (5) reassign the prisoner to appropriate evidence-based recidivism reduction programs or productive activities . . .”).
⁹ *Supra* note 7.
¹⁰ NIJ Report at 26, 51.
¹¹ See Brandon L. Garrett, Federal Criminal Risk Assessment, *Cardozo L. Rev.* Forthcoming 102, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3425183; Sharad Goel, Ravi Shroff, Jennifer L. Skeem and Christopher Slobogin, The Accuracy, Equity, and Jurisprudence of Criminal Risk Assessment, Dec. 26, 2018, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3306723;



Laurel Eckhouse, Kristian Lum, Cynthia Conti-Cook, Julie Ciccolini, Layers of bias: A unified approach for understanding problem with risk assessment, *Criminal Justice And Behavior*, Nov. 23, 2018.

¹² See also American Bar Association, Re: ABA Questions Regarding the Proposed Risk and Needs Assessment System, submitted on August 26, 2019. See enclosure.

¹³ *Id.*

¹⁴ We do know that BRAVO-R was developed by the BOP’s Office of Research and Evaluation (“ORE”) to predict three-year recidivism risk for people released from BOP facilities. BRAVO-R is a modified version of the Bureau Risk and Verification Observation (“BRAVO”) tool which BOP has used since the 1970s to predict misconduct risk for BOP custody decisions.

¹⁵ U.S. Department of Justice, Legal Resource Guide to the Federal Bureau of Prisons (2019), https://www.bop.gov/resources/pdfs/legal_guide_march_2019.pdf.

¹⁶ Matt Shipman, Research Finds Offender Risk Assessment Tools in U.S. Are Promising, but Questions Remain, North Carolina State University (Jun. 2016), <https://news.ncsu.edu/2016/06/offender-risk-assessments-2016/>.

Sarah L. Desmarais and Kiersten L. Johnson, North Carolina State University; Jay P. Singh, Global Institute of Forensic Research, Performance of Recidivism Risk Assessment Instruments in U.S. Correctional Settings”.

¹⁷ Partnership on AI, *Report on Algorithmic Risk Assessment Tools in the U.S. Criminal Justice System*, at 29. (“the training datasets, architectures, algorithms, and models of all tools under consideration for deployment must be made broadly available to all interested research communities— such as those from statistics, computer science, social science, public policy, law, and criminology, so that they are able to evaluate them before and after deployment.”)

¹⁸ NIJ should describe what instances qualify as a return to BOP custody following release within the training data.

¹⁹ Recidivism is broadly defined (i.e. arrest or return to BOP custody) could be for technical violations or false arrests or arrest that isn’t adjudicated yet and the true measure of recidivism should be conviction of a new crime.

²⁰ Partnership on AI, *Report on Algorithmic Risk Assessment Tools in the U.S. Criminal Justice System*, 16. (“Statistical validation of recidivism prediction in particular suffers from a fundamental problem: the ground truth of whether an individual committed a crime is generally unavailable, and can only be estimated via imperfect proxies such as crime reports or arrests.”)

²¹ Laurel Eckhouse, Kristian Lum, Cynthia Conti-Cook, Julie Ciccolini, Layers of bias: A unified approach for understanding problem with risk assessment, *Criminal Justice And Behavior*, Nov. 23, 2018.

²² (Noting that the report “[r]el[ies] on the AUC as the primary metric for evaluating predictive validity...” at 50.

²³ Jay P. Singh, Predictive Validity Performance Indicators in Violence Risk Assessment: A Methodological Primer, *Behav. Sci. Law* 31: 8–22 (2013).

²⁴ By definition, the dynamic factors can change based on successful completion of the previously mentioned BOP’s rehabilitative programs.

²⁵ 18 U.S.C. §3632(a)(4) “... the Independent Review Committee authorized by the First Step Act of 2018, shall develop and release publicly. . . a risk and needs assessment system . . . which shall be used to – reassess the recidivism risk of each prisoner periodically, based on factors including indicators of progress, and of regression, that are dynamic and that can reasonably be expected to change while in prison.”

²⁶ NIJ Report at 26, 51 (2019).

²⁷ NIJ Report at 26, 53-56.

²⁸ The point values detailed in this hypothetical are set forth in the NIJ Report at 53-56.

²⁹ Federal Bureau of Prisons, Directory of National Programs: A practical guide highlighting reentry programs available in the Federal Bureau of Prisons (2017), https://www.bop.gov/inmates/custody_and_care/docs/20170518_BOPNationalProgramCatalog.pdf.

³⁰ See *id.* (listing the limited number of facilities at which particular programs are available).

³¹ See, e.g., Drug Abuse Treatment Program, 81 Fed. Reg. 24,484, 24,488 (Apr. 26, 2016) (codified at 28 C.F.R. pt. 550) (noting that as of 2016, over 5000 inmates were on the waitlist for the Residential Drug Abuse Treatment Program); Federal Bureau of Prisons, Federal Bureau of Prisons Education Program Assessment: Final Report 22 (Nov. 29, 2016) (noting that as of 2016, 15,629 inmates were on the waitlist for the Bureau Literacy Program); see also *id.* at v. (finding the organizational structure of BOP’s educational programs to be “incoherent” and that “occupational training options vary by institution, are often unaccredited, and rarely lead to meaningful certifications”).

³² FY 2019 Performance Budget: Congressional Submission, United States Department of Justice Federal Prison System, available at: <https://www.justice.gov/jmd/page/file/1034421/download>

³³ See, e.g., Federal Bureau of Prisons, Federal Bureau of Prisons Education Program Assessment: Final Report 22 (Nov. 29, 2016) at 39 (finding with regard to BOP’s English-as-a-Second Language Program that “[t]he Admissions and Orientation interview is not a formal, standardized instrument or protocol” and that “[t] is unclear how the interviewer determines an inmate’s English proficiency”).

³⁴ NIJ Report at 26 (“In many cases, however, dynamic items are only incrementally predictive of criminal behavior, as compared to the static items in various RNA tools.” (citing Faye S. Taxman, *Risk assessment: Where do we go from here?*, in Handbook of Recidivism Risk/Needs Assessment Tools 271–84 (Jay P. Singh et al. eds., 2018)).

³⁵ Though the term is footnoted in the report, it appears that it is incorrectly footnoted.

³⁶ Federal inmates working custodial, maintenance, laundry, grounds keeping, and food service jobs earn between \$0.12 to \$1.15 per hour for at least 7 hours of work per day. With weekends off, inmates can earn between \$31.32 and \$307.05 per year. See Prison Policy Initiative, How much do incarcerated people earn?: State and federal prison wage policies and sourcing information, https://www.prisonpolicy.org/reports/wage_policies.html -- appendix to Wendy Sawyer, “How much do incarcerated people earn in each state?,” Prison Policy Initiative, April 10, 2017, <https://www.prisonpolicy.org/blog/2017/04/10/wages/>.

³⁷ Data for Progress leading The Justice Collaborative, “Voters Support Reducing the Use of Fines and Fees in Sentencing,” August 2019,



http://filesforprogress.org/memos/fines_and_fees.pdf (proposing renewed administrative, judicial and prosecutorial discretion in setting reduced fees or eliminating them all together for indigent people).

³⁸ Federal Bureau of Prisons, Operations Memorandum, “Provision of Feminine Hygiene Products,” August 1, 2017,

https://www.bop.gov/policy/om/001_2017.pdf

³⁹ *Id.*

⁴⁰ Davis, Lois M., Robert Bozick, Jennifer L. Steele, Jessica Saunders, and Jeremy N. V. Miles, Evaluating the Effectiveness of Correctional Education: A Meta-Analysis of Programs That Provide Education to Incarcerated Adults. Santa Monica, CA: RAND Corporation, 201.

⁴¹ Leadership Conference on Civil and Human Rights, The Use of pretrial Risk Assessment Instruments: A Shared Statement of Civil Rights Concerns, <http://civilrightsdocs.info/pdf/criminal-justice/Pretrial-Risk-Assessment-Full.pdf>.

⁴² NIJ Report at 59-60.

⁴³ NIJ Report at 53.

⁴⁴ NIJ Report at 28.

⁴⁵ Jon Kleinberg et al., Inherent Trade-Offs in the Fair Determination of Risk Scores, ARXIV (Nov. 17, 2016),

<https://arxiv.org/pdf/1609.05807.pdf>; Alexandra Chouldechova, Fair prediction with disparate impact: A study of bias in recidivism prediction instruments, ARXIV (Feb. 2017), <https://arxiv.org/abs/1703.00056>.

⁴⁶ Laurel Eckhouse, Kristian Lum, Cynthia Conti-Cook, Julie Ciccolini, Layers of bias: A unified approach for understanding problem with risk assessment, *Criminal Justice And Behavior*, Nov. 23, 2018, 6.

⁴⁷ Partnership on AI, *Report on Algorithmic Risk Assessment Tools in the U.S. Criminal Justice System*, at 21 (“Given that each of these approaches involves inherent trade-offs,³⁵ it is also reasonable to use a few different methods and compare the results between them. This would yield a range of predictions that could better inform decision-making.”)

⁴⁸ See Technical Flaws of Pretrial Risk Assessments Raise Grave Concern, Jul. 16, 2019, A letter to the Los Angeles County Board of Supervisors from researchers in the fields of statistics, machine learning and artificial intelligence, law, sociology, and anthropology in their personal capacities.

⁴⁹ Sonja B. Starr, *Evidence-Based Sentencing and the Scientific Rationalization of Discrimination*, 66 *Stan. L. Rev.* 803, 825 (2014).

⁵⁰ *Id.*; see *Craig v. Boren*, 429 U.S. 190, 202 (1976) (noting that the Supreme Court has “consistently rejected the use of sex as a decisionmaking factor even though the statutes in question certainly rested on . . . predictive empirical relationship[s]”).

⁵¹ Starr, 66 *Stan. L. Rev.* at 826.

⁵² NIJ Report at 43, 50, 63.

⁵³ *Id.* at 45, 63.

⁵⁴ *Id.* at 28, 63.

⁵⁵ Matt Zapotosky, 3,100 inmates to be released as Trump administration implements criminal justice reform, *Washington Post* (Jul. 19, 2019), https://www.washingtonpost.com/national-security/3100-inmates-to-be-released-as-trump-administration-implements-criminal-justice-reform/2019/07/19/7ed0daf6-a9a4-11e9-a3a6-ab670962db05_story.html?noredirect=on.