JUNE 2023

Federal PROBATION

a journal of correctional philosophy and practice

Federal PROBATION

a journal of correctional philosophy and practice

PUBLISHED BY
The Administrative Office of the U.S. Courts

Judge Roslynn R. Mauskopf, *Director*

John J. Fitzgerald, *Chief*
Probation and Pretrial Services Office

Federal Probation ISSN 0014-9128 is dedicated to informing its readers about current thought, research, and practice in criminal justice, community supervision, and corrections. The journal welcomes the contributions of persons who work with or study defendants and offenders and invites authors to submit articles describing experience or significant findings regarding the prevention and control of crime and delinquency. A style sheet is available from the editor.

Federal Probation is published three times yearly—in June, September, and December. Permission to quote is granted on the condition that appropriate credit is given the author and *Federal Probation*. For information about reprinting articles, please contact the editor.

Subscriptions to *Federal Probation* are available from the Superintendent of Documents of the Government Printing Office at an annual rate of $16.50 ($22.40 foreign). Please see the subscription order form on the last page of this issue for more information.

Federal Probation can also be accessed online at no charge at www.uscourts.gov.

EDITORIAL STAFF

Charles Robinson, *Executive Editor*
Ellen Wilson Fielding, *Editor*

Federal Probation
Administrative Office of the U.S. Courts
Washington, DC 20544
telephone: 202-502-1651
fax: 202-502-1677
email: Ellen_Fielding@ao.uscourts.gov

Postmaster: *Please send address changes to the editor at the address above.*

THIS ISSUE IN BRIEF

Is Pretrial Detention an Effective Deterrent? An Analysis of Failure to Appear and Rearrest Says No"

Alexander M. Holsinger
University of Missouri—Kansas City
Christopher T. Lowenkamp
University of Missouri—Kansas City
Probation and Pretrial Services Office
Administrative Office of the U.S. Courts
Travis C. Pratt
University of Cincinnati
Harris County Community Supervision and Corrections Department

ANY TIME A PERSON is arrested and accused of committing a crime, a decision has to be made during this "pretrial" stage about whether the individual facing charges is going to be released directly back into the community right away or is instead detained in jail to await the next stage of case processing (McIntyre & Baradaran, 2013; Oleson et al., 2016; Sacks, Sainato, & Ackerman, 2015). This decision is not, however, a strictly "either-or" proposition (Martinez, Petersen, & Omori, 2020). Some defendants, for instance, are incarcerated during the pretrial stage for a long time, some not at all, while others only spend a few days incarcerated before being released (Kim et al., 2018; Lowenkamp, Van Nostrand, & Holsinger, 2013; Sacks & Ackerman, 2014).

How this decision is handled is critical, because pretrial detention carries serious consequences "downstream" in the justice process (Martinez, Petersen, & Omori, 200). To be sure, research indicates that being incarcerated prior to trial is associated with an increased likelihood of being convicted (Menefee, 2018; Petersen, 2020), of being sentenced to prison (and for a longer period of time) (Donnelly & MacDonald, 2018; Williams, 2003), and even of finding it harder to find a job later (Dobbie, Goldin, & Yang, 2018; Wakefield & Anderson, 2020). It is therefore inevitable that the decision to detain someone in jail before trial—or to let them stay out on their own recognizance—is based on a complex set of factors (Dobbie, Goldin, & Yang, 2018; Viljoen et al., 2019). Concerns about community safety, the constitutional rights of justice-involved persons, and the need for individuals to appear in court all play an important role (Leslie & Pope, 2017; Oleson et al., 2016).

But when judges decide to detain someone for a stint of incarceration prior to their trial, the primary legal justification that is often invoked is rooted in the language of deterrence (D'Alessio & Stolzenberg, 1998; Pyrooz, Gartner, & Smith, 2017; Walker & Herting, 2020). To be sure, those who favor locking people up prior to their trial typically assume that pretrial detention causes those facing charges to "think twice" about failing to show up at court or committing a new crime later on, because they want to avoid being incarcerated again (Raaijmakers et al., 2017). Such a position is consistent with the long-standing belief in American jurisprudence—at least among some—that incarceration "works" as an effective deterrent (see, e.g., the discussion by Pratt, 2019). So if this is actually the case, there may be a benefit to the practice of pretrial detention with respect to public safety.

We do, however, have good reason to believe that the deterrent power of incarceration has been grossly exaggerated. Indeed, stacks of criminological literature indicate that the threat of stiffer sanctions does little to deter people from committing crimes (Apel, 2013; Pratt & Cullen, 2005; Pratt et al., 2006); that trying to lock people up more quickly in an effort to satisfy the "swiftness" element of deterrence—the key marketing strategy for the popular-yet-empirically-unsupported "swift-certain-fair" model of punishment—fares no better (Cullen, Pratt, & Turanovic, 2016; Cullen et al., 2018; Pratt & Turanovic, 2018); and that locking up lower risk people may actually end up doing more harm than good when it comes to recidivism (Ogle & Turanovic, 2019). In the most recent comprehensive assessment of this idea, Petrich et al.'s (2021) meta-analysis of over 100 studies on the effects of custodial versus community sanctions indicated that incarceration actually makes things worse for justice-involved individuals—a finding that highlights not only the compromising of public safety when incarceration is used as a "general" crime-control strategy, but also the additional cost

of incarcerating citizens unnecessarily that the public will have to bear.

Even so, little of this research is focused on the pretrial stage. Thus, although we may think of it as "criminological fact" that incarceration is neither a consistent nor effective deterrent to criminal—or even problematic—behavior (Petrich et al., 2021:353), it is still unclear whether this fact extends to the pretrial phase of case processing. It is also possible—in line with the "deterrability hypothesis"—that people vary in their response to sanctions in general and to incarceration in particular (Herman & Pogarsky, 2022; Jacobs, 2010; Maxson, Matsuda, & Henigan, 2011). Put simply, some people may be deterred by pretrial detention and "learn their lesson," others might get worse because of it, and still others might be unaffected at all (see, e.g., Braithwaite, 1989; Sherman, 1993).

Accordingly, the present study uses data on 1,487,107 individuals booked into a jail in Kentucky between 2009 and 2018 to address two research questions: (1) does being subjected to pretrial detention reduce the likelihood of a defendant failing to appear (FTA) in court? (2) does being subjected to pretrial detention (at all, and with varying length of incarceration) reduce the likelihood of a defendant acquiring a new arrest during the pretrial phase? In answering these questions, our broader purpose is to shed light on whether—or perhaps to what extent—locking people up prior to their trial represents sound public policy.

Methods

Data

The sample used for the current study includes the 1,487,107 cases that involved arrest and booking into a Kentucky jail between the years of 2009 and 2018. Data elements included demographic characteristics (sex, race, age), the actuarial risk of failure to appear for a court hearing (FTA) and risk of new criminal activity during the pretrial stage (NCA) as assessed via the Public Safety Assessment (PSA), characteristics of the booking offense, the defendant's pretrial status, and time at risk in the community.[1]

A large majority of the sample was male (71 percent) and white (80 percent) while defendants' average age was 34.3 years. The risk profile of the entire sample for FTA was 12 percent, 22 percent, 23 percent, 20 percent, 17 percent, and 5 percent for categories one through six respectively. For NCA the distribution of actuarial risk was 10 percent, 25 percent, 23 percent, 24 percent, 11 percent, and 6 percent for categories one through six respectively. Most defendants had been arrested and booked for a misdemeanor offense (61 percent), and likewise a majority (56 percent) experienced pretrial detention for less than 24 hours. For the analyses presented below, the primary variable of interest is days spent in pretrial detention (assessed first as a binary, then as an ordinal variable), with number of charges, felony charge, misdemeanor charge, violent charge, property charge, time at risk in the community, race, and sex serving as control variables.

The two outcome measures include failure to appear for at least one court hearing during the period of pretrial release (FTA – 0 = No; 1 = Yes), and rearrest for a new offense during the period of pretrial release (NCA – 0 = No; 1 = Yes). The base rates of each outcome were low (FTA = 17%; NCA = 12%) but were high enough to allow for all necessary analyses.

Analytic Strategy

In order to assess the extent to which pretrial incarceration may serve as a deterrent to missing court hearings and/or new criminal activity while released pretrial, four binary regression models were calculated. For each respective outcome (FTA and NCA), one model was calculated using days spent in pretrial detention as a binary (less than 24 hours vs. 24 hours or more) followed by a second model that used categories of time (e.g., 1 day, 2 days, 3 days) with less than 24 hours (which also included no time spent in detention) serving as the reference category. For all four models, control variables included charge characteristics that may be related to FTA or NCA (number of charges, felony charge, misdemeanor charge, violent charge, property charge), time at risk in the community, and demographic characteristics (race and sex). Risk ratios and their statistical significance were used to assess the relationship between each covariate and outcome.

Results

Table 2 (next page) presents the results from the first model, using time spent in detention as a binary (less than 24 hours vs. 24 hours or more) in order to predict FTA. In addition, the analyses were restricted to those defendants who were released to the community during the pretrial period and likewise had at least 21 but not more than 365 days of time at risk

TABLE 1.
Descriptive statistics, entire sample

Variable	N	%
Sex		
Male	1,056,966	70.92%
Female	418, 012	28.05%
Unknown	15,408	1.03%
Race		
White	1,200,295	80.54%
Black	256,054	17.18%
Unknown	29,613	1.99%
Asian	3,470	0.23%
Indian	963	0.06%
Age	34.44 years ($\bar{X}$)	
Risk of FTA (PSA)		
One	94,671	12%
Two	172,093	22%
Three	182,589	23%
Four	157,675	20%
Five	133,482	17%
Six	36,703	5%
Risk of NCA (PSA)		
One	78,014	10%
Two	196,654	25%
Three	176,446	23%
Four	188,521	24%
Five	87,148	11%
Six	47,440	6%
Booking offense		
Felony No	911,746	61.18%
Felony Yes	578,637	38.82%
Pretrial detention		
Under 24 hours	624,070	56.32%
24 hours+	484,050	43.68%
Failure to appear (FTA)		
No	923,149	83.22%
Yes	186,205	16.78%
New criminal activity (NCA)		
No	976,488	88.02%
Yes	132,886	11.98%

[1] The exact number of cases used in each analysis varies due to missing data. Sometimes this number is substantially different and sometimes it is not. For detail on changes in the number of cases included in each model, see the technical appendix at: https://osf.io/ykuqd/?view_only=6dad2630567e425d9c9636c03d6d0e37.

in the community. This decision was made in an attempt to limit the analyses to defendants and cases where there was enough time at risk to fail. The limit on the upper end was used given most cases (greater than 80 percent) are resolved within one year, and information from the agency providing the data indicated cases that stretched beyond that time frame were atypical. Neither time in detention nor sex met the criteria for statistical significance. In other words, there does not appear to be a relationship between time spent in pretrial detention and the likelihood of failure to appear when controlling for number of charges, the characteristics of the charge(s), sex, and race. Every other variable in the model (save sex, as noted above) did reveal a statistically significant relationship with failure to appear. Defendants who had more charges (as opposed to fewer), a misdemeanor charge, a property charge, and who spent more time at risk in the community, were more likely to fail to appear for one or more court hearings, as were African American defendants. Defendants charged with a felony and/or those charged with a violent offense were significantly less likely to miss one or more court dates.

Table 3 presents the results from the second model that used time spent in detention as an ordinal variable, with the interval 0 to 23 hours serving as the reference category (the same case restrictions as noted above regarding release and time at risk in the community were observed). While the model contained in Table 2 indicates that time spent in detention is *not* related to FTA, it is possible that a relationship could be revealed after a certain point, or after a specific amount of time is spent incarcerated. Each successive time interval represents an approximate additional day of time, with that amount of time spent in detention compared to the reference category (0 to 23 hours, or, less than 1 day in detention, labeled as category 0). For example, category "1" represents those defendants who spent between 24 and 47 hours in detention (i.e., between one whole day and just shy of two whole days), who are in effect compared to defendants who spent less than one whole day (less than one 24-hour period) in detention, which also includes those who spent no measurable amount of time in detention at all. Category "2" represents whose defendants who spent between 48 and 71 hours in detention (i.e., between two whole days and just shy of 3 whole days), who in turn are compared to defendants who spent less than one whole day (less than one 24-hour period) in detention, and so on.

Similar results were revealed for all the control variables that were included in the model. Once again, sex does not appear to be related to FTA, while number of charges, a misdemeanor charge, property charge, time at risk, and race (African American) all significantly increase the likelihood of FTA occurring. Likewise, as before, being charged with a felony and/or a violent crime appear to be associated with a decreased likelihood of FTA. Interestingly, none of the categories of time spent in detention were significantly related to FTA, except for 10, 11, and 12 days. In short, amounts of time spent in detention that lasted between 1+ and up to 9+ days were statistically unrelated to FTA, as were amounts of time ranging from 13+ days and higher. Despite a relationship emerging for 10, 11, and 12 days, it appears that the relationship between time spent in detention and FTA is non-existent for all intents and purposes, if not inconsistent.

Table 4 (next page) presents a similar model to that which appears in Table 2, except that Table 4 uses new arrest (NCA) as the dependent variable. Once again, analyses were restricted to those defendants who were released pretrial and who also had at least 21 days but not more than 365 days of time at risk in the community before their case was resolved. In addition, the same variables as before were used as predictors in the model (charge characteristics, time at risk, race, sex), with the primary variable of interest days spent in detention measured as a binary (less than 24 hours vs. 24 hours or more). Days spent in detention measured as a dichotomy revealed a statistically significant relationship with rearrest, with those spending more time in detention (more than 23 hours) having a lower likelihood. The number of charges was also significantly related to new arrest (more charges = lower likelihood), as was having a

TABLE 2.
Predicting FTA – Binary regression model predicting FTA with time in detention as a binary dummy variable. Limited to those released from pretrial detention and with at least 21 days of time at risk and not more than 365 days of time at risk

Variable	Risk Ratio
Number of charges	**1.038**
Felony charge	**0.704**
Misdemeanor charge	**1.340**
Against person	**0.576**
Property	**1.380**
Time at risk	**1.003**
Black	**1.198**
Male	0.997
Days in detention (under 24 hours)	0.902
Constant	**0.082**

Bold = p ≤ .001

TABLE 3.
Predicting FTA – Binary regression model predicting FTA with time in detention as an ordinal variable. Limited to those released from pretrial detention and with at least 21 days of time at risk and not more than 365 days of time at risk

Variable	Risk Ratio
Number of charges	**1.038**
Felony charge	**0.691**
Misdemeanor charge	**1.352**
Against person	**0.553**
Property	**1.383**
Time at risk	**1.003**
Black	**1.216**
Male	1.001
Days/hours in detention (under 24 hours)	
(0/23=0)	Reference
(24/47=1)	1.061
(48/71=2)	1.060
(72/95=3)	1.108
(96/119=4)	1.124
(120/143=5)	1.137
(144/167=6)	1.118
(168/191=7)	1.092
(192/215=8)	1.114
(216/239=9)	1.110
(240/263=10)	**1.186**
(264/287=11)	**1.259**
(288/311=12)	**1.185**
(312/335=13)	1.128
(336/359=14)	1.044
(360/383=15)	1.212
(384/407=16)	1.131
(408/431=17)	1.165
(432/455=18)	1.157
(456/479=19)	1.179
Constant	**0.071**

Bold = p ≤ .001

misdemeanor charge (higher likelihood), a charge for a property offense (higher likelihood), time at risk in the community pretrial (higher likelihood), and being male (higher likelihood). Being charged with a felony, a violent offense, and race were statistically unrelated to new arrest during the pretrial period.

Interestingly, things change dramatically when the expanded measure of time spent in detention is used. Table 5 presents a model similar to that displayed in Table 3, although the outcome variable is new arrest during the pretrial period. The same control variables were used (charge characteristics, time at risk, sex, race), and the same case restrictions were in place as well regarding release status and time at risk in the community. Once again, the primary variable of interest was the ordinal measure of time spent in detention, measured as described above, with each success interval of time compared to the reference category of 0 to 23 hours ("0"). Each category of time spent in detention revealed a statistically significant relationship with NCA, relative to spending the smallest amount of time (0 to 23 hours) in detention. Moreover, each coefficient was greater than 1.0, indicating that every interval of time (1+ day, 2+ days, 3+ days, and so on) had a significantly higher likelihood of rearrest relative to those defendants that spent the least amount of time in detention pretrial. Statistically significant relationships were also revealed for number of charges (more charges = less likely to be rearrested), being charged with a misdemeanor (more likely to be rearrested), being charged with a property offense (rearrest is more likely), time at risk in the community pretrial (more likely to be rearrested), and being a male defendant (arrest was more likely). Further, being charged with a felony, a violent offense, and race did not reveal a relationship with rearrest during the pretrial period.

Discussion

Pretrial detention—and the wide array of bail reform efforts that have come along with it in recent years—continues to be a source of contention in public policy circles. And a big part of the controversy has to do with whether keeping someone locked up prior to trial is helpful (i.e., that it serves as a deterrent, or at least yields a bit of an incapacitation effect), or if it instead makes things worse. So with that in mind, we took a closer look at the consequences of pretrial detention, and one rather significant—and unequivocal—conclusion is warranted.

We did not find any evidence of a consistent or reliable "deterrent effect" of pretrial detention on either the failure to appear (FTA) or recidivism. This should come as no surprise. The research literature has been clear on this issue for several decades now: getting "tough" on crime (or on recidivism, or juvenile delinquency, or school violence; pick your preferred form of misbehavior), as a "general" strategy, is a bad idea (Petrich et al., 2021). The question, of course, is: why? To answer that, evidence indicates that incarceration—even if the stint is short—can cut justice-involved people off from prosocial attachments to things like their job and their social relationships, which tends to increase the likelihood of reoffending (Maroto & Sykes, 2020). In addition, while people are incarcerated, any active criminogenic needs that are not being address by detention (e.g., deviant peer influences, antisocial attitudes), may in turn increase likelihood of rearrest (Pratt et al., 2010). So if incarceration is going to make things worse for justice-involved people, then pretrial detention appears to be an effective shortcut to experiencing a host of negative consequences.

In the end, whether our findings revealed a deterrent effect or not, it is worth noting that one of our key outcomes—failure to appear at court processing (FTA)—is a problem that is worth addressing either way (Desmarais et al., 2021). Various strategies for getting people to show up—strategies that are not rooted in a thirst to punish severely those who do not—have shown promise (e.g., text reminder programs; Zottola et al., 2023). The bottom line is that locking people up while their case is being processed—a beloved move of punishment enthusiasts—tends to do more harm than good.

TABLE 4.
Predicting NCA – Binary regression model predicting NCA with time in detention as a binary dummy variable. Limited to those released from pretrial detention and with at least 21 days of time at risk and not more than 365 days of time at risk

Variable	Risk Ratio
Number of charges	**0.982**
Felony charge	1.079
Misdemeanor charge	**1.180**
Against person	0.884
Property	**1.258**
Time at risk	**1.003**
Black	0.958
Male	**1.185**
Days in detention (under 24 hours)	**0.729**
Constant	**0.074**

Bold = p ≤ .001

TABLE 5.
Predicting NCA – Binary regression model predicting NCA with time in detention as an ordinal variable. Limited to those released from pretrial detention and with at least 21 days of time at risk and not more than 365 days of time at risk

Variable	Risk Ratio
Number of charges	**0.975**
Felony charge	1.056
Misdemeanor charge	**1.163**
Against person	0.875
Property	**1.270**
Time at risk	**1.003**
Black	0.958
Male	**1.188**
Days/hours in detention (under 24 hours)	
(0/23=0)	Reference
(24/47=1)	**1.237**
(48/71=2)	**1.289**
(72/95=3)	**1.448**
(96/119=4)	**1.432**
(120/143=5)	**1.459**
(144/167=6)	**1.428**
(168/191=7)	**1.432**
(192/215=8)	**1.484**
(216/239=9)	**1.504**
(240/263=10)	**1.528**
(264/287=11)	**1.581**
(288/311=12)	**1.500**
(312/335=13)	**1.406**
(336/359=14)	**1.407**
(360/383=15)	**1.513**
(384/407=16)	**1.456**
(408/431=17)	**1.518**
(432/455=18)	**1.490**
(456/479=19)	**1.375**
Constant	**0.055**

Bold = p ≤ .001

References

Apel, R. (2013). Sanctions, perceptions, and crime: Implications for criminal deterrence. *Journal of Quantitative Criminology, 29,* 67-101.

Braithwaite, J. (1989). *Crime, shame and reintegration*. Cambridge, UK: Cambridge University Press.

Cullen, F. T., Pratt, T. C., & Turanovic, J. J. (2016). It's hopeless: Beyond zero-tolerance supervision. *Criminology and Public Policy, 15,* 1215-1227.

Cullen, F. T., Pratt, T. C., Turanovic, J. J., & Butler, L. (2018). When bad new arrives: Project HOPE in a post-factual world. *Journal of Contemporary Criminal Justice, 34,* 13-34.

D'Alessio, S. J., & Stolzenberg, L. (1998). Crime, arrests, and pretrial jail incarceration: An examination of the deterrence thesis. *Criminology, 36,* 735-762.

Desmarais, S. L., Zottala, S. A., Clarke, S. E. D., & Lowder, E. M. (2021). Predictive validity of pretrial risk assessments: A systematic review of the literature. *Criminal Justice and Behavior, 48,* 398-420.

Dobbie, W., Goldin, J., & Yang, C. S. (2018). The effects of pretrial detention on conviction, future crime, and employment: Evidence from randomly assigned judges. *American Economic Review, 108,* 201-240.

Donnelly, E. A., & MacDonald, J. M. (2018). The downstream effects of bail and pretrial detention on racial disparities in incarceration. *Journal of Criminal Law and Criminology, 108,* 775-814.

Kim, J., Chauhan, P., Lu, O., Patten, M., & Smith, S. S. (2018). Unpacking pretrial detention: An examination of patterns and predictors of readmissions. *Criminal Justice Policy Review, 29,* 663-687.

Herman, S., & Pogarsky, G. (2022). Morality, deterrability, and offender decision making. *Justice Quarterly, 39,* 1-25.

Jacobs, B. A. (2010). Deterrence and deterrability. *Criminology, 48,* 417-441.

Leslie, E., & Pope, N. G., (2017). The unintended impact of pretrial detention on case outcomes: Evidence from New York City arraignments. *The Journal of Law and Economics, 60,* 529-557.

Lowenkamp, C. T., Van Nostrand, M., & Holsinger, A. (21013). *The hidden costs of pretrial detention*. Laura and John Arnold Foundation. www.arnoldfoundation.org

Martinez, B. P., Petersen, N., & Omori, M. (2020). Time, money, and punishment: Institutional racial-ethnic inequalities in pretrial detention and case outcomes. *Crime and Delinquency, 66,* 837-863.

Maroto, M., & Sykes, B. L. (2020). The varying effects of incarceration, conviction, and arrest on wealth outcomes among young adults. *Social Problems, 67,* 698-718.

Maxson, C. L., Matsuda, K. N., & Hennigan, K. (2011). "Deterrability" among gang and nongang juvenile offenders: Are gang members more (or less) deterrable than other juvenile offenders? *Crime and Delinquency, 57,* 516-543.

McIntyre, F., & Baradaran, S. (2013). Race, prediction, and pretrial detention. *Journal of Empirical Legal Studies, 10,* 741-770.

Menefee, M. R. (2018). The role of bail and pretrial detention in the reproduction of racial inequalities. *Sociology Compass, 12,* e12576.

Ogle, M. R., & Turanovic, J. J. (2019). Is getting tough with low-risk kids a good idea? The effect of failure to appear detention stays on juvenile recidivism. *Criminal Justice Policy Review, 30,* 507-537.

Oleson, J. C., Lowenkamp, C. T., Cadigan, T. P., Van Nostrand, M., & Wooldredge, J. (2016). The effect of pretrial detention on sentencing in two federal districts. *Justice Quarterly, 33,* 1103-1122.

Petersen, N. (2020). Do detainees plead guilty faster? A survival analysis of pretrial detention and the timing of guilty pleas. *Criminal Justice Policy Review, 31,* 1015-1035.

Petrich, D. M., Pratt, T. C., Jonson, C. L., & Cullen, F. T. (2021). Custodial sanctions and reoffending: A meta-analytic review. *Crime and Justice: A Review of Research, 50,* 353-424.

Pratt, T. C. (2019). *Addicted to incarceration: Corrections policy and the politics of misinformation in the United States.* 2nd edition. Thousand Oaks, CA: Sage.

Pratt, T. C., & Cullen, F. T. (2005). Assessing macro-level predictors and theories of crime: A meta-analysis. *Crime and Justice: A Review of Research, 32,* 373-450.

Pratt, T. C., Cullen, F. T., Blevins, K. R., Daigle, L. E., & Madensen, T. D. (2006). The empirical status of deterrence theory: A meta-analysis. Pp. 367-395 in Cullen, F.T., Wright, J. P., & Blevins, K. R. (Eds.), *Taking stock: The status of criminological theory—Advances in criminological theory,* Vol. 15. New Brunswick, NJ: Transaction.

Pratt, T. C., Cullen, F. T., Sellers, C. S., Winfree, L. T., Madensen, T. D., et al. (2010). The empirical status of social learning theory: A meta-analysis. *Justice Quarterly, 27,* 765-802.

Pratt, T. C., & Turanovic, J. J. (2018). Celerity and deterrence. Pp. 187-210 in Nagin, D. S., Cullen, F. T., & Jonson, C. L. (Eds.), *Deterrence, choice, and crime: Contemporary perspectives—Advances in criminological theory,* Vol. 23. New York: Routledge.

Pyrooz, D. C., Garner, N., & Smith, M. (2017). Consequences of incarceration for gang membership: A longitudinal study of serious offenders in Philadelphia and Phoenix. *Criminology, 55,* 273-306.

Raaijmakers, E. A. C., de Keijser, J. W., Nieuwbeerta, P., & Dirkzwager, A. J. E. (2017). Why longer prison terms fail to serve a specific deterrent effect: An empirical assessment on the remembered severity of imprisonment. *Psychology, Crime and Law, 23,* 32-55.

Sacks, M., & Ackerman, A. R. (2014). Bail and sentencing: Does pretrial detention lead to harsher punishment? *Criminal Justice Policy Review, 25,* 59-77.

Sacks, M., Sainato, V. A., & Ackerman, A. R. (2015). Sentenced to pretrial detention: A study of bail decisions and outcomes. *American Journal of Criminal Justice, 40,* 661-681.

Sherman, L. W. (1993). Defiance, deterrence, and irrelevance: A theory or the criminal sanction. *Journal of Research in Crime and Delinquency, 30,* 445-473.

Viljoen, J. L., Jonnson, M. R., Cochrane, D. M., Vargen, L. M., & Vincent, G. M. (2019). Impact of risk assessment instruments on rates of pretrial detention, postconviction placements, and release: A systematic review and meta-analysis. *Law and Human Behavior, 43,* 397-420.

Wakefield, S., & Anderson, L. H. (2020). Pretrial detention and the costs of system overreach for employment and family life. *Sociological Science, 7,* 342-366.

Walker, S. C., & Herting, J. R. (2020). The impact of pretrial juvenile detention on 12-month recidivism: A matched comparison study. *Crime and Delinquency, 66,* 1865-1887.

Williams, M. R. (2003). The effect of pretrial detention on imprisonment decisions. *Criminal Justice Review, 28,* 299-316.

Zottola, S. A., Crozier, W. E., Ariturk, D., & Desmarais, S. L. (2023). Court date reminders reduce nonappearance: A meta-analysis. *Criminology and Public Policy, 22,* 97-123.

Ten Things Risk/Needs Assessment Is Not[1]

Raymond Chip Tafrate
Damon Mitchell
Tom Hogan
Central Connecticut State University
Gina M. Vincent
UMass Chan Medical School
Natalie J. Jones
Public Safety Canada and Carleton University
Guy Bourgon
Private Practice and
Consultant Specializing in Risk Assessment and Forensic Intervention

COMMUNITY CORRECTIONS PROVIDES a host of essential services to the courts, communities, and individuals. The growing emphasis on reducing recidivism has led many agencies to infuse Risk-Need-Responsivity (RNR) principles[2] into their practices, including the adoption of standardized risk/needs assessment instruments (Taxman, 2018). The developers of the RNR model (Andrews & Bonta, 2010; Andrews et al., 1990, 2006; Bonta & Andrews, 2017) emphasized reliable and valid assessment as the cornerstone of effective supervision and case management. Instruments such as the Level of Service/Case Management Inventory (Andrews et al., 2004), Ohio Risk Assessment System (Latessa et al., 2011), Post-Conviction Risk Assessment (Lowenkamp et al., 2013), and many others are designed to identify justice-involved clients (JICs)[3] who are most likely to reoffend and the influences that contribute to that risk. By knowing who and what to focus on, probation officers (POs) can achieve the best possible results with their caseloads by working to alter JICs' criminal trajectories and ultimately improving community safety.

Because the benefits to JICs and the public are so clear, the value of good risk/needs assessment seems self-evident. In actual practice, however, something seems to get lost. Too often, the inclusion of risk/needs assessment has not sufficiently transformed supervision work. Despite the best efforts of all involved, the promised reductions in offending remain elusive. However, decades of research show that—if adhered to—the RNR principles are sound and that improved supervision outcomes are achievable (Bonta et al., 2021; Wormith & Bonta, 2018).

Collectively, the authors of this article have over 100 years of experience working with JICs, training and supervising probation and parole officers, and designing and implementing effective correctional practices. Our own grasp of the RNR model has evolved along the way. So has our understanding of how to improve real-world implementation. The lessons learned have been hard-won. We owe a debt of gratitude to the agencies, managers, and officers we have worked with, as we have learned a great deal from them. We have also been inspired by their dedication to an essential and challenging profession. In our view, the path forward begins with identifying some

[1] This work was partially funded by a Federal Grant Award (#2019-SM-BX-0006, CFDA #16.812) from the U.S. Bureau of Justice Assistance, U.S. Department of Justice, in collaboration with the State of Connecticut Judicial Branch Court Support Services Division. The opinions expressed here are those of the authors and do not necessarily reflect the views of the U.S. Department of Justice or the State of Connecticut Judicial Branch or Public Safety Canada. Correspondence concerning this article should be addressed to Raymond Chip Tafrate, Department of Criminology and Criminal Justice, Central Connecticut State University, 1615 Stanley Street, New Britain, CT 06050. E-mail: Tafrater@ccsu.edu.

[2] RNR principles include directing services to higher risk cases, criminogenic needs become the focus of supervision and community referrals, and CBT methods are used and tailored to clients.

[3] Many labels are used to describe people receiving services in community corrections (e.g., probationer, parolee, offender, juvenile delinquent, justice-involved person, etc.). Throughout this article we use the term *justice-involved client* and in the interest of brevity the abbreviation *JIC*.

of the most common misconceptions about post-conviction risk/needs assessment.[4]

Despite knowing what *should* be done, implementing risk/needs assessment in practice is not straightforward. A set of challenges exist for officers who must conduct assessments with clients who are mandated to be there, while a different set of challenges present themselves to administrators who must determine how best to train officers and use the data provided by assessment instruments to manage cases. Navigating these challenges can result in gaps between theory (what we *should* do) and practice (what we *are* doing or what we realistically *can* do).

In this paper, we first note five misconceptions that we have often observed among officers conducting assessments and offer suggestions for making the assessment process one that yields valid information relevant for supervision and case management. We then proceed to note five issues of likely concern to administrators who must oversee the agency's implementation and training of risk/needs assessment. See if you agree with our top *ten things risk/needs assessment is not*.

1. Risk/needs assessment is not based on one source of information

Leading risk/needs instruments are centered around a one-on-one interview between officer and client. This emphasis on self-report information from JICs has led agencies and POs at times to misinterpret the assessment process as *solely* based on client self-report. At one extreme, POs may conclude that information that conflicts with a JIC's self-report is to be disregarded because scoring the instrument is to be based *only* on self-report. At the other extreme, officers may assume that assessment is not to be trusted because JIC self-report is itself untrustworthy and unreliable. Either extreme is likely to lead POs to question the results and treat the assessment process as a box-checking activity rather than an essential part of case management. While the JICs' self-report is the heart of a criminal risk/needs assessment, the interview should not be the sole source of information for scoring the assessment, and self-report from the JIC should neither trump additional sources of information nor be dismissed out of hand as unreliable.

We recommend viewing client self-report as but *one* important stream of information from which to score the risk/needs instrument and inform decision-making about a case. In fact, valid assessment *requires* corroborating information from sources such as official records, as well as interviews with collateral contacts such as a JIC's intimate partner, family members, teachers, work supervisors, treatment providers, and so forth. POs will more accurately score items on risk/needs assessment instruments and generate case plans that are more targeted and practical when they consider all the best available—and not just the most readily available—information. When a JIC's description of their functioning in a criminal risk domain[5] is vague, or otherwise appears unreliable, additional sources of information should be solicited so that POs do not rely on hunches when scoring and case planning in that risk area.

When information from collateral sources conflicts with JIC self-report, risk/needs assessments rely on POs using good judgment on the credibility of that information and the extent to which it should be integrated into scoring the assessment instrument and formulating their case management plans. For example, if a PO is conducting the portion of an assessment interview focused on education, and their adolescent client reports that "school is going well," and "I don't get into trouble at school," but based on principal and teachers' reports, the PO knows that the youth is failing most classes and is repeatedly suspended, the collateral information should inform scoring and case management and not be disregarded simply because it contradicts self-report.

PO judgment is also important in evaluating the risk information that JICs may not perceive as relevant to their lifestyles. For example, a JIC may genuinely perceive a friend as a positive influence and therefore describe the person as such. Yet, the description of that friendship may indicate this friend supplies the JIC with illegal substances, encourages their use, and "has my back" when the JIC gets into trouble. In this example, the accuracy and usefulness of the resulting risk/needs assessment will rely on the PO's judgment and knowledge of RNR principles to note the companion as a potential risk rather than a positive influence.

While it is inevitable that some JICs will attempt to deceive POs, research suggests that such concerns are overstated; client self-report is largely consistent with criminal information in official records (Daylor et al., 2019; Pollock et al., 2016) and predictive of both general and violent reoffending (Mills et al., 2003). The fact that some clients will distort information to present themselves in a more favorable light, coupled with the possibility that clients can occasionally misperceive their life circumstances, reinforces the need for assessments to be based on multiple sources of information. These realities simply underscore the necessity for assessments to be conducted by POs who are skilled in interviewing and critically evaluating multiple streams of information.

2. Risk/needs assessment is not a conversation about change

Effective probation work is complicated, and POs need to possess a multifaceted knowledge base (e.g., understanding legal concepts, administrative mandates and procedures, RNR principles, and the suitability of local treatment resources). Officers must also be proficient in engaging clients, influencing behavior change, and promoting accountability. To that end, motivational interviewing (MI) skills can be valuable in supervision, and MI has established itself as a communication style that is commonly integrated into community corrections (Stinson & Clark, 2017; Tafrate et al., 2019).

MI skills assist POs in establishing rapport and guiding conversations in productive directions. MI skills also allow POs to explore and heighten JICs' awareness of the connections between lifestyle choices and subsequent losses (e.g., damaged relationships, ruined career paths, financial problems, and incarcerations), fostering motivation around changing activities in risk-relevant life areas.

Similarly, the integration of cognitive behavioral techniques (CBT) into community corrections has shown promising reductions in recidivism and contributed to several probation-oriented CBT curriculums (see STICS, Bonta et al., 2021; EPICS, Smith et al., 2012; STARR, Lowenkamp et al., 2014; Forensic

[4] These guidelines pertain to the use of post-conviction risk and needs assessment instruments to inform decisions and case planning that occur after case disposition; specifically, after conviction and sentencing. They may also be applied to the application of assessment results to inform decision-making and case planning in the context of alternative forms of criminal justice processing.

[5] Although criminal risk/needs are commonly referred to as risk factors, we prefer terms such as *criminal risk domains, criminogenic life areas,* or *risk-relevant life areas*. The term risk factor implies a single characteristic about a person that is linked with a negative outcome (e.g., high cholesterol is a risk factor for heart disease), whereas the attributes assessed on risk/needs instruments usually encompass a JIC's functioning in broad life areas.

CBT, Tafrate et al., 2018). CBT emphasizes conversations that POs can use to help JICs alter problematic thinking and behavior patterns relevant to reoffending.

POs trained in the use of MI and/or CBT might try to overuse these skills during an assessment interview, adopting a "fix as you go" approach that can unnecessarily lengthen and complicate the assessment interview. Well-meaning POs can drift into two common activities: (1) attempting to enhance a JIC's motivation to make changes in criminogenic life areas uncovered during the assessment and (2) prematurely applying active interventions such as recommending behavioral "action steps" to a JIC to alter risky behavior.

Consider this scenario: An officer wants to score an assessment item that measures a JIC's attitude towards offending. Referring to a conviction for stealing a car, the JIC states, "The owner parked it next to the store. It was dark. They left it running with the keys in it. Someone was going to take it, it just happened to be me." The PO immediately reflects, "Your view is you had a right to take the car." The client pauses, then adds, "It's not like that… I wasn't thinking… I knew it was wrong but it just kind of happened and now I'm paying for it. I can't keep doing stupid stuff like that. I tell myself, 'Keep your hands off other people's stuff.'" For some insightful and highly skilled POs, it can then be quite natural to automatically "pursue and reinforce" the JIC's Change Talk (i.e., verbalizations in favor of change). There may also be temptations to transition into "fix-it" mode and make suggestions (e.g., "It seems like you need to stay off of Lyndon Street at night"), recommend intervention referrals (e.g., "You would benefit from some cognitive skills classes"), or challenge criminogenic thinking (e.g., "How does not caring about others sometimes backfire on you?"). However, doing all this while conducting an assessment interview can make the process cumbersome and unfocused.

We recommend that the assessment interview be thought of as an exploration of the JIC's life—a baseline measure of risk and need—that sets the stage for *subsequent* motivational and change-oriented conversations. Certainly, MI skills can be used sparingly to engage JICs and navigate and achieve more depth during the assessment interview. However, adopting a pure MI style during the risk/needs assessment (e.g., having JICs take the lead about what is most important, exploring change goals, and POs strategically evoking change talk) is likely to complicate or even derail the process. Risk/needs assessment is not about enhancing motivation, evoking change talk, or building client skills for changing thinking and behavior. These types of change methodologies should not be integrated into the assessment interview and are best done "downstream" after the instrument is scored and the PO understands *all* the factors that can lead to re-offending. With this information at hand, the PO will be able to identify the most critical criminogenic life areas to work on with a particular client.

Conducting a quality risk/needs assessment does not mean doing skillful supervision or case management. Likewise, being skilled at promoting behavior change does not automatically translate to good assessment. One does not guarantee the other. It is important to distinguish risk/needs assessment from change-oriented endeavors.

3. Risk/needs assessment is not an inquiry about mental health problems

Major mental disorders are common in justice-involved populations (both in prisons and probation/parole), with prevalence rates exceeding those found in non-justice community samples (Brooker et al., 2012; Steadman et al., 2009). In terms of managing JICs on supervision, there is an intuitive appeal in the idea that criminal behavior is a byproduct of psychological distress and that alleviating symptoms will reduce reoffending. In fact, the scientific literature suggests otherwise; mental health symptoms are rarely the main drivers of criminal or violent offending, and focusing on mental health symptoms alone is unlikely to have a significant impact on future criminality (Bolaños et al., 2020; Morgan et al., 2012; Peterson et al., 2014; Skeem et al., 2015).

Screening for mental health problems should be part of the assessment, but not the primary focus of a risk/needs assessment. The interplay between mental health and criminal risk is complex. Mental health difficulties will often contribute to, exacerbate, and even be the "on-ramp" to problems in criminogenic life areas. For example, symptoms related to schizophrenic spectrum difficulties, depression and anxiety, or traumatic experiences can lead to changes in relationships, routines, and habits (e.g., withdrawal from work/school, seeking out substances to self-medicate, increasing the influence of negative peers). In such cases, a common mistake is to assume that, by itself, addressing symptoms will automatically impact entrenched patterns now driving criminal behavior. Making the shift away from a focus on symptoms of psychological distress to JICs' functioning in major criminogenic life areas can be difficult for some POs.[6] After all, messages from the media, the public, and some policymakers often attribute criminal conduct to mental illness, and the symptoms of emotional and psychological disturbance can sometimes be obvious during PO-client interactions. For many POs there is the natural desire to want to fully explore (and alleviate) psychological distress and help JICs feel better. While this goal may be worthwhile, nothing is likely to make JICs' mental health problems worse than getting incarcerated or re-incarcerated. The reality is that if POs' hearts are in the right place, then understanding those influences that put JICs at greatest risk for rearrest must be the top priority.

At the other extreme, it can be a mistake to assume that mental health symptoms are unimportant and should not be identified or addressed. There are certainly some individual cases, albeit the exception to the rule, in which a JIC's mental health symptoms relate directly to the offending behavior. In addition, addressing mental health problems is sometimes necessary to help JICs focus, attend, and participate fully in interventions that target criminogenic life areas. When mental health symptoms are particularly severe, alleviating psychological distress is important so that JICs can be less distracted and more open to working on the risk-relevant areas of their lives (e.g., maintaining employment, improving family relationships, avoiding contact with criminal peers). In this sense, mental health symptoms can be viewed as *responsivity factors* that can interfere with working on larger life areas (Bonta & Andrews, 2017).

We recommend POs prioritize the identification and exploration of criminogenic life areas rather than symptoms of psychological distress. Screening for and making referrals for mental health problems is an important adjunct to, but not a replacement for, identifying and addressing the criminogenic domains of a JIC's life.

Given that many JICs will have mental health problems, it is not uncommon for items

[6] Other practitioners such as case managers also struggle with this issue. Several of us are clinical psychologists and for years our main way of approaching cases centered on a symptom-based approach. When we began working with JICs, it took a long time to make the transition from a *symptom-based* to a *risk-based* mindset.

related to depression, anxiety, trauma, and serious mental illness to appear on risk/needs assessments. The purpose of these items is to screen for clients with mental health needs to make potential referrals. Practically, POs will have to recognize and screen for mental health disorders, locate community resources, and make appropriate treatment referrals when necessary. Additionally, understanding how symptoms may be influencing behavior in risk-relevant life areas is important for case planning purposes and can inform decisions about recommending some form of mental health treatment. However, POs should not be expected to drift too far into the roles of psychologists, social workers, or therapists. Being able to screen for mental health problems and understand how such symptoms are influencing supervision is enough for officers to tackle.

4. Risk/needs assessment is not pessimistic (or "negative labelling")

Let's face it—much of what occurs in criminal justice settings is negative. JICs sometimes commit crimes that challenge our sense of right and wrong in the world and violate the safety of people in their communities. They also might resist supervision and intervention that is mandated to them by the courts. Moreover, to many, risk models seem negatively keyed.

It can be common for POs and case managers to be reluctant to embrace a *risk reduction* perspective with JICs because it is viewed as intrinsically pessimistic and deficit-driven. Working from a risk framework may seem like emphasizing client failures and problems, and the assessment process simply a matter of adding up the number of risk factors a client has amassed. We have heard POs describe scoring a risk/needs assessment item as "dinging" a JIC. In this view, endorsing an item is perceived as a moral judgment against the client and a potential punishment.

The reality is that risk-based models, when used effectively, offer a constructive perspective on client functioning. A risk/needs assessment is more than just a checklist of JICs' shortcomings. A shift towards a different, more multifaceted way of thinking about risk/needs is often required to make these models useful. *We recommend* that officers frame risk/needs assessment as a more optimistic endeavor than it appears on the surface. RNR principles actually provide an *optimistic* view of JIC functioning because the criminogenic life areas measured are largely changeable. Most risk areas have dynamic components (the one exception being criminal history). Therefore, JICs' future criminality can be prevented if supervision assesses and addresses their functioning within the unique constellation of risk domains relevant to their lives.

It is also important to appreciate the interrelationships between criminal risk domains and the unique ways these may manifest across individuals. Given the multidetermined nature of human behavior, risk domains are often connected in ways that can be complex in how they amplify or reduce each other (Mitchell et al., 2018). For example, consider a JIC who spends the better part of his considerable free time with friends who drink heavily, smoke pot, and steal cars. His friends reinforce his unproductive beliefs about pursuing school or meaningful employment, and his cannabis use worsens his school performance or potential marketability as an employee. His friends also reinforce his excitement-seeking attitudes, his sense that he doesn't fit in with conventional school or work activities, and encourage increasing levels of impulsive behavior. Thus, the various criminogenic life areas relevant to forensic cases often influence each other in an interrelated and destructive manner. Due to such synergistic effects, the influence of risk domains can be multiplicative rather than simply additive. On the optimistic side, because risk domains are interconnected, a positive change in one of these life domains can facilitate positive changes in others. A strategic focus on one or two criminogenic life areas can often create a positive ripple effect in a JIC's life.

Framing risk/needs assessment as a thoughtful, strategic, and optimistic endeavor —requiring officer curiosity and expertise—prevents assessment from becoming an unchangeable, pessimistic, or condemning conclusion about a JIC. Having an in-depth understanding of JICs' functioning in these important life areas, rather than a list of factors to be hastily checked off a list, is at the heart of a practical, solution-focused, and individually tailored supervision strategy.

5. Risk/needs assessment is not an unstructured search for strengths

Consideration of JICs' strengths—although sometimes factored into program delivery to establish rapport with clients and improve the effectiveness of intervention—is typically not built into risk/needs assessments. In our experience, the way POs typically use strengths in case planning, if at all, is to leverage them to increase a JIC's interest in a particular intervention or activity (e.g., flex funds might be used to have a JIC attend music lessons at a community center or join a YMCA sports league). In other words, strengths are not commonly measured and incorporated into the calculation of a JIC's risk to reoffend.

Although criminal justice assessments tend to focus exclusively on empirically established risks and needs, some JICs—even those with high scores across multiple risk domains—manage to overcome these issues and ultimately become law-abiding or even thriving members of society. This leads us to borrow from the literature on *resilience* that originates from other fields, in which strengths-based factors serve to protect otherwise high-risk individuals against the onset or development of negative outcomes (Farrington, 2003; Masten, 2016; Seligman, 2002). Similarly, research on criminal desistance supports the identification of strengths that are associated with a JIC's transition out of crime (Burnett & McNeill, 2005; Maruna, 2001; Serin & Lloyd, 2009; 2017). This body of research recognizes that strengths can be present in the form of external circumstances (e.g., prosocial bonds with friends, family, or mentors), personal qualities (e.g., academic ability), or internal processes (e.g., optimism regarding one's ability to lead a prosocial life).

However, the research on strengths and their application in practice lags far behind research on criminal risk domains. There remains confusion and a lack of consensus on how to refer to, define, measure, and use information about strengths. It is important to note that a strength is not simply the opposite of, or absence of, a risk. Strengths that have value in the assessment process are those that (1) have a negative association with criminal conduct (a promotive factor), and/or (2) exert a buffering effect on overall risk level—typically in a higher risk group (a protective factor) (Wanamaker et al., 2018). In fact, strengths can exist concurrently with risk domains and can account for differences in criminal outcomes in JICs with otherwise comparable risk profiles. For example, consider two JICs who have the same number of close antisocial peers. One of the JICs also has several close prosocial peers, while the other does not. It is easy to see how the difference in these two cases can influence their supervision outcomes. To that point, there is preliminary yet compelling evidence that including strengths information—in a

structured way using validated assessment instruments—improves the prediction of recidivism *over and above* risk domains alone (Brown et al., 2020; Burghart et al., 2022; Jones et al., 2015; 2016).

Although more research around strengths is needed, what we do know is that, just as not all deficits are criminogenic in nature, not all strengths contribute to lowering risk of recidivism (e.g., "JIC has a nice smile," "…is a smooth talker," "…is street smart."). There are several validated instruments with integrated strengths domains, as well as several structured, validated strengths-based instruments tailored to JICs designed to supplement risk/needs assessment (Wanamaker et al., 2018).[7]

We recommend that officers (and agencies) interested in integrating strengths into assessments adopt a validated instrument and avoid relying on intuition or unstructured judgment. Caution is warranted against adding strengths to the assessment process in an unguided, unsystematic way, without proper training. The haphazard integration of strengths into an assessment can lead to potential errors in estimating a JIC's risk to reoffend and ultimately to inappropriate classification decisions.

Despite recent research demonstrating the incremental value of strengths in the prediction of recidivism, the practical application of strengths in case planning is still emerging. *We recommend* that officers (and agencies), interested in integrating strengths into case planning look to balance both traditional *avoidance* goals with *approach* goals (Tafrate et al., 2018). In this balanced strategy, case planning does not solely rely on JICs avoiding criminal peers or places where illegal substances might be available. Rather, approach goals assist JICs in pursuing valued life directions in ways that combat the riskiest aspects of their lives. Examples include having success in school or work, acquiring and spending time with prosocial friends, and developing more positive family or intimate relationships. The two approaches (reducing risk and enhancing strengths) can be complementary and applied concurrently. As a final caution, although the inclusion of strengths adds value to assessment and case planning, a focus on strengths should not be in lieu of activities that focus on risk reduction.

6. Risk/needs assessment is not a box-checking exercise

Jurisdictions are increasingly pressured into incorporating risk/needs assessment into their workflow, and such initiatives have been *rolled out* across the country. In some agency cultures, "rolling it out" can devolve into a process of checking boxes and recording information in the case notes, with little appreciation of how to use the assessment to guide supervision, referral, and intervention practices. In its most extreme version, the culture around risk/needs assessment is swathed in a veneer of dismissive resentment. The implication is that assessment is merely a "data collection" mechanism for the "bean counters" that is burdensome and distracting to those tasked with doing the "real work" of community supervision.[8] In a similar but less emotionally charged culture, the view is that "policy" requires assessments to be completed, and this task is primarily approached as a bureaucratic duty. In our experience, a box-checking culture is problematic as it degrades the quality of assessments and the degree to which the results guide supervision practices. The culture surrounding risk/needs assessment impacts how it gets "rolled out," administered, and used.

In a box-checking culture, risk/needs assessment is often trained in a void, with little time devoted to case planning concerns such as discussing the complexity of how individual risk domains play out in the real-world lives of JICs. Training POs to use assessment instruments before they are educated about the nature of criminogenic life areas is a cart-before-the-horse situation. Knowledge of the various ways risk domains present themselves in JICs' lives, and how those domains influence each other (as noted earlier), is foundational for conducting effective case management. Administrators may optimistically assume that effective case planning skills will automatically emerge by training POs in scoring risk/needs assessment instruments; but that rarely happens in a box-checking culture.

The following are some hallmarks of box-checking cultural practices: (1) conducting assessments in settings that lack privacy; (2) making few, or no, attempts to reduce noise and office distractions (e.g., phone calls, intercom announcements, interruptions by colleagues); (3) entering client information into the data management system during the interview, resulting in the PO paying more attention to the computer screen than the client; and (4) over-emphasizing timely completion and electronic filing of results, resulting in pressure on POs to cut corners to avoid criticism or discipline. These types of cultural practices surrounding the administration of risk/needs assessment can make the process feel more like an assembly line than an in-depth, thoughtful human service interaction, can reduce JICs' willingness to disclose sensitive personal information, and inadvertently create a culture that emphasizes the quantity and timeliness of assessments over their quality.

In terms of utilization, the results of risk/needs assessments will occupy a minor role in guiding supervision practices in a box-checking culture. Instead, probation staff will emphasize compliance with conditions over improvements in risk-relevant life areas during supervision. This focus creates an additional problem, because greater emphasis on conditions than on criminogenic risk has been associated with greater recidivism (Andrews & Bonta, 2010). Further, an excessive focus by POs on supervision conditions may tend to underestimate sources of risk in clients' lifestyles (e.g., "I've supervised this guy lots of times before. He's not a knucklehead—he shows up, he does his programs. Why do I even need to know who he hangs out with when he leaves my office?").

We recommend that agencies work to create an office culture that supports quality risk/needs assessments and the principles of RNR. Office culture can be thought of as the values, assumptions, tacit agreements, understandings, and ways of thinking and behaving that are shared by the members of an office and that are taught to new members. Because supervisors play an integral role in setting the cultural tone, agencies can improve the culture around assessment by hiring, promoting, and developing supervisors with a strong understanding and buy-in of RNR concepts. One of the most important, yet unappreciated, responsibilities of a supervisor is managing the office culture. A culture around assessment will develop whether a supervisor makes an effort to influence it or not. If not established around RNR principles, a box-checking culture is likely to fill that void.

[7] Examples include but are not limited to the Service Planning Instrument (SPIn; Wanamaker, 2003); the Inventory of Offender Risk, Needs, and Strengths (IORNS; Miller, 2006); and the Structured Assessment of Protective Factors for Violence Risk (SAPROF; de Vogel et al., 2009). For further information about these and other instruments that measure strengths, please consult Wanamaker et al. (2018).

[8] For the record, this perspective is not so different from that once held by one of the authors of this paper who was a probation officer and later became a probation supervisor.

POs will excel in skills that supervisors actively reinforce. Behaviors and skills not supported or modeled are less likely to be used by officers in daily interactions. Supervisors should be encouraged to discuss cases with POs by using risk/needs assessment results as a starting point in case reviews (e.g., begin case reviews with an overview of the relevant risk domains).

Supervisors can also emphasize case planning that targets changes in risk-relevant areas over superficial compliance with conditions. Finally, supervisors can monitor and reward both the quality *and* timeliness of assessments. Supervisors do not need to be experts at administering the instruments themselves; rather, they need to know good work when they see and hear it, be able to provide accurate feedback on areas in an assessment interview that were overlooked, and connect the dots between assessment, case planning, and subsequent supervision contacts. Will POs still have to check boxes when they conduct assessments? Of course. But in strong office cultures, they no longer disparage the process as *just* a box-checking exercise.

7. Risk/needs assessment is not implemented through a "drive-by" training

While the constraints of budgets, staffing, and time favor "one and done" training events, proficiency in risk/needs assessment is not achieved through a single training workshop. Such an approach paves the way for poor implementation of evidence-based practices. In turn, poor implementation of an instrument can lead to its being disregarded in supervision, and equally troubling, to inaccurate assessment results that do not aid in predicting risk or identifying needs (Vincent et al., 2016). All these outcomes undermine the ultimate purpose of implementing risk/needs assessment in the first place.

Alexander (2011) noted that effectively implementing evidence-based interventions in community corrections requires attending to fidelity (i.e., making sure the intervention is delivered as intended), devoting sufficient organizational resources to thoroughly train the staff who will be delivering the intervention, and providing follow-up coaching so that staff become proficient and sustain their proficiency over time. These overarching points about fidelity, proficiency, and sustainability are relevant to the implementation of risk/needs assessment.

With respect to fidelity in risk/needs assessment, *we recommend* POs receive education in the nature of RNR principles *before* they are trained in the intricacies of administering and scoring the instrument itself. Risk/needs assessment instruments are designed to quantify elements of JICs' functioning in broad-life areas linked with reoffending. However, for risk/needs assessment to meaningfully inform supervision and case management, officers must understand qualitatively how a JIC's life functioning in any given area is linked to that individual's offending behavior. For example, knowing that a client is unemployed is useful for predicting the likelihood of reoffending. However, understanding the nature of the client's employment problems from skills—to attitude—to history is going to drive case management decisions and the focus of supervision conversations. Training POs to administer a risk/needs instrument without first establishing a foundational understanding of the nature of criminogenic life areas is akin to medical students learning to administer physical exams without first understanding anatomy and physiology. A strong grounding in RNR principles sets the stage for effective case planning and supervision strategy.

With respect to proficiency and sustainability, *we recommend* equipping POs with communication skills (such as motivational interviewing; with the caveat noted in item #2) that enable them to develop rapport, encourage greater client disclosure of relevant information, and guide the pace and structure of the assessment interview. It is unrealistic to assume that communication skills will be acquired and implemented in a single training. Indeed, research indicates that follow-up or refresher training is needed for sustained, effective application of these skills (Alexander et al., 2013; Lowenkamp et al., 2012).

Having POs record their assessments and receive feedback and coaching helps them improve their performance and become proficient (Ferguson, 2002). POs often find reviewing office visit recordings intimidating at first, but later view the process as essential to improving their skills and using them routinely (Alexander et al., 2013). Recent national guidelines for risk/needs assessment recommend that POs complete a minimum of three practice cases and have them reviewed to help ensure inter-rater reliability (Desmarais et al., 2022). After POs have achieved a degree of proficiency in administering an assessment instrument, routine quality assurance or follow-up with a coach can help with reliability and prevent the development of idiosyncratic scoring errors and other signs of "drift" that lead to unreliability over time. Implementing and sustaining risk/needs assessment often means developing agency capacity to provide coaching and constructive feedback on recorded work samples.

8. Risk/needs assessment is not intrinsically biased against BIPOC (Black, Indigenous, and people of color)

There is a long-standing problem of systemic injustices experienced by certain racial/ethnic groups. That Black individuals are arrested and convicted for crime at a rate considerably higher than Whites, for example, is indisputable (e.g., Abrams et al., 2021; Hockenberry & Puzzanchera, 2020; Kim & Kiesel, 2018; Schleiden et al., 2020). At this point, most justice administrators have likely heard the recent sentiments that risk/needs assessments produce racist algorithms that merely exacerbate these long-standing racial disparities in the justice system.

The fact is, a primary intention of the development of risk/needs assessment instruments was to promote objective and accurate case management decisions, thereby reducing racial and ethnic disparities. Identifying factors known to be predictive of a negative outcome and using those factors to guide decisions *reduces* subjective and biased decision-making that is otherwise based on hunches (Vincent & Viljoen, 2020).

There is increasing evidence that when courts *do not follow* valid risk/needs assessment results provided by their probation offices, their decisions lead to racial disparities among incarcerated JICs, whereas reliance on risk/needs instruments would have eliminated those disparities (Lehmann et al., 2020; Marlowe et al., 2020). Moreover, there is evidence that use of a valid risk/needs assessment instrument before sentencing may significantly reduce racial disparities in diversion decisions (Onifade et al., 2019). Currently, there is no evidence from credible studies that use of risk/needs assessment has a disparate impact on BIPOC (Viljoen et al., 2019). In other words, there is no evidence that once a risk/needs assessment instrument is adopted for diversion, incarceration, or other such decisions, it increases rates of incarceration or other negative outcomes for BIPOC relative to their White counterparts. Indeed, to date, quite the opposite appears to be true.

Currently, racial bias has been detected in very few risk/needs assessment instruments

to date[9] and sometimes falls in the direction of working against White defendants as opposed to against BIPOC defendants (Vincent & Viljoen, 2020). If an instrument were racially biased, it would mean the instrument functioned differently for one racial/ethnic group than for another racial/ethnic group (see joint statement from the American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 2014). For example, one way in which a risk/needs instrument would be racially biased is if BIPOC individuals scored higher on the instrument, on average, than their White counterparts and yet were *not* more likely to recidivate than their White counterparts. Scores need to mean the same thing for different groups of individuals.

We—and recent national guidelines (Desmarais et al., 2022)—*recommend* that justice agencies make sure the assessment instrument they are using is tested statistically to determine whether it contains racial bias. Moreover, *we recommend* that justice agencies favor the use of instruments with items that consider a combination of JICs' actual self-reported violent and illegal behaviors and official criminal records as opposed to items based entirely on counts of events from official records, such as number of prior felonies or prior probation violations (Miller et al., 2021; Skeem & Lowenkamp, 2020; Vincent & Viljoen, 2020). We already know BIPOC individuals are significantly more likely to be apprehended than White individuals. The racial disparities in official records are far greater than the disparities in self-reported offending and violence between these groups (Loeber et al., 2015). In other words, official records can contain disparities that do not equate to JICs' *actual behavior*, making the "inputs" to the risk instrument potentially biased. In addition to the problem with the inputs, a highly sensitive risk instrument designed to predict recidivism in a particular jurisdiction where BIPOC individuals are more likely to be arrested is going to produce higher scores for BIPOC individuals because it is doing its job: predicting who will be rearrested (Skeem & Lowenkamp, 2020). In effect, in these situations, the risk instrument would not be the problem. Instead, it would be shining a light on long-standing systemic issues (Vincent & Viljoen, 2020).

[9] Racial bias may be more common in pretrial risk tools (Desmarais et al., 2021).

After justice agencies test the risk/needs instrument to ensure it is free of racial bias, w*e recommend* that they be sure to make their staff and court partners aware of the validity of their risk/needs assessment instrument for use with BIPOC. This will promote acceptance of the instrument in decision-making, which, as noted, may be essential for reducing disparities.

Finally, algorithms or structured decisions are transparent, objective, adjustable, and easily regulated. Unstructured human decision-making is not (Mullainathan, 2019). The psychology literature on implicit bias suggests that people can discriminate without intentional awareness (Gran-Ruaz et al., 2022; Greenwald & Banaji, 1995). For example, let's assume racial disparities appear to be emerging in a particular probation office where unstructured decision-making is the norm for assigning risk levels to JICs. If you were to ask POs in this office what is leading to the problem, you would get different answers (and probably defensiveness). Attempts at identifying the underlying mechanism(s) driving the disparate decisions would be unsuccessful due to the idiosyncratic nature of individual officers' attitudes, and because people often cannot explain the reasoning behind their hunches. In addition, rectifying biased decision-making in this type of scenario is complicated—usually attempted through training endeavors to combat prejudicial attitudes (Kim & Roberson, 2022; Paluck et al., 2021). On the other hand, for a validated risk/needs assessment instrument, if a problem is identified, it is rather straightforward to adjust the decision-making algorithm. An algorithmic approach makes it much easier for agencies to detect and fix problems. Removal of risk/needs instruments simply means going back to the opaque human decision-making processes that led to the grave injustices in the first place.

9. Risk/needs assessment is not to be overridden lightly

Sometimes, criteria other than risk/needs assessment scores will guide supervision practices. Professional overrides occur when formal assessment is completed and the assessed risk level of a JIC is then changed. The resulting decision-making, supervision activities, and services provided are then made congruent with the new overridden risk level, rather than the originally assessed risk level. Overrides generally occur at two different points: (1) when the PO conducting the assessment decides to override the results, and (2) when the agency's policies dictate that a specific group of JICs are mandated to predetermined risk/supervision levels.

At the officer stage, unofficial or outside information that is *not incorporated* into the risk/needs assessment may be used to justify changing the assessed risk level of a JIC. In these cases, the primary intent of the override is to provide a more accurate assessment of risk, and by extension, more appropriate levels of service and supervision.

At the organizational stage, policies or mandates may direct POs to automatically change a JIC's assessed risk level based on a pre-determined criterion. This is most often found in policies regarding JICs convicted of sexual, domestic violence, or other serious violent offenses. In almost every case, policies override the assessed risk into higher or even the highest risk level. One of the primary intents is to ensure that these JICs are intensely supervised in the community to protect the agency from controversy should there be a re-offense (i.e., "CYA").

Research has examined whether professional overrides increase or decrease the predictive accuracy of risk/needs assessment instruments. Testing the impact of professional overrides has clearly shown that the predictive accuracy of risk/needs instruments deteriorates when there is an override of the assessed level of risk. This decline occurs irrespective of the override being the decision of individual POs or organizational policy (Cohen et al., 2020; Wormith & Bonta, 2021). Unfortunately, the use of overrides has increased over the years (Wormith & Bonta, 2021).

Although there are times when overriding the assessed risk level makes sense, the impact of overrides can have serious and significant detrimental effects to an organization and the supervision of JICs. Experience over the past 20-plus years shows that resources for community supervision agencies are stretched and limited. The override of assessed risk levels results in a higher proportion of cases supervised as high-risk, placing greater demands on staff and resources. JICs end up being supervised more closely than their assessment indicates they should be, and such cases occupy services and officer time that could otherwise be spent on JICs that are truly high risk. This practice can also be of harm to lower risk clients who end up becoming more entrenched in the system (e.g., more conditions interfering with employment; increased contact with higher risk JICs).

Another unforeseen consequence of routine overrides is their impact on POs' perception of the value of the risk/needs assessment process itself (e.g., "What is the point of conducting a quality assessment when policy, not the assessment results, will direct my day-to-day supervision of the case?"). If JICs are all fated to receive the same level of supervision, then officers will attach little value to the process (e.g., "Sex offenders go to S.O. supervision, DV guys go to D.V. treatment—what's the difference how they score?"). High levels of policy-directed overrides facilitate mistrust of the usefulness of assessment in general. The agency's culture will tend to breed a dismissiveness of risk/needs assessment and RNR principles, tainting the work of well-trained officers and hindering efforts to implement these and other evidence-based practices.

We recommend that agencies track and approve PO-initiated overrides and the rationale behind them. The expectation would be that overrides should occur in less than 10 percent of cases (ideally less than 5 percent), and should be distributed across all risk levels, offense types, and other representative demographics. This will ensure that no inherent biases (e.g., cultural, racial, offense-driven, etc.) are fueling overrides. Tracking the rationales for overrides permits POs and supervisors to discuss the accuracy of scoring, reasoning behind decisions, how overrides affect the probability of reoffending (i.e., risk principle), and, if the override results in a higher risk classification, how this will be addressed in supervision. In this manner, both officer and supervisor can determine whether there is consensus on the appropriateness of the override and broaden their understanding and application of RNR principles. It might also be helpful to let actual experience with a specific case guide an override decision. For example, if, after a few months of supervision, the PO believes a JIC to be at a different level of risk due to a factor that was not part of the initial risk/needs assessment, the PO can then request an override with a better-informed rationale and information at that time. With more observable data available, it is likely that the rationale for an override would be more evidence-based and justified.

Deviations from assessed risk level are most empirically defensible during the most at-risk period for recidivism. Virtually all research on recidivism patterns (i.e., survival curves), regardless of risk level, shows that the first year under supervision is the time when JICs are most at risk for reoffending. Incorporating this evidence into policy is certainly more defensible than policies for niche groups of JICs. For example, a policy could direct that all JICs with violent (including sexual) offenses will begin their period of supervision under close observation for the first 6 months, due the harm they have caused. This 6-month period would typically allow sufficient time to fully assess the JIC with one or more risk/needs assessment instruments, develop a stronger PO-JIC relationship, and assess the JIC's initial response to supervision. At the 6-month mark, these JICs could automatically be placed on their assessed level of risk/supervision unless the PO can justify an override.

Finally, *we recommend* that correctional agencies incorporate offense-specific assessments. There are specific risk/need domains for sex offending and domestic violence cases that are not directly assessed by general risk/need instruments (see Wormith et al., 2020 for a review of violence risk assessments).[10] Adding specialty risk assessments to the protocol will require additional training and resources. However, the benefits include sound, empirically defensible decision-making, policy development, evidence-based practice, and less time spent supervising cases that unnecessarily result from overrides into higher risk levels.

10. Risk/needs assessment is not a crystal ball

Leading criminal risk/needs instruments have a moderate to high degree of accuracy in predicting reoffending, which may vary based on offender characteristics and settings (Desmarais et al., 2016). A risk/needs assessment result that classifies a JIC as low risk is not a guarantee that the client will desist from offending, just as a result that classifies a JIC as high risk is not a guarantee that the client will reoffend. In fact, sometimes low-risk JICs recidivate and high-risk JICs do not. Yet, this does not mean that criminal risk/needs instruments are generally inaccurate. In making sense of the discrepancy that can occur between a JIC's risk assessment result and outcome, it's helpful to keep in mind that (1) the intent of risk assessment is probabilistic rather than deterministic, (2) criminal risk is dynamic rather than stable and is therefore subject to change—sometimes rapidly if JICs find themselves in particular situations that are unpredictable—and therefore needs to be reassessed periodically, and (3) the purpose of risk/needs assessment is to put strategies in place that will *prevent* reoffending—meaning, JICs identified as high risk will not reoffend because the strategy was successful.

The probabilistic versus deterministic interpretation of criminal risk/needs assessment results is in many ways analogous to a screening for heart disease. Both criminal risk assessment and heart disease screenings are based on examining the client's status on a series of risk factors (or domains). The more risk factor/domains, the higher the person's likelihood of reoffending or, in the alternative scenario, developing heart disease. Those found to be at high-risk for reoffending/heart disease are not destined to reoffending or heart disease; they are just more likely to experience these negative outcomes than those with lower risk scores. Similarly, someone with only a few risk factors/domains for reoffending/heart disease may also experience these negative outcomes, but are less likely to do so than their higher-risk counterparts. Lower-risk individuals may sometimes find themselves in unpredictable situations that trigger or prompt heart-attacks/criminal offenses.

The intent of criminal risk assessment is probabilistic in nature, not diagnostic. Criminal risk assessments are not designed for, nor do they attempt to establish, dichotomous groups (e.g., will offend/won't offend) (Andrews et al., 1990). Rather, they typically group JICs into a three-, four-, or five-tiered system (depending upon the instrument), with each succeeding tier reflecting an increased likelihood of reoffending (Kroner et al., 2020). Using a tiered system, rather than a binary system, allows agencies to modulate supervision intensity more efficiently for different JIC groups.

The analogy between criminal risk/needs assessments and screening for heart disease is also relevant to understanding the dynamic nature of criminal risk and its impact on assessment results. Some of the leading risk factors for heart disease are largely static (e.g., gender, age, family history), while others are dynamic (e.g., smoking, sedentary lifestyle, blood pressure). While the static

[10] Instrument such as the STATIC-99-R (Hanson & Anderson, 2021) can be utilized for sex offenders; and the *Ontario Domestic Assault Risk Assessment* (ODARA; Hilton, 2021) and *Spousal Assault Risk Assessment* (SARA; Kropp & Gibas, 2021) for domestic violence cases. Readers looking for more information regarding training on specialty instruments are referred to *The Society for the Advancement of Actuarial Risk/Need Assessment (SAARNA) https://saarna.org/about/*

risk factors cannot change and will therefore always impact risk, the dynamic factors can be altered in ways that increase or reduce risk. If, for example, a client becomes more sedentary and doubles their cigarette intake, their risk for heart disease will increase. Conversely, the introduction of exercise and healthy eating will lower risk. In the same way, some of the leading risk factors/domains for reoffending are static (e.g., prior criminal history), whereas most are dynamic (e.g., presence of criminal peers, substance misuse, employment instability, criminogenic attitudes). Therefore, risk of recidivism can increase or decrease in the months after a criminal risk/needs assessment is conducted, based on fluctuation in a JIC's dynamic risk factors.

The dynamic nature of criminal risk means that JICs who are high-risk can become low-risk and vice versa during a period of community supervision. In this regard, POs who successfully address the needs of their high-risk JICs may have a large proportion of high-risk cases that do not recidivate. However, officers who do not address the dynamic needs of their lower or medium-risk JICs may see more reoffending in these lower risk groups than expected. The dynamic nature of criminal risk means it should be reassessed periodically over the course of supervision. As JICs' criminal risk increases or decreases, supervision intensity can be modulated as needed to allocate resources efficiently and produce better outcomes.

We recommend that agencies conduct a local validation of their risk/needs assessment instruments to ensure that instruments are adequately predicting recidivism. Important benchmarks include the following: (1) clients classified at lower tiers of risk reoffend at lower rates than those classified at higher tiers of risk, (2) the overall predictive validity of the instrument minimally achieves a 65 percent degree of discrimination between those who recidivate and those that do not, and (3) observed agreement reliability among assessors is at least 80 percent (Desmarais et al., 2022). Depending on the risk/needs instrument adopted, it may also be important to periodically reexamine the accuracy of risk/needs instruments and recalibrate cutoff scores (i.e., scores defining the different risk levels) as needed. See the recent guidelines published by the Council of State Governments on *Advancing Fairness and Transparency* for more in-depth recommendations on the validation process (Desmarais et al., 2022). Despite the best attempts of agencies to maximize predictive accuracy, predicting human behavior is complex, and there will always be error in risk/needs assessments and unpredictable circumstances. But rest assured that the proper use of these instruments produces results that are more accurate than unstructured professional judgment.

Conclusion

If you have read this far you undoubtedly have a strong interest in the most effective use of criminal risk/needs assessment in community corrections. You probably are already familiar with the use of one or more risk/needs assessment instruments and have faced your own challenges in integrating assessment into real-world practice. Perhaps you have experienced some of the stumbling blocks we've described. If so, we hope our discussion has been useful. It can be reassuring to know that these issues are common.

We have shared some hints for moving forward, but if you have found better ways of addressing these issues—or sidestepping some of these problems altogether—we would like to hear from you. One thing we have learned is that there are many creative and eager people, working on the frontlines of our field, whose passion and experience can benefit the entire profession. Also, the community corrections field is dynamic. Best practices in assessment and supervision are not set in stone and will surely be influenced in the coming years by further developments in theory and research.

If there is one overarching theme from this article, we think it is this: risk/needs assessment is the cornerstone of effective community corrections, and optimizing the use of risk/needs assessment takes time and persistent attention. Like all science-based methods, it requires careful implementation and periodic recalibration. It is doable. And it is worth doing.

References

Abrams, L. S., Mizel, M. L., & Barnert, E. S. (2021). The criminalization of young children and overrepresentation of Black youth in the juvenile justice system. *Race and Social Problems*, *13*(1), 73–84. https://doi-org.ccsu.idm.oclc.org/10.1007/s12552-021-09314-7

Alexander, M. (2011). Applying implementation research to improve community corrections: Making sure that "new" thing sticks! *Federal Probation*, *75*(2), 47–51.

Alexander, M., Palombo, L., Cameron, E., Wooten, E., White, M., Casey, M., & Bersch, C. (2013). Coaching: The true path to proficiency, from an officer's perspective. *Federal Probation*, *77*(2), 64–68.

American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational and Psychological Testing (U.S.). (2014). *Standards for Educational and Psychological Testing*. Washington, DC: AERA.

Andrews, D. A., & Bonta, J. (2010). Rehabilitating criminal justice policy and practice. *Psychology, Public Policy, and the Law, 16*, 39-55. doi: 10.1037/a0018362

Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, *17*(1), 19–52. https://doi.org/10.1177/0093854890017001004

Andrews, D. A., Bonta, J., Wormith, S. J. (2004). *Level of Service/Case Management Inventory: An offender assessment system*. Multi-Health Systems.

Andrews, D. A., Bonta, J., & Wormith, J. S. (2006). The recent past and near future of risk and/or need assessment. *Crime and Delinquency*, *52*(1), 7–27. https://doi.org/10.1177/0011128705281756

Bolaños, A. D., Mitchell, S. M., Morgan, R. D., & Grabowski, K. E. (2020). A comparison of criminogenic risk factors and psychiatric symptomatology between psychiatric inpatients with and without criminal justice involvement. *Law and Human Behavior.* Advance online publication. https://doi:10.1037/lhb0000391

Bonta, J., & Andrews, D. A. (2017). *The psychology of criminal conduct* (6th ed.). London, UK: Routledge.

Bonta, J., Bourgon, G., Rugge, T., Pedneault, C. I., & Lee, S. C. (2021). A system-wide implementation and evaluation of the Strategic Training Initiative in Community Supervision (STICS). *Journal of Criminal Justice*, *74*, N.PAG. https://doi-org.ccsu.idm.oclc.org/10.1016/j.jcrimjus.2021.101816

Brooker, C., Sirdifield, C., Blizard, R., Denney, D., & Pluck, G. (2012). Probation and mental illness. *Journal of Forensic Psychiatry & Psychology, 23*(4), 522-537. http://dx.doi.org/10.1080/14789949.2012.704640

Brown, S. L., Robinson, D., Wanamaker, K. A., & Wagstaff, M. (2020). Strengths matter: Evidence from five separate cohorts of justice-involved youth and adults across North America. *Criminal Justice and Behavior, 47*(11), 1428–1447. https://doi.org/10.1177/0093854820931549

Burghart, M., de Ruiter, C., Hynes, S. E., Krishnan, N., Levtova, Y., & Uyar, A. (2022). The structured assessment of protective factors for violence risk (SAPROF): A meta-analysis of its predictive and incremental valid-

ity. *Psychological Assessment*. https://doi-org.ccsu.idm.oclc.org/10.1037/pas0001184.supp (Supplemental)

Cohen, T. H., Lowenkamp, C. T., Bechtel, K., & Flores, A. W. (2020). Risk assessment overrides: Shuffling the risk deck without any improvements in prediction. *Criminal Justice and Behavior, 47(12)*, 1609-1629.

Daylor, J. M., Blalock, D. V., Davis, T., Klauberg, W. X., Stuewig, J., & Tangney, J. P. (2019). Who tells the truth? Former inmates' self-reported arrests vs official records. *Journal of Criminal Justice*, *63*, 49–57. https://doi-org.ccsu.idm.oclc.org/10.1016/j.jcrimjus.2019.04.002

Desmarais, S. L., D'Amora, D. A., & Tavárez, L. P. (2022). *Advancing fairness and transparency: National guidelines for post-conviction risk and needs assessment*. Council of State Governments Justice Center/ Bureau of Justice Assistance Public Safety Risk Assessment Clearinghouse. https://bja.ojp.gov/program/psrac

Desmarais, S. L., Johnson, K. L., & Singh, J. P. (2016). Performance of recidivism risk assessment instruments in US correctional settings. *Psychological Services*, *13*(3), 206–222. https://doi-org.ccsu.idm.oclc.org/10.1037/ser0000075.supp (Supplemental)

Desmarais, S. L., Zottola, S. A., Duhart Clarke, S. E., & Lowder, E. M. (2021). Predictive validity of pretrial risk assessments: A systematic review of the literature. *Criminal Justice and Behavior*, *48*(4), 398–420. https://doi-org.ccsu.idm.oclc.org/10.1177/0093854820932959

de Vogel, V., de Ruiter, C., Bouman, Y. H. A., & de Vries Robbe, M. (2009). *Guidelines for the Structured Assessment of Protective Factors for Violence Risk (SAPROF)*. Forum Educatief.

Farrington, D. P. (2003). Developmental and life-course criminology: Key theoretical and empirical issues—The 2002 Sutherland award address. *Criminology, 41*, 221–255. https://doi.org/10.1111/j.1745-9125.2003.tb00987.x

Ferguson, J. L. (2002). Putting the "what works" research into practice: An organizational perspective. *Criminal Justice and Behavior, 29*(4), 472-492. https://doi.org/10.1177/0093854802029004007

Gran-Ruaz, S., Feliciano, J., Bartlett, A., & Williams, M. T. (2022). Implicit racial bias across ethnoracial groups in Canada and the United States and Black mental health. *Canadian Psychology/Psychologie Canadienne*. https://doi-org.ccsu.idm.oclc.org/10.1037/cap0000323.supp (Supplemental)

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*(1), 4–27. https://doi.org/10.1037/0033-295X.102.1.4

Hanson, R. K., & Anderson, D. (2021). Static-99R: An empirical-actuarial risk tool for adult males with a history of sexual offending. In K. S. Douglas & R. K. Otto (Eds.), *Handbook of violence risk assessment, 2nd ed.* (pp. 106–130). Routledge/Taylor & Francis Group. https://doi-org.ccsu.idm.oclc.org/10.4324/9781315518374-8

Hilton, N. Z. (2021). *Domestic violence risk assessment: Tools for effective prediction and management, 2nd ed*. American Psychological Association. https://doi-org.ccsu.idm.oclc.org/10.1037/0000223-000

Hockenberry, S., & Puzzanchera, C. (2020, April). *Juvenile court statistics 2018*. National Center for Juvenile Justice.

Jones, N. J., Brown, S. L., Robinson, D., & Frey, D. (2015). Incorporating strengths into quantitative assessments of criminal risk for adult offenders: The Service Planning Instrument. *Criminal Justice and Behavior, 42*(3), 321–338. https://doi.org/10.1177/0093854814547041

Jones, N. J., Brown, S. L., Robinson, D., & Frey, D. (2016). Validity of the youth assessment and screening instrument: A juvenile justice tool incorporating risks, needs, and strengths. *Law & Human Behavior, 40*(2), 182–194. https://doi.org/10.1037/lhb0000170

Kim, J., & Kiesel, A. (2018). The long shadow of police racial treatment: Racial disparity in criminal justice processing. *Public Administration Review*, *78*(3), 422–431. https://doi-org.ccsu.idm.oclc.org/10.1111/puar.12842

Kim, J. Y., & Roberson, L. (2022). I'm biased and so are you. What should organizations do? A review of organizational implicit-bias training programs. *Consulting Psychology Journal*, *74*(1), 19–39. https://doi-org.ccsu.idm.oclc.org/10.1037/cpb0000211

Kroner, D. G., Morrison, M. M., & Lowder, E. M. (2020). A principled approach to the construction of risk assessment categories: The Council of State Governments Justice Center Five-Level System. *International Journal of Offender Therapy and Comparative Criminology, 64*(10-11), 1074-1090. https://doi.org/10.1177/0306624X19870374

Kropp, P. R., & Gibas, A. (2021). The Spousal Assault Risk Assessment Guide (SARA). In K. S. Douglas & R. K. Otto (Eds.), *Handbook of violence risk assessment., 2nd ed.* (pp. 389–409). Routledge/Taylor & Francis Group. https://doi-org.ccsu.idm.oclc.org/10.4324/9781315518374-20

Latessa, E. J., Lemke, R., Makarios, M., Smith, P., & Lowenkamp, C. T. (2011). The creation and validation of the Ohio Risk Assessment System (ORAS). *Criminal Justice Research Review*, *12*(3), 38–44.

Lehmann, P. S., Meldrum, R. C., & Greenwald, M. A. (2020). Upward departures from structured recommendations in juvenile court dispositions: The intersection of race, ethnicity, and gender. *JQ: Justice Quarterly*, *37*(3), 514–540. https://doi-org.ccsu.idm.oclc.org/10.1080/07418825.2018.1531143

Loeber, R., Farrington, D.P., Hipwell, A.E., Stepp, S.D., Pardini, D., & Abonen, L. (2015). Constancy and chance in the prevalence and frequency of offending when based on longitudinal self-reports or official records: Comparisons by gender, race, and crime type. *Journal of Developmental Life Course Criminology, 1*, 150-168.

Lowenkamp, C. T., Johnson, J. L., Holsinger, A. M., VanBenschoten, S. W., & Robinson, C. R. (2013). The federal Post Conviction Risk Assessment (PCRA): A construction and validation study. *Psychological Services*, *10*(1), 87–96. https://doi-org.ccsu.idm.oclc.org/10.1037/a0030343

Lowenkamp, C. T., Holsinger, A., Robinson, C. R., & Alexander, M. (2014). Diminishing or durable treatment effects of STARR? A research note on 24-month re-arrest rates. *Journal of Crime and Justice, 37(2)*, 275–283. https://doi.org/10.1080/0735648X.2012.753849

Lowenkamp, M. S., Robinson, C. R., Koutsenok, I., Lowenkamp, C. T., & Pearl, N. (2012). The importance of coaching: A brief survey of probation officers. *Federal Probation*, *76*(2), 36–40.

Marlowe, D. B., Ho, T., Carey, S. M., & Chadick, C. D. (2020). Employing standardized risk assessment in pretrial release decisions: Association with criminal justice outcomes and racial equity. *Law and Human Behavior, 44*(5), 361. http://dx.doi.org/10.1037/lhb0000413

Maruna, S. (2001). *Making good: How ex-convicts reform and rebuild their lives*. American Psychological Association.

Masten, A. S. (2016). Resilience in developing systems: The promise of integrated approaches. *European Journal of Developmental Psychology, 13*(3), 297–312. https://doi.org/10.1080/17405629.2016.1147344

Miller, H. A. (2006). *Inventory of offender risks, needs, and strengths (IORNS): Professional Manual*. Psychological Assessment Resources.

Miller, W. T., Campbell, C. A., & Larnell, T. (2021). Bias detected? An examination of criminal history using the OYAS-DIS for girls and black youth. *Journal of Ethnicity in Criminal Justice*, *19*(2), 101–119. https://doi-org.ccsu.idm.oclc.org/10.1080/15377938.2021.1932661

Mills, J. F., Loza, W., & Kroner, D. G. (2003). Predictive validity despite social desirability:

Evidence for the robustness of self-report among offenders. *Criminal Behaviour & Mental Health*, *13*(2), 140–150. https://doi-org.ccsu.idm.oclc.org/10.1002/cbm.536

Mitchell, D., Tafrate, R. C., & Hogan, T. (2018). Cognitive behavioral therapy in forensic treatment. In E. L. Jeglic & C. Calkins (Eds.), *New frontiers in offender treatment*. Springer.

Morgan, R. D., Flora, D. B., Kroner, D. G., Mills, J. F., Varghese, F., & Steffan, J. S. (2012). Treating offenders with mental illness: A research synthesis. *Law and Human Behavior*, *36*(1), 37–50. https://doi.org/10.1037/h0093964

Mullainathan, S. (2019, March 22). *Discrimination by algorithm and people* [video]. Youtube. https://www.law.uchicago.edu/recordings/sendhil-mullainathan-discrimination-algorithm-and-people

Onifade, E., Barnes, A., Campbell, C., & Mandalari, A. (2019). Risk assessment as a mechanism for reducing differential treatment of minorities by a juvenile justice system. *Journal of Ethnicity in Criminal Justice, 17(3)*, 203-227. DOI: 10.1080/15377938.2019.1619647

Orbis Partners. (2003). *Service Planning Instrument (SPIn).* Author.

Paluck, E. L., Porat, R., Clark, C. S., & Green, D. P. (2021). Prejudice reduction: Progress and challenges. *Annual Review of Psychology*, *72*, 533–560. https://doi-org.ccsu.idm.oclc.org/10.1146/annurev-psych-071620-030619

Peterson, J. K., Skeem, J., Kennealy, P., Bray, B., & Zvonkovic, A. (2014). How often and how consistently do symptoms directly precede criminal behavior among offenders with mental illness? *Law and Human Behavior, 38*, 439-449. https://doi: 0.1037/lhb0000075

Pollock, W., Hill, M., Menard, S., & Elliott, D. (2016). Predicting consistency between officially recorded and self-reported records of arrest. *American Journal of Criminal Justice*, *41*(4), 623–644. https://doi-org.ccsu.idm.oclc.org/10.1007/s12103-016-9343-9

Schleiden, C., Soloski, K. L., Milstead, K., & Rhynehart, A. (2020). Racial disparities in arrests: A race specific model explaining arrest rates across black and white young adults. *Child & Adolescent Social Work Journal*, *37*(1), 1–14. https://doi-org.ccsu.idm.oclc.org/10.1007/s10560-019-00618-7

Seligman, M. E. P. (2002). Positive psychology, positive prevention, and positive therapy. In C. R. Snyder & S. J. Lopez (Eds.), *Handbook of positive psychology* (pp. 3–9). Oxford University Press.

Serin, R. C., & Lloyd, C. D. (2009). Examining the process of offender change: The transition to crime desistance. *Psychology, Crime, & Law, 15*(4), 347–364. https://doi.org/10.1080/10683160802261078

Serin, R. C., & Lloyd, C. D. (2017). *Understanding the risk, need, and responsivity (RNR) model and crime desistance perspective and integrating them into correctional practice.* Retrieved from https://carleton.ca/cjdml/wp-content/uploads/Crime-desistance-and-RNR-final.pdf

Skeem, J., & Lowenkamp, C. (2020). Using algorithms to address trade-offs inherent in predicting recidivism. *Behavioral Sciences & the Law*, *38*(3), 259–278. https://doi.org/10.1002/bsl.2465

Skeem J. L., Steadman, H. J., & Manchak, S. M. (2015). Applicability of the Risk-Need-Responsivity model to persons with mental illness involved in the criminal justice system. *Psychiatric Services, 66*, 916-922. doi: 10.1176/appi.ps.201400448

Smith, P., Schweitzer, M., Labrecque, R. M., & Latessa, E. J. (2012). Improving probation officers' supervision skills: An evaluation of the EPICS model. *Journal of Crime and Justice, 35*(2), 189–199. https://doi.org/10.1080/0735648X.2012.674826

Steadman, H. J., Osher, F. C., Robbins, P. C., Case, B., & Samuels, S. (2009). Prevalence of serious mental illness among jail inmates. *Psychiatric Services, 60*, 761-765. https://doi: 10.1176/appi.ps.60.6.761.

Stinson, J. D., & Clark, M. D. (2017). *Motivational Interviewing with offenders: Engagement, rehabilitation, and reentry*. Guilford Press.

Tafrate, R. C., Hogan, T., & Mitchell, D. (2019). Integrating motivational interviewing with risk-need-responsivity based practice in community corrections: Collaboratively focusing on what matters most. In D. Polaschek, A. Day, and C. Hollin (Eds.), *The Wiley international handbook of correctional psychology* (pp. 603 –622). Wiley-Blackwell.

Tafrate, R. C., Mitchell, D., & Simourd, D. J. (2018). *CBT with justice-involved clients: Interventions for antisocial and self-destructive behaviors.* Guilford Press.

Taxman, F. S. (2018). Risk assessment: Where do we go from here? In J. P. Singh, D. G. Kroner, J. S. Wormith, S. L. Desmarais, & Z. Hamilton (Eds.), *Handbook of recidivism risk/needs assessment tools* (pp. 271-284). Wiley-Blackwell.

Viljoen, J. L., Jonnson, M. R., Cochrane, D. M., Vargen, L. M., & Vincent, G. M. (2019). Impact of risk assessment instruments on rates of pretrial detention, postconviction placements, and release: A systematic review and meta-analysis. *Law and Human Behavior, 43*(5), 397-420. https://doi.org/10.1037/lhb0000344

Vincent, G. M., Guy, L. S., Perrault, R. T., & Gershenson, B. G. (2016). Risk assessment matters, but only when implemented well: A multi-site study in juvenile probation. *Law and Human Behavior, 40*, 683-696.

Vincent, G. M., & Viljoen, J. L. (2020). At the forefront: Racist algorithms or systemic problems? Risk assessments and racial disparities. *Criminal Justice and Behavior*, *47*(12), 1576-1584. https://doi.org/10.1177/0093854820954501

Wanamaker, K. A., Jones, N. J., & Brown, S. L. (2018). Strengths-based assessment for use with forensic populations: A critical review. *International Journal of Forensic Mental Health, 17*(2), 202–221. https://doi.org/10.1080/14999013.2018.1451414

Wormith, J. S., & Bonta, J. (2018). The level of service (LS) instruments. In J. P. Singh, D. G. Kroner, J. S. Wormith, S. L. Desmarais, & Z. Hamilton (Eds.), *Handbook of recidivism risk/needs assessment tools* (pp. 117-145). Wiley-Blackwell.

Wormith, J. S., & Bonta, J. (2021). Risk/need assessment for adults and older adolescents: The level of service (LS) instruments. In K. S. Douglas & R. K. Otto (Eds.), *Handbook of violence risk assessment, 2nd ed.* (pp. 159–190). Routledge/Taylor & Francis Group. https://doi-org.ccsu.idm.oclc.org/10.4324/9781315518374-11

Wormith, J. S., Craig, L. A., & Hogue, T. E. (Eds.). (2020). *The Wiley handbook of what works in violence risk management: Theory, research, and practice*. Wiley Blackwell.

Building a Risk Tool for Persons Placed on Federal Post-Conviction Supervision for Child Sexual Exploitation Material Offenses: Documenting the Federal System's Past, Current, and Future Efforts

Thomas H. Cohen[1]
Probation and Pretrial Services Office
Administrative Office of the U.S. Courts

OVER THE LAST two decades, the number of persons placed on federal supervision for Child Sexual Exploitation Material (CSEM)[2] offenses has increased exponentially. The surge in CSEM supervisees can be attributed to technological changes that allow for easier access to sexually explicit materials on the internet and federal laws and enforcement mechanisms that have resulted in growing numbers of persons convicted of CSEM under federal sentencing statutes (Faust & Motivans, 2015; U.S. Sentencing Commission [USSC], 2012, 2021; Wolak et al., 2005, 2009). Specifically, the use of various technologies, including peer-to-peer networks, texting and instant messaging, cloud-based hosting services, social media platforms, and chatrooms, has created the context in which the typical person convicted of CSEM offenses will have generated voluminous collections of graphical images, including those of very young children (USSC, 2012, 2021). Moreover, federal legislation, particularly the Prosecutorial Remedies and Other Tools to End the Exploitation of Children Today Act of 2003 (The PROTECT Act), has resulted in increased penalties for persons convicted of CSEM through the addition of new enhancements and mandatory minimums to the federal sentencing guidelines (USSC, 2021). The PROTECT Act also gave federal judges the discretion to impose life supervision terms on persons convicted of federal sex offenses (Faust & Motivans, 2015; USSC, 2012). In addition to these technological and legislative changes, numerous regional taskforces and specialized units have been established by the U.S. Department of Justice to prosecute persons engaged in CSEM offenses (Wolak et al., 2005).

These trends combined have resulted in substantial increases in the number of persons prosecuted, incarcerated, and (most importantly for this research) placed on federal post-conviction supervision for CSEM offenses[3] (Faust & Motivans, 2015; U.S. Sentencing Commission [USSC], 2012, 2021). Faust and Motivans (2015) report that the number of persons placed on federal post-conviction supervision for sex offenses increased by 1,400 percent, from 321 supervisees in 1994 to 4,714 supervisees, in 2013. Much of this increase could be attributed to the prosecution of persons charged with CSEM offenses (i.e., possession, receipt, distribution, or production of child pornography). Moreover, persons convicted of CSEM offenses are increasingly being sentenced to lengthy post-conviction supervision terms in the federal system. The average term imposed on nearly 4,700 CSEM supervisees placed on federal post-conviction supervision during fiscal years 2010 through 2016 was about 98 months (see Table 1). In comparison, the average term imposed on federal supervisees in 2010 was about 43

[1] Thomas H. Cohen, Senior Social Science Analyst, Administrative Office of the U.S. Courts, Washington, D.C. The author would like to thank Scott VanBenschoten, Michelle Spidell, Stacy Merolla, Vignesh Narayanaswamy, Ann Smith, and Mariana Madrid for their helpful suggestions and comments. This publication benefited from the careful editing of Ellen Fielding. Direct correspondence to Thomas H. Cohen, Administrative Office of the U.S. Courts, One Columbus Circle NE, Washington, D.C. 20544 (email: thomas_cohen@ao.uscourts.gov).

[2] In prior research of federal sex offenders conducted by the Federal Probation and Pretrial Services Office (PPSO), the term "child pornography offender" was used to refer to persons placed on supervised release for possessing, receiving, distributing, or producing child pornography. Given the efforts to discourage the use of the word "offender," the term CSEM or CSEM supervisee was substituted for child pornography offender.

[3] Federal post-conviction supervision refers to persons sentenced to a term of community supervision following a period of imprisonment within the Federal Bureau of Prisons (18 U.S.C. §3583). Probation refers to persons sentenced to a period of federal supervision without any imposed incarceration sentence (18 U.S.C. §3561).

months (USSC, 2012).

The growth of CSEM supervisees presents serious challenges to the federal supervision system. Prior research shows many CSEM supervisees initially being designated as low risk to reoffend according to the federal Post Conviction Risk Assessment (PCRA) instrument (Cohen & Spidell, 2016; Cohen 2018); however, these persons tend to be placed through overrides into the highest supervision levels (Cohen et al., 2020). Officers' propensity to override CSEM supervisees stems from concerns about whether these persons have histories of, or are likely to engage in, offline contact sexual behavior with children (DeLisi et al., 2016). A meta-analysis focusing on the backgrounds of CSEM persons, for example, found that about 12 percent had an official arrest or conviction record of contact sexual behavior, but 55 percent admitted through self-reporting that they had prior sexual contact with children (Seto et al., 2011). Moreover, the risk tool officers used to gauge the likelihood of recidivism for federal supervisees is not calibrated to measure sexual deviance, ascertain the presence of non-official contact sex behavior, or assess the risk of sexual recidivism for sex offenders generally or CSEM supervisees in particular (Cohen & Spidell, 2016; Cohen 2018).

TABLE 1.
Descriptive Statistics of Study Sample

Descriptive factors	n	Percent or mean
Race/Ethnicity		
White, not Hispanic	5,113	88.8 %
Hispanic, any race	377	6.6
Black	157	2.7
Asian or Pacific Islander	76	1.3
American Indian or Alaska Native	30	0.5
Other	6	0.1
Average age	5,768	46.1
Type of supervision		
Term of supervised release	5,615	97.4 %
Probation	153	2.7
Most serious conviction offense/a		
Possession of child pornography	3,806	66.0 %
Distribution/receipt/transportation of child pornography	1,483	25.7
Sexual exploitation of children	462	8.0
Transfer obscene material to minors	118	2.1
Production of child pornography	39	0.7
Obscenity	6	0.1
PCRA risk levels		
Low	4,323	75.0 %
Low/moderate	1,234	21.4
Moderate	190	3.3
High	21	0.4
Average PCRA score	5,768	4.1
Supervision time imposed (Months)/b	4,695	97.7
Rearrested for any sex offense within 60 months of initial PCRA assessment	262	4.5 %
Number of supervisees	5,768	

Note: Includes 5,768 male supervisees convicted of online sex offenses placed on federal supervision between fiscal years 2010 through mid-2016 whose rearrest activity could be tracked for 60 months.
PCRA = Post Conviction Risk Assessment
a\Will not sum to 5,768 because supervisees can be convicted of multiple offenses.
b\Post-conviction supervision time imposed available for 81% of the CSEM study population.

The combination of growing numbers of persons on federal supervision for CSEM offenses, concerns about the frequency with which this population has engaged in unrecorded contact sexual behavior, and issues with using the current risk tool employed by federal probation officers to assess the risk of sexual recidivism (i.e., the PCRA) gave rise to an initiative by the Administrative Office of the U.S. Courts, Probation and Pretrial Services Office (PPSO), to construct a risk tool that could gauge the likelihood of sexual recidivism for the CSEM population.

This article documents PPSO's efforts to construct a risk tool that could be used on persons placed on federal post-conviction supervision for CSEM offenses. Initially, the article will delve into federal policies for supervising persons convicted of CSEM offenses and contrast those policies with an examination of how the CSEM population is actually being supervised. Next, it will detail PPSO's attempts to build a risk tool based on the Child Pornography Offender Risk Tool (CPORT) (see Eke et al., 2018; 2019), PCRA (see Johnson et al., 2011; Lowenkamp et al., 2013; 2015) and FBI criminal history records to assess the risk of sexual recidivism for CSEM supervisees. Specifically, the paper will describe the methods, data, and principal findings stemming from PPSO's efforts to use the CPORT and an amalgamation of fields obtained from the CPORT, PCRA, and FBI criminal history files to predict sexual recidivism for CSEM supervisees. Additional work involving the use of machine learning to gauge the likelihood of sexual reoffending for CSEM supervisees will also be detailed. Ultimately, as will be shown, none of these efforts were successful in creating a risk tool that officers could use for CSEM supervisees. The paper will conclude by discussing the implications of PPSO's efforts to build a CSEM-specific risk tool and suggest possible alternatives for

future research.

Federal Policies and Practices for Supervising Persons Convicted of CSEM Offenses

PPSO has responded to the growing number of persons convicted of CSEM offenses under federal supervision and the concerns that some CSEM supervisees might be involved in hands-on offending by issuing guidance for federal officers charged with supervising these persons. Under current policy, officers are instructed to use information gleaned from both the PCRA and other sources of information, including presentence reports, polygraphs, and psychosexual evaluations, to conduct an initial risk assessment evaluation. The PCRA is a dynamic actuarial instrument developed for federal probation officers that classifies supervisees into a matrix containing 12 risk categories (Administrative Office of the U.S. Courts (AO), 2018: see page 14). These categories provide crucial information about a supervisee's likelihood of committing any or violent offenses both during and after the supervisee has completed the supervision term (for more information about the PCRA, see Johnson et al., 2011; Lowenkamp et al., 2013; Lowenkamp et al., 2015; Luallen et al., 2016; Serin et al., 2016).

Although the PCRA provides crucial information about a supervisee's propensity for reoffending, it is not geared towards CSEM supervisees, constructed to assess their likelihood of sexual reoffending, or designed to measure sexual deviance. Additionally, nearly all CSEM supervisees (97 percent) are classified as low or low/moderate risk according to the PCRA (Cohen & Spidell, 2016). The lack of any official contact record for many CSEM supervisees, combined with their low-risk classification status, initially produced a policy in which officers were required to place all CSEM supervisees into the highest supervision levels regardless of their original risk classification (Cohen & Spidell, 2016; Cohen, 2018). Officers responded to this policy by applying overrides as a means of supervising nearly all CSEM supervisees at the highest supervision levels (Cohen et al., 2016; 2020).

This policy underwent a revision several years ago. Specifically, officers are no longer required to place all CSEM supervisees into the highest supervision levels through overrides. Rather, in December 2017, the policy was changed to acknowledge the risk principle and account for variations of risk within this population. As a result, officers are now encouraged to consider a combination of factors when designating the levels of supervision intensity. At the onset of supervision, officers may have limited case information and rely on known recidivism rates of CSEM individuals and suggested PCRA risk levels. However, during the course of supervision, as information related to the risk and needs of the case change, officers should respond by adjusting supervision levels as necessary.[4]

While policy no longer mandates that officers place CSEM supervisees into higher risk categories, evidence suggests that officers continue to use overrides to elevate the supervision levels for these persons. An examination of nearly 6,900 CSEM supervisees who received PCRA assessments between fiscal years 2017-21 showed officers overriding 96 percent of these persons and placing nearly all of them into the moderate or high supervision categories (data not shown). Moreover, officers typically keep these supervisees in the highest supervision categories even after multiple assessments. The continued use of supervision overrides for CSEM supervisees and the intensity of resources and staff directed at CSEM supervision provided the impetus for PPSO to develop an actuarial tool that could be used to supervise this specific population of sex offenders. The remainder of this paper details PPSO's effort to use the CPORT and a combination of CPORT, PCRA, and criminal history risk factors to construct a tool that could be used to assess the risk of sexual recidivism for CSEM supervisees.

Using the CPORT to Assess the Risk of Sexual Recidivism for CSEM Supervisees

In order to address the challenges inherent in supervising persons convicted of CSEM offenses, PPSO decided to attempt to assess whether the CPORT alone, or in combination with the PCRA and criminal history files, could be used to provide officers with a means of accurately gauging a CSEM supervisee's risk of engaging in sexual recidivism. PPSO selected the CPORT because of a growing literature showing its efficacy in differentiating the risk of sexual recidivism for persons convicted of CSEM offenses (Black, 2018; Eke et al., 2018, 2019; Pilon, 2016; Savoie et al., 2022; Seto & Eke, 2015; Soldino et al., 2021). Specifically, the CPORT was created to gauge the risk of any sexual recidivism among a population of adult males convicted of CSEM offenses (Eke et al., 2018, 2019; Seto & Eke, 2015; Soldino et al., 2021).[5] This risk instrument was originally constructed using a sample of 266 males convicted of CSEM offenses in Canada whose arrest activity could be followed for a period of five years and then validated on an additional sample of 80 men (Seto & Eke, 2015; Eke et al., 2019; Soldino et al., 2021). The CPORT's developers showed that this tool was effective at predicting any sexual recidivism (Area Under the Curve (AUC =. 74)) as well as sexual recidivism for CSEM subpopulations with histories of contact sexual behavior (AUC = .80) or backgrounds of general criminal activity not involving contact sexual offending (AUC = .69) (Seto & Eke, 2015). The tool's predictive efficacy, however, degraded when predicting sexual recidivism for CSEM subpopulations with only a history of child pornography offenses (AUC = .63) (Seto & Eke, 2015).

Subsequent CPORT studies showed the tool manifesting mixed effectiveness in terms of its ability to predict sexual recidivism for persons convicted of CSEM offenses. In a study conducted on 141 adult CSEM males in Scotland, the authors found that the CPORT significantly predicted various forms of reoffending behavior, including recidivism for any offenses (AUC = .81), any sexual offenses (AUC = .78), and CSEM offenses (AUC = .74) (Savoie et al., 2022). Other studies, however, produced results that did not replicate the CPORT's original predictive effectiveness. Using a truncated version of the CPORT scale[6] on 279 persons convicted of CSEM offenses in Canada and with a follow-up period of over

[4] It should be noted that about 1 out of 5 persons on federal supervision for CSEM offenses has a valid Static-99/R score (Cohen & Spidell, 2016). The Static-99/R is an actuarial risk prediction instrument that estimates the probability of sexual and/or violent reconviction for adult males who have already been charged with or convicted of at least one contact sexual offense against a child or non-consenting adult (Hanson et al., 2016; Helmus & Hanson, 2007). The Federal Bureau of Prisons attempts to score the Static-99/R on all sex offenders; however, valid scores are calculated for only those persons with current or prior arrest/conviction records for contact sex offending. Should the CSEM supervisee have a Static-99/R score, policy mandates that the officer default to the risk tool (i.e., PCRA or Static-99/R) that recommends the highest levels of supervision intensity.

[5] For a complete overview of the CPORT items, see Eke et al. (2018), as well as this paper's methods section.

[6] The truncated version omitted two items (questions #6 and #7) measuring the content of boy vs. girl material in the collections of persons convicted of CSEM offenses.

three years, Pilon (2016) generated results in the mediocre predictive range (AUC = .56). Another study conducted by Black (2018) using a shortened version of the CPORT scale[7] covering 547 persons with CSEM convictions in New Zealand and tracking their arrest activity for a period ranging from 2 to 19 years, found effect sizes ranging from the small (AUC = .60) to large (AUC >= .80), depending upon the arrest outcome examined. Last, Soldino et al. (2021) examined the CPORT's predictive efficacy on a sample of 304 men arrested for CSEM offenses in Spain and tracked for a duration of 5 years. The Soldino et al. (2021) study used the complete CPORT scale as well as the Correlation of Admission of Sexual Interest in Children (CASIC)[8]; overall, the results were mixed, with the CPORT total scores mostly producing AUC values of below .60 irrespective of the presence or absence of missing data. The authors, however, were able to generate AUC scores of .70 when applying the CASIC to a subset of the study population (Soldino et al., 2021).

Conducting a Pilot Test of the CPORT

Given the CPORT's potential effectiveness as a risk classification tool for CSEM supervisees, PPSO decided to ascertain whether this instrument could be integrated into the federal supervision system. The effort to integrate the CPORT occurred through two initiatives. Initially, PPSO attempted to conduct a pilot test of the CPORT by bringing in about 20 probation officers to manually code the CPORT on a random sample of 200 persons placed on federal post-conviction for CSEM offenses. For the pilot effort, PPSO contracted with one of the CPORT developers (Doctor Angela Eke) and she, along with Detective Sergeant Monica Denreyer, trained the federal probation officers on how to accurately code this risk tool. As a result of this training, CPORT and CASIC data were coded for 195 CSEM supervisees placed on federal post-conviction supervision between fiscal years 2011 and 2012.[9] The coding primarily involved examining presentence reports (PSRs) and other materials produced at supervision intake. While a great deal of information was learned from the pilot, unfortunately the officers had difficulty coding the CPORT items measuring boy to girl content in the child pornography material (CPORT item #6) and nude/other material (CPORT item #7). Moreover, officers were unable to code most of the CASIC items to determine a CSEM supervisee's sexual interest in children or teenagers. The combination of high rates of missing data for several CPORT and CASIC items, along with a relatively low rate of sexual reoffending for the pilot sample (only 9 of the 195 persons sexually reoffended), resulted in AUC scores in the mediocre to poor range (AUC = .54) for this instrument.

As a result of the pilot's poor performance, PPSO decided to rethink how to empirically test the CPORT's predictive performance for persons placed on federal supervision for CSEM offenses. Ultimately, PPSO decided to conduct a larger test of the CPORT using a population of 5,768 male supervisees placed on federal post-conviction supervision between fiscal years 2010 through mid-2016. Rather than have officers manually code the CPORT items, PPSO contracted with the MITRE Corporation (hereafter MITRE) to conduct a text mining endeavor aimed at collecting the CPORT elements. The efforts the MITRE project entailed, along with the data elements collected and analyzed, are further detailed in the methods section.

Method

CPORT and CASIC elements

Extracting both the CPORT and CASIC elements from PPSO's case management system (i.e., The Probation and Pretrial Services Automated Case Tracking System or PACTS) is problematic, because many of the risk factors scored in these instruments are not readily available for electronic data extraction. Persons attempting to score the CPORT, for example, are required to mark the following items as present or absent: (1) age at the time of index investigation, 35 or younger; (2) any prior criminal history; (3) any failure on conditional release, including charge at index; (4) any contact sexual offending, including a charge at index; (5) indication of pedophilic or hebephilic interests; (6) more boy than girl content in the child pornography material; (7) more boy than girl content in the nude/other child material (Eke et al., 2018). It's also important to note the instrument allows scorers to substitute the CASIC as a method for assessing CPORT item #5 (indication of pedophilic or hebephilic interests). The CPORT's developers suggested using the CASIC in lieu of attempting to directly ascertain the presence of sexual interest in children or teenagers because of concerns that many persons being scored on this instrument would not readily admit to these deviant forms of sexual behavior (Eke et al., 2018; Soldino et al., 2021).

The CASIC measures whether the CSEM supervisee manifests key characteristics associated with admission of pedophilic or hebephilic sexual interests (Eke et al., 2018; Seto & Eke, 2017). In the CASIC, six items are coded as being present or absent: (1) never married; (2) had child pornography videos; (3) had child pornography text stories; (4) child pornography material spanning two or more years; (5) volunteering in a role with high access to children; (6) engaging in online sexual communications with a minor or undercover officer posing as a minor (Eke et al., 2018; Seto & Eke, 2017). CASIC scores of 3 or higher are indicative that the person being scored is sexually interested in children or teenagers and hence should receive a score for CPORT item #5.

Several of the CPORT and CASIC items are stored in PPSO's case management system (i.e., PACTs) in a format that allows for further analysis. For example, the CPORT items measuring age and criminal history (CPORT items #1 through #4) are available in PACTS and can be readily extracted to generate a truncated CPORT score. The remaining CPORT items (items #5 through 7), however, are not entered into the PACTs system in a structured format that can be easily retrieved, assuming they are entered at all. For example, a CSEM supervisee's admission of sexual interest in children or preference in boys over girls might be manifested in the text embedded in a presentence report or psychosexual assessment uploaded into PACTs, but this information is typically stored in unstructured PDF files or images; none of these items are entered into specific numeric fields.

Hence, any attempt to obtain these data would involve officers having to read through case files and manually code the CPORT items measuring sexual interest in children or teenagers or preference in boys over girls.[10] The

[7] This version omitted the CPORT's last three items, including question #5 (indication of pedophilic or hebephilic interests) and questions # 6 and #7 (measuring boy vs. girl content).

[8] The CASIC is used to assess a CSEM person's sexual interest in children or teenagers (see Seto & Eke, 2015).

[9] Five supervisees were removed from the sample because subsequent data obtained from PPSO's case management system showed they did not meet the criteria of persons who should be scored on the CPORT.

[10] Obtaining data for the CASIC is even more challenging; only 1 of the 6 items (never married) could be readily extracted through PACTs.

level of time, effort, and resources involved in obtaining this information through a review of PDFs or scanned documents uploaded into PACTs could be enormous given the number of CSEM cases officers would potentially have to code. Generally, the rates of sexual recidivism for CSEM supervisees are relatively low. Cohen and Spidell (2016) showed about 3 percent of CSEM supervisees being rearrested for sexual offenses within three years of their supervision start dates. Given the low base rates of sexual re-offending for this population, any effort to validate the CPORT on this population would involve collecting CPORT information on potentially thousands of CSEM supervisees. The challenge, therefore, was to devise a way to collect CPORT data through a mechanism that minimized having officers individually go through case files while simultaneously extracting the CPORT elements from as many cases as possible. Ultimately, the AO's Department of Technology Services contracted with MITRE to engage in a proof-of-concept project on the feasibility of applying natural language processing and machine learning techniques to retrieve CPORT data elements from thousands of CSEM supervisees. The MITRE project and its results are detailed below.

The MITRE Data Collection Effort

The MITRE project's primary goal was to construct an algorithm for extracting information from various documents to complete the CPORT risk tool from an initial list of 8,896 male CSEM supervisees placed on federal post-conviction supervision between fiscal years 2011 through 2018.[11] MITRE extracted unstructured text data from numerous sources, including PSRs, with a particular emphasis on the sections containing information on charges and convictions, mental and emotional health, personal and family data, offense conduct, and victim impact statements; polygraph reports; and psychosexual assessments and psychological evaluations. In total, MITRE processed an estimated 11,000 PSR documents,[12] 60,000 psychological and psychosexual assessments, and 55,000 polygraph reports. The process resulted in the analysis of about 126,000 PDF and scanned documents containing over 8 million sentences. To these 8 million sentences, MITRE applied a combination of content extraction, natural language processing, and artificial intelligence reasoning capacities to automatically produce responses that could be used to complete the CPORT risk instrument.[13] The entire automated process took about 12 days to complete. In comparison, if PPSO had opted for manual data collection, and if the amount of time required to complete the instrument were similar to the pilot effort (about one hour per CSEM supervisee), it is estimated that it would have taken four full-time staff about one year to manually code the CPORT for the same 8,896 CSEM supervisees.

It is important to note that while the MITRE effort produced results that mostly adhered to the CPORT data elements, there was some divergence between the MITRE-generated and CPORT fields. CPORT element #5 (indication of pedophilic or hebephilic interests), for example, was split into two elements measuring the presence of pedophilic or hebephilic interests separately (see Table 3). In addition, CPORT element #6 (more boy than girl content in the child pornography material) and CPORT element #7 (more boy than girl content in the nude/other child material) were combined into one field measuring whether the CSEM supervisee evidenced greater sexual interests in boys over girls. Moreover, the MITRE project attempted to gather several additional elements that could be associated with sexual recidivism. This effort involved measuring the presence or absence of the following elements: evidence of deviant sexual interests (a catchall category created by MITRE); lives with lover or partner for less than 2 years; engaged in online communications for illicit purposes; any prior non-contact sexual offenses; and any prior violent (non-sexual) offenses.

Though the MITRE project involved a novel initiative to transform unstructured text files into structured datasets for nearly 9,000 CSEM persons on federal supervision, some limitations about this project should be noted. First, it's important to acknowledge that MITRE relied on admissions, rather than on an examination of actual child pornography collections, to gauge preferences for boys over girls. Admissions, and not CASIC, were also used to ascertain the presence of pedophilic or hebephiliac interests. The use of admissions over these other forms of obtaining the CPORT data could explain some of the differences in the study's primary findings compared to previous CPORT research.

Inclusion of Elements From PCRA and Rap Sheets as Additional Predictors of Sexual Recidivism

In addition to the factors extracted through the MITRE project, this effort attempted to determine whether any other factors collected by the federal post-conviction risk assessment tool (i.e., PCRA), the generalized assessments conducted on all supervisees, or the criminal history data embedded in rap sheets were correlated with sexual recidivism. An effort was made to examine these other factors because of the concerns that the CPORT might not be predictive of sexual recidivism, given the low rates of reoffending activity among CSEM supervisees (Cohen & Spidell, 2016; USSC, 2021). The specific non-CPORT elements identified through this effort, the processes for selecting these elements, and their predictive efficacy are further detailed in the findings section.

Outcome Measure

The primary outcome of interest involves whether the CSEM supervisee was rearrested for any new sexual offenses. Rearrests for new criminal activity were obtained from the National Crime Information Center (NCIC) and Access to Law Enforcement System (ATLAS). ATLAS is a software program used by the AO that provides an interface for performing criminal record checks through a systematic search of official state and federal rap sheets (Baber, 2010). Sexual recidivism was defined to include arrests for any sexual offenses—either violent or non-violent but excluding prostitution offenses—within a fixed five-year time frame from the supervision start date. Similar to other CPORT validation studies, an attempt was made to distinguish contact from non-contact sexual recidivism events; however, there were so few CSEM supervisees arrested for contact sex crimes (less than 1 percent) that it was ultimately not practicable to separate out these arrest outcomes in the extant study.

The five-year follow-up period aligns with the tracking time used in the CPORT development study (Eke et al., 2019; Seto & Eke, 2015) as well as subsequent CPORT validation efforts (Soldino et al., 2021). The decision to use a five-year fixed follow-up period resulted

[11] The challenges inherent in obtaining CASIC items necessitated that we focus solely on the CPORT for this project. An effort was made, however, to collect the CASIC field measuring online communications with a minor.

[12] Some CSEM supervisees have multiple PSRs.

[13] For a more in-depth overview of the processes MITRE applied to data-mine the judicial system's text documents, see Megerdoomian et al. (2019), which discusses this effort for a related PPSO-sponsored project.

in 3,128 of the 8,896 CSEM supervisees whose CPORT data were collected by MITRE being removed from the analysis because their arrest outcomes could not be tracked for a minimum of five years. The remaining cohort of 5,768 male CSEM supervisees who were included in the current study, however, constitute one of the largest samples attempting to validate the CPORT ever conducted.

Analytical Approach

The statistical techniques applied to this analysis involved a combination of descriptive techniques, chi-square tests, and AUC-ROC scores. The AUC-ROC scores were primarily used to assess the predictive accuracy of the CPORT risk tool as well as the risk tool PPSO constructed, which combined elements from the CPORT, PCRA, and rap sheets. In addition to these techniques, an attempt was made to apply machine learning approaches to predict sexual recidivism. Specifically, random forest machine learning approaches were employed to assess whether these novel methods could substantially improve prediction compared to traditional risk assessment approaches. The random forest analyses are further discussed in the findings section.

Findings

Validating the CPORT

Descriptive information about the study sample is provided in Table 1. Of the 5,768 CSEM supervisees in the study population, nearly 90 percent were non-Hispanic whites, while the remainder were a combination of Hispanics, Blacks, Asian/Pacific Islanders, or American Indians or Alaska Natives. The average age was about 46 years, and almost the entire population (97 percent) were placed on post-conviction supervision through a term of supervised release, meaning that these supervisees had served time in federal prison before being released. Over 90 percent of the study population were convicted of CSEM offenses involving child pornography possession (66 percent) or the distribution, receipt, or transportation of child pornography (26 percent). About 1 percent were convicted of actually producing child pornography materials. The study population skewed low risk, with three-quarters receiving a low-risk classification from the PCRA; about 4 percent were assessed as moderate or high risk. By comparison, about one-fourth of the general federal supervision population are classified as moderate or high risk at initial assessment (Johnson & Baber, 2015). The rates at which CSEM supervisees recidivated for sexual offenses were also relatively low. Approximately 5 percent of the 5,768 CSEM supervisees were rearrested for any sexual offenses within 60 months of their supervision start date. In contrast, 43 percent of all federal supervisees were rearrested for any new offenses within 5 years from supervision commencement (Markman et al., 2016). The low-risk distribution skew for the CSEM study population, combined with their minimal rates of sexual recidivism, gives rise to various challenges for risk assessment construction, development, and validation that are subsequently detailed.

Information about the PCRA's capacity to predict sexual recidivism among CSEM supervisees is provided in Table 2. Overall, the PCRA's capacity to predict sexual recidivism for the 5,768 CSEM supervisees placed on federal supervision is in the weak range (AUC = .61. 95% CI [.58 - .64]). The PCRA's inability to differentiate CSEM supervisees by risk is especially apparent when examining the sexual recidivism rates for low/moderate and moderate CSEM supervisees, which are essentially the same (8.1 percent vs. 7.9 percent). These findings further illustrate the need to move beyond the PCRA and apply other tools (e.g., CPORT) in attempting to distinguish the risk of sexual recidivism for the federal CSEM population.

TABLE 2.
Association Between Post Conviction Risk Assessment (PCRA) Risk Levels and Any Sexual Recidivism for Online Sex Offenders

PCRA risk levels	n	Percent sexually recidivated
All supervisees	5,768	4.5 %
Low	4,323	3.3 %
Low/moderate	1,234	8.1
Moderate	190	7.9
High	21	28.6
AUC-ROC	0.61 [0.58 - 0.64]	

Note: Includes 5,768 male supervisees convicted of online sex offenses placed on federal supervision between fiscal years 2010 through mid-2016 whose rearrest activity could be tracked for 60 months.
PCRA = Post Conviction Risk Assessment

The presence of the CPORT and other risk factors generated by MITRE are provided in Table 3 (next page) in a sorted format. The MITRE data collection effort showed over a third of the CSEM population evidencing sexual interests in children or teenagers through admissions to officers, treatment providers, or polygraph administrators and nearly two-fifths were 35 years or younger at the time of index investigation. Approximately one-fifth manifested any criminal history, but only 2 percent were determined by MITRE to have a background of contact sexual offending. The rates of prior contact sex offending are lower than those reported in other studies of CSEM supervisees (see Cohen & Spidell, 2016) and ultimately resulted in an effort to supplement the criminal history backgrounds of these persons with FBI rap sheet data (see next section). Last, MITRE identified 6 percent of CSEM supervisees evidencing greater sexual interests in boys over girls.

Information about the presence of other (non-CPORT) risk factors generated by MITRE is also provided in Table 3. The most common other risk factors included evidence of deviant sexual interests (57 percent) and lives with lover or partner for less than two years (30 percent). About 12 percent of the study population engaged in online communication for illicit purposes and less than 5 percent had an arrest record for non-contact sexual or violent offenses.

Data on the bivariate associations between the MITRE-generated risk factors—both CPORT and other—and the five-year sexual recidivism rates are provided in Table 4 (next page). Several of the CPORT risk factors were shown to be significantly associated with sexual recidivism ($p < .05$), including age at index investigation, presence of pedophilic interests, presence of previous criminal history, any failure on conditional release, and presence of contact sexual reoffending. Of all the CPORT risk factors, any failure while on conditional

release and presence of contact sexual offending had the strongest associations with sexual recidivism; CSEM supervisees with these characteristics were about three times more likely to sexually recidivate compared to the overall baseline sexual recidivism rates. Interestingly, the CPORT factors measuring the presence of hebephiliac interests and greater sexual interests in boys over girls were not associated with sexual recidivism. Among the other risk factors produced by MITRE, only those measuring the presence of prior non-contact sex offenses and violent offenses manifested significant associations with sexual recidivism. Over 15 percent of CSEM supervisees with these characteristics were rearrested for sexual offenses.

The predictive effectiveness of the CPORT risk tool and various modified versions of this tool for CSEM supervisees are detailed in Table 5 (next page). Initially, an attempt was made to ascertain the CPORT's efficacy by assigning scores of 0 or 1 to each CPORT risk factor and summing the scores into a total score; the scores were included in the sum irrespective of whether they were significantly associated with sexual recidivism (see Table 4). Using this approach generated some differentiation in the sexual recidivism rates. For example, the percentage of CSEM supervisees who sexually recidivated increased somewhat incrementally from 2 percent of supervisees with no CPORT risk criteria (score = 0) to 8 percent of supervisees with at least three CPORT risk factors (score = 3). CSEM supervisees with five or more CPORT risk factors were 10 times more likely to be rearrested for sexual offenses (20 percent rearrested) compared to their counterparts with zero CPORT risk factors (2 percent rearrested). Despite these promising patterns, the overall AUC scores for the CPORT are in the mediocre predictive range (AUC = .62. 95% CI [.58 - .65]). The low-risk skew of the CSEM population—56 percent manifested CPORT scores of ranging from to 0 to 1—provides a partial explanation for these poor prediction metrics.

Attempts were made to evaluate whether the CPORT's predictive effectiveness could be enhanced by modifying this risk tool. Specifically, the modifications involved constructing an assessment score that included all the CPORT and other risk factors generated by MITRE regardless of their significant association with sexual recidivism as well as constructing a truncated assessment score that used only those CPORT and other risk factors significantly associated with sexual

TABLE 3.
Presence of CPORT or Other Risk Factors Associated with Sexual Recidivism for Online Sex Offenders

Risk items	Number	Percent
All supervisees	5,768	
CPORT risk factors		
35 years or younger at time of index investigation	2,168	37.6 %
Presence of indication of pedophilic interest	2,151	37.3
Presence of indication of hebephiliac interest	2,018	35.0
Presence of previous criminal history	1,276	22.1
Evidences greater sexual interest in boys over girls	350	6.1
Any failure on conditional release/a	236	4.1
Presence of contact sexual offending/a	130	2.3
Other risk factors		
Evidence of deviant sexual interests	3,265	56.6 %
Lives with lover or partner for less than 2 years	1,745	30.3
Engaged in online communication for illicit purpose	705	12.2
Presence of prior non-contact sex offenses	122	2.1
Presence of prior violent (non-sexual) offenses	101	1.8

Note: Includes 5,768 male supervisees convicted of online sex offenses placed on federal supervision between fiscal years 2010 through mid-2016 whose rearrest activity could be tracked for 60 months.
CPORT = Child Pornography Offender Risk Tool
The CPORT and other risk factors shown in table were generated by MITRE.
a/Includes previous or instant offenses

TABLE 4.
Association Between Individual CPORT or Other Risk Factors Associated with Sexual Recidivism for Online Sex Offenders

Risk items	n	Percent sexually recidivated
All supervisees	5,768	4.5 %
CPORT risk factors		
35 years or younger at time of index investigation	2,168	6.2 %***
Presence of indication of pedophilic interest	2,151	5.3 *
Presence of indication of hebephiliac interest	2,018	3.7
Presence of previous criminal history	1,276	7.9 ***
Evidences greater sexual interest in boys over girls	350	4.9
Any failure on conditional release/a	236	15.7 ***
Presence of contact sexual offending/a	130	14.6 ***
Other risk factors		
Evidence of deviant sexual interests	3,265	4.7 %
Lives with lover or partner for less than 2 years	1,745	4.9
Engaged in online communication for illicit purposes	705	4.1
Presence of prior non-contact sex offenses	122	18.9 ***
Presence of prior violent (non-sexual) offenses	101	15.8 ***

Note: Includes 5,768 male supervisees convicted of online sex offenses placed on federal supervision between fiscal years 2010 through mid-2016 whose rearrest activity could be tracked for 60 months.
CPORT = Child Pornography Offender Risk Tool
The CPORT and other risk factors shown in table were generated by MITRE.
Chi-square used to indicate statistical significance * P <.05; ** P < .01; *** P < .001
a/Includes previous or instant offenses

reoffending (see Table 4 for information about the types of risk factors significantly associated with sexual recidivism). The approach employing all the risk factors constructed by MITRE also produced sub-par predictive indices (AUC = .61. 95% CI [.57 - .64]). Conversely, employing a technique where only those risk factors significantly associated with sexual recidivism were included in the assessment calculations produced the highest AUC scores (AUC = .65. 95% CI [.62 - .69]) and patterns of sexual reoffending that increased somewhat *monotonic*ally by risk score. Though promising, even this method failed to generate predictive AUC scores in the high effect size range (e.g., AUC score > .70) (Rice & Harris, 2005).

TABLE 5.
Association Between CPORT Risk Scores and Risk Scores Using Other Factors with Sexual Recidivism for Online Sex Offenders

Risk scores	n	Percent sexual recidivated
All supervisees	5,768	4.5 %
Only CPORT risk factors		
0	1,314	2.4 %
1	1,907	4.0
2	1,526	4.2
3	765	7.7
4	210	10.0
5 plus	46	19.6
AUC-ROC	0.62 [0.58 - 0.65]	
CPORT plus other risk factors		
0	724	2.5 %
1	1,054	3.1
2	1,233	4.2
3	1,225	4.2
4	865	5.2
5	463	9.1
6	150	8.0
7	37	13.5
8 plus	17	23.5
AUC-ROC	0.61 [0.57 - 0.64]	
CPORT and other risk scores (reduced)/a		
0	1,748	2.4 %
1	2,407	3.7
2	1,187	6.0
3	322	10.6
4	84	22.6
5 plus	20	35.0
AUC-ROC	0.65 [0.62 - 0.69]	

Note: Includes 5,768 male supervisees convicted of online sex offenses placed on federal supervision between fiscal years 2010 through mid-2016 whose rearrest activity could be tracked for 60 months.
The scores calculated in table were based on factors generated by MITRE.
CPORT = Child Pornography Offender Risk Tool
95% confidence intervals shown.
a/ Reduced risk characteristics selected from CPORT and other factors significantly associated with any sexual recidivism at .05 level.

Building a CSEM Risk Instrument Based on CPORT, PCRA, and Criminal History Factors

Given the issues pertaining to sex offender prediction using the CPORT, PPSO decided to rethink its approach to developing a risk tool for CSEM supervisees. Specifically, an attempt was made to ascertain whether an in-house risk tool could be developed using data elements from a multitude of sources including the MITRE-generated CPORT and other risk factors, the risk elements collected by officers when conducting PCRA assessments, supervisee characteristics generated from officer assessments, and the FBI criminal history data. Elements were selected from these sources if they were associated with an increase of over three percentage points in the likelihood of sexual recidivism occurring within five years of the supervision start date. Though selecting elements through this approach might be viewed as less rigorous compared to selecting elements that are statistically significant, given the low base rates of sexual recidivism (4.5 percent), this method seemed to offer the best means for building a risk tool that could predict sexual recidivism among CSEM supervisees. In order to avoid the pitfall of generating a risk tool that overfits the data, and hence might not be useable when applied to a new group of CSEM supervisees, the database was randomly split into a training and testing data file. The variables associated with an increase of over three percentage points with sexual rearrest activity were selected from the training database and then applied to the testing file for the purpose of assessing this instrument's potential predictive efficacy.

The specific variables selected for CSEM risk construction and development are detailed in Table 6 (next page). In the training dataset, the following elements were selected from the MITRE-generated factors: presence of previous criminal history, evidences greater sexual interest in boys over girls, any failure on conditional release, presence of contact sexual offending, presence of prior non-contact sexual offenses, and presence of prior violent (non-sexual) offenses. All of these factors—with the exception of evidences greater sexual interest in boys over girls—were associated with significantly higher likelihoods of sexual recidivism in Table 4. Though the boy over girl content was not statistically associated with higher rearrest rates, in the training data this variable was correlated with a more than three percentage point increase in

the likelihood of sexual recidivism and hence was included as a potential predictor variable.

Several non-MITRE risk factors also were associated with an increase of more than three percentage points in the likelihood of sexual recidivism. Many of these factors hailed from the PCRA and included officer scores measuring whether a supervisee manifested social problems associated with drug use or negative attitudes towards supervision or had a record of institutional adjustment. Another factor, denial of harm, hails from the Psychological Inventory of Criminal Thinking Styles section of the PCRA and essentially measures if the supervisee either rationalizes or minimizes the harm their criminal lifestyle might have done to others (Walters, 2013). In addition to these factors, an assessment indicating that the supervisee had a record of domestic violence was also associated with sexual recidivism. Last, the presence of an FBI record indicating that the supervisee had an arrest history for sex offenses (prostitution excluded) was shown to be associated with a more than three percentage point increase in sexual reoffending. The FBI criminal history records augmented the MITRE criminal history data, since MITRE recorded relatively few CSEM supervisees having any arrest histories for sexual offenses. The factors listed in Table 6 all received scores of 0 or 1 depending upon whether their presence was recorded for the CSEM supervisee, and their individual scores were summed to generate a total score. The predicative effectiveness of these total scores for both the training and testing data are shown in the next table and figure.

Results from the hybrid approach explicated above are provided in Table 7. Overall, the AUC-ROC scores approach acceptable levels for the training data (AUC = .68. 95 percent CI [.63 - .73]); however, there is a slight though not significant deterioration when moving to the testing data (AUC = .65. 95 percent CI [.60 - .70]). Among both the training and testing samples, CSEM supervisees with higher risk scores were more likely to sexually recidivate compared to their counterparts who scored lower on the assessment instrument. For example, the percentage of CSEM supervisees in the testing sample rearrested for sexual offenses manifested the following incremental increases: 3 percent (score = 0), 6 percent (score = 2), 19 percent (score = 4), and 41 percent (score = 5). The rearrest rates do fall off when moving to scores of 6 or above; however, that pattern is partially explained by the small number of CSEM supervisees (n = 5)

TABLE 6.
MITRE and non-MITRE Generated Factors Used to Predict Sexual Recidivism for Online Sex Offenders

	Training data		Testing data	
Risk items	Number	Percent	Number	Percent
MITRE generated factors/a				
Presence of previous criminal history	635	22.0 %	641	22.2 %
Evidences greater sexual interest in boys over girls	186	6.5	164	5.7
Any failure on conditional release	121	4.2	115	4.0
Presence of contact sexual offending	63	2.2	67	2.3
Presence of prior non-contact sex offenses	53	1.8	69	2.4
Presence of prior violent (non-sexual) offenses	48	1.7	53	1.8
Non-MITRE factors/b				
PCRA - Social problems associated with drug use	346	12.0 %	327	11.3 %
PCRA - Negative attitudes towards supervision	417	14.5	391	13.6
PCRA - Institutional adjustment	359	12.5	365	12.7
PCRA - Denial of harm	36	1.4	29	1.1
Assessment - Domestic violence	102	3.5	84	2.9
RAP Sheets -Prior arrests for sex offenses				
None	2,571	89.2	2,610	90.5
One or more	313	10.9	274	9.5
All supervisees	2,884		2,884	

Note: Includes 5,768 male supervisees convicted of online sex offenses placed on federal supervision between fiscal years 2010 through mid-2016 whose rearrest activity could be tracked for 60 months.
a/MITRE generated factors.
b/Information obtained from PCRA, PICTS, Assessments, and Rap sheets.
Unless otherwise noted, data available for 99% of supervisees with exception of denial of harm factor.
Data on denial of harm factor available for 93% of supervisees.

TABLE 7.
Association Between Calculated Risk Scores Using MITRE and Other Risk Factors with Any Sexual Recidivism Using Training and Testing Data

	Training data		Testing data	
Risk scores	n	Percent sexually recidivated	n	Percent sexually recidivated
All	2,652	4.2 %	2,643	4.9 %
0	1,241	2.1 %	1,266	2.8 %
1	781	3.7	769	4.7
2	370	7.0	359	6.4
3	156	5.8	164	8.5
4	68	17.7	63	19.1
5	27	25.9	17	41.2
6 plus	9	33.3	5	20.0
AUC-ROC	0.68 [0.63-0.73]		0.65 [0.60 - 0.70]	

Note: MITRE and other risk factors used to calculate risk scores selected from factors associated with a three percentage point increase in sexual recidivism.
Includes 5,295 male supervisees convicted of online sex offenses whose risk scores could be calculated and whose rearrest activity could be tracked for 60 months.

receiving these high scores.

Figure 1 highlights the predictive efficacy of the PPSO-generated risk tool with the combined training and testing data. Results show somewhat incremental increases in the sexual rearrest rates by risk score. In general, the sexual recidivism rates rise from 2.5 percent to 4.2 percent when moving from scores of 0 to 1; afterwards they plateau at about 7 percent between scores 2 and 3 and then move up again to 18 percent and then 32 percent for persons scoring 4 and 5, respectively. The combined data produces predictive metrics that approach (AUC = .67. 95% CI [.63 - .70]) but do not meet, nor exceed, the acceptable range for most risk instruments (AUC > .70).

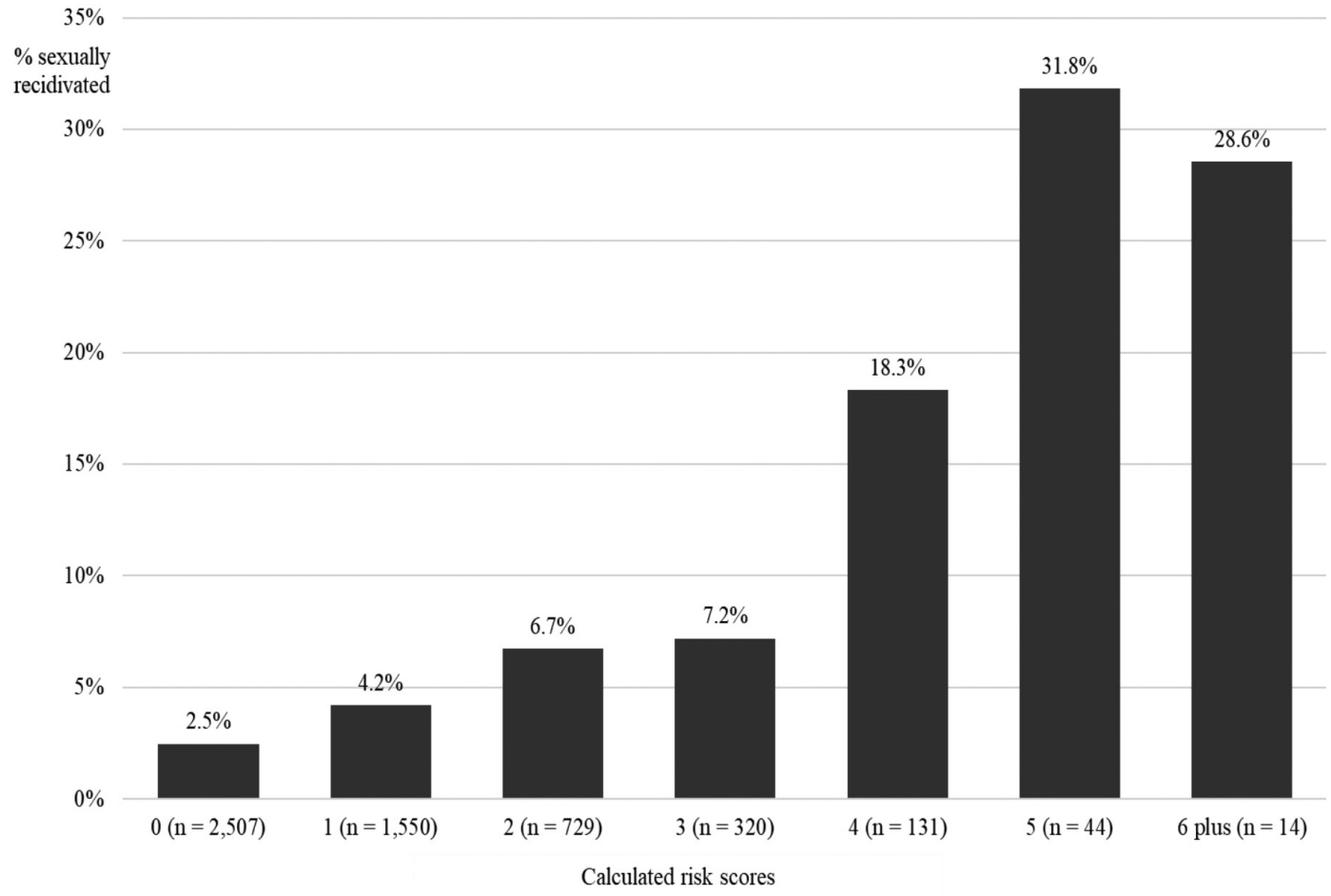
Using Machine Learning Approaches for CSEM Prediction

Though attempting to produce an in-house risk instrument geared to CSEM supervisees generated results that almost met the acceptable range of most criminal justice and sex offender risk assessments, this attempt fell short. Ultimately, PPSO was unable to construct a risk tool that could appreciably differentiate the risk of sexual recidivism among CSEM supervisees by using either the CPORT or a combination of CPORT and other risk factors embedded within the PCRA, officer assessments, and FBI criminal history records. In light of these results, PPSO made an additional effort to ascertain the feasibility of constructing an in-house risk assessment through the use of machine learning techniques. A brief description of PPSO's effort to apply machine learning applications to the problem of CSEM risk prediction is subsequently provided.

Machine learning is essentially an area of artificial intelligence that operates under the concept that a computer program can learn and adapt to data without the need for human intervention in the analytical process (Burkov, 2019). Over the past 20 years, machine learning has become increasingly used in the area of prediction, including investing, advertising, lending, fraud detection (Burkov, 2019) and, for purposes of this research, criminal justice risk assessment (Berk et al., 2019). An effort was made to apply random forests, which is a commonly used supervised machine learning approach. The random forest model works by growing algorithms called multiple decision trees,[14] which are then merged together for a more accurate prediction (Hartshorn, 2016). Specifically, multiple uncorrelated models

[14] The decision trees are basically algorithms used to classify data through a flowchart-type format. Each tree starts at a single point and then branches in two or more different directions, with each branch incorporating a variety of decisions until a final outcome is achieved (Hartshorn, 2016).

FIGURE 1.
Association Between Calculated Risk Scores Using MITRE and Other Risk Factors with Any Sexual Recidivism Combining Training and Testing Data



Note. Includes 5,295 male supervisees convicted of online sex offenses whose risk scores could be calculated and whose rearrest activity could be tracked for 60 months. AUC-ROC score for all supervisees is 0.67 with confidence interval of 0.63 - 0.70.

(e.g., decision trees) are applied to generate predictions superior to what would occur if only one decision tree was applied. Applying the random forests method results in each tree giving a classification or vote, and the forests picks the average of all outputs or trees (Hartshorn, 2016). In the current research, a total of 75 variables extracted from PPSO's case management system were used in the random forest models. The random forest models were configured to incorporate 1,400 trees with a maximum depth of 40 branches.

Results from the random forest models also failed to generate predictive indices that met the appropriate criminal justice risk assessment benchmarks (AUC > .70). When applied to the testing data, the random forest models generated AUC scores in the mediocre range (AUC = .54) (data not shown). Moreover, the true positive rate, or the percentage of CSEM supervisees arrested for sexual offenses who were predicted by the model to garner a new arrest, was 54 percent. The remaining 46 percent constituted false negatives, meaning that the model failed to accurately predict that these persons would be rearrested for sexual offenses. These suboptimal metrics of prediction remained constant even when differing random forest applications, including gradient boosting and other machine learning applications, were applied.

Discussion

This article documents PPSO's efforts to construct a risk assessment tool specifically geared towards predicting sexual recidivism among CSEM supervisees. Initially, the endeavor attempted to gauge whether the CPORT could be used for CSEM prediction. One of the challenges in using the CPORT involved the problem of coding several elements, including sexual interests in children and teenagers and preference for boys over girls, that are not readily extractable from PPSO's case management system. PPSO attempted to address this issue by employing MITRE, which used natural language processing for the purpose of text mining 126,000 PDF and scanned documents and, through this method, constructed a dataset composed of a modified version of the CPORT's elements as well as several additional factors believed to be associated with sexual recidivism. The construction of a structured dataset from a myriad of unstructured files embedded within PSRs, polygraph reports, and psychosexual assessments represented a novel effort to use many of the text files generated by federal probation officers during the course of supervision and is suggestive that many of the emerging data science techniques might be directed toward making PPSO's unstructured data more useful for research purposes.

Although MITRE was able to successfully transform unstructured files into structured data, regrettably this effort fell short of being able to construct and deploy a risk tool that could be used on CSEM supervisees. Overall, the modified version of the CPORT risk tool failed to adequately differentiate CSEM supervisees by their likelihood of sexual re-offending and produced AUC scores indicative of mediocre prediction (AUC = .62). While an effort to apply a truncated version of the CPORT performed somewhat better, it still resulted in predictive metrics (AUC = .65) that did not approach those reported by the CPORT's developers (AUC =. 78) (Seto & Eke, 2015).

In light of these results, PPSO attempted to build its own CSEM risk tool that was based on a combination of MITRE-generated factors and elements obtained from the PCRA and FBI criminal history records. This approach performed somewhat better at distinguishing a supervisee's risk of sexual recidivism and produced AUC values approaching the acceptable range for the training data (AUC =. 68), but there was some fall-off in prediction when moving to the testing data (AUC = .65). While PPSO's efforts geared toward building a CSEM risk tool from a combination of factors was somewhat more favorable, this approach produced predictive indices that did not meet the standard benchmarks of many criminal justice risk assessment instruments (AUC > .70). Finally, PPSO attempted to employ machine learning techniques (i.e., random forests) in order to evaluate whether these approaches might assist with CSEM risk prediction. In findings mirroring other analyses discussed in this report, the machine learning approach failed to provide an effective method for ascertaining a CSEM supervisee's likelihood of sexual recidivism.

In general, these findings were disappointing, given the level of effort PPSO expended in attempting to use the CPORT or build its own risk tools for CSEM risk prediction. The results should not be taken, however, as a denigration of the CPORT, which has been shown to be predictive in several studies assessing this risk instrument (Eke et al., 2019; Savoie et al., 2022; Seto & Eke, 2015). A variety of reasons could explain why the current research failed to replicate prior efforts highlighting the CPORT's predictive efficacy. First, MITRE's use of text mining and natural language processing precluded the generation of CPORT factors in a manner similar to that used by Seto and Eke (2015). Specifically, Seto and Eke (2015) combed through the collections of CSEM supervisees to assess the extent to which these collections indicated preferences of boys over girls. Moreover, Seto and Eke (2015) recommended using the CASIC to gauge a CSEM supervisee's sexual interests in children and teenagers. Unlike the approach taken by the CPORT's developers, the limited information available on the types or characteristics of the child pornography collections, the length of time engaged in child pornography activity, or the extent to which CSEM supervisees volunteered in roles with high access to children precluded the CASIC from being used to gauge pedophilic or hebephilic interests or the child pornography collections from being employed to ascertain sexual interests in boys over girls. Ultimately, MITRE relied on admissions to officers, treatment providers, and polygraph administrators to address the CPORT items related to boy over girl preferences or sexual interests in children or teenagers, and this reliance on admissions could have resulted in a diminishment in the predictive efficacy of the CPORT tool.

Other potential explanations for the study's results include the lower base rates for sexual recidivism for the federal CSEM sample (4.5 percent sexually recidivated) compared to study sample used by Seto and Eke (2015) to construct the CPORT (16 percent sexually recidivated). The low-risk skew of the federal CSEM population was also problematic. Over half the population had CPORT risk scores of 0 or 1, and about half received a score of 0 using the risk tool constructed by PPSO. The fact that so many CSEM supervisees garner few if any points using the various risk tools employed in this study and that relatively few sexually recidivated produces various challenges when it comes to developing and deploying an effective risk tool. In addition to these issues, differences between the U.S. and Canadian CSEM populations and the typical degradation in effect sizes when moving from the development to validation samples could explain the study's results (Copas, 1983; Soldino et al., 2021). Last, similar to other studies (see Soldino et al., 2021), the divergence in data quality between the Seto and Eke's (2015) CPORT development study and PPSO's data collection efforts might also explicate these findings. Basically, text mining

126,000 PDF and scanned documents cannot approximate in quality the work conducted by the CPORT's developers to manually code the instrument through a careful review of the case files. While text mining may have potential future applications in PPSO's research, it is possible that some types of information are better obtained through manual (i.e., non-machine) methods.

Future Directions for CSEM Research

While this initial attempt to develop a CSEM-based risk tool failed to generate an instrument that officers could use to supervise this key subpopulation of sex offenders, the research suggests several directions for future risk assessment development. First, several factors embedded within PPSO's risk tool (e.g., PCRA) were identified as being correlated with sexual recidivism for the CSEM population, including social problems associated with drug use, negative attitudes towards supervision, institutional adjustment, presence of criminal thinking style indicating denial of harm, and an assessment for domestic violence. Moreover, the presence of prior criminal behavior and in particular an arrest history for sex offenses were associated with sexual recidivism. At the very least, CSEM supervisees possessing one or more of these characteristics should be subjected to higher levels of supervision intensity compared to their CSEM counterparts without any of these attributes. In addition to these factors, PPSO has begun collecting data that might prove valuable for future efforts aimed at CSEM prediction. The fields currently include prior arrests for any type of sexual assault or production of child pornography, stranger victimization during any type of violent or sex offense, sexual assault of an unrelated male under the age of 17, and presence of valid Static-99 scores. Moreover, officers are being asked to collect information on whether the CSEM supervisee admitted to any hands-on sexual behavior irrespective of any arrests associated with this conduct. Information about the number of victims associated with this behavior is also being collated. The endeavor currently underway to obtain information on admissions of contact sex behavior represents a first-time national level effort to measure the extent to which CSEM supervisees have a history of contacting sexual offending that did not result in an official arrest. Future research efforts conducted by PPSO will attempt to ascertain whether these newly collected risk factors in conjunction with factors already scored by the PCRA might be combined to generate a new risk tool centered on CSEM supervisees.

Regarding the CPORT and CASIC, the viability of any future efforts aimed at using this risk tool depend upon the availability of information that is currently not being systematically collected during the supervision terms for persons convicted of CSEM offenses. Specifically, greater resources would be required at the sentencing stage to obtain information on the details of the child pornography collections gathered by CSEM supervisees. This information could then be used to address the CPORT and CASIC questions pertaining to the nature of the child pornography collections. Additionally, more methodical approaches would be required to address CASIC questions about volunteering in a role with high access to children and engaging in online sexual communications with minors. Purposefully attempting to extract the CASIC elements would enhance the feasibility of accurately addressing the CPORT question concerning sexual interests in children and teenagers. PPSO is exploring the viability of making changes to its case management system in order more uniformly and comprehensively to obtain data measuring the CPORT and CASIC elements.

Last, relying on FBI criminal history files to track the sexual recidivism behavior of CSEM supervisees has serious limitations. Essentially, the literature shows more than half of persons convicted of CSEM offenses engaging in contact sex behavior that never resulted in an actual arrest via admissions (Seto et al., 2011). Given the potential of many CSEM supervisees to engage in behavior that remains unknown to law enforcement officials, it might be advisable to move away from relying on official criminal history records and instead use polygraphs to track any self-reported behavior involving new sex crimes committed while on federal supervision. The practicability of using self-reporting methods should be more fully explored by federal probation.

Conclusion

This report sought to document PPSO's efforts to develop an actuarial tool that could be used to gauge the risk of sexual recidivism for persons convicted of CSEM offenses placed on federal supervision. The report delved into PPSO's attempts to employ the CPORT, including an explication of the challenges inherent in extracting the CPORT data elements and the efforts to overcome these challenges by contracting the data collection process to MITRE. While MITRE was able to successfully extract the CPORT factors for nearly 5,800 CSEM supervisees using text mining and natural language extraction methods, the instrument produced through this process failed to generate predictive indices similar to those reported by its developers (Eke et al., 2019; Seto & Eke, 2015). Given these findings, PPSO then detailed its efforts to construct its own in-house CSEM risk tool using various elements from the CPORT, PCRA, assessment fields, and criminal history files as well as applying machine learning to CSEM risk prediction. These in-house efforts, while somewhat successful, ultimately fell short of PPSO's goal of constructing a risk tool that could effectively differentiate CSEM supervisees by their levels of risk. In light of these findings, at this time PPSO cannot recommend using an actuarial tool outside the PCRA and policy guidelines related to supervising CSEM supervisees. PPSO will continue to engage in the problem of CSEM risk prediction, with particular emphasis on assessing whether some of the new risk factors currently being collected by officers can be combined with the PCRA elements to construct a risk tool that officers could apply to CSEM supervisees. Finally, PPSO will explore the feasibility of more uniformly and systematically collecting information that can be used to re-examine the CPORT's predictive effectiveness. We hope that these approaches will result in a risk tool that officers can use to effectively and judiciously supervise persons convicted of CSEM offenses on federal supervision.

References

Administrative Office of the U.S. Courts (2018). *An overview of the federal Post Conviction Risk Assessment*. Administrative Office of the U.S. Courts. https://www.uscourts.gov/sites/default/files/overview_of_the_post_conviction_risk_assessment_0.pdf.

Baber, L. (2010). Results-based framework for post-conviction supervision recidivism analysis. *Federal Probation*, *74*(3), 5.

Berk, R., Berk, D., & Drougas D. (2019). *Machine learning risk assessments in criminal justice settings* (pp. 1-178). Springer.

Black, C. (2018). *Predicting recidivism among an adult male child sexual abuse imagery offender population with the Child Pornography Offender Risk Tool Short Version (CPORT-SV): A New Zealand validation study*. [Unpublished master's thesis]. University of Canterbury.

Burkov, A. (2019). *The hundred-page machine learning book* (Vol. 1, p. 32). Quebec City, Canada: Andriy Burkov.

Cohen, T. H., & Spidell, M. C. (2016). How dangerous are they? An analysis of sex offenders under federal post-conviction supervision. *Federal Probation, 80*(2), 21- 32.

Cohen, T. H. (2018). Predicting sex offender recidivism: Using the federal Post Conviction Risk Assessment instrument to assess the likelihood of recidivism among federal sex offenders. *Journal of Empirical Legal Studies, 15*(3), 456-48.

Cohen, T. H., Lowenkamp, C. T., Bechtel, K., & Flores, A. W. (2020). Risk assessment overrides: Shuffling the risk deck without any improvements in prediction. *Criminal Justice and Behavior*, *47*(12), 1609–1629.

Copas, J. B. (1983). Regression, prediction and shrinkage. *Journal of the Royal Statistical Society: Series B (Methodological)*, *45*(3), 311-335.

DeLisi, M., Caropreso, D. E., Drury, A. J., Elbert, M. J., Evans, J. L., Heinrichs, T., & Tahja, K. M. (2016). The dark figure of sexual offending: New evidence from federal sex offenders. *Journal of Criminal Psychology, 6*(1), 3-15.

Eke, A. W., Helmus, L. M., & Seto, M. C. (2018). *Scoring guide for the Child Pornography Offender Risk Tool (CPORT)*: Version 2. https://www.researchgate.net/project/Child-Pornography-Offender-Risk-Tool-CPORT.

Eke, A. W., Helmus, L. M., & Seto, M. C. (2019). A validation study of the Child Pornography Offender Risk Tool (CPORT). *Sexual Abuse, 31*(4), 456–476.

Faust, E., & Motivans, M. (2015). Sex offenders in the federal correctional system: The consequence of heightened attention on increased certainty and severity of punishment. *Justice Research and Policy, 16*(1), 81–98.

Hanson, R. K., Thornton, D., Helmus, L.-M., & Babchishin, K. M. (2016). What sexual recidivism rates are associated with Static-99R and Static-2002R Scores? *Sexual Abuse, 28*(3), 218–252.

Hartshorn, S. (2016). *Machine learning with random forests and decision trees: A mostly intuitive guide, but also some python*.

Helmus, L. M., & Hanson, R. K. (2007). Predictive validity of The Static-99 and Static-2002 for sex offenders on community supervision. *Sexual Offender Treatment*, *2*(2), 1-14.

Johnson, J. L., Lowenkamp, C. T., VanBenschoten, S. W., & Robinson, C. R. (2011). The construction and validation of the federal Post-Conviction Risk Assessment (PCRA). *Federal Probation*, *75*(2), 16–29.

Johnson, J. L., & Baber, L. M. (2015). State of the system: Federal probation and pretrial services. *Federal Probation*, *79*(2), 34–40.

Lowenkamp, C. T., Johnson, J. L., Holsinger, A. M., VanBenschoten, S. W., & Robinson, C. R. (2013). The federal Post Conviction Risk Assessment (PCRA): A construction and validation study. *Psychological Services, 10*(1), 87–96.

Lowenkamp, C. T., Holsinger, A. M., & Cohen, T. H. (2015). PCRA revisited: Testing the validity of the federal Post Conviction Risk Assessment (PCRA). *Psychological Services, 12*(2), 149–157

Luallen, J., Radakrishnan, S., & Rhodes, W. (2016). The predictive validity of the Post-Conviction Risk Assessment among federal offenders. *Criminal Justice and Behavior, 43*(9), 1173–1189.

Markman, J. A., Durose, M. R., Rantala, R. R., & Tiedt, A. D. (2016). Recidivism of offenders placed on federal community supervision in 2005: Patterns from 2005 to 2010. *Bureau of Justice Statistics*, *2*(1), 1-16.

Megerdoomian, K., Branting, K., Horowitz, C., Marsh, A., Petersen, S., & Scott, E. (2019). Automated narrative extraction from administrative records. In *AIAS@ ICAIL* (pp. 38-48).

Pilon, A. J. M. (2016). *The predictive validity of general and offence-specific risk assessment tools for child pornography offenders' reoffending*. [Unpublished master's thesis]. University of Saskatchewan.

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. *Law and Human Behavior*, *29*, 615-620.

Savoie, V., Quayle, E., Flynn, E., & O'Rourke, S. (2022). Predicting risk of reoffending in persons with child sexual exploitation material offense histories: The use of child pornography offender risk tool in a Scottish population. *Sexual Abuse*, *34*(5), 568-596.

Serin, R. C., Lowenkamp, C. T., Johnson, J. L., & Trevino, P. (2016). Using a multi-level risk assessment to inform case planning and risk management: Implications for officers. *Federal Probation, 80*(2), 10–15.

Seto, M. C., Karl Hanson, R., & Babchishin, K. M. (2011). Contact sexual offending by men with online sexual offenses. *Sexual Abuse, 23*(1), 124–145.

Seto, M. C., & Eke, A. W. (2015). Predicting recidivism among adult male child pornography offenders: Development of the Child Pornography Offender Risk Tool (CPORT). *Law and Human Behavior*, *39*(4), 416.

Seto, M. C., & Eke, A. W. (2017). Correlates of admitted sexual interest in children among individuals convicted of child pornography offenses. *Law and Human Behavior*, *41*(3), 305.

Soldino, V., Carbonell-Vayá, E. J., & Seigfried-Spellar, K. C. (2021). Spanish validation of the Child Pornography Offender Risk Tool. *Sexual Abuse*, *33*(5), 503-528.

United States Sentencing Commission. (2012). *2012 Report to the Congress: Federal child pornography offenses*. United States Sentencing Commission.

United States Sentencing Commission. (2021). *Federal sentencing of child pornography non-production offenses*. United States Sentencing Commission.

Walters, G. D. (2013). The psychological inventory of criminal thinking styles (PICTS) professional manual. *Center for Lifestyle Studies*.

Wolak, J., Finkelhor, D., & Mitchell, K. J. (2005). *Child pornography possessors and the internet: A national study*. National Center for Missing & Exploited Children.

Wolak, J., Finkelhor, D., & Mitchell, K. J. (2009). *Trends in arrests of "online predators."* Crimes Against Children Research Center.

Burnout Prevention for Federal Probation and Pretrial Services Officers

Robert McMackin
Aspire Health Alliance
Joseph LaFratta
Scelte Consulting

THE TERM BURNOUT has become ubiquitous, used freely to describe an assortment of conditions or presentations. Yet a point of agreement among its various uses is that it refers to a negative physical and/or emotional state marked by exhaustion. This is an exhaustion that is beyond fatigue and will not be cured by a good night's sleep. An exhaustion that can be bone-crushing, leaving the sufferer weary, depleted, and feeling alone. This sense of exhaustion is also noted in the burnout professional literature. In the models of burnout put forward by the principal theorists in the field (Maslach & Leiter, 1997, 2005; Demerouti & Bakker, 2007; Demerouti et al., 2002), exhaustion is the first factor noted in both burnout models. Both groups agree that it is an exhaustion brought on by intense cognitive, emotional, and physical job demands.

Maslach and Leiter (2021) describe two other burnout factors: 1) cynicism and a depersonalization or mental distancing from one's job and 2) reduced professional efficacy. Demerouti and Bakker have the additional factor of disengagement in their model of burnout, stating that they do not include a professional efficacy dimension as they view that as an outcome of burnout rather than a "core dimension" (Demerouti & Bakker, 2007).

For our work we have adopted the model of burnout put forward by Demerouti and Bakker for two reasons:

1. It can be measured by the Oldenburg Burnout Inventory, which is open access, and we were able to administer it to federal probation and pretrial services officers.
2. It is linked to their Jobs Demands-Resources theory, which we have modified as a component in a model for a district level burnout intervention.

Stress, Trauma Exposure, and Burnout

The impact of stress on health has long been recognized, being linked to a range of physiological and psychological states. Stress, particularly acute stress, sets off a strong hormonal response, which is a normal part of a stress adaptation. This activation typically "causes secretion of glucocorticoids, which act on multiple organ systems to redirect energy resources to meet real or anticipated demand" (Herman et al., 2016, p. 1). While such a stress reaction is adaptive when one has to gear up for an immediate response, such as fight or flight, it is maladaptive to have one's body frequently bathed in such hormones. *Allostatic load* is the term used to describe the cumulative stress we may experience. Arline Geronimus describes the effects of constant or repeated exposure to stress, including the stresses of poverty and racism, as leading to a premature aging or "weathering" of one's body (Geronimus et al., 2006).

It should also be remembered that our brains are a type of tissue, and Agnese Mariotti points out that "chronic stress is linked to macroscopic changes in certain brain areas, consisting of volume variations and physical modifications of neuronal networks" (Mariotti, 2015, p. 2). Some stress is adaptive and helps us respond to threats or challenges we encounter, but as the 16th century physician Paracelsus explained—*anything in excess is poison*—and that is most assuredly true for stress.

As noted, the job of a federal probation and pretrial services officer as well as other frontline law enforcement officers is critical and by its nature exposes one to stress. Figley, who has researched the impact of stress on health care providers, identified "compassion fatigue" as an outcome of cumulative stress and described it as "the cost of caring" (Figley, 1995). For probation and pretrial services officers to be effective in their positions, they must pay the price of caring. The capacity for empathy allows one to connect with another in a human manner, to place oneself in the shoes of another. Yet this can also be a two-edged sword if one does not have good boundaries or good supervision. One's empathic capacity can be a contributor to or down payment on that cost of caring, leading to the exhaustion and disengagement of burnout.

Figley speculated that the exhaustion and disengagement one experiences in stressful professions may be protective coping mechanisms to help manage the emotional costs of working in such difficult situations. He called this "compassion fatigue." For example, an officer completing a presentence report on an individual charged with child pornography or sexual assault may have to view some of the evidence in the case, including videos of children being raped. It is not hard to imagine

how much tension viewing such material would create. A natural response would be for the officer to disengage or disassociate from the activity as a type of protective mechanism, or to simply become exhausted by the process. Additionally, viewing such material may trigger an understandably angry response in the officer, flooding the officer's body with the stress hormones mentioned above.

None of us live in a safe, frictionless world. Trauma exposure affects almost all of us. The Adverse Childhood Events (ACES) population study showed 61 percent of the U.S. population experiences an ACES, such as experiencing or witnessing abuse or neglect or having a close relative commit suicide (CDC, 2021) and the World Health Organization reported the general trauma exposure rate at over 70 percent (Kessler et al., 2017). The majority of us have some trauma exposure in our lives, and when one works in law enforcement there is an overlay of unavoidable trauma exposures. These exposures are unavoidable because they are part of the job. In a way they are exposures each person who entered the field signed up for when they took the job, likely without fully knowing their severity at the time. They are the unavoidable frictions of the profession, and at times the heat caused by such frictions may ignite.

A variety of terms have been used, frequently interchangeably, to describe the impact of being exposed to trauma through work. These terms generally include secondary trauma, vicarious trauma, and post-traumatic stress disorder (PTSD). An additional term not seen in the work stress or work trauma exposure literature, which we nonetheless consider important, is complex post-traumatic stress disorder (C-PTSD), which refers to multiple trauma exposures. While often used interchangeably, these terms represent distinct psychological phenomena, and we propose they fall into two distinct exposure groupings: Indirect trauma exposure (secondary trauma & vicarious trauma) and direct trauma exposure (PTSD & C-PTSD).

Secondary trauma and vicarious trauma are considered indirect exposures because they entail being exposed to a traumatic event via the experience of another, where the professional develops similar symptoms to the clients. Most descriptions of secondary trauma stress the professional mirroring the client's PTSD symptoms, while vicarious trauma literature often notes enduring changes in the professional's cognitive or affective state. The descriptions of secondary trauma and vicarious trauma strongly overlap, which is why many view the differences as semantic rather than actual.

PTSD and C-PTSD result from direct trauma exposure where one is exposed to potentially traumatic events either once or repeatedly through one's work. As mentioned earlier, if a probation officer is preparing a presentencing report on an individual who sexually assaulted a child, the officer may be required to view the evidence against the offender. Hearing a person describe being raped as a child is vastly different from viewing a ten-minute video of a child being raped or viewing two or ten or twenty such videos. Similarly, for a police officer to hear from another officer at shift change about a fatal accident they responded to is vastly different from that officer responding to a fatal accident.

The fatigue related to the indirect or direct trauma exposure related to work is gradual, as is the more enduring burnout resultant from those same exposures. It is the very gradualness of this progression which makes it so easy to miss for both supervisors and the officers themselves. We have all used the phrase describing someone as "a bit crisp around the edges." Unbeknownst even to ourselves, we are making an informal assessment of that colleague's burnout risk. It is an assessment generally made in jest, yet it is exceedingly, even deadly, serious, with law enforcement officers being 54 percent more likely to die of suicide than those in other professions (Voilanti & Steege, 2021).

The model of burnout we use in our work is based on the research of Demerouti et al. This model identifies two components to burnout: Exhaustion and Disengagement. They state, "Each burnout dimension is differentially related to specific short-term consequences of strain: Exhaustion is primarily related to mental fatigue, whereas disengagement is primarily related to satiation and the experience of monotony" (Demerouti et al., 2002, p. 423). Importantly, they identify four factors that are antecedents or precursors to burnout. These factors are:

1. *Mental Fatigue:* "The impairment of mental and physical functional efficiency, depending on the intensity, duration, and temporal pattern of the preceding strain." This impairment can eventually lead to poor performance, loss of concentration, and exhaustion.
2. *Monotony:* "A state of reduced activation (within the individual) which may occur during repetitive task performance with a narrow field of attention under monotonous job conditions." Monotony can disappear with a change of activity.
3. *Satiation:* "A state of nervously unsettled, strongly emotional rejection of a (structurally) repetitive task or situation in which the experience is of 'marking time' or 'not getting anywhere.'" Like monotony, satiation can disappear with a change of activity.
4. *Stress Sensations:* These are "complex psycho-physiological reactions to unacceptable, conflicting, or especially threatening demands that may result from a perceived over- or under-load (e.g., time pressure), causing frustration of personal goals and aversive consequences." Long stress sensation can lead to "chronic stress sensations, exhaustion, shifts of the aspiration level, and finally to health impairments" (Demerouti et. al., 2002, p. 425).

These antecedents to burnout can be seen in any profession, from a law enforcement officer to a factory worker. As with any precursor to illness, *an ounce of prevention is worth a pound of cure.* These potential burnout precursors provide us with a partial road map to the issues that should be addressed in a burnout prevention plan, which will be discussed later in this paper.

Burnout Risk Among Federal Probation and Pretrial Services Officers

The Oldenburg Burnout Inventory (OLBI) is a sixteen-item questionnaire that examines the two-factor model of burnout (Demerouti & Bakker, 2007). Each item on the scale is composed of a statement that subjects respond to on a four-point scale (strongly agree, agree, disagree, and strongly disagree). Responses to eight statements form a disengagement subscale, and the other eight form an exhaustion subscale. The results of all sixteen items taken together form a full burnout scale. An example of a disengagement item is: "Lately, I tend to think less at work and do my job almost mechanically." An example of an exhaustion item is "During my work, I often feel emotionally drained" (MDApp, 2020).

In addition to the burnout scales, some OLBI items consider a person's degree of positive work engagement, which Schaufeli and Bakker view as the antithesis of burnout:

"burnout and engagement are considered each other's opposites, particularly as far as exhaustion and vigor, and cynicism and dedication are concerned" (Schaufeli & Bakker, 2004, p. 296). Items that tie into a subject's positive engagement in work include: "I find my work to be a positive challenge" and "When I work, I usually feel energized" (MDApp, 2020).

As part of two Federal Judicial Center trainings the authors presented on developing a trauma-informed wellness program for U.S. probation and pretrial officers in 2021, participants were asked to complete the OLBI at the conclusion of the training. Participants came from federal probation and pretrial districts throughout the country, and no identifying information was asked except whether the person was an officer or a supervisor. Eighty-nine staff persons completed the inventory, including 48 officers and 41 supervisors. On the OLBI, a score below 1.63 represents low burnout risk, 1.64 to 2.67 represents moderate risk, and above 2.68 represents high burnout risk. Table 1 shows the OLBI scores for officers and supervisors.

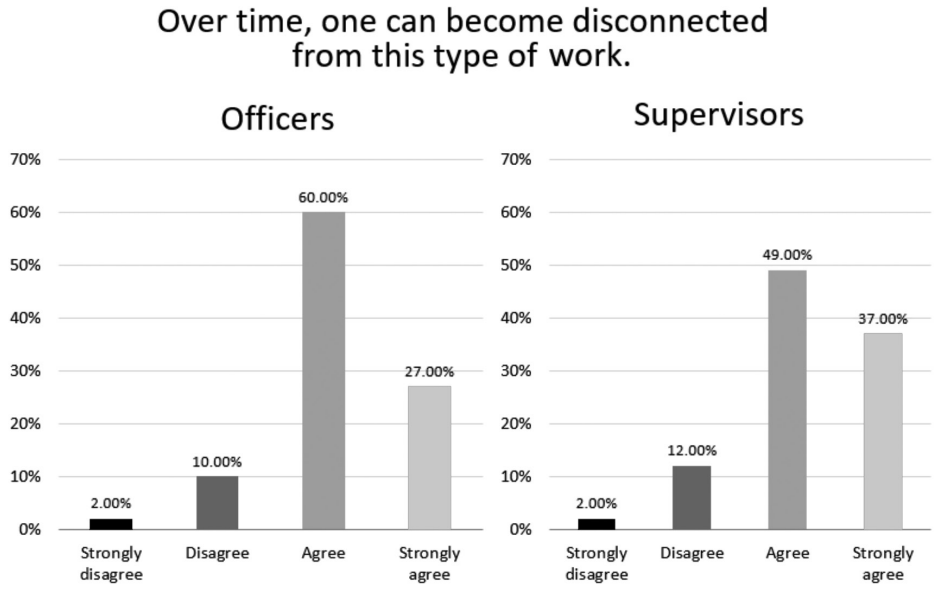
TABLE 1
Oldenburg Burnout Inventory Scores

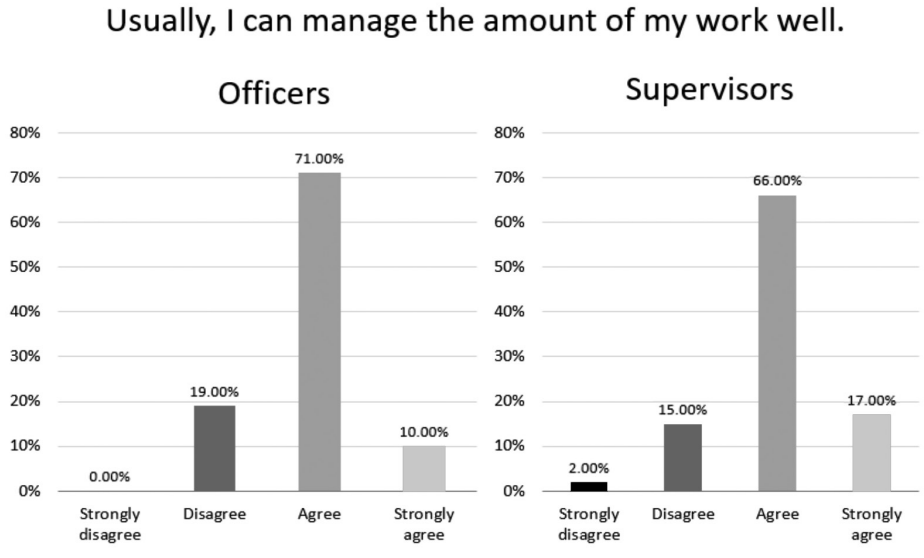
	Officers	Supervisors
Full Scale	2.52	2.45
Exhaustion	2.59	2.53
Disengagement	2.45	2.38

The Oldenburg scores clearly show both officers and supervisors scoring at the upper end of the moderate burnout risk area. Yet even more striking was the response pattern on individual items. On the Exhaustion Scale, a strong majority of officers (87 percent) and supervisors (83 percent) agreed or strongly agreed with the following statement, *"Over time one can become disconnected from this type of work."* Additionally, 58 percent of officers and 61 percent of supervisors agreed or strongly agreed with, *"After my work, I usually feel worn out and weary."* The respondents almost universally agreed that the responsibilities of an officer can take a toll, that they may be subject to that toll to some degree, and that some already feel the price. Yet equally striking was how positively engaged most were in their work, with 85 percent of officers and 75 percent of supervisors agreeing with *"I can tolerate the pressure of my work very well,"* and 81 percent of officers and 83 percent of supervisors agreeing with *"Usually, I can manage the amount of my work well."*

The results on the Disengagement Scale were equally striking. A majority of respondents acknowledged some degree of disengagement, with 54 percent of officers and 44 percent of supervisors agreeing with the statement *"It happens more and more often that I talk about my work in a negative way,"* and 44 percent of officers and 59 percent of supervisors agreeing with the statement *"I feel more and more engaged in my work."* Yet, while acknowledging how one can become disengaged from the work of a federal probation and pretrial services officer, 72 percent of officers and 95 percent of supervisors agreed that *"I find my work a positive challenge,"* and only 28 percent of officers and 44 percent of supervisors agreed with the statement, *"Lately, I tend to think less at work and do my job almost mechanically."*

The overall scores of federal probation and pretrial services officers show a group of individuals in a high-risk job that can take a toll on them both physically and mentally. Nonetheless, the vast majority of officers are highly committed to their positions and find their work a positive challenge in which they can take pride. Still, it must be remembered that all things distribute normally; while the majority of officers appear to do well in their positions from a burnout perspective, a sizeable minority, likely in the 10 percent range, are struggling to some degree. It should also be remembered that while the majority appear to be doing well, all remain at risk due to the stresses of the job and all can benefit from a work environment that actively tries to prevent burnout.





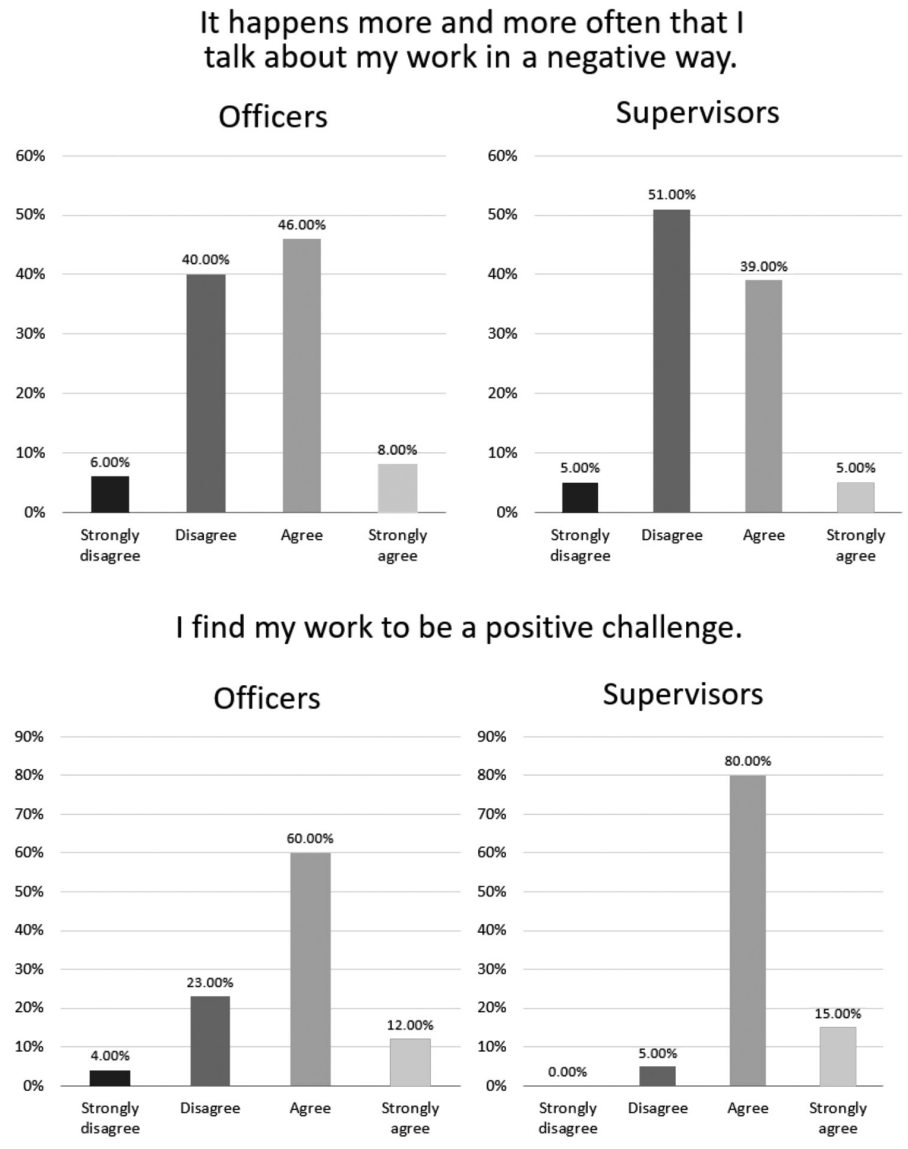
The Job Responsibilities-Resources Model

Nearly all jobs have performance expectations for employees, and hopefully there are also resources available to help the employees meet those expectations. Expectations and resources are in a delicate balance that is not always in synch—at times the demands may be too high and the resources too low or vice versa.

The Job Demands–Resources (JD-R) model suggests that all professions have demands that can become risk factors, leading to job stress or strain, as well as resources that can become protective factors, mitigating against such stress. Job demands may be broadly viewed as "The physical, psychological, social, and organizational aspects of a job that requires sustained physical, cognitive, and emotional effort and skill" (Bakker & Demerouti, 2007, p. 312). Job resources are those "physical, psychological, social and organizational aspects of the job that are either/or: functional in achieving work goals; reduce job demands and the associated physiological and psychological costs; stimulate personal growth, learning, and development" (Bakker & Demerouti, 2007, p. 312). Demands can be viewed as a positive challenge or a negative hindrance. Resources should help one meet demands, thus lowering the potentially toxic effects of demands.

Generally, when people accept a job, they also agree to a job description that outlines their responsibilities. By willingly accepting that job, the person also accepts all of its related responsibilities. As a result, the onus of fulfilling those responsibilities falls upon the employee, with the employer having the responsibility of providing the resources required in order for the employee to be successful in fulfilling those duties. Therefore, we have modified the Job Demands-Resources (JD-R) model to the Job Responsibilities-Resources (JR-R) model. We view this as giving the employee greater personal agency as one who is fulfilling responsibilities, not meeting demands. It is akin to those who have experienced significant trauma viewing themselves as survivors rather than as victims of trauma. Additionally, we believe this model places the employer and employee on a more equal footing, with shared responsibilities and accountability.

The JR-R model does not reduce the risk of stress on a job and potential burnout. Responsibilities with inadequate resources remain a recipe for disaster: a type of unfunded mandate that an employee may never be able to get out from under and the weight of which may eventually be crushing.



Developing a Burnout Prevention Plan

The two-factor model of burnout identifies exhaustion and disengagement as the dominant features in burnout. This model identifies four precursors to burnout: mental fatigue, monotony, satiation, and stress sensations. Unsurprisingly, research has indicated a strong correlation of exhaustion with mental fatigue and stress sensations, and disengagement with satiation, monotony and stress sensations, although all four burnout antecedents contribute to each factor. Figure 1 (next page) shows the relationship of the two burnout factors to the four burnout antecedents (Demerouti et al., 2002).

Understanding the relationship of the four antecedents to burnout is important, as it provides a partial blueprint for the areas that should be addressed in a burnout prevention plan. The remaining elements of that blueprint are provided by examining those job resources that mitigate the negative impacts of the four burnout antecedents, specifically social support, autonomy, a positive supervisory relationship, and constructive performance feedback (Bakker et al., 2005). It is our experience that a clear and shared mission among all staff, officers and supervisors is an additional stress-mitigating factor for those in the law enforcement field.

Bakker et al. have noted that autonomy is the protective factor that most fully buffers job demands and "The level of exhaustion and of

cynicism was elevated particularly when job demands (work overload, emotional demands, unfavorable work conditions, and work-home interference) were high and job resources (autonomy, social support, high quality relationship with the supervisor, and performance feedback) were lacking" (Bakker et al., 2005, p. 176). While Bakker et al. refer to job demands, we prefer to focus on job responsibilities, seeing all the job demands mentioned above, except for work-home interference, as related to the employee's job responsibilities and factors the employee should have some ability to address within a healthy work environment. The employee should also be able to address work-home interference, but we consider that a factor that may be addressed as part of a personal wellness program.

At times wellness programming can be seen as a way to promote a healthy workplace environment and reduce the risk of burnout. While we agree with this, we view most wellness programs as highly individually based and often containing only one factor directly related to one's work (occupation) and up to seven factors not related to one's work environment (e.g., finances, environment, social, etc.). We believe positive wellness should be promoted as a complement to a workplace burnout prevention plan.

We will outline the five steps we consider critical to developing a burnout prevention plan for a federal probation and pretrial district based on the two-factor model of burnout. We recommend consultation be used to implement this plan.

- Step 1: The first step is a full department training (officers, supervisors, & support staff) on the stresses inherent in their jobs. This training would look at the responsibilities of all staff, skills or competencies required to perform the various jobs, types of trauma and stress that one encounters in their role, an understanding of direct and indirect trauma exposure, ways of managing stress related to such exposure, and an understanding of how the workplace can be affected. At the conclusion of the training, the group will be asked to complete the Oldenburg Burnout Inventory.
- Step 2: The second step is a training with the same group. The two-factor model of burnout is presented with a discussion of the OLBI results for the group. The mission of the district and the responsibilities of all staff are discussed. The four antecedents that contribute to burnout are discussed, as are the four mitigating factors and the importance of a clear mission. The group considers and lists resources they believe can be implemented to reduce stress related to the four burnout precursors.
- Step 3: The third step is a meeting with supervisory staff to evaluate which resources could be enhanced to reduce the burnout risk for staff, allowing them to fulfill their responsibilities most effectively.
- Step 4: The fourth step is for a burnout prevention plan to be presented to the full staff based on the feedback from support staff, officers, and supervisors. Modifications to existing protocols are discussed (e.g., a possible change in the on-call schedule, caseload expectations).
- Step 5: The final step calls for a quarterly meeting with supervisors and officers to review how the burnout plan is being implemented. An annual on-site review meeting with the full staff is recommended.

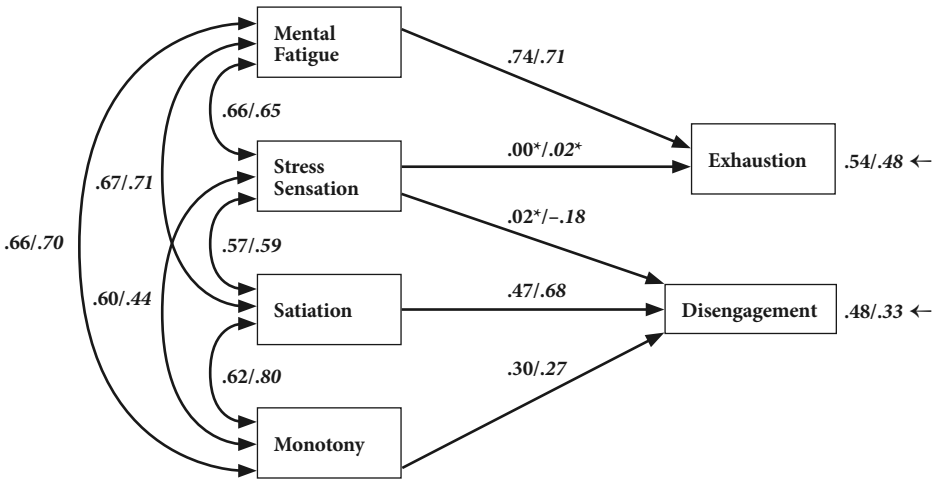
The process outlined above is designed to be both informative and inclusive. For it to be successful, there must be a common understanding among all staff of the stresses of the job, the responsibilities and expectations for both officers and supervisors, and the resources available to fulfill such responsibilities.

Conclusion

Federal probation and pretrial services officers and others in law enforcement serve in critical occupations designed to support the welfare and betterment of the wider community. They also serve in occupations where they are exposed to traumatic material that is unavoidable, as it is a function of their jobs. It is not a situation where *what does not break you makes you stronger.* None of us are immune to the impacts of traumatic exposures, and cumulative exposures only intensify the effects; they do not mitigate them. It is normal to want to reduce the stress of such exposures, and this can be done in a positive manner by talking with friends and colleagues, focusing on the positives of the job and the successes, good supervision, or healthy humor. But it can also be done in unhealthy manners such as by excessive substance use, aggressive behavior, cynicism, or dark humor.

Those who have entered the fields mentioned above have voluntarily assumed the responsibilities inherent in their jobs, yet they also deserve the resources required to do their job and to not become physically or psychologically damaged in the process. Burnout is a term referring to the exhaustion and disengagement individuals may develop over time due to the stresses of certain jobs. While this burnout for many may be a protective mechanism to dampen the impact of the traumas to which one is exposed, it is not a healthy coping mechanism. In law enforcement, it is the responsibility of the supervising agencies for whom officers work and for the entire communities whom they serve and protect to develop interventions and resources that lessen the impact of those factors known to contribute to burnout. This paper outlines a

FIGURE 1.



Standardization solution of the model of short-term effects of strain and burnout for human service professionals (N = 149) and production workers (N = 145; in italics). All parameters, except those marked with an asterisk (*), are significant at the $p < .05$ level.

burnout protection plan for federal probation and pretrial services officers, but it can also be adapted for other law enforcement or courthouse occupations that tend to be high stress.

References

Demerouti, E., Bakker, A. B., Nachreinre, F., & Ebbinghaus M., (2002). From mental strain to burnout. *European Journal of Work and Organizational Psychology, 4,* 423-441.

Demerouti, E., & Bakker, A. B. (2007). *The Oldenburg Burnout Inventory: A good alternative to measure burnout (and engagement).* Utrecht University and Erasmus University Rotterdam.

Bakker, A. B., & Demerouti, E. (2006). The Job Demands-Resources model: State of the art. *Journal of Managerial Psychology*, *22*(3), 309-328.

Bakker, A. B., Demerouti, E., & Euwema, M. C. (2005). Job resources buffer the impact of job demands on burnout. *Journal of Occupational Health Psychology, 10*(2), 170-180.

Centers for Disease Control and Prevention (April 6, 2021). Preventing adverse childhood experiences |Violence Prevention| Injury Center| CDC. www.cdc.gov/violenceprevention/aces/fastfact.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fviolenceprevention%2Facestudy%2Ffastfact.html+%29.

Figley, C. R. (1995). Compassion fatigue as secondary traumatic stress disorder: An overview. In C. R. Figley (Ed.), *Compassion fatigue: Coping with secondary traumatic stress disorder in those who treat the traumatized* (pp. 1-20). New York, NY: Brunner-Routledge.

Geronimus, A. T., Hicken, M., Keene, D., & Bound, J. (2006). "Weathering" and age patterns of allostatic load scores among Blacks and Whites in the United States. *American Journal of Public Health, 96*(5), 826-833.

Herman, J. P., McKlveen, J. M., Ghosal, S., Kopp, B., Wulsin, A., Makinson, R., Scheimann, J. Myers, B. (2016). Regulation of the hypothalamic-pituitary-adrenocortical stress response. *Comparative Physiology, 6*(2), 603–621.

Kessler, R. C., Aguilar-Gixiola, S., Alanso, J., Benjet, C., Bromet, E., & Cardoso, G. (2017). Trauma and PTSD in the WHO World Mental Health Surveys. *European Journal of Psychotraumatology,* Vol. 8.

Mariotti, A. (2015). The effects of chronic stress on health: New insights into the molecular mechanisms of brain–body communication. *Future Science OA*, *1*(3).

Maslach, C., & Leiter, M. P. (1997). *The truth about burnout.* San Francisco: Jossey-Bass.

Maslach, C., & Leiter, M. P. (2005). Understanding the burnout experience: Recent research and its implications for psychiatry. *World Psychiatry*, *15*(2), 103-111.

MDApp. (2020). Oldenburg Burnout Inventory. https://www.mdapp.co/oldenburg-burnout-inventory-olbi-calculator-606/.

Schaufeli, W.B., & Bakker, A.B. (2004). Job demands, job resources and their relationship with burnout and engagement: A multi-sample study. *Journal of Organizational Behavior*, *25*, 293-315.

Voilanti, J. M., & Steege, A. (2021). Law enforcement worker suicide: An updated national assessment. *Policing, 44*(1), 18–31.

Self-directed Workbooks: Evaluating Their Efficacy in a U.S. Probation Setting

Ralph C. Serin and Stephanie M. Biro
Carleton University

NOTWITHSTANDING A SLIGHT decrease of .9 percent in the adult population of parolees or probationers in 2019, according to the Bureau of Justice Statistics, there remain just under 4.4 million adults under community supervision in the 50 states and Washington, D.C. (Oudekerk & Kaeble, 2021). Presented differently, this represents that 1 in 59 adults in the U.S. report to probation or parole officers and must abide by certain supervision conditions to avoid incarceration. Probation is over-represented, accounting for about 80 percent of those under community supervision, compared to parolees, who represent the remaining 20 percent.

The supervision of clients within the community after sentencing has been shown to be significantly less costly than incarceration, with incarceration costing eight times more (i.e., $34,770 annually per incarcerated individual on average versus $4,392 per community supervised individual; U.S. Courts, 2017). Given the proportion of individuals under community supervision, as well as the difference in cost when compared to incarceration, the continual improvement of community interventions to promote prosocial behavior change should continue to be a major focus for correctional agencies and policymakers. Based on these numbers, even a minimal increase in the effectiveness of both case management techniques and intervention strategies will produce significant cost savings for correctional agencies, enhancing decarceration efforts.

Over the last few decades, considerable research has provided a theoretical perspective to understanding criminal behavior, as well as a set of principles that intend to guide the management and treatment of correctional clients (Bonta & Andrews, 2017), including those under community supervision (e.g., Bourgon & Gutierrez, 2012). Importantly, it seems that the transition into a criminal lifestyle and the transition out of such a lifestyle are quite different processes (e.g., Serin, Lloyd, & Hanby, 2010). While current interventions are successful at reducing the risk of recidivism (Chadwick et al., 2015; Robinson et al., 2012), there is considerable room for improvement when it comes to understanding what motivates an individual to change behavior (i.e., lead a prosocial life), as well as the interventions that agencies can provide to assist in this process.

Beyond cost savings from decarceration, both ideology and research have led to an evolution in community supervision practice. In terms of philosophy, over the past decade, supervision practices have evolved to a greater emphasis on officers being change agents (Bourgon et al., 2011) or coaches (Lovins et al., 2018), balancing the more traditional role of surveillance (Viglione et al., 2017). Against this backdrop, the recent COVID-19 pandemic has significantly changed community supervision client contact, suggesting self-directed workbooks may have appeal in supporting and facilitating change-related work by officers. Of note, previous research has suggested journaling is an effective intervention to target general recidivism among incarcerated individuals (e.g., Proctor et al., 2012). The present paper describes the findings from a small random assignment pilot study in a U.S. probation site comparing client outcomes when officers did and did not use self-directed workbooks.

Evidence-Based Practice and Policy

EBP is the notion that policy and practice (i.e., including decision-making) should align with current empirical research in order to best achieve the desired outcomes and to make the most efficient use of financial resources (Taxman, 2012). Specific intervention skills are encompassed in the concepts of EBPs, often referred to as core correctional practices (CCPs) in community corrections (Dowden & Andrews, 2004). Briefly, models of community supervision that adhere to evidence-based practices attempt to move away from surveillance-based and brokerage of services activities toward a model where officers serve as an active participant in the delivery of rehabilitative services. This emphasis on officer involvement in rehabilitative work has often been referred to as being a change agent (Bourgon et al., 2011) or coach (Lovins et al., 2018).

Encouragingly, evidence-based supervision is associated with reductions in recidivism compared to the status quo training that is provided to community supervision officers (e.g., Robinson et al., 2012). A recent meta-analysis of training programs aimed at enhancing the use of evidence-based practices in community supervision found that clients supervised by

[1] Correspondence concerning this article should be addressed to Ralph C. Serin, Ph.D., C.Psych., Department of Psychology, Carleton University, 1125 Colonel By Drive, Ottawa ON K1S 5B6, Tel: 613-520-2600, ext. 1557, ralph.serin@carleton.ca

officers trained in Core Correctional Practice demonstrated an approximately 13 percent reduction in recidivism, compared to clients supervised by officers who did not receive the supplemental training (Chadwick et al., 2015). Unsurprisingly, such research and changing philosophy about community corrections has led to increased interest in the development of standardized training curricula for community supervision.

Successful Reintegration

When individuals under community supervision fail to abide by the conditions assigned by the courts or parole boards, they are at risk of being returned to jail or prison. These returns to jail, or unsuccessful exits, are significant drivers of incarceration (Oudekerk & Kaeble, 2021). For probationers, about 16 percent of the unsuccessful exits resulted in reincarceration, and 10 percent were otherwise unsatisfactory. For parolees, 29 percent returned to incarceration (Oudekerk & Kaeble, 2021). The Council of State Governments (CSG) similarly highlighted the impact of supervision violation behavior on prison admissions, noting that on any given day there are 280,000 people in prison because of a supervision violation, which is nearly 1 in 4 (CSG, 2019). In addition, estimates indicate that approximately 48 percent of probation violations and 64 percent of parole violations are for technical reasons (e.g., breach of conditions and not necessarily new criminal behavior on its own), amounting to costs of $2.8 billion to the criminal justice system (CSG, 2019).

The high number of individuals that fail to complete their community supervision (i.e., commit a new offense or incur a technical violation) is of concern, as it results in many new crimes being committed, increased victimization rates, and an increase in spending through court costs and incarceration costs. This situation suggests that there is room for improvement regarding the efficacy of current supervision practice. One potential option is to use interactive journals in the reentry process, as they have demonstrated their effectiveness with incarcerated individuals (e.g., Proctor et al., 2012).

Influencing Client Change

Interventions that align with the RNR principles have continually been demonstrated to significantly reduce recidivism (e.g., Bonta & Andrews, 2017; Lowenkamp et al., 2006; refer to Smith et al., 2009, for a systematic review). Criminogenic needs (Andrews & Bonta, 2017) have been widely accepted to be key targets whereby attenuation of these needs improves client outcomes and would therefore seem to be important topics to be included in client workbooks. In addition, staff's ability to build strong working relationships with clients (Ross et al., 2008) and high fidelity of the intervention (Andrews & Dowden, 2004; Lowenkamp et al., 2006) are also important to influence client change.

While many structured group-based programs are didactic or psychoeducational, strategies such as self-directed journaling and experiential disclosure have been used to influence client change across a variety of settings and for different client concerns (Frattaroli, 2006, Proctor et al., 2012, Richards et al., 2000). Of these two strategies, the experiential disclosure is much more unstructured than self-directed journaling. A structured and experiential writing process known as Interactive Journaling®, based on the Transtheoretical Model, cognitive behavioral therapy, and motivational interviewing principles, aims to reduce substance abuse and substance-related behaviors (e.g., recidivism) by motivating and guiding individuals towards positive life change. Interactive Journaling® is also included in SAMHSA's National Registry of Evidence-based Programs and Practices (SAMHSA, 2014). Research has indicated that Interactive Journaling® with incarcerated individuals may be successful at reducing criminal thinking (Folk et al., 2016), substance use (Scheck et al., 2013), and recidivism (Proctor et al., 2012), as well as receiving positive feedback from participants (e.g., Scheck et al., 2013). Structured journaling has been used as both a self-administered resource or as a complement to individual or group counselling (Davidson et al., 2008). It is against this background that we developed self-directed workbooks, wanting to create materials that had structured content regarding criminogenic needs and that required clients to complete written work, optimally prompting greater self-awareness.

This project consisted of developing and piloting 5 self-directed workbooks for use by probation officers. The primary research goal was to examine their effectiveness at reducing recidivism and technical violations. The workbooks were designed for use with lower risk clients, consistent with the Risk and Need principles (Andrews & Bonta, 2017) or as preparation for higher risk clients to participate in formal intervention. The hypothesis was that clients who participate in self-directed, criminogenically relevant efforts would have lower rates of community supervision failure.

Methods

Participants

For the purpose of the current study, a sample of 32 probation clients from a probation site in Texas was recruited in person between January 2017 and April 2017. Participants were low-moderate risk probationers who had previously been assessed using the Texas Risk Assessment System (Criminal Justice Connections, 2015), which combines and interprets an individual's criminal history and criminogenic needs to create the most effective case management plan. As such, this is a conservative test of the efficacy of self-directed workbooks. Each participant was randomly assigned to either the control (i.e., current probation practices only) or experimental group (i.e., self-directed workbooks alongside current probation practices).

Of the final sample (n = 32), seven identified as female and the remaining 25 identified as male. The participants' ages ranged from 22 to 59 (M = 36.97, SD = 10.15). The sample comprised approximately 84 percent who identified as Caucasian (n = 27), while the other 16 percent (n = 5) identified as other. As well, 52 percent (17) of the sample identified their ethnicity as Hispanic. Approximately 68 percent of participants indicated that they had a grade twelve education or less (M = 12.50 years), and 59 percent had a previous offense that was a felony.

(See Table 1, next page)

Measures

Intervention. A set of five self-directed workbooks (i.e., clients complete the exercises within each workbook at their own pace and with little staff contact) that make up the Client Handbook Series was used as the intervention in the current study. The workbooks were based on criminogenic needs identified through previous research (e.g., Bonta & Andrews, 2017), and each workbook targets a different factor (e.g., motivation, anger, criminal peers, criminal attitudes, and substance abuse) essential to managing criminal behavior. These workbooks are designed to assist clients in reflecting on the different choices and thought processes that have led them to their involvement in the criminal justice system. Clients are to work through them at their own pace. The goal of the workbooks is to provide alternative, more prosocial ways

of thinking and behaving in a variety of situations specific to the individual that result in successful community reintegration (i.e., the individual desists from crime). Workbooks present information and then provide reflective opportunities for the client to apply this information to the client's situation as a written exercise. The format is introduction, key issues, examples and worksheets, summary.

The workbooks were developed according to the Flesch-Kincaid scale, so that individuals with lower level reading skills would be able to complete them (see Table 2). Scores computed by this formula range from 0 to 100, where higher scores indicate reading material that is easier to read. Workbooks vary slightly in length (Motivation – 15 pages; Anger – 12 pages; Criminal Attitudes – 8 pages; Peer Relationships – 11 pages; Substance Abuse – 12 pages).

Within each workbook, the content is organized hierarchically from basic to more advanced. More specifically, the first few pages of the workbooks help the clients learn the core concepts and reflect on the choices and behaviors that have led to their current situation. In the next set of pages, the clients apply the core concepts to their specific life situations through a variety of activities (e.g., "make a list of three factors that hold the highest risk for you" or "what are your reasons for abstaining completely"). The clients are encouraged to reflect on these experiences and their responses. Finally, each workbook ends with a summary of what they've learned. Skill development is a process (i.e., awareness of new concepts, learning those new concepts, applying those new concepts to one's everyday activities), and these workbooks attempt to help to build the skills these individuals require to remain crime free in the community.

Outcome data. Initially, both probation sites were to provide a de-identified dataset that would list all of the charges each client had acquired approximately four months after the implementation of the workbooks; this time frame was extended to seven months post-implementation. The number of charges for each individual was expected to vary, so the Cormier/Lang method, which assigns a weight to each charge type, was going to be used to code the most serious charge (Harris, Rice, Quinsey, & Cormier, 2015). Next, the charges were to be recoded into four new variables: (1) technical violations (e.g., breach of supervision restrictions), (2) general recidivism, (3) violent recidivism, and (4) any recidivism (i.e., technical violations or a new charge). For the current study, general recidivism was to include all charges for drug-, driving- or property–related offenses, while violent recidivism was to include any charges related to assault, sexual assault, domestic abuse, robbery and armed robbery, or manslaughter and homicide. However, given the extremely small sample size and the dataset that was provided, the data were recoded into: 1) technical violations and 2) any new charges.

Finally, time at risk was to be calculated using the supervision start date and the date of the new charge. For those individuals who did not receive a new charge, time at risk was to be calculated to the end of the follow-up period (i.e., December 2017). Unfortunately,

TABLE 1.
Client Characteristics

		Condition		
Demographics		**Control**	**Experimental**	**Total**
Age (years)	Mean (*SD*)	38.50 (10.35)	36.27 (10.22)	37.97 (10.15)
	Minimum	27	22	22
	Maximum	59	55	59
Age (grouped)				
	20 - 29	20% (2)	32% (7)	28% (9)
	30 - 39	40% (4)	36% (8)	38% (12)
	40 - 49	20% (2)	23% (5)	22% (7)
	50 -59	20% (2)	9% (2)	12% (4)
Gender				
	Male	60% (6)	86% (19)	78% (25)
	Female	40% (4)	14% (3)	22% (7)
Race				
	Caucasian	90% (9)	82% (18)	84% (27)
	Other	10% (1)	18% (4)	16% (5)
Ethnicity				
	Hispanic	40% (4)	59% (13)	53% (17)
	Non-Hispanic	60% (6)	41% (9)	47% (15)
Level of Education (years)				
	Mean (*SD*)	12.50 (2.84)	12.50 (2.04)	12.50 (2.27)
	Minimum	7	10	7
	Maximum	16	16	16
Level of Education (grouped)				
	Less than grade 12	30% (3)	36% (8)	34% (11)
	Grade 12	40% (4)	32% (7)	34% (11)
	Any higher education	30% (3)	32% (7)	31% (10)
Previous Offence				
	Misdemeanor	50% (5)	36% (8)	41% (13)
	Felony	50% (5)	64% (14)	59% (19)
Total		**10**	**22**	**32**

TABLE 2.
Readability of the client self-directed workbooks

	Flesch-Kincaid Scale	
Workbook Topic	**Grade Level**	**Reading Ease[a]**
Motivation	5.2	79.7
Anger	4.5	80.9
Criminal Attitudes	6.2	72.6
Peer Relationships	5.3	75.7
Substance Abuse	6.0	73.0

[a] *Higher scores indicate easier readability. Scores of 65 indicate plain English.*

while the supervision start date was provided in the dataset, the date of the new charge was not, meaning time at risk was unable to be calculated.

Procedure

Following ethics approval, an email recruitment notice was sent to the POs that also included an informed consent form and recruitment script. Consenting participants were randomly assigned to either the control or experimental group using an online randomizer. If the participant was assigned to the experimental group, they were provided with the self-directed workbooks to use alongside the current probation services. At this location, POs acted as a support for using the workbooks—if clients had questions or wished to discuss the workbooks with their PO, they were encouraged to do so. If the participant was assigned to the control group, the participant followed the current community supervision of that site only. All clients who chose to participate received 10 hours of Community Service Restitution (CSR) credits, whether they were assigned to the control or the experimental group. Clients were then debriefed through an internal bulletin board notice that was posted partway through participant recruitment.

Finally, outcome data (i.e., new offenses and/or technical violations) was collected for all participants at the site seven months after implementation was complete (i.e., December 2017). As stated previously, the probation site provided a de-identified dataset that would list charges each client had acquired since the sharing of the workbooks.

Results

Data Preparation

Missing data. First, key variables were screened for missing values, and there were no out-of-range values on any of the key variables. Despite the small sample size ($N = 32$), violations or normality were not a concern, nor were there univariate outliers.

Differences Between Groups

In order to examine whether the individuals in the workbook group differed on demographic characteristics compared to those in the control group, independent samples *t*-tests and chi-square tests were conducted. Odds ratios and Cohen's *d* were used to examine effect size.

There were no significant differences between groups in terms of age $t(30, N = 32) = .57$, $p > .01$, $d = .22$, 95% CI [-.96, .54], although the experimental groups was slightly younger.

The relationship between age and condition was not significant, $t(30, N = 32) = .57$, $p > .01$, $d = .22$, 95% CI [-.96, .54].

The relationship between gender and condition was not significant ($N = 32$, $p > .01$, two-tailed). In comparison to the control condition, the workbook condition had a higher percentage of males (i.e., 86 percent compared to 60 percent) and a lower percentage of females (i.e., 14 percent compared to 40 percent). Furthermore, men were 4.22 times more likely to be in the experimental group (OR = 4.22, 95 percent CI [0.73, 24.44]).

The relationship between race and condition was not significant ($N = 32$, $p > .01$, two-tailed). In comparison to the control condition, the workbook condition had a lower percentage of Caucasian participants (82 percent compared to 90 percent) and a higher percentage of participants who identified as other (18 percent compared to 10 percent). Individuals who identified as Caucasian were .50 times more likely to be in the experimental group (OR = .50, 95 percent CI [.05, 5.15).

The relationship between ethnicity and condition was not significant ($N = 32$, $p > .01$, two-tailed). The percentage of Hispanic participants was higher in the workbook condition compared to the control condition (i.e., 59 percent versus 40 percent). Individuals who identified as Hispanic were 2.17 times more likely to be in the experimental group (OR = 2.17, 95 percent CI [.47, 9.95]).

The relationship between education and condition was not significant, $t(30, N = 32) = .00$, $p > .01$, $d = .00$, 95 percent CI [-.75, .75]).

The relationship between previous offense and condition was not significant ($N = 32$, $p > .01$, two tailed). The workbook condition had a higher percentage of participants whose previous offense was a felony (i.e., 64 percent compared to 50 percent in the control condition) and a lower percentage of those with a misdemeanor (i.e., 36 percent compared to 50 percent). Individuals whose previous offense was a felony were 1.75 times more likely to be in the experimental group (OR = .1.75, 95 percent CI [.39, 7.95]).

Overall, the experimental group was male, younger, and had a more serious criminal history.

Differences Between the Conditions on Recidivism

To examine whether the individuals in the workbook condition significantly differed in technical violations or any new charges compared to those in the control group, Fisher's exact tests were conducted. Given the small sample size and the nature of the data that was received from the probation site, an examination of the time to failure across the conditions was unable to be conducted.

Comparisons were made between those in the experimental condition and those in the control condition for both technical violations and any new charges. Fisher's exact test was used to examine technical violations across the experimental and control conditions and then to examine the difference between the two conditions on any new charges. As stated previously, this procedure is typically used on a 2 X 2 contingency table (i.e., two variables, each with two levels) with a small sample.

A Fisher's exact test was performed to examine the relationship between technical violations and condition (see Table 3, next page). The relationship between technical violations and condition was not significant ($N = 32$, $p > .01$, two tailed). An examination of the relative frequencies was conducted next (see Figure 1, next page). The workbook condition had a lower percentage of participants who experienced a technical violation post-treatment (i.e., 50 percent compared to 80 percent in the control condition). Furthermore, individuals in the control group were 4.00 times more likely to have a technical violation than those in the experimental group (OR = 4.00, 95 percent CI [.69, 23.26]).

A Fisher's exact test was performed to examine the relationship between any new charges and condition (see Table 4, next page). The relationship between any new charges and condition was not significant ($N = 32$, $p > .01$, two tailed). An examination of the relative frequencies was conducted next (see Figure 2). The workbook condition had a lower percentage of participants who received a new charge post-treatment (i.e., 5 percent compared to 10 percent in the control condition). Furthermore, individuals in the control group were 2.33 times more likely to have a new charge than those in the experimental group (OR = 2.33, 95 percent CI [.13, 41.46]).

Discussion

The aim of this study was to explore the effectiveness and utility of a set of self-directed criminogenic-focused workbooks in a community supervision setting. While similar workbooks have been used in a variety of contexts with in-custody populations, this is

one of the first studies to explore the use of workbooks with a community sample and their effect on recidivism. Thus, the descriptive nature of this study makes it the first step in determining the utility of these workbooks, and the results of each research question will be discussed separately. Practical implications and limitations will then be discussed. Suggestions for future research will be discussed more generally at the end.

Summary and Implications of Findings

We hypothesized that there would be no major demographic differences between the workbook and control conditions, given that participation was voluntary and a process was used to ensure random assignment to groups. While there were no statistically significant differences between groups, the experimental group (i.e., workbook condition) was more likely to be younger and to have a previous felony conviction than those in the control condition. Efforts to ensure random assignment were insufficient, perhaps due to the small sample.

We also hypothesized that the individuals in the workbook groups would have lower rates of recidivism (i.e., either technical violations or new charges) and longer time to failure in comparison to the control groups. Again, while there were no statistically significant differences on either outcome variable across the workbook and control conditions, there was some variation when the relative frequencies and effect sizes (odds ratios) were examined. This demonstrated a very modest treatment effect; workbook use was related to slightly improved outcomes for the experimental group.

Individuals in the control group were 4.00 times more likely to have a technical violation and 2.33 times more likely to have any new charges, even though individuals in the workbook group were likely somewhat higher risk (i.e., younger, male, previous offense is a felony). This difference is encouraging, as some technical violations are related to factors that the workbooks target (e.g., avoiding substance use as a condition of probation). While the results of the current study are insufficient to fully support the efficacy of the set of workbooks in a community supervision setting, participants did not reject their use, nor did their use yield iatrogenic effects in this very small pilot study.

Replication with a larger sample could demonstrate the effectiveness and usefulness of these workbooks, which would give supervision agencies reason to consider their implementation in the future. Considering the success of similar workbooks with in-custody populations (e.g., Proctor et al., 2012), it is not unrealistic to suggest that future research with these workbooks may produce favorable results for probationers.

Limitations

Originally, this study was to be implemented at two probation sites in the United States. Unfortunately, organizational changes at one probation site led to implementation delays that ultimately resulted in the study being dropped.

The next limitation was the small sample size and resulting inadequate power, which prohibited the use of most inferential statistical procedures. The observed findings should be considered very preliminary, and further investigation with a larger sample is necessary. Furthermore, this resulted in limiting the generalizability of any findings to the broader population of clients under community supervision.

Another limitation is that information regarding the clients' motivation and readiness to change and perceived self-efficacy prior to workbook implementation, which are important factors to consider, were not examined. Scheck et al. (2013) observed a weak correlation between knowledge and attitude, suggesting that education alone does not allow anyone to infer an adequate level of motivation to promote successful behavior change, at least in the context of substance use behaviors. Controlling for prior programming and supervision experience and motivation

TABLE 3.
Results of Independent Samples t-tests and Fisher's Exact Tests for Age, Gender, Race, Ethnicity, Education, and Previous Offense Across Condition

Demographics	*t*	*p*	*d / OR*	*95% CI*
Age (years)	.57[a]	.57	.22[c]	[-.96, .54]
Gender				
Male	–	.17[b]	4.22[d]	[0.73, 24.44]
Female				
Race				
Caucasian	–	.99[b]	.50[d]	[.05, 5.15]
Other				
Ethnicity				
Hispanic	–	.45[b]	2.17[d]	[.47, 9.95]
Non-Hispanic				
Level of Education (years)	.00[a]	.00	.00[c]	[-.75, .75]
Previous Offense				
Misdemeanor	–	.70[b]	1.75[d]	[.39, 7.95]
Felony				

[a] *t-statistic.* [b] *p value for Fisher's exact test.* [c] *Cohen's d value.* [d] *Odds ratio.*

TABLE 4.
Results of Fisher's Exact Tests for Technical Violations and Any New Charges Across Condition

	Condition					
Outcome	Control	Experimental	*Total*	*p*[a]	*OR*	*95% CI*
Technical violation						
Yes	80% (8)	50% (11)	59% (19)	.14	4.00	[.69, 23.26]
No	20% (2)	50% (11)	41% (13)			
Any new charges						
Yes	10% (1)	5% (1)	6% (2)	.53	2.33	[.13, 41.46]
No	90% (9)	95% (21)	94% (30)			

[a] *p value for Fisher's exact test.*

level could be important in understanding the incremental utility of journaling.

Despite efforts to have the workbooks at a high ease of reading level, the reading level of clients was not formally assessed prior to the implementation of the workbooks. Of note, no clients requested that the PO read aloud the consent form during the recruitment process of the current study. All text included in the workbooks is intended for individuals with a reading level between the grades of four and six. Furthermore, the text is broken up into short, easy-to-digest sections, and there are a variety of graphics to accompany the key concepts, mitigating the potential challenge of reading ability.

It is also important to note that we were unable to complete a client feedback survey that may have been instructive. While participants received compensation in the form of CSR credits for their involvement in phase one, they did not receive any compensation for their involvement in completion of the survey.

FIGURE 1.
Relative Frequencies of Technical Violation Across Condition.

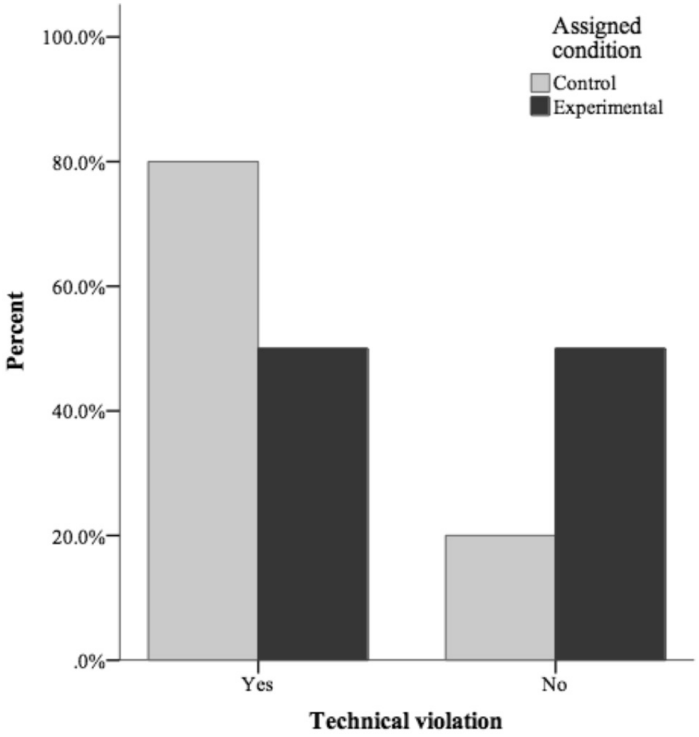
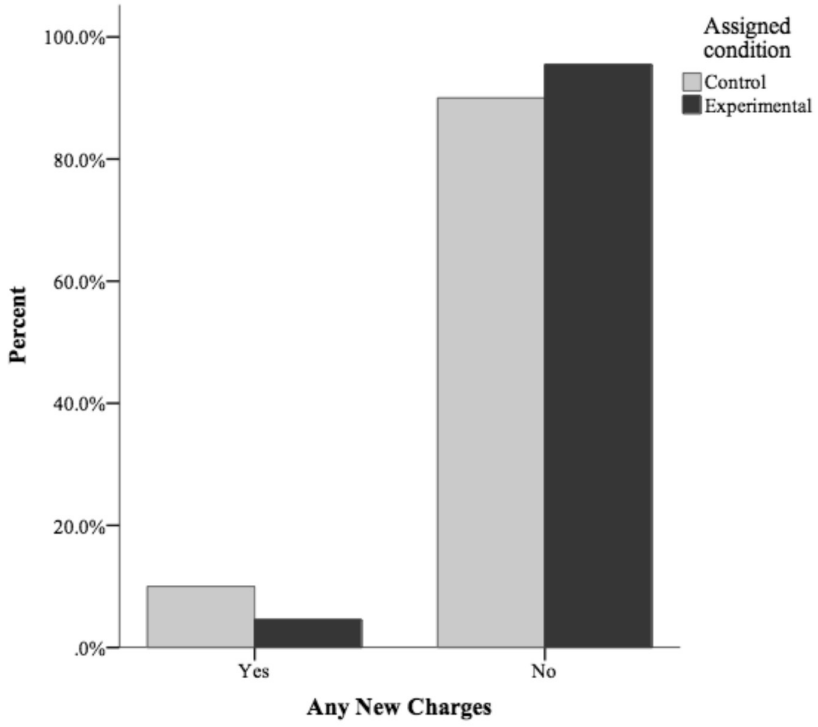


FIGURE 2.
Relative Frequencies of Any New Charges Across Condition.



Future Research

Given these limitations, we present several over-arching suggestions for future research on the self-directed workbooks. First, in order to prevent having such a small sample size, it may be more effective to use a matched sample based on either exact matching or propensity matching instead of a control group, as this will allow for a larger experimental group without significantly compromising the conclusions that could be made.

Second, running a focus group in order to assess the participants' views on the usability of the workbooks may prevent a no-response situation, as well as allow for more detailed responses from the participants. With a higher level of detailed responses in a semi-structured interview setting, a thematic analysis of the responses and suggestions for changes to the workbooks could be conducted. It is also possible that just providing clients with a paper copy of the survey would have increased responses.

Recently we added reentry and trauma workbooks into our suite of workbooks to broaden their utility and address emerging concerns. Our experience regarding their use in a remand center indicated that many clients enthusiastically completed all the workbooks, potentially out of boredom or to present to the courts when adjudicated. Ideally, a risk and need assessment should be used to match the workbook(s) to client needs. Nonetheless, it is possible that the workbooks may have application in the pretrial world.

Finally, given the high rates of mental health diagnoses in this population (see Prins [2014] for a systematic review), future research should consider the impact of major mental

health disorders on the efficacy of this intervention. Studies that evaluate the potential mediators and moderators of efficacy could further refine our understanding of the merits of self-directed workbooks and journaling.

Conclusion

Despite the limitations, especially the small sample and non-significant findings, self-directed workbooks still may have potential to reduce both technical violations and recidivism among individuals supervised in the community. More research is necessary to make stronger conclusions that could inform policy and practice; however, this is a sufficient first step or proof of concept to warrant expanding this type of intervention, especially to lower risk clients. Most notably, there is no indication they have an iatrogenic effect, they are minimally invasive, and they have no financial cost, supporting their inclusion in the community supervision arsenal.

References

Andrews, D. A., & Dowden, C. (2004). Managing correctional treatment for reduced recidivism: A meta-analytic review of programme integrity. *Legal and Criminological Psychology, 10*, 173-187. doi:10.1348/135532505X36723

Bonta, J., & Andrews, D. A. (2017). *The psychology of criminal conduct* (7th ed.), New York, NY: Routledge.

Bourgon, G., Gutierrez, L., & Ashton, J. (2011). The evolution of community supervision practice: The transformation from case manager to change agent. *Irish Probation Journal, 8*, 28-48.

Bourgon, G., & Gutierrez, L. (2012). The general responsivity principle in community supervision: The importance of probation officers using cognitive intervention techniques and its influence on recidivism. *Journal of Crime and Justice, 35*(2), 149-166. doi: 10.1080/0735648X.2012.674816

Chadwick, N., DeWolf, A. H., & Serin, R. C. (2015). Effectively training community supervision officers: A meta-analytic review of the impact on offender outcome. *Criminal Justice and Behavior, 42*, 977-989. doi: 10.1177/0093854815595661

Council of State Governments. (2019). Confined and costly: How supervision violations are filling prisons and burdening budgets. *Council of State Governments*. Retrieved from: https://csgjusticecenter.org/wp-content/uploads/2020/01/confined-and-costly.pdf

Criminal Justice Connections (2015, January/February). *The Teas Risk Assessment System: A new direction in supervision planning, 22* (3). Retrieved from: https://www.tdcj.state.tx.us/connections/JanFeb2015/Images/JanFeb2015_agency_TRASS.pdf

Davidson, J. T., Crawford, R., & Kerwood, E. (2008). Constructing an EBP post-conviction model of supervision in United States Probation, District of Hawaii: A case study. *Federal Probation, 72*(2), 22-28.

Dowden, C., & Andrews, D. A. (2004). The importance of staff practice in delivering effective correctional treatment: A meta-analytic review of core correctional practices. *International Journal of Offender Therapy and Comparative Criminology, 48*(2), 203-214. doi: 10.1177/0306624X03257765

Folk, J. B., Blasko, B. L., Warden, R., Schaefer, K., Ferssizidis, P., Stuewig, J., & Tangney, J. P. (2016). Feasibility and acceptability of an impact of crime group intervention with jail inmates. *Victims & Offenders, 11*(3), 426-454. doi: 10.1080/15564886.2014.982777

Frattaroli, J. (2006). Experimental disclosure and its moderators: A meta-analysis. *Psychological Bulletin, 132*(6), 823-865. doi: 10.1037/0033-2909.132.6.823

Lovins, B. K., Cullen, F. T., Latessa, E. J., & Jonson, C. L. (2018). Probation officer as a coach: Building a new professional identity. *Federal Probation 82(1),* 13-20.

Lowenkamp, C. T., Latessa, E. J., & Smith, P. (2006). Does correctional program quality really matter? The importance of adhering to the principles of effective intervention. *Criminology and Public Policy, 5,* 575–594, http://dx.doi.org/10.1016/j.avb.2012.09.002

Oudekerk, B., & Kaeble, D. (2021). Probation and parole in the United States, 2019. *Bureau of Justice Statistics.* Retrieved from: https://bjs.ojp.gov/sites/g/files/xyckuh236/files/media/document/ppus19.pdf

Prins, S. J. (2014). The prevalence of mental illnsesses in U.S. state prisons: A systematic review. *Psychiatric Services, 65*(7), 862-872. doi: 10.1176/appi.ps.201300166

Proctor, S. L., Hoffmann, N. G., & Allison, S. (2012). The effectiveness of interactive journaling in reducing recidivism among substance-dependent jail inmates. *International Journal of Offender Therapy and Comparative Criminology. 56*(2), 317-332. doi: 10.1177/0306624X11399274

Richards, J. M., Beal, W. E., Seagal, J. D., & Pennebaker, J. W. (2000). Effects of disclosure of traumatic events illness on behavior among psychiatric prison inmates. *Journal of Abnormal Psychology 109*(1), 156-160.

Robinson, C. R., Lowenkamp, C. T., Holsinger, A. M., VanBenschoten, S., Alexander, M., & Oleson, J. C. (2012). A random study of Staff Training Aimed at Reducing Re-arrest (STARR): Using core correctional practices in probation interactions, *Journal of Crime and Justice, 35:2*, 167-188, DOI: 10.1080/0735648X.2012.674823

Ross, E. C., Polaschek, D. L. L., & Ward, T. (2008). The therapeutic alliance: A theoretical revision for offender rehabilitation. *Aggression and Violent Behavior, 13,* 462–480, doi:10.1016/j.avb.2008.07.003

Scheck, A. M., Hoffmann, N. G., Proctor, S. L., & Couillou, R. J. (2013). Interactive journaling as a brief intervention for Level-II DUI and DWI offenders. *Journal of Alcohol & Drug Education, 57*(3), 66-85.

Serin, R. C., Lloyd, C. D., & Hanby, L. J. (2010). Enhancing offender re-entry: An integrated model for enhancing offender re-entry. *European Journal of Probation, 2*(2), 53-75.

Smith, P., Gendreau, P., & Swartz, K. (2009). Validating the principles of effective intervention: A systematic review of the contributions of meta-analysis in the field of corrections. *Victims & Offenders, 4*(2), 148-169, doi: 10.1080/15564880802612581

Substance Abuse and Mental Health Services Administration. (2014). *Interactive Journaling.* National Registry of Evidence-based Programs and Practices. Retrieved from: http://www.nrepp.samhsa.gov

Taxman, F. S. (2012). Crime control in the twenty-first century: Science-based supervision. *Journal of Crime and Justice, 33*(2), 135-144. https://doi.org/10.1080/0735648X.2012.686583

U.S. Courts. Administrative Office of the U.S. Courts (2017). Incarceration costs significantly more than supervision. Retrieved from: https://www.uscourts.gov/news/2017/08/17/incarceration-costs-significantly-more-supervision

Viglione, J., Rudes, D. S., & Taxman, F. S. (2017). Probation officer use of client-centered communication strategies in adult probation setting. *Journal of Offender Rehabilitation, 56*, 38-60. https://doi.org/10.1080/10509674.2016.1257534

Low/Moderate Risk Policy Change White Paper

Administrative Office of the United States Courts
Probation and Pretrial Services Office

An Overview of the Expansion of Low-Risk Supervision Standards
February 1, 2023

Editor's Note: *In March 2020, the Administrative Office of the U.S. Courts – Probation and Pretrial Services Office (PPSO) enacted a temporary change to its supervision practices in response to growing caseloads, budgetary pressures, and the COVID-19 pandemic. This change involved an expansion of the population eligible to be supervised under the low-risk supervision standards (LRSS). LRSS is geared towards supervisees classified on the lower end of the recidivism risk classification continuum and hence, persons placed on LRSS receive less monitoring and fewer restrictions and interventions compared to persons on regular supervision. The following report constitutes a White Paper that details the policy change with particular emphasis on the extent to which the federal supervision system's practices changed in response to this reform while simultaneously dealing with the COVID-19 pandemic. Importantly, the report also examines whether expansion of the LRSS population endangered community safety. We are publishing the report here in the belief that it will be of substantial interest to researchers, policymakers, and legal personnel (such as defenders, U.S. attorneys, and judges) involved in the federal supervision system.*

Part A: Introduction

The preceding fiscal years (i.e., 2020, 2021, and 2022) presented many challenges for the federal probation and pretrial services system. Specifically, research conducted on our system showed that budgetary reductions resulted in fewer officers being able to do the work involved in supervising persons placed on federal post-conviction supervision. In addition to reductions in officer staffing levels, there were substantial increases in the number of persons being placed on federal supervision because of early releases stemming from implementation of the First Step Act. There were also impacts from the COVID-19 pandemic, with officers dramatically reducing their in-person contacts with supervisees to reduce their potential exposure to this virus. The combined effects of budget cuts, staffing reductions, expanded caseloads, and the pandemic presented unparalleled challenges for the federal supervision system.

To alleviate these pressures on officers, the Judicial Conference Committee on Criminal Law (Criminal Law Committee) recommended that the federal supervision system increase the number of persons to be supervised using low-risk supervision standards (LRSS). Persons placed on LRSS receive less monitoring and fewer restrictions and interventions compared to persons on regular supervision. LRSS has the benefit of improving officers' ability to strategically shift time and resources to higher risk supervisees who pose the greatest danger to the community, while simultaneously allowing officers to take on low-risk caseloads at higher volumes. This can help alleviate workload pressures on probation offices.

This initiative to implement the revised LRSS policy was enacted in March 2020. In response to this implementation, the Administrative Office of the U.S. Courts (AO) engaged in a research effort to assess (1) whether the federal system's practices changed in response to the LRSS policy and (2) whether these changes endangered community safety. Findings from this research show that federal supervision practices changed during the period examined, with in-person contacts for the LRSS group registering less of an increase in 2021 compared to the other risk classification categories. In addition, results show that community safety was not negatively impacted by the implementation of LRSS expansion. The remainder of this report details how this research was conducted and covers the principal findings and conclusions.

Part B: Criteria for Expanding Persons Eligible for Low-Risk Supervision

Before this expansion, LRSS was available only to supervisees whom the Post Conviction Risk Assessment (PCRA) categorized as low risk and whose supervision history showed no more than a low-severity violation.[1] In March 2020, the AO initiated a change in its supervision policies, allowing a new subset of supervisees to be placed on LRSS supervision. Specifically, the AO recommended that the federal supervision system increase the number of supervisees on LRSS supervision if they met the following criteria:

- The risk instrument used by federal

[1] For an overview of the PCRA and the original low-risk supervision standards, see *An Overview of the Federal Post Conviction Risk Assessment Instrument* (June 2018), AO. https://www.uscourts.gov/sites/default/files/overview_of_the_post_conviction_risk_assessment_0.pdf

probation officers (i.e., the PCRA) designated them as low/moderate risk.
- Their probability of committing violent crimes was low (i.e., category 1 violence).
- Their PCRA scores placed them on the lower end of the low/moderate-risk continuum (i.e., PCRA raw scores of 6 or 7).
- They did not manifest high levels of criminal thinking.
- They were not convicted of sex offenses.
- They did not have an instant conviction offense for a violent felony and had not been previously convicted of two or more prior violent felonies.
- Their overall risk classifications had not increased by the second PCRA assessment. Typically, the second assessment takes place six months after the initial risk assessment.

Using these criteria, the Probation and Pretrial Services Office (PPSO) estimated that an additional 13,655 supervisees, or about 12 percent of the federal supervision population, could be placed on LRSS. They were recommended for LRSS supervision because they have relatively low recidivism rates. For example, about 8 percent were arrested for any offense within one year of their supervision start date, while fewer than 2 percent were arrested for violent crimes. In comparison, supervisees scoring on the higher end of the low/moderate-risk continuum manifested rearrest rates ranging from 13 to 14 percent for any offenses and approximately 2 percent for violent offenses. Moreover, supervisees who are designated as moderate or high risk witnessed general recidivism rates of 18 to 37 percent and violent recidivism rates for 6 to 13 percent.

With the advent of the LRSS expansion, it was important to assess the extent to which federal supervision practices changed in response to the new policy and to ascertain whether the change endangered community safety. A finding that officers contacted people meeting the LRSS eligibility criteria less frequently in 2020 and 2021, compared to earlier years when the expansion was not in effect (i.e., 2017, 2018, and 2019), supports the fact that the federal system changed in response to this expansion. Moreover, results showing that recidivism behavior for those placed on LRSS supervision after the expansion was similar to, or perhaps lower than, it was for people meeting the LRSS criteria before the expansion could indicate that community safety was not endangered. Hence, below are questions that form the main components of this analysis:

- What percentage of persons under federal supervision are eligible for LRSS supervision under the expansion? Are officers treating supervisees differently when they are eligible for LRSS supervision after the expansion (e.g., contacted less), compared to people who met the LRSS criteria before the enactment of this change?
- What does the recidivism behavior look like for supervisees who are eligible for LRSS supervision under the expansion? Are supervisees who are eligible for LRSS supervision after the expansion recidivating at elevated, similar, or lower levels compared to people who met the LRSS criteria before enactment of these changes?

Part C: Data and Method

Impact of COVID-19 Pandemic on LRSS Research

Before delving into this research, it is important to note the challenge that the COVID-19 pandemic presents to the current analysis. Specifically, the LRSS expansion was implemented at the start of the pandemic in March 2020. This made any pre/post assessment of supervisees meeting the revised LRSS thresholds somewhat problematic, since any changes in the contact or rearrest patterns of this group post expansion might be driven by the pandemic rather than by changes in officer supervision practices or supervisee criminal behavior. Research conducted by the AO shows that the pandemic was associated with substantial declines in officer in-person contact patterns and supervisee violation activity.[2]

Hybrid Pre-/Post-Analytical Approach

The AO's research analysts addressed the methodological challenges presented by the COVID-19 pandemic by using a hybrid pre-/post-analytical framework. Specifically, AO researchers evaluated the contact and recidivism patterns not only for the LRSS group pre- and post-policy change, but also for the other PCRA risk groups pre- and post-policy change.

We anticipated that using this approach would show both the contact and recidivism rates declining across all PCRA risk categories at the pandemic's onset in 2020. However, we hypothesized that the LRSS group would experience contact and rearrest patterns that differed from the other PCRA risk groups as the system emerged from the pandemic in 2021. The specific hypotheses that oriented this research follow:

- The average number of monthly contacts between officers and supervisees should decline in 2020 for all PCRA risk groups. But in 2021, they should increase more slowly for the LRSS group compared to the other PCRA risk categories.
- The recidivism outcomes (including noncompliance, revocations, and rearrests) should decline across all PCRA risk groups in 2020. But in 2021, they should rise more slowly or not at all for the LRSS group compared to the other PCRA risk categories.

Before delving into the study's findings, it is important to understand the PCRA's risk classification groupings. For some background, the PCRA uses the following five-color-ordered risk scheme to measure a supervisee's likelihood of recidivism: blue, green, yellow, orange, and red. The degree of predicted risk increases with each change in color, with blue supervisees having the lowest failure probability and red supervisees having the highest failure probability.

Before the enactment of the LRSS expansion, only the PCRA blue group qualified for LRSS supervision. The expanded LRSS group is within the PCRA green category, with about half of the PCRA greens meeting LRSS eligibility and hence qualifying for low-risk supervision under the revised program.

Table 1 (next page) provides information about the PCRA risk groups analyzed pre- and post-LRSS expansion, including the number and percentage of supervisees who met the LRSS eligibility criteria by fiscal year of case supervision.

Ordinarily, most pre- and post- studies would place the fiscal years into specific groups. For example, the 2017-2019 cohort would be in the pre group, and the 2020-2021 cohort would be in the post group. However, as will be shown, the pandemic dominated officer contact activity so much in 2020 that placing that year and 2021 into one group is problematic. Hence, this hybrid pre- and post-approach examines officer contact activity and supervisee violation rates for each fiscal year separately.

Population Examined

Several important aspects of the population examined should be noted. First, since the PCRA color-coded risk schematic was not implemented until early 2017, we decided to remove all supervisees received on supervision before that fiscal year. Supervisees who were

[2] For information about the pandemic's impact on the federal supervision system, see the June 2021 special edition of *Federal Probation*.

TABLE 1.
Percentage of Persons Under Supervision by PCRA Risk Levels, Supervision Year, and LRSS Classification

	Pre-LRSS Expansion						Post-LRSS Expansion			
PCRA	2017		2018		2019		2020		2021	
Codes	Number	Percent	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Total	18,985	100.0 %	55,923	100.0 %	92,020	100.0 %	117,086	100.0 %	128,210	100.0 %
Blue	4,054	21.4 %	11,615	20.8 %	18,777	20.4 %	23,562	20.1 %	25,281	19.7 %
Green-LRSS	**2,940**	**15.5**	**8,369**	**15.0**	**13,385**	**14.6**	**16,355**	**14.0**	**16,579**	**12.9**
Green-No LRSS	2,339	12.3	6,892	12.3	11,461	12.5	15,007	12.8	17,226	13.4
Yellow	4,145	21.8	12,544	22.4	20,872	22.7	26,964	23.0	29,977	23.4
Orange	2,829	14.9	8,237	14.7	13,723	14.9	17,734	15.2	19,824	15.5
Red	2,678	14.1	8,266	14.8	13,802	15.0	17,464	14.9	19,323	15.1

Note: Includes people received on supervision between fiscal years 2017 and 2021 with a PCRA 2.0 assessment. Those received on supervision before 2017 were omitted from the study. Each year encompasses persons under supervision with a PCRA assessment at any time during that fiscal year.
Persons under supervision for less than six months were omitted from the analysis.
Bold denotes that the supervisee was eligible for LRSS supervision under the new low-risk supervision category.

placed on supervision before fiscal year 2017 were omitted, because this study attempted to measure the average number of monthly contacts from the supervision start date until case closure or, if the case was still open, an anchor date of April 23, 2022. Including people who began supervision before 2017 would be problematic because the PCRA color-coded risk categories, including those meeting LRSS eligibility, were unavailable until 2017.[3]

In addition to removing people who were placed on federal supervision before fiscal year 2017, several other exclusionary criteria were employed. Since the color-coded PCRA risk schematic was implemented through a rolling deployment, not all people who were received on supervision during fiscal year 2017, and even 2018 had an initial PCRA assessment using the revised PCRA color-coded groups. Supervisees with an original assessment involving the older non-color-coded PCRA categories were removed for reasons similar to those that were used to exclude people who were placed on supervision before 2017. Also, the study cohort included only those with a minimum of two or more PCRA assessments. A baseline of two or more assessments was required because one of the LRSS criteria was that there be no increase in a supervisee's overall risk classification between assessments.

Since the LRSS group required two or more risk assessments, it was important for consistency for the other color-coded risk groups to have similar assessment criteria. Moreover, cases had to be under supervision for six months or more for the purposes of following contact and recidivism activity. Last, all people who received supervision overrides were omitted. Removing supervisees who are placed into higher risk categories than originally classified allows for a more robust and clearer comparison of officer contact activity between the risk groups.[4]

Measures

This study uses two primary measures of interest: officer contacts and recidivism outcomes. Officer contacts with supervisees are used as a proxy to measure the extent to which the federal supervision system changed after implementation of the LRSS expansion. Three types of contacts were measured, including in-person, other-person, and collateral. In-person contacts include any contact between the officer and supervisee that involves an in-person interaction. These contacts typically take place between the officer and supervisee in the supervisee's home, neighborhood, place of employment, or federal probation office. An other-person contact means that the officer contacted the supervisee through other—usually electronic—means, such as by telephone, voicemail, text message, and email. Last, collateral contacts refer to contacts between the officer and collateral sources, typically a treatment provider, employer, or law enforcement officer.

Recidivism outcomes are used to measure the extent to which the supervisee failed during or after supervision. The recidivism outcomes of interest for this study include any forms of noncompliance (e.g., positive drug tests, technical violations of supervision conditions), revocations from supervision, and rearrest for any new crimes or violent offenses.

Part D: Results

Characteristics of the LRSS Population

Table 2 (next page) provides descriptive information about people who were eligible for LRSS supervision before and after enactment of the policy expansion. A total of 21,259 people who were received on supervision during fiscal years 2017-2021 met the revised LRSS supervision criteria. More than half of these cases (54 percent) were convicted for drug offenses. The other most frequent conviction offenses included property (19 percent) and weapons/firearms (11 percent). Seventy-seven percent of the LRSS group comprised males; Hispanics, Whites, and Blacks accounted for relatively similar proportions of supervisees in the LRSS group.

Under the expansion, supervisees are not eligible for LRSS supervision unless their risk levels remain unchanged for a minimum of two PCRA assessments. Using these standards, many supervisees will have met the LRSS eligibility criteria by their second

[3] Before 2017, the PCRA used a four-tier risk classification scheme of low, low/moderate, moderate, and high. The PCRA was modified in 2017 to include a violence trailer, which resulted in the risk instrument generating the five color-coded risk categories. For more information about the original PCRA and the deployment of the PCRA violence trailer, see Johnson et al. (2011), *The Construction and Validation of the Federal Post Conviction Risk Assessment (PCRA)*, and Serin et al. (2016), *Using a Multi-Level Risk Assessment to Inform Case Planning and Risk Management: Implications for Officers*.

[4] The PCRA gives officers discretion to depart from the risk instrument's original classification scheme. For more information about the role of overrides, see Cohen et al. (June 2016), *Examining Overrides of Risk Classifications for Offenders on Federal Supervision.*

assessment. However, others might not be eligible for LRSS supervision until their third or fourth assessment.

Figure 1 provides information about the assessment number when a supervisee became eligible for LRSS supervision. As expected, a majority of LRSS supervisees (72 percent) met the eligibility standards (meaning no changes in their risk levels) by their second assessment. However, about 24 percent of LRSS supervisees were not eligible for LRSS supervision until their third or fourth PCRA assessment. For example, these supervisees might have started in a higher supervision category at their initial PCRA assessment (e.g., yellow or orange) and moved into the LRSS green risk category by their second assessment. To qualify for low-risk supervision, the risk profiles for these people would need to remain unchanged from the second to the third PCRA assessment.

Changes in Officer Contact Patterns with Supervisees Resulting from the LRSS Expansion

Figures 2a-2c (next page) provide information on the average number of monthly contacts during a supervisee's first 12 months of supervision by PCRA risk levels and fiscal year of case activation. It should be noted that this approach examines contacts that occurred only during a person's first supervision year. The monthly contact numbers are calculated by summing the number of times that officers contacted supervisees within the first year of supervision and then dividing that total by 12. These calculations were performed separately for each of the fiscal years examined (i.e., cases activated in 2017-2021).

These figures illuminate trends in monthly in-person contacts (see Figure 2a), monthly other-person contacts (see Figure 2b), and monthly collateral contacts (see Figure 2c). In general, they show substantial declines in the average number of monthly in-person contacts for all supervisees when the pandemic started in 2020; these in-person contacts subsequently rebounded in 2021. Conversely, the average number of monthly other-person contacts increased for all risk levels in 2020 and then proceeded to decline. Last, collateral monthly contacts remained relatively stable during the time period examined. Though interesting, these figures do not support the contention that supervision practices changed for the LRSS population after implementation of the LRSS expansion.

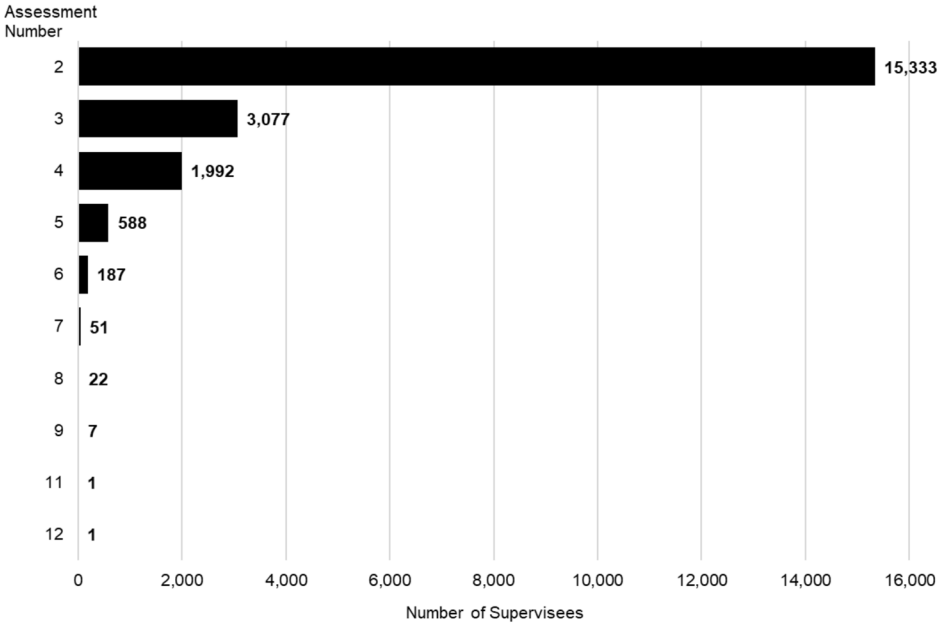
A more illuminating way to measure the

TABLE 2.
Characteristics of Supervisees Eligible for LRSS Supervision Under the LRSS Expansion

Characteristics	Number	Percent or Average
Most Serious Conviction Offense		
Drugs	11,448	53.9 %
Property	4,032	19.0
Weapons/Firearms	2,398	11.3
Immigration/Customs	1,788	8.4
Violence	607	2.9
Traffic/DWI	493	2.3
Obstruction/Escape	279	1.3
Public Order	204	1.0
Gender		
Male	16,399	77.1 %
Female	4,860	22.9
Race and Ethnicity		
Hispanics, Any Race	7,091	33.5 %
Whites, Non-Hispanic	6,810	32.1
Blacks, Non-Hispanic	6,481	30.6
Asians/Pacific Islanders	416	2.0
American Indians/Alaska Natives	322	1.5
Other	78	0.4
Age	21,256	40.3
Citizenship		
U.S. or Naturalized Citizen	20,183	95.1 %
Legal Alien	605	2.9
Illegal Alien	426	2.0

Note: Includes only supervisees eligible for LRSS supervision under the expansion received on supervision between fiscal years 2017 and 2021.
A total of 21,259 supervisees qualify for LRSS supervision.

FIGURE 1.
PCRA Assessment Number in Which Supervisee Became Eligible for LRSS Supervision Under the LRSS Expansion, Fiscal Years 2017–2021



LRSS expansion's potential effects on the federal supervision system is to examine the percentage change in the average number of monthly contacts between the fiscal years of case activation (see Figures 3a-3c, next page). Figure 3a, for example, shows the percentage changes in the average number of in-person contacts for fiscal years 2017-18, 2018-19, 2019-20, and 2020-21 across the PCRA risk levels. Not surprisingly, this figure shows substantial declines in the average number of in-person contacts, irrespective of risk, during 2020. For example, the average number of in-person contacts for 2020 declined by 27 percent for the green LRSS group, but similar declines were witnessed for the green no-LRSS group (-27 percent) and the yellow group (-26 percent).

However, during 2021, in-person contacts for the green LRSS category diverged slightly from the other PCRA risk groups (see Figure 3a). Specifically, the green LRSS group witnessed smaller rises in the average in-person contact numbers (+4 percent) compared to the other PCRA risk groups, which saw their in-person contacts increase in the range of 8-9 percent (except for the PCRA blues). The fact that the green LRSS group manifested less of an increase in the in-person contacts compared to the other PCRA risk categories provides some evidence in support of the LRSS expansion's impact on federal supervision practices.

Examining the other contact types (i.e., other-person and collateral) presents a mixed picture in terms of implementing LRSS expansion. The percentage change in other-person contacts was not appreciably different for the green LRSS group compared to the other PCRA risk categories (see Figure 3b). Regarding collateral contacts, the green LRSS supervisees were the only group witnessing declines in their average monthly collateral contacts during 2020 (-3 percent), while the other risk categories saw no changes in their monthly collateral contacts (PCRA blues) or increases in their monthly collateral contacts (green no-LRSS, yellow, orange, or red) (see Figure 3c).

Figures 4a-4c (page 51) provide information about the percentage change in the number of in-person, other-person, and collateral contacts in a somewhat different format. Specifically, they illuminate changes in contacts by supervision year rather than in the first 12 months under supervision. Using supervision year allows us to count contacts for all persons under federal supervision for that particular year examined, regardless of their start date or the amount of time they were under supervision during that year. Hence, a person who started supervision in 2017 and was still under supervision in 2020 would have monthly contact numbers counted for each individual year while under federal supervision (i.e., 2017, 2018, 2019, and 2020).

In the above example, this person's per-year contact numbers would be calculated by totaling the number of contacts made by officers for each year and then dividing the total number of contacts by 12 per year. Unlike the prior approach, which counted contacts only during the first 12 months of supervision, this method can ascertain a supervisee's monthly contact numbers for a more extended time period.

Examining the percentage changes for in-person contacts by supervision year lends further support to the contention that officer supervision practices changed for people who were placed on LRSS supervision. In Figure

FIGURE 2A.
Average Number of In-Person Monthly Collateral Contacts Within 12 Months of Case Activation, by PCRA 2.0 Risk Levels, Fiscal Years 2017–2021

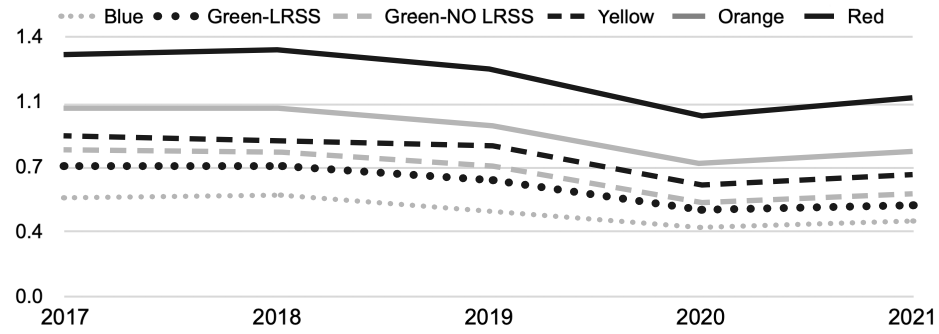


FIGURE 2B.
Average Number of Other-Person Monthly Collateral Contacts Within 12 Months of Case Activation, by PCRA 2.0 Risk Levels, Fiscal Years 2017–2021

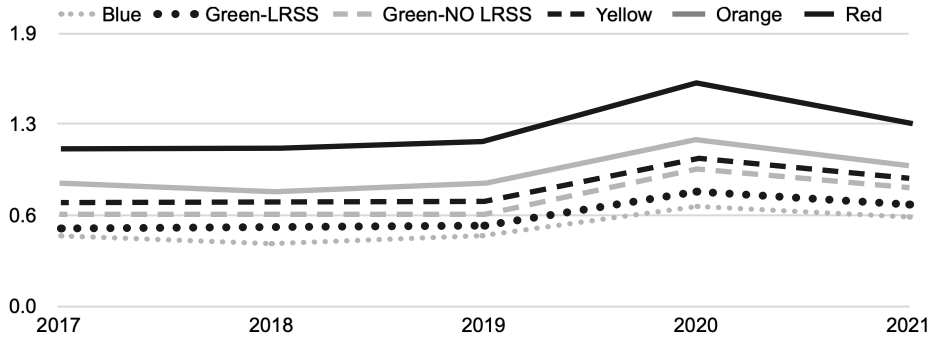
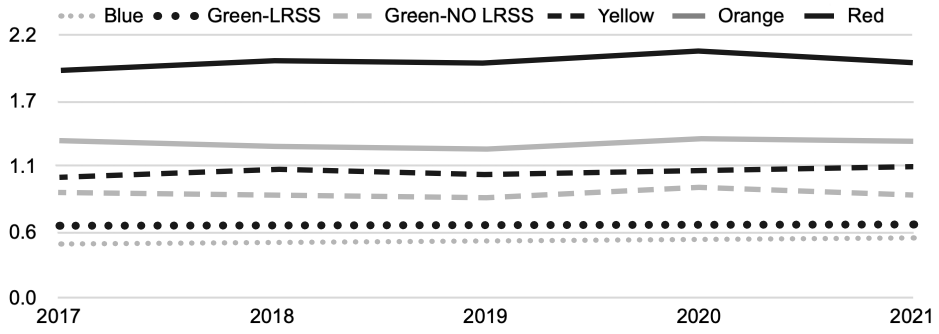


FIGURE 2C.
Average Number of Monthly Collateral Contacts Within 12 Months of Case Activation, by PRCA 2.0 Risk Levels, Fiscal Years 2017–2021



4a, for example, the average number of in-person monthly supervision contacts declined for all risk levels in 2020. However, 2021 showed larger declines for the green LRSS group (-15 percent) compared to the other risk categories, where declines for in-person contacts ranged from -3 percent to -6 percent.

Similar to the prior analysis, the patterns of other-person contacts for the green LRSS group did not differ appreciably from those of the other PCRA risk categories. Last, the monthly collateral contacts for green LRSS supervisees manifested declines in 2020 and 2021 that exceeded nearly all PCRA risk categories, except for the PCRA blue group.

In addition to highlighting yearly contact numbers, it can be interesting to examine monthly contact patterns. A month-to-month analysis of contacts can highlight how people who were eligible for LRSS supervision before the enactment of the expansion were treated compared to similarly situated persons post expansion.

Figures 5a and 5b (page 52) provide information about the average number of monthly in-person contacts for the fiscal year before the enactment of the low-risk expansion (2019) and for the fiscal year after the expansion's enactment (2021). During 2019, the green LRSS group manifested monthly in-person contact rates similar to those of the green no-LRSS group. This pattern of contact activity is expected since, before the expansion, officers had no reason to treat supervisees who were classified in the green PCRA category differently, irrespective of whether they met the LRSS eligibility standards. In 2021, however, those who met the green LRSS criteria manifested contact patterns that diverged from the green no-LRSS group and somewhat mirrored supervisees with a PCRA blue risk classification.

Analysis of the LRSS Expansion's Effect on Community Safety

In the remaining part of this analysis, we examine whether enactment of the low-risk expansion endangered community safety. This analysis was conducted by exploring the recidivism activity of federal supervisees across all PCRA risk categories yearly. Recidivism includes any form of noncompliance, revocations from supervision, and rearrests for any crime or violent crimes. Supervision year encompasses any form of noncompliance, revocations, or rearrests that occurred for persons under supervision for the specific year examined.

The percentage of supervisees who engaged in noncompliance, separated into PCRA risk levels and supervision year, is highlighted in Table 3 (page 53). For all fiscal years examined, people who were eligible for LRSS supervision have noncompliance rates higher than the PCRA blues but lower than the other PCRA risk categories, including the green no LRSS, yellow, orange, or red. As anticipated, the noncompliance rates declined in 2020 and then rose irrespective of the PCRA risk levels. While the percentage of persons with noncompliance increased for the green LRSS category in 2021, the reported increase was similar to that manifested by some of the other risk groups, including the green no LRSS, yel lows, and reds.

The percentage of people revoked from supervision, separated into fiscal year and PCRA risk classification, is reported in Table

FIGURE 3A.
Percentage Change for In-Person Monthly Contacts Within 12 Months of Case Activation

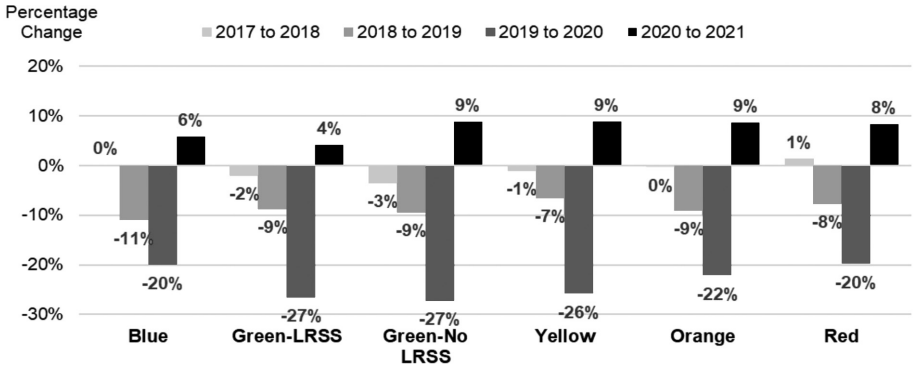


FIGURE 3B.
Percentage Change for Other-Person Monthly Contacts Within 12 Months of Case Activation

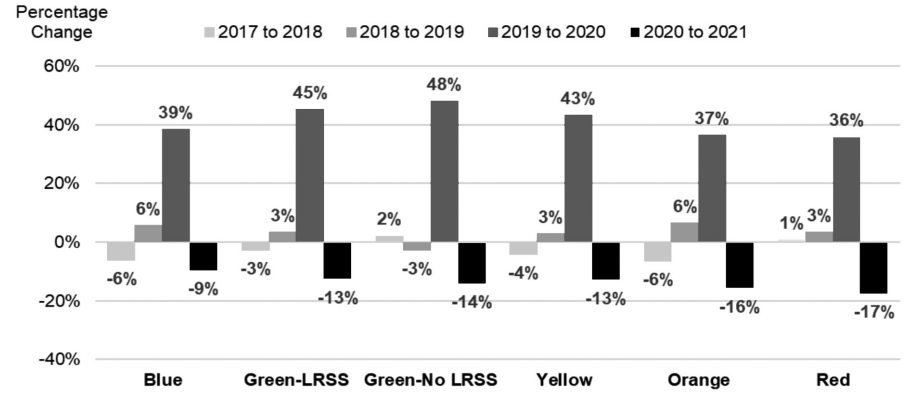
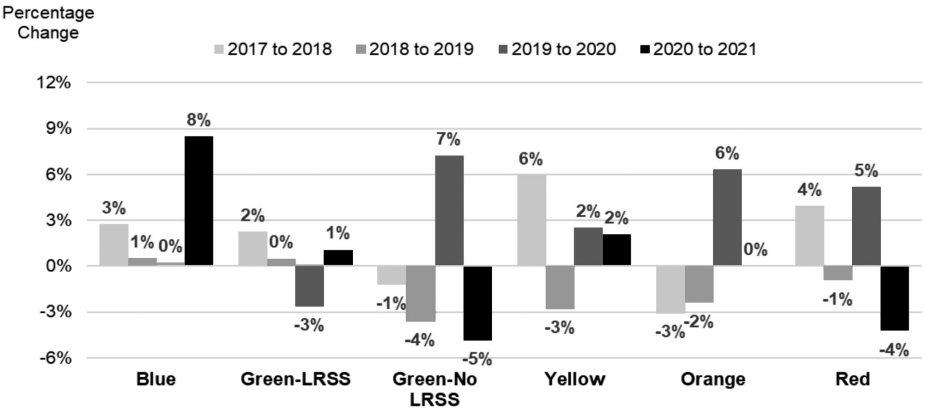


FIGURE 3C.
Percentage Change for Monthly Collateral Contacts Within 12 Months of Case Activation



4 (page 53). Similar to the noncompliance analysis, the revocation rates for people meeting the LRSS classification criteria were lower than all PCRA risk categories, except for the lowest risk classification group (PCRA blues). Although the revocation rates for the green LRSS group rose by 1 percentage point between 2020 and 2021, this increase was smaller than that manifested in the PCRA yellow, orange, and red groups. Among these three risk categories, the revocation rates increased by two to three percentage points between 2020 and 2021.

Rearrest Analysis

Information on the recidivism rates of people eligible for LRSS supervision is presented in Tables 5, 6, and 7. Tables 5 and 6 present information on the percentage of supervisees rearrested for any offense or violent offenses within the first 12 and 24 months of their supervision start dates. Rearrest rates are shown for all PCRA risk categories, though the discussion will focus on people with a green LRSS classification.

In general, results show that the rearrest rates for LRSS supervisees have remained markedly stable, even after implementation of the new low-risk expansion. For example, the percentage of LRSS supervisees who were rearrested for any offenses within 12 months of their supervision start date declined from 8 percent for people placed on supervision in 2017 to 6 percent for those placed on supervision in 2021 (see Table 5, page 53).

The 24-month rearrest rates (any offense) for this group also declined from 14 percent to 12 percent (see Table 6, page 54). The violent rearrest rates for the LRSS group never exceeded 3 percent, irrespective of the follow-up time or the supervision start year examined (see Tables 5 and 6). Last, LRSS supervisees recidivated at rates lower than nearly all PCRA risk groups, except for PCRA blues.

Table 7 (page 54) covers the rearrest activity for persons under federal supervision. Unlike the prior tables (i.e., 5 and 6) that examined the percentage of persons rearrested within 12 and 24 months after their supervision start dates, this analysis explores the percentage of persons under supervision for a particular fiscal year who were rearrested for any offense or violent offenses.

Overall, results continue to show stability in the rearrest rates pre- and post-expansion of the LRSS supervision group. Specifically, the percentage of LRSS-eligible people who were rearrested for any offense was essentially unchanged, at about 5 to 6 percent for each supervision year examined. Moreover—and perhaps more important—relatively few LRSS-eligible people (about 1 percent) were rearrested for violent offenses during the supervision years examined.

The remaining analyses (see Figures 6 and 7, page 55) explore the types of offenses for which LRSS-eligible supervisees were rearrested. Two-thirds of LRSS supervisees who recidivated were rearrested for public order (27 percent), drug (21 percent), or property (18 percent) offenses, while 12 percent were rearrested for crimes of violence (see Figure 6).

Among the LRSS-eligible supervisees who were rearrested for violent offenses, 87 percent

FIGURE 4A.
Percentage Change for In-Person Monthly Contacts by Supervision Years

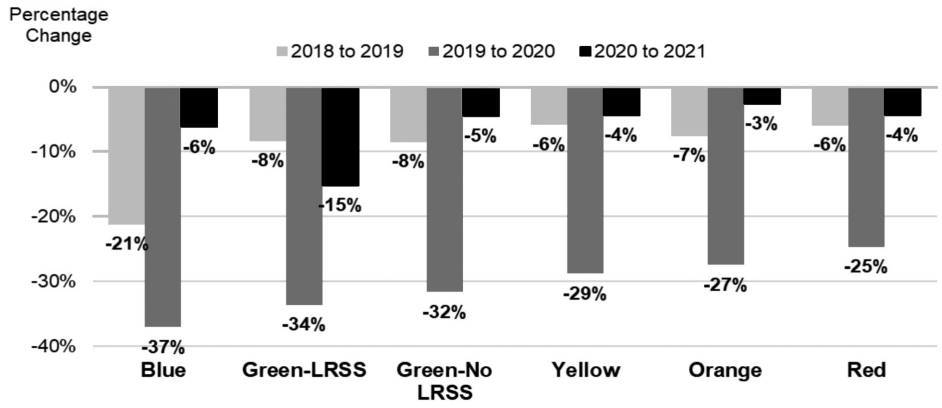


FIGURE 4B.
Percentage Change for Other-Person Monthly Contacts by Supervision Years

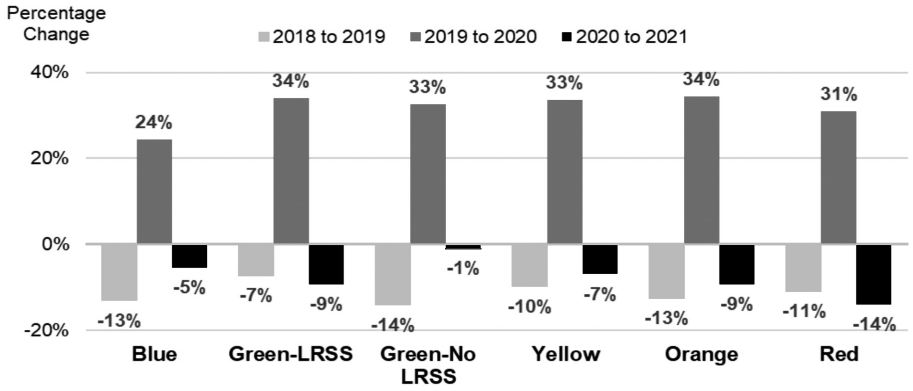
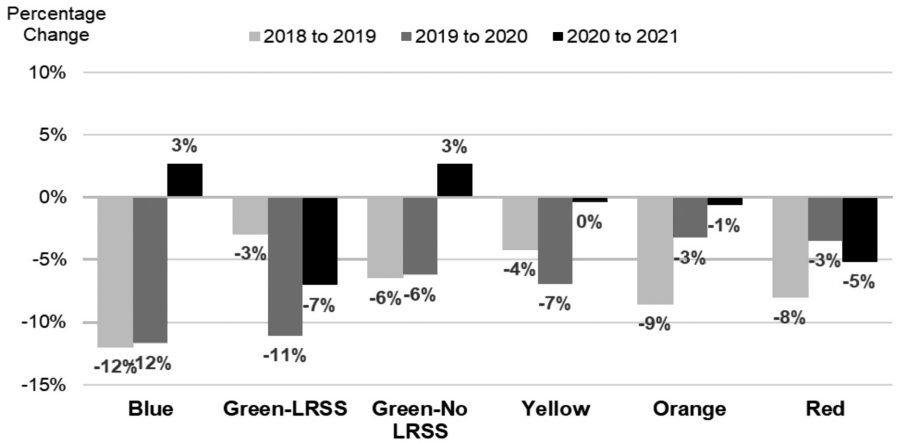


FIGURE 4C.
Percentage Change for Monthly Collateral Contacts by Supervision Years



were arrested for assault. The remaining 13 percent involved a combination of attempted or actual homicide, kidnapping, robbery, etc.

Part E: Conclusions

This study sought to examine expansion of the LRSS group, which the AO instituted in March 2020. In general, evidence produced by this study somewhat supports the contention that federal supervision practices changed for the LRSS group after enactment of the low-risk expansion. Specifically, the pattern of monthly in-person contacts for the LRSS supervisees diverged from that of the other PCRA risk groups in two substantial ways during the period after the expansion went into effect and the federal system began to recover from the pandemic.

First, our examination of the average number of in-person contacts during a person's first supervision year shows that these contacts increased less substantially for the LRSS group compared to the other risk groups in 2021. Second, our analysis of in-person contacts by supervision year demonstrates that LRSS supervisees manifested larger declines post expansion—especially in 2021—than people in the other risk categories. Officer contacts with collateral sources also manifested more substantial declines for the LRSS group once the expansion took effect.

Though these results are promising, it is important to note that the results that support system change are somewhat mixed. Unlike the in-person and collateral contacts, contacts involving electronic means of communication (e.g., other-person) did not differ appreciably between the LRSS group and the other groups. Moreover, while the post-expansion pattern of in-person and collateral contacts for the LRSS group differed somewhat from that of the other PCRA risk categories, the differences were not as extensive as initially anticipated. Additional years of officer contact activity will be required to assess whether any of these observed changes are transitory or permanent.

Given that there seems to be some indication that the federal system changed in response to enactment of the LRSS expansion, the next crucial factor to be examined was whether this expansion resulted in threats to community safety. Here the results are less ambiguous. Essentially, there is no evidence that community safety was endangered by implementation of the LRSS expansion. Specifically, while there was a slight increase in the noncompliance and revocation rates for the LRSS group post expansion, these increases were relatively negligible and did not supersede those of the other risk categories. Moreover, and this is important, the rates of noncompliance and revocations for the LRSS group were consistently lower than those of all other PCRA risk classifications, except for the blue classification group. Stated differently, changing the way that the LRSS group was supervised did not generate any appreciable increases in failure rates beyond those already predicted by the PCRA.

An examination of rearrest activity for LRSS-eligible supervisees also showed no evidence that this expansion put the community's safety at risk. Overall, rearrest rates remained markedly stable, even after this expansion was implemented. For example, the percentage of LRSS supervisees who were rearrested for any offenses within 12 or 24 months after their supervision start dates declined slightly between the pre- and post-expansion periods. Perhaps more important, the percentage of LRSS supervisees who were rearrested for violent offenses was essentially unchanged during the 2017-2021 time period and never exceeded 3 percent.

FIGURE 5A.
Average Number of In-Person Monthly Contacts by Supervision Month, Fiscal Year 2019

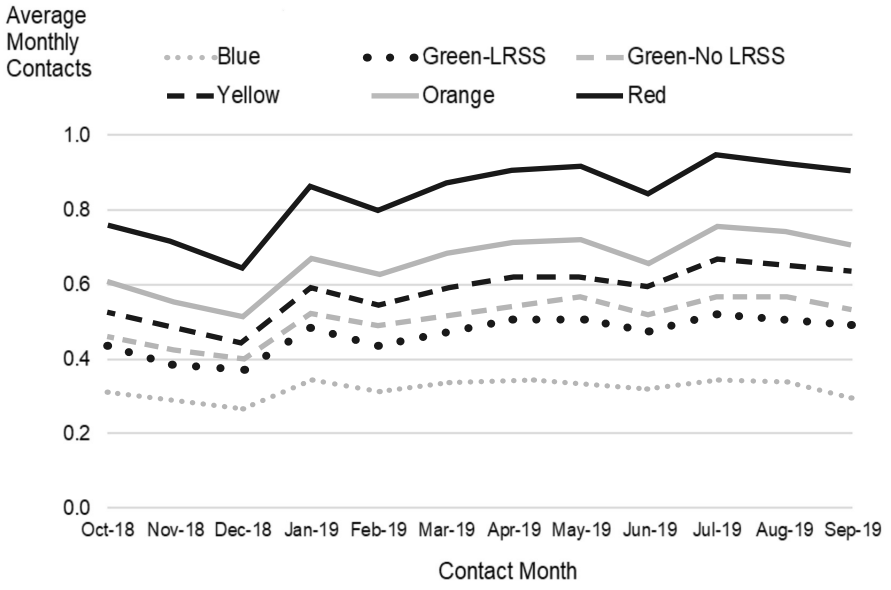


FIGURE 5B.
Average Number of In-Person Monthly Contacts by Supervision Month, Fiscal Year 2021

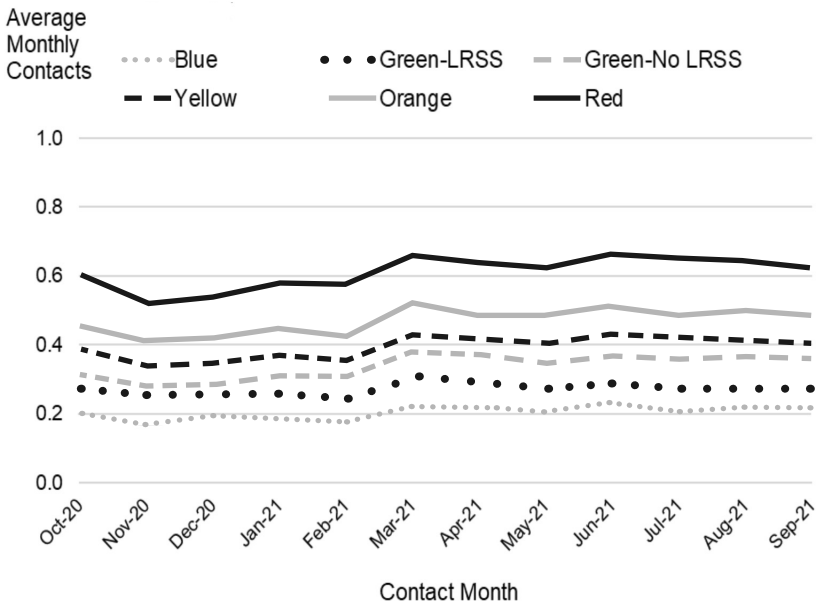


TABLE 3.
Percentage of Supervisees with Noncompliance by Fiscal Year of Case Supervision and PCRA Risk Levels

	2018			2019			2020			2021		
		Noncompliance			Noncompliance			Noncompliance			Noncompliance	
	Number	Any	Average Number	Number	Any	Average Number	Number	Any	Average Number	Number	Any	Average Number
All	53,268	40.7 %	6.9	87,405	38.9 %	6.9	110,903	34.1 %	6.4	120,812	35.3 %	6.5
Blue	9,407	19.7 %	4.1	14,769	17.4 %	4.3	18,099	14.8 %	4.1	18,632	15.4 %	4.2
Green-LRSS	**7,897**	**29.8**	**4.5**	**12,788**	**28.7**	**4.9**	**16,000**	**23.1**	**4.8**	**17,053**	**24.2**	**5.0**
Green-No LRSS	7,709	40.0	6.1	12,692	37.0	6.0	16,293	31.5	5.9	17,968	33.1	6.0
Yellow	12,235	43.8	6.9	20,387	41.9	6.9	26,206	37.1	6.3	28,959	37.7	6.4
Orange	7,739	50.3	7.7	12,956	48.3	7.5	16,830	42.6	6.8	18,872	43.9	6.8
Red	8,281	62.3	9.0	13,813	59.9	8.7	17,475	54.0	7.8	19,328	54.2	7.8

Note: Includes persons under supervision anytime during fiscal years 2018, 2019, 2020, and 2021. Noncompliance information available for 95% of people placed on supervision between fiscal years 2018 and 2021. Information on people placed on supervision in fiscal year 2017 not shown because there were too few persons under supervision at that time period, and the follow-up times were too short.

TABLE 4.
Percentage of Supervisees Revoked by Fiscal Year of Case Supervision and PCRA Risk Levels, Fiscal Years 2019–2021

	2019		2020		2021	
	Number	Percent Revoked	Number	Percent Revoked	Number	Percent Revoked
All	83,714	6.4 %	106,199	5.7 %	115,750	7.3 %
Blue	13,910	1.2 %	17,031	1.1 %	17,582	1.3 %
Green-LRSS	**13,163**	**1.7**	**16,102**	**1.7**	**16,351**	**2.7**
Green-No LRSS	10,433	6.2	13,588	5.4	15,577	6.1
Yellow	19,513	6.8	25,257	6.1	28,111	8.0
Orange	12,902	8.3	16,768	7.3	18,819	9.4
Red	13,793	13.7	17,453	11.6	19,310	14.5

Note: Includes persons under supervision anytime during fiscal years 2019, 2020, and 2021.
Revocation information available for 91% of people placed on supervision between fiscal years 2019 and 2021.
Revocation information for people placed on supervision in fiscal years 2017 and 2018 not shown because there were too few persons under supervision at that time period, and the follow-up times were too short.

TABLE 5.
Percentage of Supervisees Rearrested Within 12 Months of Fiscal Year Supervision Start Date for any Offense or Violent Offenses by PCRA Risk Levels

	2017 - Start Date			2018 - Start Date			2019 - Start Date			2020 - Start Date			2021 - Start Date		
		Percent Arrested			Percent Arrested			Percent Arrested			Percent Arrested			Percent Arrested	
	Number	Any Offense	Violent	Number	Any Offense	Violent	Number	Any Offense	Violent	Number	Any Offense	Violent	Number	Any Offense	Violent
All	20,019	14.1 %	3.1 %	38,979	14.3 %	3.1 %	42,206	13.4 %	3.0 %	39,980	12.5 %	3.0 %	9,343	11.4 %	2.6 %
Blue	4,054	5.8 %	0.5 %	7,564	6.1 %	0.4 %	7,872	6.0 %	0.6 %	7,097	5.0 %	0.4 %	1,678	5.8 %	0.6 %
Green-LRSS	**2,940**	**7.9**	**1.1**	**5,430**	**8.0**	**0.9**	**5,339**	**7.3**	**1.0**	**4,205**	**6.3**	**0.8**	**842**	**6.1**	**0.6**
Green-No LRSS	2,339	15.9	2.4	4,555	13.3	2.3	5,089	12.1	1.8	5,235	10.8	2.1	1,325	9.1	1.5
Yellow	4,145	15.1	3.4	8,406	15.9	3.3	9,169	14.7	3.0	9,027	13.5	3.1	2,094	10.8	2.1
Orange	2,829	19.2	4.9	5,409	19.8	5.5	6,107	17.7	4.7	6,049	15.5	4.3	1,420	14.5	4.4
Red	2,678	26.3	7.7	5,594	25.8	7.4	6,414	24.1	7.7	6,320	23.7	7.5	1,488	20.6	6.4

Note: Includes persons received on supervision for fiscal years 2017, 2018, 2019, 2020, and 2021. Arrest activity tracked for 12 months from supervision start date for any offense or violent offenses. Arrest activity for people received on supervision for fiscal year 2021 is limited because fewer persons could be followed for up to 12 months.
Totals include a category of supervisees—green LM1s—who were ineligible for LRSS supervision because they did not meet certain criteria (e.g., did not have two assessments, sex offender). The rearrest rates for these people are not shown.

Nearly 9 out of 10 LRSS supervisees who committed crimes of violence were arrested for assault offenses. Last, LRSS supervisees consistently recidivated at levels lower than those in the other PCRA risk categories, except for the PCRA blues.

While the study's findings are promising concerning implementation of the LRSS expansion, some limitations to this research should be noted. First, the study is relatively exploratory, meaning that descriptive statistics served as a basis for this research. And while this study is informative, more robust techniques, including multivariate analyses, will be required to further assess the expansion's efficacy.

Another issue involves the relatively short time period during which the LRSS expansion has been in effect (i.e., about 24 months) and the occurrence of the pandemic at the same time as the expansion (i.e., 2020), potentially diluting the results. Additional years of contact and recidivism data will be required to assess the permanency of this report's findings.

While important, these caveats do not take away from the principal findings highlighted in this research:

1. That the federal supervision system changed in response to the LRSS expansion.
2. That this expansion did not threaten community safety.

The AO will continue to track the LRSS policy expansion's implementation and monitor impacts on community safety.

TABLE 6.
Percentage of Supervisees Rearrested Within 24 Months of Fiscal Year Supervision Start Date for Any Offense or Violent Offenses by PCRA Risk Levels

	2017 - Start Date			2018 - Start Date			2019 - Start Date			2020 - Start Date		
		Percent Arrested			Percent Arrested			Percent Arrested			Percent Arrested	
	Number	Any Offense	Violent	Number	Any Offense	Violent	Number	Any Offense	Violent	Number	Any Offense	Violent
All	20,019	24.1 %	6.0 %	38,979	23.4 %	6.0 %	42,206	22.1 %	6.0 %	10,406	22.3 %	6.4 %
Blue	4,054	9.8 %	1.1 %	7,564	9.3 %	0.9 %	7,872	8.4 %	1.1 %	1,901	7.7 %	0.9 %
Green-LRSS	**2,940**	**14.3**	**2.8**	**5,430**	**13.9**	**2.1**	**5,339**	**12.6**	**2.3**	**1,193**	**12.4**	**1.7**
Green-No LRSS	2,339	26.5	5.1	4,555	22.6	4.6	5,089	20.3	3.8	1,348	20.1	4.8
Yellow	4,145	27.5	6.7	8,406	26.3	6.6	9,169	24.5	6.1	2,305	24.6	6.8
Orange	2,829	33.0	10.0	5,409	33.3	10.4	6,107	30.3	9.9	1,522	28.5	9.9
Red	2,678	42.2	13.5	5,594	40.9	13.9	6,414	39.0	14.1	1,598	41.7	14.9

Note: Includes people received on supervision for fiscal years 2017, 2018, 2019, and 2020. Arrest activity tracked for 24 months from supervision start date for any offense or violent offenses. Arrest activity for people received on supervision for fiscal year 2020 is limited because fewer persons could be followed for up to 24 months.
Persons received onto supervision in fiscal year 2021 not shown because none could be followed for 24 months.
Totals include a category of supervisees—green LM1s—who were ineligible for LRSS supervision because they did not meet certain criteria (e.g., did not have two assessments, sex offender). The rearrest rates for these people are not shown.

TABLE 7.
Percentage of Supervisees Rearrested for Any Offense or Violent Offenses by Fiscal Supervision Year and PCRA Risk Levels

	2018 - Supervision Year			2019 - Supervision Year			2020 - Supervision Year			2021 - Supervision Year		
		Percent Arrested			Percent Arrested			Percent Arrested			Percent Arrested	
	Number	Any Offense	Violent	Number	Any Offense	Violent	Number	Any Offense	Violent	Number	Any Offense	Violent
All	58,978	9.8 %	2.0 %	97,087	10.8 %	2.3 %	123,546	9.2 %	2.3 %	135,338	8.6 %	2.3 %
Blue	11,615	3.9 %	0.3 %	18,777	4.4 %	0.4 %	23,562	3.1 %	0.3 %	25,281	2.7 %	0.4 %
Green-LRSS	**8,369**	**5.3**	**0.8**	**13,385**	**6.0**	**0.8**	**16,355**	**5.3**	**0.9**	**16,579**	**5.3**	**1.0**
Green-No LRSS	6,892	10.1	1.5	11,461	10.4	1.8	15,007	8.3	1.4	17,226	7.7	1.7
Yellow	12,544	10.9	2.1	20,872	12.2	2.5	26,964	10.5	2.6	29,977	9.8	2.5
Orange	8,237	13.5	3.3	13,723	15.0	3.9	17,734	13.1	3.8	19,824	12.1	3.7
Red	8,266	18.0	5.2	13,802	19.2	5.4	17,464	17.4	5.7	19,323	15.6	5.0

Note: Includes people supervised during fiscal years 2018, 2019, 2020, and 2021. Arrest activity covers whether a person was arrested for each supervision year shown for any or violent offenses.
Totals include a category of supervisees—green LM1s—who were ineligible for LRSS supervision because they did not meet certain criteria (e.g., did not have two assessments, sex offender). The rearrest rates for these people are not shown.

FIGURE 6.
Most Common Offense Types for LRSS People Who Recidivated

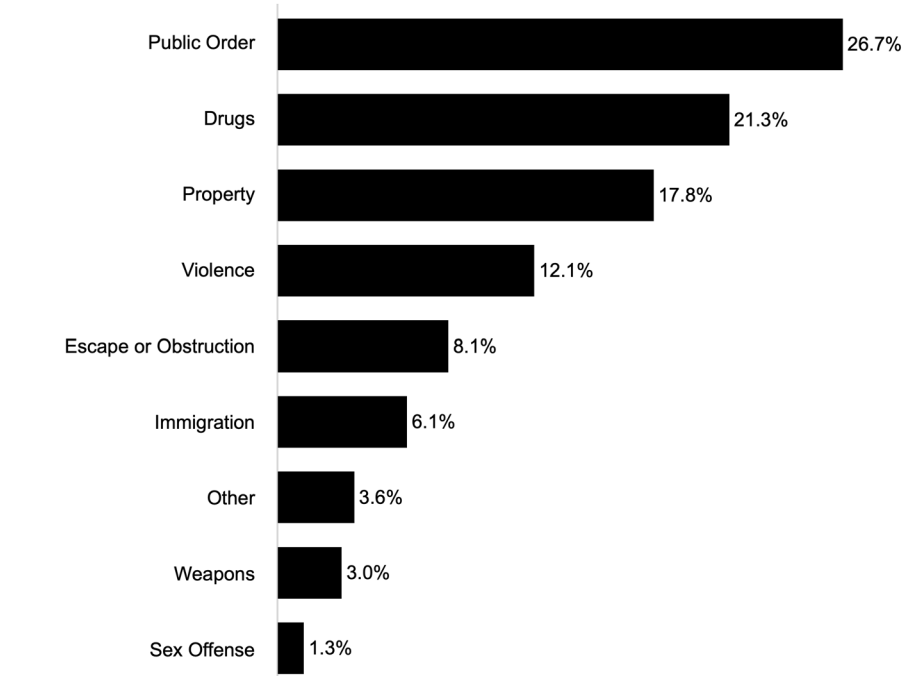
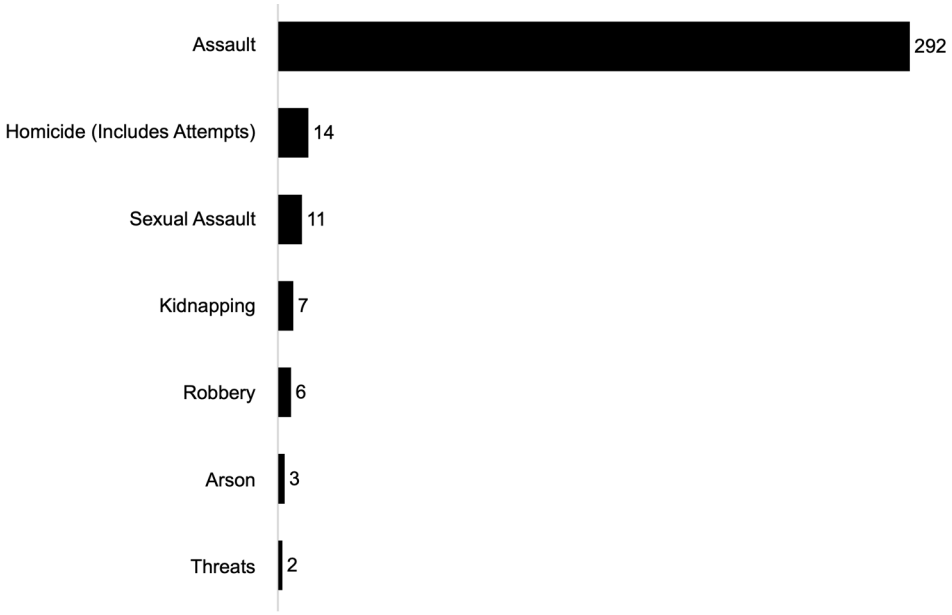


FIGURE 7.
Types of Violent Offenses Committed by People Eligible for LRSS Supervision



CONTRIBUTORS
To This Issue

Stephanie M. Biro
Senior Research Officer, Correctional Service Canada. M.A. in Psychology, Carleton University. Co-author of "Understanding Overdose Incidents in Canadian Federal Custody," *International Journal of Drug Policy*, 92 (2021).

Guy Bourgon
Psychologist, self-employed. Previously, Senior Researcher, Public Safety Canada. Ph.D., University of Ottawa. Co-author of "A System-Wide Implementation and Evaluation of the Strategic Training Initiative in Community Supervision (STICS)," *Journal of Criminal Justice, 74* (2021).

Thomas H. Cohen
Senior Social Science Analyst, Probation and Pretrial Services Office, Administrative Office of the U.S. Courts. Previously, Statistician, Bureau of Justice Statistics. Ph.D., Rutgers University School of Criminal Justice. J.D., University of Maryland School of Law. Co-author of "Examining Federal Pretrial Release Trends Over the Last Decade," *Federal Probation* (September 2018).

Joseph LaFratta
Joseph LaFratta, Principal, Scelte Consulting. Previously, Assistant Deputy Chief U.S. Probation Officer, District of Massachusetts. M.S., Northeastern University. Co-author of "Developing a Trauma-Informed Wellness Program," *Federal Probation* (2021).

Tom Hogan
Training and Program Specialist, Central Connecticut State University. B.A., Public Administration (Social Systems and Criminal Justice). Co-author of "Integrating Motivational Interviewing with Risk-Needs-Responsivity-Based Practice" in the *Wiley International Handbook of Criminal Psychology*" (2019).

Alexander M. Holsinger
Associate Dean—School of Humanities and Social Sciences and Professor of Criminal Justice, University of Missouri - Kansas City. Ph.D., University of Cincinnati. Author of "The Mentor" in the journal *Victims and Offenders*, 16(7)(2021).

Natalie J. Jones
Research Manager, Public Safety Canada. Previously, Senior Evaluation Officer, Social Sciences & Humanities Research Council of Canada. Ph.D., Carleton University. Co-author of "The Gendered Nature of Criminogenic Thinking Patterns Among Justice-Involved Clients: A Pilot Study," *Criminal Justice and Behavior*, 48(4), 442–458 (2021).

Christopher T. Lowenkamp
Probation and Pretrial Services Office, Administrative Office of the U.S. Courts. Ph.D., University of Cincinnati. Co-author of "Understanding Racial Disparities in Pretrial Detention Recommendations to Shape Policy Reform," *Criminology & Public Policy*, 22(2), 233-262.

Robert McMackin
Senior Psychologist, Aspire Health Alliance. Previously, Chief of Psychology, Lemuel Shattuck Hospital in Boston, MA. Ed.D., Boston University. Co-author of *Trauma Therapy in Context: The Science and Craft of Evidence-Based Practice.* American Psychological Association Press, Washington, D.C. (2011).

Damon Mitchell
Professor, Dept of Criminology & Criminal Justice, Central Connecticut State University. Ph.D., Kent State University.

Travis C. Pratt
Fellow, University of Cincinnati Corrections Institute. Ph.D., University of Cincinnati. Author of *Addicted to Incarceration* (2019).

Ralph C. Serin
Professor, Carleton University. Ph.D. in Psychology, Queen's University.

Raymond Chip Tafrate
Professor, Department of Criminology & Criminal Justice, Central Connecticut State University. Ph.D., Hofstra University. Author of "Criminogenic Thinking among Justice-Involved Persons: Practice Guidelines for Probation Staff" (2022).

Gina M. Vincent
Professor and Co-Director, Law & Psychiatry Program, UMass Chan Medical School. Ph.D., Simon Fraser University. Co-author of "The Impact of COVID-19 on Forensic Practitioners' Telepsychology Use and Perceptions," *Law and Human Behavior, 5,* 468-480 (2021).

Federal Probation Style Sheet

A Journal of Correctional Philosophy and Practice

Federal Probation–published three times a year by the Probation and Pretrial Services Office, Administrative Office of the U.S. Courts–has been in print for more than 80 years. Articles published in the journal reflect current thought, practice, and research in the community supervision, corrections, and criminal justice fields. Authors include criminologists, academicians, researchers, and legal specialists, but also "front-line" probation, pretrial services, and corrections officers.

U.S. probation and pretrial services officers constitute the bulk of recipients of the approximately 5,500 recipients of *Federal Probation* in hard copy. Other readers include federal judges; correctional and criminal justice practitioners at the federal, state, and local levels; and educators and students. Most of our outside subscriptions are for university libraries; in addition, *Federal Probation* is available and widely accessed online at www.uscourts.gov.

The editors seek material that is either directly related to the work and professional interests of the readers or of significant peripheral interest. Articles of interest include those that describe or evaluate a new or innovative program or approach, discuss how a policy evolved, offer insight into an issue, propose a change, or provide historical perspective. The journal publishes articles on a wide spectrum of topics; articles in recent issues have focused on topics as diverse as risk assessment, evidence-based practices in community supervision, pretrial detention, disparity, prisoner reentry, location monitoring, officer safety training, and program implementation.

The editors of *Federal Probation* are looking for original material. Manuscripts submitted to *Federal Probation* should not be previously published or under simultaneous consideration elsewhere.

Authors preparing and submitting manuscripts should:

- In general, keep manuscript length at about 3500–5000 words, although we do publish longer articles, depending on the topic and treatment.
- Double-space manuscripts, allowing standard margins.
- Submit two or three title suggestions, and a brief abstract of your article.
- Type all quoted matter of more than three lines single-spaced and indented on both sides.
- Use tables and figures intentionally, and cite them in the text. That is, they should be used to support, illuminate, or substantiate the author's narrative argument or presentation.

Email to: Ellen Wilson Fielding, Editor, *Federal Probation* (email: Ellen_Fielding@ao.uscourts.gov), Administrative Office of the U.S. Courts, One Columbus Circle NE, Washington, DC 20544 (ph: 202-502-1651).

United States Government

INFORMATION

Order Processing Code

***5876**

Federal PROBATION

a journal of correctional philosophy and practice

Please send me _____ subscription(s) to *Federal Probation* at $16.50 each ($22.40 foreign) per year. The total cost of my order is $_____.

Price includes shipping and handling and is subject to change.

_____

(PLEASE TYPE OR PRINT) NAME OR TITLE

_____

COMPANY NAME ROOM, FLOOR, OR SUITE

_____

STREET ADDRESS

_____

CITY STATE ZIP CODE + 4

_____

DAYTIME PHONE INCLUDING AREA CODE

_____

PURCHASE ORDER NUMBER (OPTIONAL)

Method of Payment

❑ Check payable to: Superintendent of Documents

❑ GPO Deposit Account _____

❑ Credit Card ❑ Visa ❑ MasterCard ❑ Discover

_____

CARD NUMBER EXPIRATION DATE

_____

AUTHORIZED SIGNATURE

ORDER ONLINE FROM

http://bookstore.gpo.gov

FAX YOUR ORDER TO

202-512-2104

PHONE YOUR ORDER TO

202-512-1800 or 866-512-1800

MAIL YOUR ORDER TO

Superintendent of Documents
P.O. Box 979050
St. Louis, MO 63197-9000

IMPORTANT!

Please include your completed order form with your remittance

thank you for your order!

FEDERAL PROBATION

Administrative Office
of the United States Courts

Washington, DC 20544