Removal of the Non-scored Items from the Post-Conviction Risk Assessment Instrument: An Evaluation of Data-driven Risk Assessment Research within the Federal System

Thomas H. Cohen[1]
Probation and Pretrial Services Office
Administrative Office of the U.S. Courts
Kristin Bechtel
Laura and John Arnold Foundation

BEYOND THE GENERATIONAL improvements observed with risk assessments, agencies have devoted a substantial amount of focused effort to develop, implement, and revise their own instruments. The preference to develop rather than adopt is often attributed to several factors, including the agency's target population, existing data, agency research capacity, staff needs, and costs. It is certainly a benefit to have a tool created specifically for an agency's population, but one potential limitation is that the instrument is developed using existing data, which may not include risk factors that research would suggest also be examined for possible inclusion in the assessment. To address this limitation, additional risk factor items can be collected but not scored; when sufficient data are available, these factors can be analyzed and, if substantial improvements in prediction are found, a revised risk assessment can be introduced.

In 2009, the Administrative Office of the U.S. Courts (AOUSC) sought to develop a dynamic risk assessment instrument comprising both risk and needs factors using existing data from the federal supervision data systems. There were several historical reasons for this shift. First, the initial risk assessments used by federal probation officers in the 1980s, the Risk Prediction Scale - 80 (RPS-80) and the United States Parole Commission's Salient Factor Score (SFS), were found to have limited predictive validity. In response to this issue, the Federal Judicial Center created and deployed the Risk Prediction Index (RPI) in the late 1990s. Although the RPI outperformed the RPS-80 and the SFS, this tool had two primary limitations. The RPI was static, which limited the federal probation officer's ability to reassess risk, and the instrument could not be used for case planning, since it lacked dynamic risk factors to target for change (AOUSC, 2011; Johnson, Lowenkamp, VanBenschoten, & Robinson, 2011; VanBenschoten, 2008). As a result, multiple commercially available instruments were considered and vetted, including the Level of Service Inventory-Revised and NorthPointe's COMPAS. Ultimately, however, the decision was made to develop the Post Conviction Risk Assessment (PCRA), using readily available federal probation data. A primary benefit of this decision was the AOUSC's ability to continuously evaluate the performance of the PCRA and, when appropriate, use the data to improve upon the assessment tool's predictive validity.

The PCRA risk score is calculated through the scoring of 15 items (located in the Officer Section of the PCRA) that have been empirically shown to be correlated with recidivism (AOUSC, 2011). The Officer Section of the PCRA also contains 15 non-scored items that prior research has suggested should predict recidivism but that, at the time of instrument development, were unavailable for analytical purposes in the AOUSC's case management systems (AOUSC, 2011). The current study seeks to examine if these 15 non-scored items improve the predictive accuracy of the instrument or if they can be removed without affecting its predictive accuracy.

Literature review

Risk prediction has undergone extensive improvements within the criminal justice field. Starting in 1954, Meehl's meta-analysis found that when reviewing multiple studies comparing actuarial and professionally derived instruments, the actuarial assessments had stronger predictive accuracy than instruments derived from professional judgment alone. Multiple subsequent studies produced similar results, leaving a lasting conclusion that risk prediction is most accurately done with actuarial risk assessment instruments rather than relying solely on professional judgment (Ægisdóttier, White, Spengler, Maugherman, Anderson, & Cook, 2006; Andrews, Bonta, &

[1] Correspondence concerning this article should be addressed to Thomas H. Cohen, Administrative Office of the U.S. Courts, One Columbus Circle, NE, Washington DC 20544. Email: thomas_cohen@ao.uscourts.gov. This publication benefited from the careful editing of Ellen Fielding.

Wormith, 2006; Grove, Zald, Lebow, Snitz, & Nelson, 2000; Latessa & Lovins, 2010; Meehl, 1954).

Four generations of risk assessment have emerged over the past 60 years. The first generation, which was guided by professional judgment, involved both correctional practitioners and clinicians making decisions about offender risk based on a review of official records, unstructured client interviews, and their professional and educational experience (Andrews & Bonta, 1998; Bonta, 1996; VanVoorhis & Brown, 1996). This first-generation risk assessment had several limitations, including lack of standardization, the potential for bias, and the inability to demonstrate inter-rater agreement among practitioners in assessing offender risk (Bonta & Andrews, 2007; Monahan, 1981; VanVoorhis & Brown, 1996). Although the first generation of risk prediction was unstandardized and often considered subjective, the process for gathering and reviewing information through interviews and a review of official records has been retained even with advances in risk assessment. What is evident in the evolution of risk assessments is that each generation of risk assessment has improved upon the previous generations' tools (Bonta, 1996).

Recognizing that one of the strongest predictors of future behavior is past behavior, formulators of the second generation of risk assessments achieved a substantial improvement by focusing on evaluating an offender's risk based on criminal history records and other official sources within a standardized and objective instrument (Bonta & Andrews, 2007). Second-generation tools incorporate primarily static risk factors, such as prior convictions, prior incarcerations, history of violence, and history of substance abuse, which are often found to be predictive of recidivism but are not necessarily derived from criminological theory (Bonta & Andrews, 2007). A well-known second-generation risk assessment, the Salient Factor Score (SFS), has been shown to be predictive of recidivism, and a primary benefit of the SFS and other second-generation tools is that the criminal history items and other static risk factors are often readily accessible within the criminal justice data systems. Further, these static risk factors have face validity, so the challenges with buy-in and professionals supporting the implementation of such instruments is often minimal, since the review of criminal history records was a common approach to decision-making within first-generation tools. However, since the second-generation instruments are composed of static items, they have limited potential for reassessment and targeting risk factors for interventions and programming (Bonta & Andrews, 2007).

Third-generation risk assessments, such as the Level of Service Inventory (LSI) and Level of Service Inventory-Revised (LSI-R), were developed in response to the inability of second-generation risk assessments to identify dynamic risk factors that could be targeted for change through programming and interventions and to reassess offenders' risk to recidivate (Andrews & Bonta, 1995; Bonta & Andrews, 2007). Since research has shown that both static and dynamic risk factors are predictive of recidivism, third-generation risk assessments continue to collect information about an offender's criminal history and other static risk factors, but also incorporate theoretically-based dynamic risk factors, or criminogenic needs, into the tools (Andrews & Bonta, 1998; Andrews & Robinson, 1984; Bonta & Wormith, 2007). With this advancement in risk assessment, offender reassessment is possible; in addition, the risk assessment can inform supervision practices and interventions based on an offender's risk and needs (Bonta & Andrews, 2007).

Although third-generation risk assessments mark a substantial gain in managing risk and identifying and targeting needs, the ability to collectively use this information to reduce risk within a formal and individualized process was not readily apparent to the field. Fourth-generation risk assessments were developed in response to this issue. Instruments such as the Level of Service/Case Management Inventory (LS/CMI) integrate the static and dynamic risk factors found within third-generation instruments, but also incorporate a formal case management process and include a systematic method for collecting information regarding responsivity factors and specific individual characteristics, such as patterns of domestic violence and incidents of institutional violence (Andrews, Bonta, & Wormith, 2006; Bonta & Andrews, 2007; Kane, Bechtel, Revicki, McLaughlin, & McCall, 2011). Fourth-generation tools are considered more comprehensive than their predecessors, since they add to the benefits of third-generation assessments a process by which this information can be thoroughly reviewed, addressed through individualized case management, and then subsequently reassessed.

The evolution of risk assessment has continuously drawn upon the benefits of prior generations and incorporated more rigorous methods to advance risk prediction (Bonta & Wormith, 2007). With more recent research, the field continues to stress the value of improving upon risk assessment instruments and practices (VanBenschoten, 2008). A fundamental objective within the federal system has been to continuously examine the use and predictive validity of its risk assessments. Empirical evaluations of prior second- and third-generation instruments within the federal system led to the most recent advancement, the development and validation of the fourth-generation PCRA.

The PCRA was initially developed and validated using three samples and comprised both scored and unscored items based on existing data and prior research (AOUSC, 2011). The original construction sample (N=51,428) and validation sample (N=51,643) contained individuals on supervised release or probation starting in October 2005. The second validation sample included 193,586 probation clients (AOUSC, 2011; Johnson et al., 2011) who started supervision between October 2005 and August 2009. The predictive accuracy of these three samples produced initial AUC-ROC values of .709 (construction), .712 (initial validation), .734 (second validation) and .783 (for long-term follow-up), suggesting that the PCRA's overall performance was good in terms of predicting recidivism (Desmarais & Singh, 2013; Doyle & Dolan, 2002; Rice & Harris, 2005). Subsequent reviews of the PCRA have demonstrated the consistent predictive accuracy of the instrument, with AUC-ROC values ranging from .70 to .77 (Lowenkamp, Johnson, Holsinger, VanBenschoten, & Robinson, 2013; Lowenkamp, Holsinger, & Cohen, 2015).

The PCRA is administered through the scoring of two sections. The first section (the Officer Section) is scored by probation officers, while offenders under supervision are responsible for completing the Offender Section of the PCRA. Since scores from the Officer Section of the PCRA are used to assess an offender's risk classification and encompass the primary items of concern for this study, we detail this section of the PCRA below.

Officer Section of the PCRA

At present, there are 15 scored items on the PCRA that measure an offender's risk characteristics on the following domains: criminal history, education/employment, substance abuse, social networks, and cognitions (e.g., supervision attitudes).[2] The criminal history domain contains six predictors that measure the number of prior felony and misdemeanor arrests, prior violent offense activity, prior varied (e.g., more than one offense type) offending pattern, prior revocations for new criminal behavior while under supervision, prior institutional adjustment while incarcerated, and offender's age at the time of supervision. The education/employment domain includes three predictors officers use to assess an offender's educational attainment, current employment status, and work history over the past 12 months. In regards to the substance abuse domain, officers score offenders on two predictors that measure whether an offender has a current alcohol or drug problem. The social network domain includes three predictors that measure an offender's marital status, presence of an unstable family situation, and the lack of any positive prosocial support networks. Last, cognitions scores an offender on one predictor that assesses an offender's attitude towards supervision and change (AOUSC, 2011).

Officers are responsible for scoring each of the 15 PCRA risk categories by interviewing offenders, reviewing relevant documents, and examining the presentence reports at the beginning of the supervision period. The PCRA scoring process uses a Burgess approach, in which each of the 15 scored predictors is assigned a value of 1 if present and 0 if absent. The exceptions include number of prior arrests (3 potential points) and age at intake (2 potential points).[3] In theory, offenders can receive a combined PCRA score ranging from 0 to 18, and these continuous scores translate into the following four risk categories: low (0-5), low/moderate (6-9), moderate (10-12), or high (13 or above). These risk categories inform officers about an offender's probability of reoffending and provide guidance on the intensity of supervision that should be imposed on a particular offender (AOUSC, 2011; Johnson et al., 2011; Lowenkamp et al., 2013).

The Officer Section of the PCRA also contains 15 additional items that are rated but not currently scored by the officer. These rated but non-scored items were included in the instrument because other empirical research—and officer input—suggested that they should be correlated with offender recidivism activity and assist officers in their case management efforts. However, at the time of instrument deployment, the AOUSC did not have the data to substantively assess whether these factors contributed to the PCRA's risk prediction accuracy outside the scored factors (AOUSC, 2011).

The non-scored factors were integrated into the PCRA domains of criminal history (1 unscored item measuring prior juvenile arrest history), education/employment (2 unscored items measuring the number of employers over the last 12 months and whether the offender was employed over 50 percent of the time during the previous two years), substance abuse (4 unscored items measuring whether drug or alcohol abuse resulted in disruptions at work, school, or home; whether the offender used drugs or alcohol in physically hazardous conditions; whether drug use continued despite social or interpersonal problems; or whether legal problems have occurred because of drug or alcohol abuse), and social networks (3 unscored items measuring whether the offender lives with a spouse or children; whether the offender has any family support; and whether the offender associates with positive or negative peers). For the cognitions domain, there was one unscored item assessing whether the offender had antisocial attitudes. Other unscored factors include four items measuring an offender's residential stability, criminal risks at home, financial situation, and level of engagement in prosocial activities (AOUSC, 2011).

It should be noted that the cognitions domain also extracts information from the Offender Section of the PCRA on the different types of criminal thinking styles that an offender might manifest. Since this study focuses solely on the scored and non-scored items contained in the Officer Section of the PCRA, we omit discussing the contribution to assessment made by the Offender Section of the PCRA. Further details on the PCRA's assessment of an offender's criminal thinking styles appear in studies published by Walters and Lowenkamp (2016) and Walters and Cohen (2016).

When the PCRA was initially implemented, it was decided to empirically explore whether these non-scored factors should eventually be incorporated into the instrument's scoring mechanism by testing whether they contributed to risk prediction above that of the scored factors (Lowenkamp et al., 2013). As we will discuss below, most of these non-scored items did not contribute to the PCRA's risk prediction effectiveness and hence will be removed from the instrument, making room for a new trailer to assess the probability of an offender being involved in a violent crime (Serin, Lowenkamp, Johnson, & Trevino, 2016).

Method

Research Agenda

In the current study we sought to explore whether the non-scored items could be removed from the Officer Section of the PCRA without compromising the instrument's risk prediction effectiveness. Specifically, we examined whether combining the 15 scored and 15 non-scored items in the PCRA's risk prediction algorithm resulted in an instrument capable of predicting offender recidivism behavior to a greater extent than the current algorithm containing only the 15 scored items. Results showing either no or negligible improvements provide empirical support for the decision to remove these non-scored items. Conversely, findings demonstrating substantial improvements in risk prediction from use of the non-scored items would indicate that the AOUSC should consider integrating these non-scored items into the risk calculation.

Our analysis of the non-scored items proceeded through several stages. Initially, we examined whether the non-scored items were more likely to be found among the high- compared to the low-risk offenders.[4] Next, we explored the bivariate correlation between the non-scored items and offender recidivism outcomes involving any or violent offenses. Afterwards, we investigated whether combining the 15 scored and 15 non-scored items into a new prediction score resulted in an improvement in recidivism prediction over that already achieved by the actual scores currently generated by federal probation officers. Finally, we evaluated whether the presence of any of the factors measured by the individual non-scored items were significantly correlated with offender rearrest activity (e.g., any or violent) while simultaneously controlling for all scored PCRA items, and

[2] See Appendix Table 1 for an overview of the scored and non-scored risk factors.

[3] Assigning scores ranging from 0 to 3 may seem counterintuitive to current trends that involve the development of weighted risk assessments; however, there is significant evidence to support the argument that this method still outperforms clinical approaches and is more robust across time and sample variations (Gottfredson & Snyder, 2005; McEwan, Mullen, & Mackenzie, 2009).

[4] Prior research (Cohen & VanBenschoten, 2014) has shown the factors measured by the scored items being present to a greater extent among high-risk compared to low-risk offenders.

(if any significant associations were found) whether the inclusion of these specific non-scored items significantly improved the instrument's overall predictive efficacy.

Study Population

The study population includes all PCRA assessments that occurred during an offender's first term of post-conviction supervision[5] whose recidivism outcomes could be tracked for a minimum of 12 months (N=196,460). These initial assessments occurred during the period spanning November 2009 through January 2015. Recidivism is defined as the arrest of an offender for either a felony or misdemeanor offense (excluding arrests for technical violations) within one year after the PCRA assessment date. In addition to measuring any arrests, we also identified arrests for violent offenses committed within one year after the initial PCRA assessment. For violent arrests we used the definitions from the National Crime Information Center (NCIC), which includes homicide and related offenses, kidnapping, rape and sexual assault, robbery, and assault (Lowenkamp et al., 2015). The recidivism data were gathered through the NCIC and Access to Law Enforcement System databases (ATLAS).[6]

As stated previously, the study population included offenders with initial PCRA assessments whose recidivism outcomes could be followed for a minimum of 12 months (N = 196,460). The 12-month follow-up period allows us to track whether offenders were arrested for any or violent offenses within 12 months of receiving their first PCRA assessment. We also included follow-up periods encompassing 24 months (N = 157,169) and 36 months (N = 116,014). Examining the non-scored PCRA items for different follow-up periods allowed us to assess whether any predictive enhancements from the non-scored items might be obtained for offenders whose recidivism outcomes could be tracked for multiple-year time periods.[7]

Measuring the Unscored PCRA Items

To reiterate, the PCRA's non-scored items are the items that are rated but not scored on the PCRA worksheet. These non-scored items were integrated into the PCRA domains of criminal history (1 unscored item), education/employment (2 unscored items), substance abuse (4 unscored items), social networks (3 unscored items), and cognitions (1 unscored item). Other unscored items include 4 items measuring an offender's residential stability, criminal risks at home, financial situation, and level of engagement in prosocial activities (AOUSC, 2011; Johnson et al., 2011). The prior section discussing the Officer Section of the PCRA and Appendix Table 1 provides a fuller description of the values assigned to both the non-scored and scored items on the PCRA worksheet.[8] With the exception of the items associated with positive/negative peers item, which has four values,[9] all the non-scored items are measured using dichotomous scales.

In addition to examining whether these non-scored items individually improved risk prediction, we transformed the scored and non-scored items into predicted risk scales to investigate whether including the non-scored items in the risk algorithm could significantly enhance recidivism prediction. The PCRA scoring process generates a raw risk score ranging from zero to 18 that is then used to classify offenders into one of four risk categories (i.e., low, low/moderate, moderate, or high) (AOUSC, 2011; Johnson et al., 2011). We compared the predictive effectiveness of these raw scores with risk scores generated by using all 30 scored and non-scored items that were also scaled to range from zero to 18. This method, which will be more fully explicated in the findings section, allowed us to analyze whether integrating the non-scored items into the risk calculation resulted in a demonstrably superior risk prediction scale.

Analysis Plan

We assessed whether the non-scored items improved the PCRA's risk prediction effectiveness through several stages. First, we used bivariate statistics (including means, cross tabulations, and chi-square statistics) to examine these non-scored items by risk level and determine whether the non-scored items were correlated with offender recidivism outcomes. Next, we employed multivariate approaches, specifically logistic regression, to investigate whether combining the 15 scored and 15 non-scored items into a revised risk scale enhanced the PCRA's risk prediction capabilities above those already achieved by the officer-calculated raw risk scores. Finally, we used stepwise logistic regression methods to assess which of the individual non-scored items were significantly correlated with offender recidivism outcomes. We also calculated zero-order correlations and area under the receiver curve operating characteristics (AUC-ROC) scores to evaluate whether the non-scored items significantly enhanced this instrument's risk-scoring capabilities or whether these items could be removed without compromising the PCRA's predictive effectiveness.

Results

Study Cohorts

Table 1 (next page) shows the raw PCRA risk distributions of offenders followed for different time periods in the study cohort, including 12 months, 24 months, and 36 months. We show the raw risk scores rather than the risk categories because these scores will be used as the primary means for assessing risk prediction in the extant study.[10] Although the raw PCRA score can reach a maximum value of 18, these values were recoded into a score of 17, as relatively few offenders (N=10) received the maximum score. In general, there were relatively negligible differences in the risk scores for the different follow-up groups. The

---

[5] Post-conviction supervision encompasses offenders sentenced to either supervised release or probation. Supervised release refers to offenders sentenced to a term of community supervision following a period of imprisonment within the Federal Bureau of Prisons (18 U.S.C. §3583), while probation refers to offenders sentenced to a period of supervision without any imposed incarceration sentence (18 U.S.C. §3561).

[6] ATLAS is a software program used by the Administrative Office of the U.S. Courts that provides an interface for performing criminal record checks through a systematic search of official state and federal rap sheets. It is widely used by probation and pretrial services officers to perform criminal record checks on defendants and offenders for supervision and investigation purposes (Baber, 2010).

[7] But see Flores, Holsinger, Lowenkamp, and Cohen (2016) for a discussion of the methodological usefulness of following offenders for time periods exceeding one year.

[8] The non-scored measures shown in Appendix Table 1 were recoded into numeric values for analytical purposes.

[9] It should be noted that we recoded the associates with negative peers or no friends item from four values to three as the recidivism rates for the "no friends" score (11 percent) was relatively similar to the recidivism rates for the "occasional association with negative peers" score (13 percent).

[10] It should be noted that since the study's sole focus was to assess whether the non-scored items increased the PCRA's predictive efficacy, we omitted variables on offender race/ethnicity/gender that have been included in other PCRA validation studies. For a discussion of the PCRA's capacity to predict recidivism across various offender demographic categories see Lowenkamp at al., 2015; Skeem & Lowenkamp, 2016; and Skeem, Monahan, & Lowenkamp, 2016.

overall mean PCRA scores decreased slightly from 6.5 for the 12-month follow-up to 6.4 for the follow-up groups in the 24- and 36-month range. The percentages of offenders classified in the moderate- or high-risk categories (i.e., who received scores of 10 points or more) were also similar across the three follow-up groups, spanning from 21 percent for the 12-month follow-up to 19 percent for the 36-month follow-up.

Table 2 explores the presence of the non-scored PCRA risk items by an offender's initial risk classification. Average scores for each of the non-scored items, with standard deviations in parentheses, are shown. With the exception of the "associates with negative peers or no friends" variable, all these mean scores could be converted into percentages, as they are binary values with scores of 0/1. Not surprisingly, this table shows that the non-scored risk items are more likely to be present among offenders initially classified into the higher risk categories by the PCRA. According to these non-scored items, offenders classified as higher risk by the PCRA are more likely to manifest juvenile criminal histories, job instability, substance abuse problems, weak social networks, and negative antisocial attitudes/values than lower risk offenders. Moreover, the non-scored items showed that higher risk offenders were more likely to lack any permanent residence, have criminal risks present at home, deal with financial stressors, and fail to engage in prosocial activities to a greater extent than their lower risk counterparts. Overall, the distribution of these non-scored items provides empirical evidence supporting the proposition that the PCRA can distinguish even among risk factors that are currently not included in the actual PCRA risk calculations.

TABLE 1.
Distribution of Post Conviction Risk Assessment (PCRA) categories, by offender follow-up period

	12 months		24 months		36 months	
Raw PCRA scores	Number	Percent	Number	Percent	Number	Percent
0	5,011	2.6%	3,907	2.5%	2,881	2.5%
1	10,160	5.2%	8,119	5.2%	6,073	5.2%
2	14,187	7.2%	11,361	7.2%	8,467	7.3%
3	16,027	8.2%	13,038	8.3%	9,734	8.4%
4	16,662	8.5%	13,489	8.6%	10,067	8.7%
5	17,615	9.0%	14,442	9.2%	10,838	9.3%
6	19,270	9.8%	15,747	10.0%	11,894	10.3%
7	20,034	10.2%	16,358	10.4%	12,395	10.7%
8	19,409	9.9%	15,570	9.9%	11,600	10.0%
9	17,417	8.9%	13,820	8.8%	10,156	8.8%
10	14,035	7.1%	11,009	7.0%	7,940	6.8%
11	10,269	5.2%	8,001	5.1%	5,612	4.8%
12	7,281	3.7%	5,527	3.5%	3,799	3.3%
13	4,610	2.4%	3,485	2.2%	2,443	2.1%
14	2,577	1.3%	1,895	1.2%	1,244	1.1%
15	1,229	0.6%	905	0.6%	564	0.5%
16	506	0.3%	375	0.2%	246	0.2%
17	161	0.1%	121	0.1%	61	0.1%
Mean score	6.49		6.44		6.36	
	(3.50)		(3.46)		(3.41)	
Number of offenders	196,460		157,169		116,014	

Note: Percentages may not sum to totals due to rounding error. The PCRA 18s have been recoded into 17s because relatively few offenders (N=10) obtained scores of 18. Standard deviations in parentheses.

Relationship Between Non-scored Factors and Recidivism Outcomes

Tables 3 and 4 examine the relationship between the non-scored risk items and rearrest activity for any or violent offenses at the bivariate level. The 12-month follow-up group was used, as this group had the largest number of offenders (N=196,460) among the three study cohorts, and chi-square tests were employed to assess whether the recidivism rates significantly increased for offenders with any of these non-scored risk characteristics. The bivariate analysis shows all the 15 non-scored items being significantly associated with increases in offender recidivism rates involving arrests for any felony or misdemeanor offenses at the .001 level. For example, the percent of offenders rearrested within 12 months of their initial assessment increases from 7 percent for those with good support networks to 19 percent for offenders with more than occasional association with negative peers. All the non-scored items were also significantly correlated with violent recidivism. This analysis showing that offenders characterized by issues measured by the non-scored items (including serious criminal histories, job instability, substance abuse issues, poor social networks, negative social attitudes or other issues associated with residential or financial instability) were more likely to recidivate should not be too surprising given the extensive literature showing the correlation between these factors and criminal conduct (Andrews & Bonta, 2010). The large study population of almost 200,000 offenders also makes probable findings of statistical significance. Whether these non-scored factors contributed to the PCRA's overall predictive capacities above that currently achieved by the 15 scored factors is an issue further explored in the next section.

Contribution of Non-scored Factors to Risk Prediction

The remaining tables and figures investigate whether inclusion of the non-scored items both substantially and significantly improved the PCRA's risk prediction accuracy. Basically, this analysis tests whether the PCRA's predictive accuracy can be improved by using both the 15 scored and 15 non-scored items to redistribute offenders by their probability of recidivism (any or violent). We conducted this analysis by employing logistic regression models to calculate a predictive probability of any or violent recidivism for offenders in the different population follow-up groups (e.g., 12 months, 24 months, or 36 months).[11] Using

[11] Logistic regression is a commonly used statistical technique applied when examining the effects of

TABLE 2.
Mean scores for non-scored Post Conviction Risk Assessment (PCRA) items, by initial risk classification

		Offenders by initial risk classification			
Non-scored PCRA items	All offenders	Low	Low/Moderate	Moderate	High
Criminal history					
Juvenile arrest	0.27 (0.45)	0.08 (0.27)	0.32 (0.47)	0.53 (0.50)	0.66 (0.47)
Education & employment					
Multiple jobs past year	0.51 (0.50)	0.38 (0.49)	0.53 (0.50)	0.70 (0.46)	0.83 (0.38)
Employed less than 50% over past two years	0.50 (0.50)	0.32 (0.47)	0.53 (0.50)	0.77 (0.42)	0.91 (0.29)
Drugs & alcohol					
Drug use related to disruption at work, school, or home	0.27 (0.44)	0.11 (0.31)	0.30 (0.46)	0.47 (0.50)	0.65 (0.48)
Drug use in physically hazardous conditions	0.21 (0.41)	0.10 (0.30)	0.24 (0.43)	0.33 (0.47)	0.44 (0.50)
Drug use led to legal problems	0.40 (0.49)	0.19 (0.40)	0.46 (0.50)	0.65 (0.48)	0.79 (0.41)
Drug use continued despite social problems	0.30 (0.46)	0.12 (0.32)	0.35 (0.48)	0.54 (0.50)	0.71 (0.45)
Social networks					
Lives with spouse and/or children	0.65 (0.48)	0.53 (0.50)	0.69 (0.46)	0.79 (0.41)	0.84 (0.36)
Lacks family support	0.09 (0.29)	0.05 (0.22)	0.09 (0.28)	0.16 (0.37)	0.31 (0.46)
Associates with negative peers or no friends	0.66 (0.90)	0.34 (0.72)	0.69 (0.89)	1.12 (0.95)	1.61 (0.82)
Cognitions					
Harbors antisocial attitude/ values	0.13 (0.34)	0.05 (0.22)	0.11 (0.32)	0.27 (0.44)	0.56 (0.50)
Other factors					
Lacks permanent residence	0.37 (0.48)	0.25 (0.43)	0.39 (0.49)	0.52 (0.50)	0.66 (0.47)
Criminal risks present in home	0.11 (0.32)	0.06 (0.24)	0.11 (0.31)	0.20 (0.40)	0.36 (0.48)
Financial stressors present	0.32 (0.47)	0.18 (0.38)	0.31 (0.46)	0.57 (0.50)	0.82 (0.38)
Does not engage in pro-social activities	0.26 (0.44)	0.15 (0.36)	0.26 (0.44)	0.43 (0.50)	0.68 (0.47)
Number of offenders	196,460	79,662	76,130	31,585	9,083

Note: Includes offenders followed for a period of 12 months. Standard errors shown in parentheses.

a model-driven approach entails generating a predicted probability for each offender in the study population being arrested that can theoretically range from 0 to 1. A 0 means that the offender has no predicted chance of being arrested, while a 1 would imply a 100 percent chance of recidivating. These predicted arrest probabilities contrast with the original officer-generated PCRA scores, which range from 0 to 18. Although the arrest probabilities produced by the logistic regressions differ from the raw PCRA risk scores, these predicted probabilities can be re-scaled through a ranking process into a scoring distribution mirroring that of the PCRA scales. Specifically, we compared the logistic regression-predicted arrest probabilities to those using the natural PCRA risk scale by dividing the ranked predictions into revised risk scores of the same size as the natural risk scores for each estimated follow-up group.

multiple independent variables on a dichotomous dependent variable (Hilbe, 2009).

As an example, suppose that for a given sample, the following percentage of offenders received PCRA raw scores of 0 (2.6 percent), 1 (5.2 percent), and 2 (7.2 percent). In this case, we ranked offenders by their predicted recidivism values—least to most risky—and selected the bottom 2.6 percent to have predicted PCRA risk scores of 0, the next 5.2 percent to have predicted PCRA risk scores of 1, and the following 7.2 percent of offenders to have predicted PCRA risk scores of 2. This procedure was followed until all offenders were redistributed by their predicted PCRA risk scores, which could range from 0 to 18. To the extent that offenders with predicted PCRA risk scores of 0, 1, or 2 comprise different groupings than offenders with original PCRA risk scores of 0, 1, or 2, rearrest rates may differ across the two groups. Moreover, the revised PCRA risk groupings might manifest higher AUC-ROC values and correlations than the original PCRA risk distributions.

Results from this analysis are presented in Table 5 and Figure 1 (any recidivism) and Table 6 and Figure 2 (violent recidivism). In general, these results show that the PCRA's risk prediction capabilities were only marginally improved by incorporating the 15 non-scored items into a revised prediction index. These marginal improvements can be viewed through an analysis of the AUC-ROC scores. The AUC-ROC score is frequently used to assess risk assessment instruments and is often preferred over a correlational analysis because it is not impacted by low base rates (Lowenkamp et al., 2013). Essentially, the AUC-ROC measures the probability that a score drawn at random from one sample or population (e.g., offenders with a rearrest) will be higher than that drawn at random from a second sample or population (e.g., offenders with no rearrest) (Lowenkamp et al., 2013; Rice & Harris, 2005). Values for the AUC-ROC range from .0 to 1.0, with values of .70 or greater indicating that the actuarial instrument does fairly well at prediction (Andrews & Bonta, 2010).

For the 12-month follow-up group, the AUC-ROC scores were higher for the rescaled prediction score (AUC-ROC = 0.73), but only slightly so, compared to the originally calculated PCRA score (AUC-ROC = 0.72). Though the confidence intervals show significant differences between the rescaled and actual PCRA scores, a 0.01 increase in the AUC-ROC score indicates that the rescaled scores were not substantively different in terms of their risk prediction capacities than

TABLE 3.

Percent of offenders arrested within 12 months of initial assessment for any offense for the non-scored Post Conviction Risk Assessment (PCRA) items

	12 month rearrest rates by recorded score			
Non-scored PCRA items	**0**	**1**	**2**	**Chi-square**
Criminal history				
Juvenile arrest	7.6%	17.0%		3800.0***
Education & employment				
Multiple jobs past year	8.2%	12.0%		796.1***
Employed less than 50% over past two years	7.3%	12.9%		1700.0***
Drugs & alcohol				
Drug use related to disruption at work, school, or home	8.5%	14.6%		1500.0***
Drug use in physically hazardous conditions	9.2%	13.8%		747.0***
Drug use led to legal problems	7.6%	13.9%		2100.0***
Drug use continued despite social problems	8.1%	14.7%		2000.0***
Social networks				
Lives with spouse and/or children	7.1%	11.7%		1000.0***
Lacks family support	9.7%	13.9%		325.5***
Associates with negative peers or no friends	7.0%	13.0%	19.3%	3400.0***
Cognitions				
Harbors antisocial attitude/values	9.1%	16.7%		1500.0***
Other factors				
Lacks permanent residence	8.6%	12.8%		875.7***
Criminal risks present in home	9.5%	14.6%		568.2***
Financial stressors present	8.0%	14.5%		2000.0***
Does not engage in prosocial activities	8.7%	14.2%		1300.0***

*Note: Includes offenders followed for a period of 12 months. For the associates with negative peers item, the values for no friends were recoded into occasional association with negative friends as the rearrest rates for the no friends (11%) was similar to the occasional association with negative peers (13%). *p < .05; **p < .01; ***p < .001*

scores actually generated by federal probation officers. These patterns in AUC-ROC scores held across the 24- and 36-month follow-up groups. For example, the actual PCRA scores produced AUC-ROC values that were relatively stable at the 24-month (0.72) and 36-(0.72) month follow-ups, while the rescaled PCRA indices showed improvements in risk prediction, with the AUC-ROC scores increasing to 0.74 at the 36-month follow-up. The divergence in the AUC-ROC scores between the actual (0.72) and rescaled (0.74) PCRA scores at the 36-month follow-up nevertheless reveals only negligible improvements resulting from the inclusion of all 15 non-scored items in the risk score calculation. In addition to the AUC-ROC scores, an analysis of the zero-order correlations reveals relatively small improvements when moving from the actual to rescaled PCRA scores.

Another way of examining whether the non-scored items could enhance risk prediction is to analyze the relationship between the recidivism rates for the actual and rescaled PCRA scores. While this analysis is provided in Table 5, Figure 1 presents a clearer picture, visualizing the functional form associated with the recidivism rates for the actual and rescaled PCRA scores. An examination of the functional form between recidivism and the PCRA scores shows the rearrest rates being essentially the same for both the actual and rescaled PCRA indices from values 0 through 11; afterwards, the rearrest rates diverge, with the rescaled PCRA scores manifesting higher arrest rates than the actual PCRA scores. The rearrest rates then begin re-converging at the highest PCRA values. This pattern shows the rescaled PCRA scores providing a better metric for identifying offender recidivism events, but only for those PCRA values at the higher end of the risk distribution. There were relatively negligible differences in the capacity to detect rearrest activity for the lower PCRA scores.

A similar pattern of marginal improvements in prediction using the rescaled PCRA scores held when examining violent recidivism outcomes at the 12-, 24-, and 36-month follow-up intervals. Specifically, the AUC-ROC scores manifested some improvements in recidivism prediction for violent offenses; moreover, the violent rearrest rates for the actual and rescaled PCRA scores were relatively similar for the PCRA values ranging from 0 through 13, after which they diverge, with the rescaled PCRA scores evidencing improved capacities to detect violent rearrests compared to the officer-generated scores.

Although improvements in recidivism prediction demonstrated in the previous analyses might be seen as a rationale for including the 15 non-scored items in the risk prediction calculation, it is important to note that in part these findings result from comparing predictions between actual and model-generated PCRA scores. Some recent research has suggested that risk scores generated through a Burgess scoring approach of the type used by the PCRA could produce inferior prediction scales compared to model-generated scores (Kim & Duwe, 2017). Hence, the modest improvements in prediction might be the result of using model-based approaches in addition to including the 15 non-scored risk items in the rescaled PCRA score. One way around this issue involves comparing the predictive indices produced from logistic regression models containing only the 15 scored PCRA risk items with those of models containing both the scored and non-scored PCRA items. This approach also allows us to assess which of the non-scored PCRA items might be correlated with recidivism when the scored PCRA items are statistically controlled and whether inclusion of any of these non-scored PCRA items significantly improves the model's capacity to predict recidivism.

In the analysis presented in Tables 7 and 8, we used backward stepwise logistic regression models to examine which of the non-scored items were significantly correlated with recidivism outcomes involving any or violent offenses, while controlling for the scored PCRA items using the 12-month follow-up group. The backward stepwise approach uses an iterative process that systematically identifies and removes variables that do not improve the model's overall fit (Field, 2013). This method works by initially placing all 15 non-scored items in the model and then

TABLE 4.
Percent of offenders arrested within 12 months of initial assessment for violent offenses for the non-scored Post Conviction Risk Assessment (PCRA) items

Non-scored PCRA items	12 month violent rearrest rates by recorded score 0	1	2	Chi-square
Criminal history				
Juvenile arrest	1.3%	3.9%		1400.0***
Education & employment				
Multiple jobs past year	1.5%	2.5%		227.7***
Employed less than 50% over past two years	1.3%	2.7%		455.2***
Drugs & alcohol				
Drug use related to disruption at work, school, or home	1.6%	3.0%		392.3***
Drug use in physically hazardous conditions	1.8%	2.8%		172.2***
Drug use led to legal problems	1.5%	2.8%		426.9***
Drug use continued despite social problems	1.6%	3.0%		448.1***
Social networks				
Lives with spouse and/or children	1.4%	2.3%		186.8***
Lacks family support	1.9%	2.8%		72.6***
Associates with negative peers or no friends	1.3%	2.7%	4.0%	766.4***
Cognitions				
Harbors antisocial attitude/values	1.8%	3.6%		396.5***
Other factors				
Lacks permanent residence	1.7%	2.6%		210.2***
Criminal risks present in home	1.9%	3.0%		134.1***
Financial stressors present	1.5%	3.0%		437.7***
Does not engage in prosocial activities	1.7%	2.9%		265.5***

*Note: Includes offenders followed for a period of 12 months. For the associates with negative peers item, the values for no friends were recoded into occasional association with negative friends as the violent arrest rates for both values were similar. *p < .05; **p < .01; ***p < .001*

calculating the contribution of each item by analyzing whether it meets criteria for inclusion specified by the user. The variable with the weakest explanatory power per the user's criteria is removed and the model is then re-estimated. This iterative process repeats itself until all the remaining covariates in the model meeting the user-specified criteria remain (Field, 2013).

For this analysis, the user-specified criteria involved retaining all non-scored items with p-values of 0.01. We selected this p-value by using the Bonferroni criterion, which entailed dividing the p-value of 0.05 into the number of non-scored variables being tested (N=15) (Allison, 2015). It is important to note that we employed stepwise deletion approaches only on the non-scored PCRA items. In other words, the 15 scored PCRA items were forced into the model, while the remaining non-scored items were subjected to exclusion through the backward stepwise regression models. This approach provides a parsimonious method for ascertaining which of the non-scored items were significantly correlated with recidivism when the PCRA factors were statistically controlled. We also provide AUC-ROC scores and sensitivity statistics to ascertain whether inclusion of the significant non-scored factors enhanced the model's overall predictive accuracy.[12] The sensitivity statistics were based on the 12-month rearrest rate for any (10.1 percent) or violent (2.0 percent) offenses. Finally, recidivism outcomes were modeled for the 12-month follow-up group, as that cohort had the largest number of offenders.

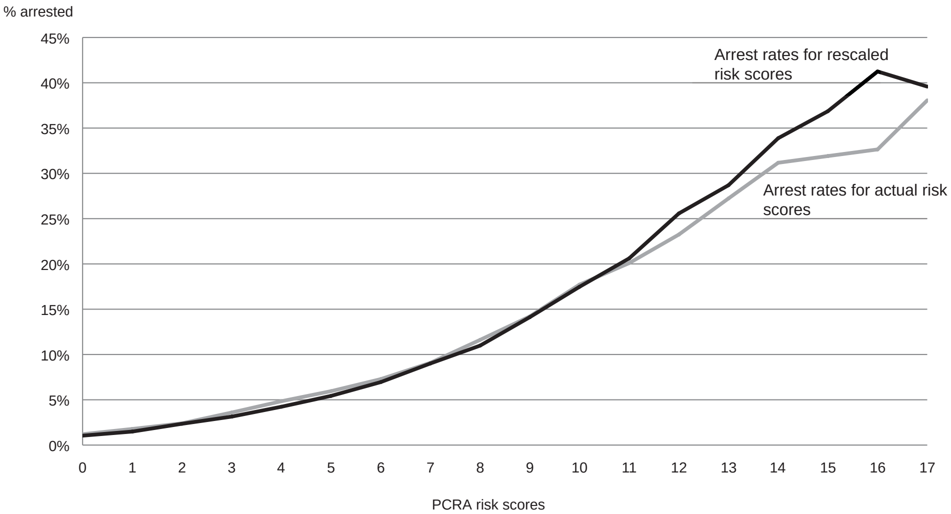
Results show several non-scored items being significantly correlated with recidivism outcomes involving any or violent offenses. The variables that were significantly correlated with any or violent rearrest behavior at the 0.01 level, net of the PCRA controls, include prior juvenile arrest, employed less than 50 percent over the past two years, does not live with spouse or children, associates with negative peers, and financial stressors. In addition, the non-scored PCRA item of drug use led to legal problems was significantly correlated with general but not violent recidivism. Interestingly, while the model containing only the scored items shows all 15 of these factors being significantly associated with general rearrests when the non-scored items were included in the regression model, some of the scored items—including prior varied offending pattern, good work assessment, current alcohol problem, and unstable family situation—witness a weakening or loss of their significant association with the any recidivism outcome. This finding should not be too surprising, as bringing the non-scored variables into the model should result in some of the original scored items becoming less significantly associated with the dependent variable.[13]

The key issue, however, involves whether adding these non-scored items significantly improves the model's efficacy at predicting any forms of recidivism. An analysis from the model-generated AUC-ROC and sensitivity statistics shows no significant improvement in prediction resulting from inclusion of the non-scored items. For example, the AUC-ROC values increased from 0.731 to 0.734 when the six non-scored items significantly associated with any recidivism were added to the logistic regression model. Because the confidence intervals associated with the AUC-ROC scores for both models overlapped, these differences were not statistically significant. Moreover, a sensitivity analysis showed no discernible improvement in the identification

[12] While there is extensive literature cautioning against the use of stepwise methods because of their reliance on computer algorithms over theory, we employed this approach because our models use variables that have been both theoretically and empirically shown to predict recidivism (Andrews & Bonta, 2010). Moreover, we attempted to minimize the problem of suppressor effects and type II errors associated with these approaches by using backward as opposed to forward stepwise regression methods (Field, 2013).

[13] We examined the variance inflation factors (VIFs) to check for the possibility of multicollinearity, as some of the non-scored PCRA items measured characteristics similar to the scored items. None of the variables—scored or unscored—in the model manifested VIFs in the range (3 or above) that would evidence serious problems with multicollinearity.

FIGURE 1.
Arrest distributions (any offense) for actual and predicted Post Conviction Risk Assessment (PCRA) risk scores



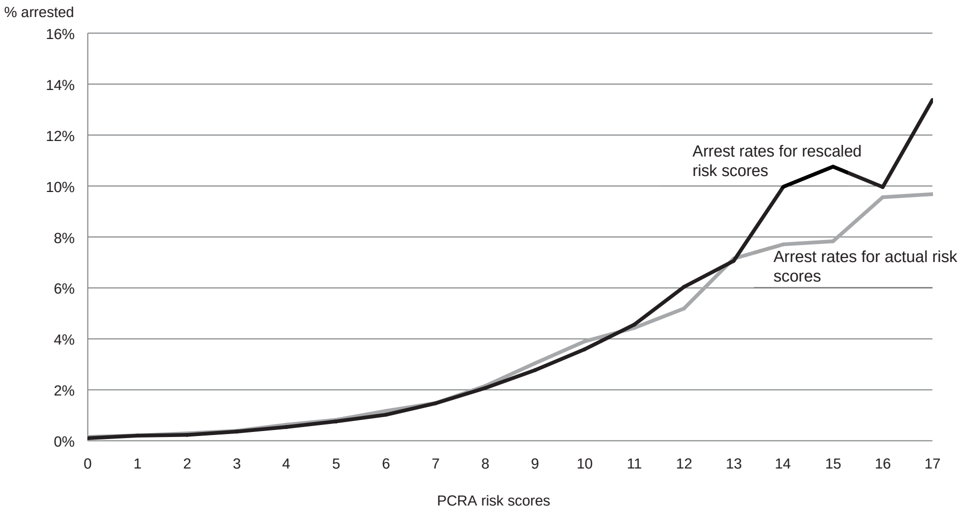
Note: Arrest rates shown include offenders followed for 12 month follow-up period.

TABLE 5.
Comparing offender recidivism rates (any offense) between actual and predicted Post Conviction Risk Assessment (PCRA) risk scores, by different follow-up periods

	12 month follow-up		24 month follow-up		36 month follow-up	
Raw PCRA scores	Actual PCRA scores	Rescaled PCRA scores	Actual PCRA scores	Rescaled PCRA scores	Actual PCRA scores	Rescaled PCRA scores
0	1.2%	1.0%	2.0%	1.7%	3.0%	2.6%
1	1.8%	1.5%	3.2%	2.7%	4.7%	4.1%
2	2.4%	2.4%	4.6%	4.3%	6.5%	6.2%
3	3.6%	3.1%	6.9%	6.0%	9.5%	8.5%
4	4.8%	4.2%	9.1%	8.2%	13.2%	12.0%
5	6.0%	5.4%	11.5%	10.2%	15.7%	14.4%
6	7.3%	7.0%	13.9%	14.2%	19.2%	19.1%
7	9.1%	9.0%	17.4%	17.1%	23.6%	23.7%
8	11.6%	11.0%	21.6%	20.7%	29.7%	28.5%
9	14.2%	14.1%	25.1%	25.2%	33.8%	34.3%
10	17.7%	17.5%	29.8%	30.1%	39.3%	40.3%
11	20.1%	20.6%	33.8%	36.0%	43.8%	46.1%
12	23.2%	25.6%	37.5%	41.1%	48.2%	53.3%
13	27.2%	28.7%	43.7%	45.1%	54.2%	55.8%
14	31.2%	33.9%	47.5%	50.2%	57.2%	60.2%
15	31.9%	36.9%	50.4%	55.5%	64.8%	66.2%
16	32.6%	41.3%	52.7%	58.5%	61.6%	71.9%
17	38.1%	39.6%	53.9%	60.5%	64.3%	69.2%
AUC-ROC	0.718 (0.714-0.721)	0.733 (0.729-0.736)	0.719 (0.715-0.722)	0.734 (0.730-0.737)	0.722 (0.718-0.725)	0.737 (0.733-0.740)
r	0.23	0.25	0.29	0.32	0.33	0.35
Number	188,542		150,405		110,240	

Note: The PCRA 18s have been recoded into 17s because relatively few offenders (N=10) obtained scores of 18. The percentage of offenders included in regression models by follow-up cohort ranges from 95%-96% of total sample. About 4%-5% of offenders omitted from analysis because they were missing values for either the scored or non-scored items.

FIGURE 2.
Arrest distributions (violent offenses) for actual and predicted Post Conviction Risk Assessment (PCRA) risk scores



Note: Arrest rates shown include offenders followed for 12 month follow-up period.

TABLE 6.
Comparing offender violent recidivism rates between actual and predicted Post Conviction Risk Assessment (PCRA) risk scores, by different follow-up periods

	12 month follow-up		24 month follow-up		36 month follow-up	
Raw PCRA scores	Actual PCRA scores	Rescaled PCRA scores	Actual PCRA scores	Rescaled PCRA scores	Actual PCRA scores	Rescaled PCRA scores
0	0.2%	0.1%	0.3%	0.2%	0.4%	0.4%
1	0.2%	0.2%	0.4%	0.3%	0.6%	0.5%
2	0.3%	0.2%	0.5%	0.5%	0.8%	0.8%
3	0.4%	0.4%	0.9%	0.8%	1.5%	1.1%
4	0.6%	0.5%	1.2%	1.0%	1.8%	1.6%
5	0.8%	0.8%	1.7%	1.7%	2.7%	2.4%
6	1.2%	1.0%	2.4%	2.2%	3.6%	3.4%
7	1.5%	1.5%	3.5%	3.2%	4.9%	4.5%
8	2.2%	2.1%	4.3%	4.0%	6.4%	6.0%
9	3.0%	2.8%	5.6%	5.3%	7.7%	7.9%
10	3.9%	3.6%	7.0%	6.9%	9.6%	9.5%
11	4.4%	4.6%	8.2%	8.6%	11.1%	12.4%
12	5.2%	6.0%	9.1%	10.9%	12.2%	14.0%
13	7.2%	7.1%	12.1%	12.6%	16.4%	16.7%
14	7.7%	10.0%	12.6%	14.7%	16.2%	18.3%
15	7.8%	10.8%	13.5%	15.2%	18.9%	22.1%
16	9.6%	10.0%	13.5%	17.1%	15.7%	22.0%
17	9.7%	13.4%	13.0%	17.5%	12.5%	18.0%
AUC-ROC	0.750 (0.743-0.757)	0.767 (0.760-0.774)	0.738 (0.732-0.744)	0.755 (0.749-0.761)	0.729 (0.723-0.735)	0.747 (0.741-0.753)
r	0.13	0.14	0.16	0.17	0.18	0.19
Number	188,542		150,405		110,240	

Note: The PCRA 18s have been recoded into 17s because relatively few offenders (N=10) obtained scores of 18. The percentage of offenders included in regression models by follow-up cohort ranges from 95%-96% of total sample. About 4%-5% of offenders omitted from analysis because they were missing values for either the scored or non-scored items.

TABLE 7.
Stepwise logistic regression analysis of non-scored factors on odds of any arrest within 12 months of Post Conviction Risk Assessment (PCRA)

	Model 1 - Scored items only			Model 2 - Scored & non-scored items		
		Confidence interval			Confidence interval	
PCRA factors	Odds Ratio	Lower	Upper	Odds Ratio	Lower	Upper
Scored items						
Number of prior arrests	1.47***	1.42	1.51	1.43***	1.38	1.47
Prior violent offense	1.18***	1.13	1.23	1.15***	1.11	1.20
Prior varied offending pattern	1.06*	1.00	1.13	1.05	0.99	1.11
Prior revocation/arrest while on supervision	1.32***	1.26	1.38	1.28***	1.22	1.34
Prior institutional adjustment	1.27***	1.21	1.33	1.22***	1.17	1.28
Age at intake to supervision	1.96***	1.89	2.03	1.88***	1.82	1.95
Less than high school or has only GED	1.18***	1.13	1.23	1.14***	1.09	1.19
Currently unemployed	1.23***	1.18	1.28	1.12***	1.07	1.17
Good work assessment over past 12 months	1.12***	1.08	1.16	1.05*	1.01	1.09
Current alcohol problem	1.09**	1.03	1.16	1.07*	1.00	1.13
Current drug problem	1.18***	1.12	1.24	1.11***	1.06	1.17
Single, divorced, separated	1.22***	1.16	1.29	1.09**	1.03	1.15
Unstable family situation	1.10***	1.05	1.15	1.05*	1.01	1.11
Lacks positive pro-social support	1.20***	1.14	1.26	1.11***	1.06	1.17
Attitude toward supervision and change	1.23***	1.17	1.30	1.19***	1.13	1.26
Non-scored items						
Juvenile arrest				1.17***	1.13	1.22
Employed less than 50% over past two years				1.10***	1.06	1.15
Drug use led to legal problems				1.09***	1.05	1.13
Lives with spouse and/or children				1.21***	1.16	1.27
Associates with negative peers or no friends				1.08***	1.06	1.11
Financial stressors present				1.16***	1.11	1.20
Constant	0.01	0.01	0.02	0.01	0.01	0.01
AUC-ROC	0.731	0.727	0.734	0.734	0.730	0.737
Sensitivity	69.9%			69.6%		
Log pseudolikelihood	-55774.5			-55598.8		
Number of offenders	188,542			188,542		

*Note: Backward stepwise logistic regression used to assess which non-scored risk items to include in second model. Only non-scored items associated with arrest outcomes at the .01 level were included in final model. Variable ordering coincides with that of appendix table 1. About 4% of offenders omitted from analysis because they were missing values for either the scored or non-scored items. *p < .05; **p < .01; ***p < .001*

TABLE 8.
Stepwise logistic regression analysis of non-scored factors on odds of violent arrest within 12 months of Post Conviction Risk Assessment (PCRA)

	Model 1 - Scored items only			Model 2 - Scored & non-scored items		
		Confidence interval			Confidence interval	
PCRA factors	**Odds Ratio**	**Lower**	**Upper**	**Odds Ratio**	**Lower**	**Upper**
Scored items						
Number of prior arrests	1.49***	1.40	1.59	1.44***	1.35	1.54
Prior violent offense	2.00***	1.84	2.18	1.94***	1.79	2.10
Prior varied offending pattern	1.06	0.94	1.20	1.04	0.93	1.17
Prior revocation/arrest while on supervision	1.29***	1.16	1.43	1.24***	1.12	1.38
Prior institutional adjustment	1.40***	1.29	1.52	1.34***	1.23	1.46
Age at intake to supervision	2.00***	1.87	2.14	1.89***	1.76	2.03
Less than high school or has only GED	1.26***	1.17	1.35	1.21***	1.12	1.30
Currently unemployed	1.17***	1.08	1.26	1.07	0.98	1.16
Good work assessment over past 12 months	1.14***	1.07	1.22	1.06	0.99	1.14
Current alcohol problem	1.26***	1.14	1.40	1.26***	1.13	1.39
Current drug problem	1.05	0.94	1.16	1.01	0.91	1.12
Single, divorced, separated	1.08	0.97	1.20	0.99	0.89	1.10
Unstable family situation	1.07	0.99	1.17	1.04	0.96	1.13
Lacks positive prosocial support	1.16**	1.06	1.28	1.09*	1.00	1.20
Attitude toward supervision and change	1.10	0.96	1.25	1.06	0.93	1.21
Non-scored items						
Juvenile arrest				1.27***	1.19	1.37
Employed less than 50% over past two years				1.14**	1.03	1.25
Lives with spouse and/or children				1.15**	1.04	1.26
Associates with negative peers or no friends				1.09**	1.03	1.15
Financial stressors present				1.14**	1.05	1.25
Constant	0.00	0.00	0.00	0.00	0.00	0.00
AUC-ROC	0.766	0.759	0.773	0.769	0.762	0.776
Sensitivity	73.9%			73.8%		
Log pseudolikelihood	-16669.5			-16626.5		
Number of offenders	188,542			188,542		

Note: Backward stepwise logistic regression used to assess which non-scored risk items to include in second model. Only non-scored items associated with violent arrest outcomes at the .01 level were included in final model. Variable ordering coincides with that of appendix table 1. About 4% of offenders omitted from analysis because they were missing values for either the scored or non-scored items.

of recidivists; both models correctly identified 70 percent of offender recidivists. In addition to these findings, the regression models examining arrests for violent offenses showed similar patterns of negligible differences in the predictive statistics between the models with the scored and non-scored PCRA risk items.

Discussion and Conclusion

Summary of Findings

In this study we sought to investigate whether incorporation of the 15 non-scored items currently rated by officers into the PCRA's risk algorithm could significantly enhance the instrument's predictive accuracy. In general, findings show that inclusion of the non-scored items results in relatively small improvements in the PCRA's capacity to predict recidivism. Specifically, the AUC-ROC values and correlations were somewhat higher for the rescaled rather than original risk scores, but the differences were not substantive enough that the AOUSC should definitely consider integrating the non-scored items into the risk prediction tool. Moreover, the actual and rescaled risk scores essentially manifested similar rearrest rates, with the exception that the rescaled scores at the upper end of the risk spectrum captured rearrest activity to a slightly greater extent than the original scores. Finally, a comparison of logistic regression models shows essentially no differences in the predictive indices (i.e., AUC-ROC, sensitivity scores) between the models using only the 15 scored PCRA items and the models using both the scored and the non-scored PCRA items. These findings provide further support that the non-scored items can be removed from the instrument's worksheet without compromising the tool's predictive effectiveness.

Implications for the Field

As a result of this research, the AOUSC decided to remove several of these non-scored items from the Officer Section of the PCRA. These include prior juvenile arrest history, number of employers in the last 12 months, offender employed less than 50 percent of the time during the previous two years, legal problems related to drug use during the past 12 months, lives with spouse and/or children, current lack of family support, antisocial attitudes, offender's residential stability, criminal risks at home, financial situation, and level of engagement in prosocial activities (AOUSC, 2016). Some of the non-scored items, however, will continue to be rated but not scored, as they could be very helpful for research purposes. These include several of the substance abuse items assessing disruption at work, school, or home resulting from substance abuse; drug use in physically hazardous conditions; and continued drug use despite social/interpersonal problems. Also, the negative companions item will remain because of its strong correlation with recidivism. While officers will continue to rate these non-scored items, they will not be incorporated into the PCRA risk algorithm. Although these items will not impact the overall score and risk level, they may inform case planning and elicit opportunities to teach the offender coping skills and problem-solving techniques.

Removal of the non-scored items has allowed the AOUSC to develop and implement a violence trailer (Serin et al., 2016). While the PCRA has been shown to be a strong predictor of general recidivism (Johnson et al., 2011; Lowenkamp et al., 2013; Lowenkamp et al., 2015), the instrument was not originally developed to predict violent recidivism. To address this, the AOUSC conducted additional research and found that there are 14 violence flags predictive of violent rearrest. These flags comprise scales from the Psychological Inventory of Criminal Thinking Styles (PICTS), which are measured in the Offender Section of the PCRA, and existing data related to violence. In addition to the PCRA score and the four PICTS scales (Power Orientation, Denial of Harm, Entitlement, Self Assertion), the violence flags include prior violent arrests, current violence offense, plans violence, age at first arrest, prior stalking, history of treatment noncompliance, gang membership, weapon use ever, prior or current domestic violence and stranger victimization. In terms of predicting violent and domestic violence rearrest, the construction sample (N=1,154) produced an AUC-ROC value of .79 when examining both the PCRA and violence flags, and the validation sample (N=1,154) had a slightly higher AUC-ROC value of .82 (Serin et al., 2016). This multilevel risk assessment process of conducting the PCRA 2.0, administering the violence trailer, and directing case management efforts and interventions to address the needs of probation clients will be the next stage of implementation and continuous improvement to the risk assessment process within the federal system.

Conclusion

This study sought to explore whether the non-scored items could be removed from the PCRA without hindering the instrument's predictive effectiveness and hence free up space for the incorporation of a trailer capable of assessing whether an offender will become involved in a catastrophically violent event. Through this research, we show that incorporating the 15 non-scored items into the PCRA's risk prediction algorithm resulted in negligible improvements in this tool's risk prediction capacities and that the AOUSC need not consider retaining these items while enhancing this tool through adoption of a violence trailer. As a result of adherence to a data-driven approach, the PCRA has witnessed two substantive improvements. First, the AOUSC has been able to field an updated risk assessment tool—PCRA 2.0—while retaining only a few select non-scored items for further research and case planning. These results ensure that officers are focusing on the strongest predictors of general and violent recidivism for their target population. Second, with the removal of these non-scored items and the integration of the violence flags, the risk assessment process within the federal supervision system will now have the capacity to alert officers about an offender's proclivity towards violence and allow officers to take actions to protect the community and safeguard the public.

References

Administrative Office of the U.S. Courts (AOUSC) (2011). *An overview of the federal Post Conviction Risk Assessment*. Washington, D.C.: Administrative Office of the U.S. Courts.

Administrative Office of the U.S. Courts (AOUSC) (2016). Federal Post-Conviction Risk Assessment 2.0 Training Manual. Washington, D.C.: Administrative Office of the U.S. Courts.

Ægisdóttir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., & Cook, R. S. (2006). The meta-analysis of clinical judgement project: Fifty six years of accumulated research on clinical versus statistical prediction. *The Counseling Psychologist, 34,* 341-382.

Allison, P. D. (2015). *Linear regression analysis*. Philadelphia, PA: Statistical Horizons.

Andrews, D., & Bonta, J. (1995). *The Level of Service Inventory-Revised*. Toronto: Multi-Health Systems.

Andrews, D., & Bonta, J. (1998). *The psychology of criminal conduct*. Cincinnati, OH: Anderson Press.

Andrews, D., & Bonta, J. (2010). *The psychology of criminal conduct (5th edition)*. Cincinnati, OH: Anderson Publishing.

Andrews, D. A., Bonta, J., & Wormith, S. J.

(2006). The recent past and near future of risk/need assessment. *Crime and Delinquency, 52,* 7–27.

Andrews, D., & Robinson, D. (1984). *The Level of Supervision Inventory: Second report*. A report to Research Services (Toronto) of the Ontario Ministry of Correctional Services.

Baber, L. (2010). Results-based Framework for Post-Conviction Supervision Recidivism Analysis. *Federal Probation, 74,* 5-10.

Bonta, J. (1996). Risk-needs assessment and treatment. In A. T. Harland (Ed.), *Choosing correctional options that work: Defining the demand and evaluating the supply* (pp. 18-32). Thousand Oaks, CA: Sage.

Bonta, J., & Andrews, D. (2007). *Risk-Need-Responsivity Model for offender assessment and rehabilitation.* (User Report 2007-06). Ottawa: Public Safety Canada.

Bonta, J., & Wormith, S. (2007). Risk and need assessment. In Gill McIvor & Peter Raynor (Eds.), *Developments in social work with offenders*. London, England: Jessica Kingsley Publishers.

Cohen, T., & VanBenschoten, S. (2014). Does the risk of recidivism for supervised offenders improve over time? Examining changes in the dynamic risk characteristics for offenders under federal supervision. *Federal Probation, 78*, 41-52.

Desmarais, S., & Singh, J. (2013). *Risk assessment instruments validated and implemented in correctional settings in the United States*. Lexington, KY: Council of State Governments.

Doyle, M., & Dolan, M. (2002). Violence risk assessment: Combining actuarial and clinical information to structure clinical judgments for the formulation and management of risk. *Journal of Psychiatric and Mental Health Nursing, 9(6),* 649-657.

Field, A. (2013). *Discovering statistics using IBM SPSS Statistics, 4th ed.,* University of Sussex: Sage.

Flores, A., Holsinger, A., Lowenkamp, C., & Cohen, T. (2016). Time-Free effects in predicting recidivism using both fixed and variable follow-up periods. Do different methods produce different results. *Criminal Justice and Behavior, 44,* 121-137.

Gottfredson, D., & Snyder, H. (2005). *The mathematics of risk classification: Changing data into valid instruments for the juvenile courts*. National Center for Juvenile Justice, Office of Juvenile Justice and Delinquency Prevention (NCJ 209158).

Grove, W., Zald, D., Lebow, B., Snitz, B., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis, *Psychological Assessment, 12,* 19-30.

Hilbe, J. (2009). *Logistic regression models.* Boca Raton, Florida: CRC Press.

Johnson, J., Lowenkamp, C., VanBenschoten, S., & Robinson, C. (2011, September). The construction and validation of the Post Conviction Risk Assessment (PCRA). *Federal Probation, 75,* 1-32.

Kane, M., Bechtel, K., Revicki, J., McLaughlin, E., & McCall, J. (2011). *Exploring the role of responsivity and assessment with Hispanic and American Indian offenders*. Crime and Justice Institute at Community Resources for Justice: Boston, MA.

Kim, K., & Duwe, G. (2017). Improving the performance of risk assessments: A case study on the prediction of sexual offending among juvenile offenders. In F. Taxman (Ed.), *Handbook on risk and need assessment: Theory and practice* (pp. 114-139). New York, New York: Routledge.

Latessa, E., & Lovins, B. (2010). The role of offender risk assessment: A policy maker guide. *Victims and Offenders, 5*, 203-219.

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

Lowenkamp, C., Johnson, J., Holsinger, A., VanBenschoten, S., & Robinson, C. (2013). The Federal Post Conviction Risk Assessment (PCRA): A construction and validation study. *Psychological Services, 10*, 1-14.

McEwan, T., Mullen, P., & Mackenzie, R. (2009). A study of the predictors of persistence in stalking situations. *Law & Human Behavior, 33*, 149-158.

Meehl, P. (1954). *Clinical vs. statistical prediction.* Minneapolis: University of Minnesota Press.

Monahan, J. (1981). *Predicting violent behavior: An assessment of clinical techniques.* Beverly Hills: Sage.

Rice, M., & Harris, G. (2005). Comparing effect sizes in follow-up studies: ROC area, Cohen's d, and r. *Law and Human Behavior, 29,* 615-620.

Serin, R., Lowenkamp, C., Johnson, J., & Trevino, P. (2016). Using a multi-level risk assessment to inform case planning and risk management: Implications for officers. *Federal Probation, 80,* 10-14.

Skeem, J., & Lowenkamp, C. T. (2016). Risk, race, and recidivism: Predictive bias and disparate impact. *Criminology, 54(4),* 680-712.

Skeem, J., Monahan, J., & Lowenkamp, C. T. (2016). Gender, risk assessment, and sanctioning: The cost of treating women like men. *Law and Human Behavior, 40(5),* 580-593.

VanBenschoten, S. (2008). Risk/Needs assessment: Is this the best we can do? *Federal Probation*, *72,* 38-42.

VanVoorhis, P., & Brown, K. (1996). *Risk classification in the 1990s*. U.S. Department of Justice, National Institute of Corrections.

Walters, G., & Cohen, T. (2016). Criminal thought process as a dynamic risk factor: Variable- and person-oriented approaches to recidivism prediction. *Law and Human Behavior, 40,* 411-419.

Walters, G., & Lowenkamp, C. (2016). Predicting recidivism with the Psychological Inventory of Criminal Thinking Styles (PICTS) in community-supervised male and female federal offenders. *Psychological Assessment, 28,* 652-659.

APPENDIX TABLE 1.
Descriptions of Items in the officer assessment of the Post Conviction Risk Assessment (PCRA)

PCRA items	Item Description	Answers	Scored
Criminal history			
1.1	Juvenile arrest	A = No; B = Yes	N
1.2	Number of prior arrests	0 = None; 1 = One or two; 2 = Three through seven; 3 = Eight or more	Y
1.3	Prior violent offense	0 = No; 1 = Yes	Y
1.4	Prior varied offending pattern	0 = 1 offense type; 1 = 2 or more	Y
1.5	Prior revocation/arrest while on supervision	0 = No; 1 = Yes	Y
1.6	Prior institutional adjustment	0 = No or NA; 1 = Yes	Y
1.7	Age at intake to supervision	0 = 41+; 1 = 26 to 40; 2 = 25 or less	Y
Education & employment			
2.1	Less than high school or has only GED	0 = High school or higher; 1 = Less than high school or GED only	Y
2.2	Currently unemployed	0 = Employed PT/FT, disabled and receiving benefits; 1 = Student, homemaker, unemployed, or retired but able to work	Y
2.3	Multiple jobs past year	A = 1; B = None or more than 1	N
2.4	Employed less than 50% over past two years	A = Employed 12 months or more; B = Employed less than 12 months	N
2.5	Good work assessment over past 12 months	0 = Yes; 1 = No	Y
Drugs & alcohol			
3.1	Drug use related to disruption at work, school, or home	A = No; B = Yes	N
3.2	Drug use in physically hazardous conditions	A = No; B = Yes	N
3.3	Drug use led to legal problems	A = No; B = Yes	N
3.4	Drug use continued despite social problems	A = No; B = Yes	N
3.5	Current alcohol problem	0 = No; 1 = Yes	Y
3.6	Current drug problem	0 = No; 1 = Yes	Y
Social networks			
4.1	Single, divorced, separated	0 = Married; 1 = Not Married	Y
4.2	Lives with spouse and/or children	A = No; B = Yes	N
4.3	Lacks family support	A = Support Present; B = No Support	N
4.4	Unstable family situation	0 = No; 1 = Yes	Y
4.5	Associates with negative peers or no friends	A = Good support; B = Occasional association with negative peers; C = More than occasional association with negative peers; D = No friends	N
4.6	Lacks positive prosocial support	0 = No; 1 = Yes	Y
Cognitions			
5.1	Harbors antisocial attitude/values	A = No; B = Yes	N
5.2	Attitude toward supervision and change	0 = Motivated; 1 = Not motivated	Y
Other factors			
6.1	Lacks permanent residence	A = 1 address in last 12 months; B = > 1 address last 12 months; no permanent address	N
6.2	Criminal risks present in home	A = No risks at home; B = Risks at home	N
6.3	Financial stressors present	A = Adequate income to manage debts; concrete financial plans; B = No plan in place; expenses exceed income	N
6.4	Does not engage in prosocial activities	A = Engages in prosocial activities; B = Has no interests; does not; or recreation presents criminal risk	N