Less Is More: Using Static-2002R Subscales to Predict Violent and General Recidivism Among Sexual Offenders

Kelly M. Babchishin[1,2], R. Karl Hanson[1], and Julie Blais[1,3]

Abstract

Given that sexual offenders are more likely to reoffend with a nonsexual offense than a sexual offense, it is useful to have risk scales that predict general recidivism among sexual offenders. In the current study, we examined the extent to which two commonly used risk scales for sexual offenders (Static-99R and Static-2002R) predict violent and general recidivism, and whether it would be possible to improve predictive accuracy for these outcomes by revising their items. Based on an aggregated sample of 3,536 adult male sex offenders from Canada, the United States, and Europe (average age of 39 years), we found that a scale created from the Age at Release item and the General Criminality subscale of Static-2002R predicted nonsexual violent, any violent, and general recidivism significantly better than Static-99R or Static-2002R total scores. The convergent validity of this new scale (Brief Assessment of Recidivism Risk–2002R [BARR-2002R]) was examined in a new, independent data set of Canadian high-risk adult male sex offenders (*N* = 360) where it was found to be highly correlated with other risk assessment tools for general recidivism and the Psychopathy Checklist–Revised (PCL-R), as well as demonstrated similar discrimination and calibration as in the development sample. Instead of using total scores from the Static-99R or Static-2002R, we recommend that evaluators use the BARR-2002R for predicting violent and general recidivism among sex offenders, and for screening for the psychological dimension of antisocial orientation.

[1]Public Safety Canada, Ottawa, Ontario, Canada
[2]Royal Ottawa Hospital, Ontario, Canada
[3]Carleton University, Ottawa, Ontario, Canada

Corresponding Author:
Kelly M. Babchishin, Royal Ottawa Hospital, 1175 Carling Avenue, Ottawa, Ontario, Canada K1Z 7K4.
Email: kelly.babchishin@theroyal.ca

Risk assessment informs nearly every decision made about offenders, including security classification, conditional release, and treatment and supervision intensity. Risk assessment is especially important for sex offenders, given the potential of serious harm to victims (Beitchman et al., 1992; R. Roberts, O'Connor, Dunn, & Golding, 2004). The primary concern for sex offenders has been sexual recidivism and a number of risk tools provide moderate predictive accuracy for this outcome (for review, see Hanson & Morton-Bourgon, 2009). Sex offenders, however, are more likely to reoffend with a nonsexual offense than with a sexual offense (Hanson & Bussière, 1998; Prentky, Lee, Knight, & Cerce, 1997). Consequently, there is increasing recognition of the importance of assessing the risk of both general criminality and sex crime specific criminality among sex offenders (National Policing Improvement Agency, 2010; Social Work Inspection Agency, 2009).

Specialized risk tools designed to predict sexual recidivism also predict nonsexual recidivism. However, they have lower accuracy than risk tools specifically created for predicting general recidivism (Hanson & Morton-Bourgon, 2009). Indeed, there are a multitude of risk tools designed to predict general and violent recidivism among non-sex offenders (e.g., Level of Service/Case Management Inventory [LS/CMI], Andrews, Bonta, & Wormith, 2004; Statistical Information on Recidivism–Revised 1 [SIR-R1], Nafekh & Motiuk, 2002; Violence Risk Appraisal Guide–Revised [VRAG-R], G. T. Harris, Rice, & Quinsey, 1993; Rice, Harris, & Lang, 2013), and these risk tools have also been found to predict general and violent recidivism among sex offenders (Hanson & Morton-Bourgon, 2009; Wormith, Hogg, & Guzzo, 2012). Many of these risk tools, however, require substantial resources (e.g., training, time) to score. If resources are limited, it would be efficient to use items already coded in an existing sexual offender risk tool to assess general and violent recidivism. Such a tool could be used as a screening measure to identify sex offenders who would benefit from more detailed time-consuming measures such as the LS/CMI.

Although violent recidivism rates tables are available for Static-99 (A. Harris, Phenix, Hanson, & Thornton, 2003) and Static-99R (www.static99.org), these tables may not be optimal. Recent studies on the construct validity of sexual offender risk scales suggest that items of the STATIC sexual offender risk tools could be reorganized to improve the prediction of general and violent recidivism (Babchishin, Hanson, & Helmus, 2012b; Lehmann et al., 2013). There would be considerable efficiency in using items already coded for predicting sexual recidivism to predict general recidivism among individuals with a history of convictions for sexual offenses.

Understanding the Source of Sex Offender Risk

Risk scales predict recidivism because the items are markers for risk-relevant propensities (Mann, Hanson, & Thornton, 2010). Many risk tools, however, have

been created using atheoretical empirical correlates and, consequently, we know remarkably little about the constructs they assess. Yet there are significant advantages to understanding the constructs of empirically derived risk tools. Different risk tools frequently provide different results (Barbaree, Langton, & Peacock, 2006a), even highly correlated measures can add incrementally to the prediction of risk (e.g., Babchishin et al., 2012b; Lehmann et al., 2013), and none of the existing risk tools can claim to measure all of the relevant risk factors. However, it is difficult to resolve disagreements between scales, or to judge whether an additional, external risk factor should be considered in the overall evaluation without knowing the constructs already assessed by the different tools.

Research on the widely used Static-99R and Static-2002R risk assessment scales (Hanson & Thornton, 2000, 2003; Helmus, Thornton, Hanson, & Babchishin, 2012) has found that sexual recidivism is predicted by the three broad constructs of age, general criminality, and sex crime specific criminality (e.g., Babchishin et al., 2012b; Lehmann et al., 2013). For the prediction of nonsexual and general recidivism, however, sexual criminality has little relevance. Consequently, it should be possible to simplify the prediction of (largely) nonsexual outcomes by only considering items related to age and general criminality.

General Criminality

Evidence is required to justify a collection of items as a measure of general criminality. General criminality includes a global propensity for rule violation, meanness, and impulsivity, and overlaps with the constructs of antisocial personality disorder (American Psychiatric Association, 2013), psychopathy (Patrick, 2005), and antisocial personality pattern (Andrews & Bonta, 2010). There is now substantial evidence that these characteristics are distributed dimensionally (Guay, Ruscio, Knight, & Hare, 2007; Patrick, Fowles, & Krueger, 2009), meaning that all individuals, including sex offenders, can be described as more or less antisocial.

Factor analyses are an important source of evidence for construct validity. Recent factor analyses of the Static-99/R and Static-2002/R have found at least two factors, one related to sexual criminality and another related to general criminality (e.g., Barbaree, Langton, & Peacock, 2006b; Brouillette-Alarie, Babchishin, Hanson, & Helmus, in press; Brouillette-Alarie, Proulx, & Benbouriche, 2013; Janka, Gallasch-Nemitz, & Dahle, 2011; C. F. Roberts, Doren, & Thornton, 2002). The factor analytic results have been particularly clear for the Static-2002R General Criminality subscale, which contains five items related to criminal history (any prior arrests, number of prior offenses, violation of conditional release, any prior violent offenses, and time-free prior to index offense).

Regardless of the other items included in the analyses, the Static-2002R general criminality items have consistently loaded onto a common dimension (Boughner, 2010; Brouillette-Alarie, 2013; Brouillette-Alarie et al., in press; Ennis, Choy, Jung, Corabian, & Hook, 2011; Langton, Barbaree, Hansen, Harkins, & Peacock, 2007). Furthermore, the Static-2002R general criminality factor added incrementally to age

for the prediction of violent and general recidivism among sexual offenders (Babchishin et al., 2012b). Building on these findings, the purpose of the current study was to develop and validate a risk scale for violent and general recidivism for sex offenders based on Static-2002R items.

Overview of Current Article

This article describes two studies in which we develop and validate the Brief Assessment of Recidivism Risk–2002R (BARR-2002R), a risk scale for predicting violent and general recidivism among sexual offenders. The first study presents the development of the BARR-2002R and the second study examines the construct validity and generalizability of the scale in an independent data set. Predictive accuracy was indexed by both discrimination (the extent to which recidivists were different from nonrecidivists) and calibration (correspondence between expected recidivism rates presented in Study 1 and observed recidivism rates from Study 2).

Study 1

Overview

The purpose of Study 1 was to describe how the age and the general criminality items of the Static-2002R can be used as a stand-alone scale to predict general and violent recidivism among sexual offenders. The psychometric properties of the BARR-2002R are presented.

Method

Samples. The development analyses of the BARR-2002R were conducted on a large sample of sex offenders drawn from the STATIC renorming project (Helmus et al., 2012). To be included, samples needed to have Static-2002 items and at least one of the recidivism types included in the current study. Tables 1 and 2 present the characteristics of the samples ($k = 9$, $N = 3,536$). Most samples were drawn from Canada ($k = 4$) or the United States ($k = 2$), followed by single samples from Denmark, Germany, and the United Kingdom.

Four types of recidivism were examined: sexual recidivism, any violent (including sexual) recidivism, nonsexual violent recidivism, and general (any new offense) recidivism. One sample reported only sexual recidivism and all other samples ($k = 8$) reported on all four recidivism outcomes. Recidivism was measured using official criminal records, with five samples using charges as the recidivism criteria and four samples using convictions. Follow-up time ranged from 0.1 to 26.5 years ($M = 9.2$, $SD = 4.0$).

Measures

Static-99R. Static-99R (Hanson & Thornton, 2000; Helmus et al., 2012) is a 10-item actuarial measure that assesses recidivism risk of adult male sexual offenders. The 10

Table 1. Description of Samples Included in Study 1.

Study	N	Age M (SD)	Recidivism criteria	Type of sample	Country	Release period	Mdn year release
Bengtson (2008)	308	32.5 (10.4)	Charges	Forensic psychiatric evaluations	Denmark	1978-1995	1986
Bigras (2007)	452	42.8 (12.0)	Charges	Routine CSC	Canada	1995-2004	1999
Boer (2003)	296	41.2 (12.5)	Convictions	Routine CSC	Canada	1976-1994	1990
Haag (2005)	190	36.7 (9.7)	Convictions	Detained until end of sentence	Canada	1995	1995
Hanson, Harris, Scott, and Helmus (2007)	710	41.6 (13.3)	Charges	Routine community supervision	Canada	2001-2005	2002
Harkins and Beech (2007)	182	43.5 (12.7)	Convictions	Prison and community treatment	The United Kingdom	1994-1998	1995
Knight and Thornton (2007a)	251	34.4 (10.1)	Charges	Referred for commitment evaluation	The United States	1957-1986	1967
Knight and Thornton (2007b)	215	38.0 (12.5)	Charges	Civilly committed	The United States	1959-1983	1972
Lehmann et al. (2013)	932	38.2 (11.6)	Convictions	All sentenced offenders	Germany	1994-2009	2000
Total	3,536	39.1 (12.3)	—	—	—	1957-2009	1998

Note. CSC = Correctional Service Canada.

Table 2. Risk Scores and Recidivism Rates of the Samples.

| Study | N | Static-99R | Static-2002R | BARR-2002R | Years follow-up | Recidivism rates (%) | | | |
		M (SD)				Sexual	Any violent	Nonsexual violent	Any
Bengtson (2008)	308	3.8 (2.4)	4.6 (2.4)	2.8 (2.0)	16.2 (4.2)	34.1	52.3	23.4	64.6
Bigras (2007)	452	2.1 (2.4)	3.5 (2.5)	3.0 (2.5)	4.6 (1.9)	5.8	14.8	10.0	23.5
Boer (2003)	296	2.8 (2.8)	3.9 (2.7)	3.1 (2.7)	13.3 (2.1)	8.8	23.3	17.6	48.3
Haag (2005)	190	4.1 (2.2)	5.7 (2.3)	4.8 (2.0)	7.0 (0.0)	24.7	—	—	—
Hanson, Harris, Scott, and Helmus (2007)	710	2.4 (2.4)	3.5 (2.5)	2.6 (2.4)	7.5 (1.9)	11.5	22.5	14.8	33.1
Harkins and Beech (2007)	182	2.1 (2.6)	3.6 (2.8)	2.3 (2.2)	10.4 (1.1)	13.7	20.9	10.4	35.2
Knight and Thornton (2007a)	251	4.0 (2.4)	5.6 (2.6)	3.8 (2.0)	8.7 (2.5)	20.7	34.7	23.1	51.4
Knight and Thornton (2007b)	215	5.3 (2.2)	6.8 (2.3)	4.2 (1.9)	8.4 (2.7)	36.7	44.2	27.0	59.1
Lehmann et al. (2012)	932	3.4 (2.2)	4.6 (2.4)	2.8 (2.0)	9.7 (3.2)	15.9	33.9	21.7	61.5
Total	3,536	3.2 (2.5)	4.3 (2.5)	3.0 (2.2)	9.2 (4.0)	16.7	29.7	18.3	47.1

Note. Static-99R scores range from −3 to 12, Static-2002R from −2 to 14, and BARR-2002R from −2 to 8. BARR-2002R = Brief Assessment of Recidivism Risk–2002R.

items cover demographics (age at release, relationship history), sexual criminal history (prior sexual offenses, any male victims, any unrelated victims, any stranger victims, any noncontact sexual offenses), and general criminal history (prior sentencing dates, index nonsexual violence, prior nonsexual violence). Total scores range between −3 and 12.

Static-2002R. Static-2002R (Hanson & Thornton, 2003; Helmus et al., 2012) is a 14-item actuarial measure that assesses sexual recidivism risk of adult male sex offenders based on commonly available demographic and criminal history information. The 14 items are grouped into five main subscales: Age at Release, Persistence of Sex Offending, Sexual Deviance, Relationship to Victims, and General Criminality. The total score (ranging from −2 to 14) can be used to place offenders in one of five risk categories: low (−2 to 2), low–moderate (3, 4), moderate (5, 6), moderate–high (7, 8), and high (9+). The Static-2002R items are identical to those in Static-2002 with the exception of updated age weights (Helmus et al., 2012). Static-99R and Static-2002R have similar levels of predictive accuracy (Area Under the Curve [AUC] values around .70; Babchishin et al., 2012b).

General criminality factor. The general criminality factor is comprised of the sum of the five items in the Static-2002R Criminality subscale. Previous research on a related data set found moderate internal consistency (α = .78, n = 2,569; Brouillette-Alarie, 2013).

BARR-2002R. The BARR-2002R was constructed and validated in the present study. The BARR-2002R is comprised of the General Criminality subscale (raw subscale score, see above) and the age item of the Static-2002R (using the weights from Helmus et al., 2012). The items are presented in Appendix A.

Plan of Analyses

Incremental validity. Incremental validity was examined using Cox regression (Allison, 1984) and conducted in SPSS Version 20. Each sample was used as a stratum to allow separate baseline hazard functions (i.e., recidivism rates), effectively removing potential effects of base rate variability across samples (Allison, 1984). The analyses provide the Wald statistic that, if significant, indicates that the scale added incremental validity to the other scale included in the model. The analyses also provide a hazard ratio (e^B), indicating the scale's relationship with recidivism. For example, a hazard ratio of 1.40 indicates that each 1-point increase on the scale increases the hazard rate of recidivism by a factor of 1.40, or 40%.

Age weights. The applicability of the existing Static-2002R age (at release) weights for the prediction of violent recidivism were examined using Cox regression analyses (with samples as strata and age at release coded as a continuous variable). The principles for selection of age weights were as follows: (a) The units should be integers (whole numbers) and (b) each unit increase in the age item should approximate the

relative risk metric for the other Static-2002R items (i.e., hazard ratio = 1.4; Babch-ishin, Hanson, & Helmus, 2012a; Hanson & Thornton, 2003).

Discrimination. Discrimination describes the extent to which recidivists score higher (or lower) than nonrecidivists. The Area Under the Receiver Operating Characteristic Curve (AUC ROC) is the most common method of assessing discrimination (Pintea & Moldovan, 2009; Rice & Harris, 1995; Swets, Dawes, & Monahan, 2000). The first set of analyses used fixed-effect and random-effects meta-analyses to compute the weighted AUCs and their 95% confidence intervals (CIs) for the two risk scales (i.e., BARR-2002R and Static-2002R). When the analysis includes a small number of studies ($k < 30$), greater weight should be given to interpreting the fixed-effect rather than random-effects analyses because the estimate of the between-study variability parameter (tau) required for random-effects analyses is imprecise (Schulze, 2007). To test the variability of findings across studies, we used Cochran's Q statistic and the I^2 statistic (Borenstein, Hedges, Higgins, & Rothstein, 2009). The Q statistic provides a significance test for variability, whereas the I^2 is a measure of the magnitude of the effect and can, therefore, be compared across analyses. As a rough heuristic, I^2 values of 25%, 50%, and 75% can be considered as low, moderate, and high variability, respectively (Higgins, Thompson, Deeks, & Altman, 2003). Note that I^2 values have little precision when the number of studies is small (Schulze, 2007).

To test the extent to which the BARR-2002R, Static-2002R, and Static-99R differed in their level of discrimination, differences between AUCs were meta-analyzed using the Delong method for computing the standard error of the difference (Delong, Delong, & Clarke-Pearson, 1988) using the pROC procedure in R (Robin et al., 2011). If the 95% confidence interval of the difference between measures includes zero, the difference in discrimination between the two scales is not statistically significant.

Results

Incremental validity. Table 3 presents the incremental validity analyses for age at release (as a continuous variable), general criminality, and sexual criminality. All three factors (age, general criminality, and sexual criminality) added incrementally to the prediction of sexual recidivism. Specifically, higher scores on sexual criminality and general criminality factors as well as younger age were associated with greater risk of sexual recidivism. For general recidivism, sexual criminality did not add incrementally after controlling for age and general criminality, Wald = 0.12, $p = .732$. For nonsexual violent recidivism, the incremental effect of sexual criminality was negative, meaning that *lower* scores on sexual criminality was associated with *higher* rates of nonsexual violent recidivism (after controlling for age and general criminality). Consequently, we proceeded to create a risk scale for violent and general recidivism based only on the age and general criminality items from the Static-2002R.

Age weights. The test of nonlinear effects (continuous age squared) did not significantly improve the model for the prediction of nonsexual violent recidivism, $\chi^2 = 0.23$,

Table 3. Incremental Validity of Age, the Static-2002R General Criminality Subscale, and the Sexual Criminality Items From Static-2002R.

	Exp(B)	95% CI	Wald
Sexual recidivism (*n* = 3,533)			
Age	0.980	[0.973, 0.987]	27.9***
General criminality	1.214	[1.152, 1.278]	54.0***
Sexual criminality	1.253	[1.95, 1.314]	87.1***
General recidivism (*n* = 3,345)			
Age	0.962	[0.957, 0.967]	238.9***
General criminality	1.438	[1.393, 1.484]	504.3***
Sexual criminality	1.006	[0.974, 1.039]	0.12
Any violent recidivism (*n* = 3,344)			
Age	0.961	[0.955, 0.967]	155.1***
General criminality	1.377	[1.324, 1.433]	252.4***
Sexual criminality	1.069	[1.028, 1.112]	11.0**
Nonsexual violent recidivism (*n* = 3,163)			
Age	0.943	[0.935, 0.952]	164.8***
General criminality	1.480	[1.405, 1.559]	217.4***
Sexual criminality	0.908	[0.859, 0.960]	11.6**

Note. Analyses conducted separately for each comparison, using Cox regression with each sample entered as strata and the three variables entered simultaneously in the model. Continuous age at release entered as a predictor. General criminality was the sum of the five items of the Static-2002R General Criminality subscale. Sexual criminality included the Static-2002R items comprised in the Persistence of Sexual Offending, Deviant Sexual Interests, and Relationship to Victim subscales (nine items). Sample sizes fluctuate because of missing data and cases censored before earliest event. CI = confidence interval. **p < .01. ***p < .001.

df = 1, p = .630; any violent (including sexual) recidivism, χ^2 = 1.13, df = 1, p = .287; or general (any) recidivism, χ^2 = 3.63, df = 1, p = .054. In the prediction of general recidivism, the units for the Static-2002R age at release item, hazard ratio = 1.44, approximated the general criminality units, hazard ratio = 1.42, and, hence, the units of the Static-2002R age at release item were used. Consequently, the BARR-2002R was created by summing the general criminality items and the age item from Static-2002R in their existing units (total scores ranged of −2 to 8). The lowest score would be a first-time offender released after the age of 60; the highest score would be an offender under the age of 35 with an extensive criminal history (>14 prior sentencing occasions, including convictions for violence and breaches of community supervision).

Table 3 presents the average scores on the Static-99R, Static-2002R, and BARR-2002R in the current samples. In the aggregated data set, the Static-2002R and BARR-2002R scales were correlated, r = .76, sharing 58% of their variance. The correlation between age and the BARR-2002R was −.55, and −.45 with Static-2002R.

Discrimination. Table 4 presents the meta-analysis of the predictive accuracy (AUCs) of the BARR-2002R, Static-2002R, and Static-99R for the four recidivism outcomes.

Table 4. Differences in Discrimination Between BARR-2002R and STATIC Scales.

Recidivism Outcome and Scale	Fixed effect		Random effects		Q	I^2	N (k)
	AUC	95% CI	AUC	95% CI			
General recidivism							
BARR-2002R	.77	[.75, .79]	.76	[.71, .81]	52.90***	86.8%	3,346 (8)
Static-2002R	.72	[.71, .74]	.71	[.67, .75]	33.44***	79.1%	3,346 (8)
Static-99R	.70	[.68, .72]	.69	[.65, .74]	37.32***	81.2%	3,346 (8)
Any violent recidivism							
BARR-2002R	.72	[.71, .74]	.73	[.69, .76]	24.44***	71.4%	3,346 (8)
Static-2002R	.69	[.67, .71]	.69	[.66, .72]	17.16*	59.2%	3,346 (8)
Static-99R	.69	[.67, .71]	.69	[.65, .72]	18.09*	61.3%	3,346 (8)
Nonsexual violent recidivism							
BARR-2002R	.74	[.72, .76]	.74	[.71, .78]	19.30**	63.7%	3,346 (8)
Static-2002R	.66	[.64, .68]	.66	[.62, .71]	28.59***	75.5%	3,346 (8)
Static-99R	.67	[.64, .69]	.67	[.64, .70]	15.22*	54.0%	3,346 (8)
Sexual recidivism							
BARR-2002R	.65	[.62, .67]	.65	[.62, .69]	18.78*	57.4%	3,536 (9)
Static-2002R	.69	[.66, .71]	.68	[.65, .72]	17.96*	55.4%	3,536 (9)
Static-99R	.68	[.66, .70]	.68	[.63, .73]	26.28***	69.6%	3,536 (9)

(continued)

196

Table 4. (continued)

Difference in discrimination	Fixed effect		Random effects		Q	I²	N (k)
	Difference	95% CI	Difference	95% CI			
BARR-2002R vs. Static-2002R							
General recidivism	.051	[.038, .064]	.051	[.038, .064]	3.08	0.0%	3,346 (8)
Any violent recidivism	.031	[.017, .046]	.031	[.017, .046]	3.85	0.0%	3,346 (8)
Nonsexual violent recidivism	.079	[.063, .094]	.079	[.063, .094]	6.02	0.0%	3,346 (8)
Sexual recidivism	−.038	[−.056, −.019]	−.035	[−.057, −.014]	10.11	20.9%	3,536 (9)
BARR-2002R vs. Static-99R							
General recidivism	.063	[.049, .078]	.063	[.049, .078]	5.35	0.0%	3,346 (8)
Any violent recidivism	.032	[.017, .048]	.032	[.017, .048]	4.85	0.0%	3,346 (8)
Nonsexual violent recidivism	.070	[.053, .088]	.070	[.053, .088]	6.39	0.0%	3,346 (8)
Sexual recidivism	−.031	[−.051, −.012]	−.028	[−.053, −.003]	11.99	33.3%	3,536 (9)

Note. A positive difference score indicates that the BARR-2002R had better discrimination than the Static-2002R or Static-99R. BARR-2002R = Brief Assessment of Recidivism Risk–2002R; AUC = Area Under the Curve; CI = confidence interval.
*p < .05. **p < .01. ***p < .001.

197

Table 5. Meta-Analysis of the Logistic Regression Coefficients for Estimating the 5-Year Recidivism Rates of the BARR-2002R.

	Fixed effect					
	M	95% CI	SE	Q	I^2	N (k)
General recidivism						
B_1 (OR)	1.75	[1.62, 1.88]	0.038	4.12	51.4	1,112 (3)
B_0 (%)	18.2%	[15.4, 21.4]	0.101	6.47*	69.1	1,112 (3)
Any violent recidivism						
B_1 (OR)	1.53	[1.42, 1.65]	0.039	0.50	0.0	1,114 (3)
B_0 (%)	12.3%	[10.1, 15.0]	0.116	9.91**	79.8	1,114 (3)

Note. In logistic regression, B_1 (the slope) is an estimate of relative predictive accuracy, or the average change in recidivism rates for each one-unit increase in risk scores, expressed as a log odds ratio. The B_0 is an estimate of the recidivism base rate for offenders. The BARR-2002R was centered on a score of 2 and, as such, the B_0 was an estimate of the recidivism base rate for offenders with the median score of 2 (sex offenders in the middle of the risk distribution). For ease of interpretation, the B_0 was transformed into probabilities (p), where $p = e^{LOGIT}/(1 + e^{LOGIT})$. The *SEs*, however, are in the original metric (logits). BARR-2002R = Brief Assessment of Recidivism Risk–2002R; CI = confidence interval; OR = odds ratio.
*$p < .05$. **$p < .01$.

Compared with the Static-2002R, the BARR-2002R was a significantly better predictor of general recidivism (AUC of .77 vs. .72), any violent (including sexual) recidivism (.72 vs. .69), and nonsexual violent recidivism (.74 vs. .66). Compared with the Static-99R, the BARR-2002R was a significantly better predictor of general recidivism (AUC of .77 vs. .70), any violent (including sexual) recidivism (.72 vs. .69), and nonsexual violent recidivism (.74 vs. .67). Both Static-2002R (AUC = .69) and Static-99R (AUC = .68) were better predictors of sexual recidivism than the BARR-2002R (AUC = .65). All differences were statistically significant based on a series of Delong tests (see Table 4). There was moderate to large variability across studies in the predictive accuracy for all three scales (I^2 ranged from 54.0% to 86.8%); however, the differences in discrimination between the scales were consistent across studies (I^2 of 0.0% to 33.3%).

Five-year recidivism rate estimates for BARR-2002R scores were derived from the three Canadian samples ($N = 1,112$ for general recidivism and $N = 1,114$ for violent recidivism; see Appendix B). The German sample was excluded because of meaningful differences in the distribution of scores compared with the Canadian samples. As with Static-2002 (Hanson, Thornton, & Helmus, 2010), the estimates were calculated using averaged logistic regression parameters (fixed-effect meta-analysis; see Table 5). The relative risk parameter (B_1) was consistent across the studies (nonsignificant Q values), whereas the base rate parameter varied across samples (significant Q values; I^2 of 69% for general recidivism and 80% for violent recidivism). Overall, the expected recidivism rates varied between 2% and 55% for violent recidivism, and between 2% and 78% for general recidivism.

Study 2

Overview

Using a new, independent sample of sexual offenders, we examined the stability of the findings of Study 1, as well as the construct validity of the Static-2002R General Criminality subscale and the BARR-2002R. If the BARR-2002R is a measure of general criminality, then it should show moderate to high correlations with other measures of antisocial orientation, such as the LS/CMI, SIR-R1, and Psychopathy Checklist–Revised (PCL-R). The risk factors for general criminality are well established and surprisingly stable across samples (Andrews & Bonta, 2010; Bonta, Blais, & Wilson, 2014; Hanson, 2009). These factors include a history of rule violation, antisocial associates, procriminal attitudes, substance abuse, and poor use of leisure time. These risk factors are well represented, for example, by the subscales of the LS/CMI (Andrews et al., 2004). Although we expected moderate correlations with individual risk factors, we expected stronger correlations with total scores because, following standard psychometric theory, total scores should be more reliable than individual items. We also expected relatively strong correlations between the BARR-2002R and risk scales derived primarily from criminal history records, such as the SIR-R1 and the Criminal History subscale of the LS/CMI. Not only do these measures address the same latent construct, they would also be expected to share method variance because they are based on the same type of information.

In contrast, we did not expect the BARR-2002R to be related to offenders' life problems that are unrelated to criminal recidivism risk (noncriminogenic needs). In particular, factors related to internalizing psychological problems (e.g., suicide attempts) and major mental illness represent a distinct set of life problems expected to have little relationship to general criminality (Bonta et al., 2014). We also did not expect the BARR-2002R to correlate with sexual deviancy. Even though sexual deviancy is a risk-relevant propensity for sexual offenders, previous research (Babchishin et al., 2012b; Hanson & Morton-Bourgon, 2005) and the findings of Study 1 suggest that it is a different construct from general criminality.

Method

Sample. The sample consisted of 360 male sex offenders under provincial or federal jurisdiction in Canada, with most offenders sampled from Ontario, British Columbia, and Saskatchewan. Three groups were included: flagged offenders ($n = 233$), long-term offenders (LTOs; $n = 87$), and dangerous offenders (DOs; $n = 40$). Flagged offenders represented offenders identified as high risk by Canada's National Flagging System (NFS) between 2004 and 2008 (Public Safety Canada, 2011; see Blais & Bonta, 2014, for more detailed sample information). LTOs and DOs were offenders designated under Part XXIV of the Criminal Code of Canada, the majority of which were designated between 2006 and 2008. DOs (i.e., those committed to indefinite sentences) were excluded from the recidivism analyses (none were released into the

community), but included in the construct validity analyses. This sample has been utilized as part of a study that examined predictors of detention (Blais & Bonta, 2014) and provided a useful opportunity to validate findings from Study 1, which were primarily drawn from routine samples, on a higher risk sample.

As expected by the sample selection process, the sample was higher than average risk, with an average Static-2002R score of 6.3 ($SD = 2.5$), BARR-2002R score of 4.7 ($SD = 2.2$), and general criminality factor score of 4.0 ($SD = 1.7$). This sample of sex offenders scored on the 88th percentile in terms of risk of sexual recidivism as assessed by the Static-2002R (Hanson, Lloyd, Helmus, & Thornton, 2012) and on the 82nd percentile in terms of risk of general recidivism as assessed by the BARR-2002R (Babchishin, Hanson, & Blais, 2013). Of sex offenders who were released from prison ($N = 310$), offenders had an average age at release of 40.8 ($SD = 11.5$).

Measures

SIR-R1. The SIR-R1 (Nafekh & Motiuk, 2002) is a 15-item evaluation tool created to predict general reoffending among male, non-Aboriginal offenders within 3 years after release. The scale is a slightly modified version of the General Statistical Information on Recidivism (GSIR; Nuffield, 1982). The scale combines demographic and criminal history characteristics to produce a total score ranging from −30 to 27 (higher scores are indicative of lower risk), which can then be classified into one of five risk categories (very low to very high). Each risk category is associated with a probability rating for recidivism. Meta-analytic reviews have demonstrated that the earlier version of the SIR (Nuffield, 1982) is moderately predictive of violent (e.g., $d = 0.81$; Campbell, French, & Gendreau, 2009; Yang, Wong, & Coid, 2010) and sexual recidivism (e.g., $d = 0.64$; Hanson & Morton-Bourgon, 2009).

PCL-R. The PCL-R (Hare, 2003) was designed to assess the traits and behaviors associated with psychopathy. Psychopathy refers to a personality disorder with interpersonal, affective, and behavioral components. The scale consists of 20 items scored on a 3-point scale (0, 1, 2) resulting in a total score ranging from 0 to 40 (standard error of measurement [SEM] = 2.90 for single ratings). Higher scores are indicative of more psychopathic features. The PCL-R is usually scored using a semistructured interview and file information; however, given enough file information it can also be scored without the interview (Hare, 2003). Research into the factor structure of the PCL-R has consistently found two overarching factors: Factor 1 (Interpersonal/Affective) and Factor 2 (Lifestyle/Antisocial); however, only total scores were available in the current study. The PCL-R was not originally created as a risk assessment tool; nevertheless, it is commonly used in applied risk assessments and has shown moderate relationships with violent recidivism (e.g., $d = 0.63$) and sexual recidivism (e.g., $d = 0.40$; Campbell et al., 2009; Hawes, Boccaccini, & Murrie, 2013; Yang et al., 2010).

LS/CMI. The LS/CMI (Andrews et al., 2004) is a risk assessment tool that emphasizes factors important for case management and treatment. Section 1 of the LS/CMI is based on the Level of Service Inventory–Revised (LSI-R; Andrews & Bonta, 1995)

and comprises items representing the Central Eight risk/need factors: Criminal History, Education/Employment, Family/Marital, Leisure/Recreation, Companions, Alcohol/Drug Problems, Procriminal Attitude/Orientation, and Antisocial Personality Pattern. Offenders can be placed in one of five risk categories based on their total Section 1 score (very low, low, medium, high, very high) each associated with probability estimates for recidivism. Section 2 focuses on case management and responsivity concerns (e.g., mental health variables, family history variables, etc.). Section 1 scores are moderately predictive of general (e.g., $d = 0.85$), violent (e.g., $d = 0.72$), and sexual recidivism (e.g., $d = 0.45$; Campbell et al., 2009; Gendreau, Goggin, & Smith, 2002; Hanson & Morton-Bourgon, 2009).

Additional variables for comparison. To further examine the discriminant validity of the Static-2002R General Criminality subscale and the BARR-2002R, several additional variables were examined that were not expected to be correlated to general criminality. The following treatment variables were coded from the offender management system of Correctional Service Canada (CSC), an administrative database for Canadian federal offenders (those serving sentences of 2 years or more): offender accountability (low—that is, offender does not accept responsibility and has high level of denial and cognitive distortions; medium—that is, offender has not fully accepted responsibility but recognizes problems and displays some guilt; and high—that is, offender accepts full responsibility and displays guilt and empathy), offender takes responsibility for actions (no, yes), motivation for treatment (low, medium, high), and engagement with treatment (no, yes). Variables were also taken from the case management section of the LS/CMI: poor social skills, health problems, physical/learning disability, fetal alcohol spectrum disorder, mental illness, and victim of family sexual violence. Finally, the presence of diagnosed sexual deviancy was also coded based on offender file information. The coding manual for these additional variables is available upon request.

Procedure. Offender files for all three groups were obtained from NFS coordinators across Canada as part of a larger evaluation of the flagging system. File information varied considerably; however, the majority of files contained case summaries, court transcripts, psychological assessments, risk assessments, and arrest information. The LS/CMI and Static-2002R were coded for research purposes retroactively from the information provided in the files. The SIR-R1 and PCL-R were copied directly from the files, or from the offender management system when available. The files were coded by a team of four coders. All coders underwent training for the scoring of the LS/CMI by one of the scale developers and Static-2002R by a certified trainer.

After completing a number of training cases, the coders were randomly paired to code 40 files (20 flagged, 5 DOs, and 15 LTOs) for interrater reliability. The reliability of categorical variables was assessed using the kappa statistic, and the reliability of ordinal or continuous variables was assessed using a two-way random-effects model intraclass correlation coefficient (ICC; absolute agreement). When kappa could not be calculated, percent agreement between raters was calculated. The majority of variables

showed acceptable reliability with kappa statistics ranging from .66 to 1.00 (*Mdn* = .75), ICCs ranging from .65 to 1.00 (*Mdn* = .87), and percent agreement ranging from 72.5% to 100% (*Mdn* = 82.5%). When variables did not meet reliability standards (<.65), the coding rules for those variables were clarified, and a consensus coding was achieved between one of the coders and the third author. For these variables, all files were recoded and verified by the third author. With the exception of the LS/CMI total scores (ICC = .94), rater reliability was not available for the other scales because the totals were either computed from coded variables or copied directly from the files.

Recidivism. Recidivism information was obtained from the Canadian Police Identification Centre (CPIC) records, a Canadian national database. Recidivism was defined as any new charges or convictions from official records. Follow-up ranged from 1.2 months to 10.6 years (*M* = 4.6 years, *SD* = 2.1). Five types of recidivism were examined. Nonsexual nonviolent recidivism included all property offenses (e.g., break and enter, theft), narcotic offenses, liquor/traffic violations, and public order and probation/parole violations. Nonsexual violent recidivism included offenses against the person (e.g., assault, robbery), crimes against property with violence (e.g., arson), and resist arrest. Sexual recidivism included any offense that was sexual by name (e.g., sexual assault, indecent exposure) including offenses against public morals that were sexual in nature (e.g., procuring/soliciting/prostitution). Any violent (including sexual) recidivism included all offenses from the violent and sexual recidivism categories. General recidivism was defined as any new charges or convictions.

CPIC coding was completed by two coders (including the third author). A total of 75 cases (60 flagged offenders, 15 LTOs) were chosen for interrater reliability. All variables showed acceptable reliability with kappa statistics ranging from .61 to 1.00 (*Mdn* = .91), ICCs ranging from .65 to 1.00 (*Mdn* = .91), and percent agreement ranging from 72% to 100% (*Mdn* = 99%). Given that all DOs were sentenced to indeterminate terms and none had been released into the community at the time of data collection, CPICs were not requested for this group.

Plan of Analysis

Convergent/discriminant validity. Depending on the measurement scaling of the variables, different correlation coefficients were presented. Specifically, Pearson's correlation coefficients were presented when both variables were continuous, and polychoric correlation coefficients were presented when both variables were ordinal or when one variable was dichotomous and one variable was ordinal. According to Cohen (1992), values of *r* for small, medium, and large effects are .1, .3, and .5.

Calibration. The fit between the expected and observed recidivism rates was examined using the *E* / *O* index (expected number of recidivists divided by observed number of recidivists). The expected recidivism rates were derived from Study 1 and are presented in Appendix B. Perfect calibration is indicated by an *E* / *O* index of 1.0. *E* / *O* values less than 1 indicate that the BARR-2002R underestimated recidivism rates and *E* / *O* values greater than 1 indicate an overestimation of recidivism rates. Following

Table 6. Convergent Validity of the BARR-2002R and General Criminality Factor Score.

	r		
	BARR-2002R	General criminality factor	Sample size
Age at release	−.62	−.16	320
PCL-R total score	.51	.53	133
SIR-R1[a]	.70	.74	214
LS/CMI total score	.61	.59	270
Leisure/recreation	.28	.24	304
Companions	.47	.43	266
Alcohol/drug	.40	.40	274
Procriminal attitude/orientation	.45	.45	307
Antisocial pattern	.58	.57	283
Criminal history	.73	.80	325
Education/employment	.40	.38	219
Family/martial	.24	.35	201

Note. All Pearson correlation coefficients are significant at $p < .01$. BARR-2002R = Brief Assessment of Recidivism Risk–2002R; PCL-R = Psychopathy Checklist–Revised; SIR-R1 = Statistical Information on Recidivism–Revised 1; LS/CMI = Level of Service/Case Management Inventory.
[a]SIR-R1 was reversed scored so that higher scores are associated with greater risk of general recidivism.

Rockhill, Byrne, Rosner, Louie, and Colditz (2003), the 95% CIs for the E/O indices were computed as follows:

$$95\% \text{ CI } (E/O) = (E/O)\exp^{\left(\pm 1.96\sqrt{1/O}\right)}.$$

The expected number of recidivists was based on the 5-year general recidivism rates. Calibration was also examined by determining the extent to which logistic regression intercept values (centered on the BARR-2002R median score of 2) of the current sample differed from those reported in Study 1 (Table 5). Fixed-effect meta-analysis (Hanson & Broom, 2005) was used to examine the difference between the recidivism base rates, as well as potential differences in discrimination.

Results

Convergent validity. Table 6 presents the correlations between the BARR-2002R, the general criminality factor scores, and measures of similar constructs. Excluding age (discussed below), correlation coefficients ranged from .24 to .73 (*Mdn* = .47) for the BARR-2002R and from .24 to .80 (*Mdn* = .45) for the general criminality factor score. The measures that were most highly correlated with the BARR-2002R and general criminality factor were the Criminal History subscale of the LS/CMI, $r_{\text{BARR-2002R}} = .73$

Table 7. Discriminant Validity of the BARR-2002R and General Criminality Factor Score (Polychoric Correlations).

	r		
	BARR-2002R	General criminality factor	Sample size
Treatment factors			
Accountability	−.16	−.17	116
Responsibility	.18	.22	116
Motivation	−.17	−.10	116
Engagement	−.11	−.10	116
Poor social skills	.08	.03	275
Health problems	−.12	.12	275
Physical disability	−.25**	−.14	275
Learning disability	.23**	.17	275
Fetal alcohol spectrum disorder	.25*	.20	275
Suicide attempts or threats	.07	.05	275
Diagnosis of serious mental disorder	.02	−.03	275
Victim of sexual family violence	.14	.13	275
Sexual deviance	−.27***	−.25***	275

Note. Treatment factors were on a 3-point scale, and all other items were scored as present or absent. BARR-2002R = Brief Assessment of Recidivism Risk–2002R.
*p < .05. **p < .01. ***p < .001.

and $r_{\text{General Criminality}}$ = .80; the LS/CMI total score, $r_{\text{BARR-2002R}}$ = .61 and $r_{\text{General Criminality}}$ = .59; the SIR-R1, $r_{\text{BARR-2002R}}$ = .70 and $r_{\text{General Criminality}}$ = .74; and the PCL-R, $r_{\text{BARR-2002R}}$ = .51 and $r_{\text{General Criminality}}$ = .53. As expected, age had a large correlation with the BARR-2002R (−.62) and a smaller correlation with general criminality (−.16).

Discriminant validity. Table 7 presents the correlations between the BARR-2002R, the general criminality factor scores, and measures of constructs we would not expect to be highly correlated with general criminality (e.g., suicide attempts). The general criminality factor was not significantly correlated with any of the variables with the exception of a small negative correlation with sexual deviancy, r = −.25, n = 275. The BARR-2002R also showed a small negative correlation with sexual deviancy (−.27), reinforcing the distinction between general criminality and sex crime specific risk factors. The BARR-2002R showed small positive correlations with learning disability (.23) and fetal alcohol spectrum disorders (.25), and a negative correlation with physical disability (−.25).

Discrimination. The BARR-2002R demonstrated moderate to large discrimination for general recidivism, AUC = .76, 95% CI = [.70, .82], n = 258; nonsexual violent recidivism, AUC = .74, 95% CI = [.66, .81], n = 258; any violent (including sexual) recidivism,

Table 8. Incremental Validity Between BARR-2002R and Other Scales.

	Exp(*B*)	95% CI	Wald
General recidivism			
Comparison 1 (*n* = 185)			
BARR-2002R	1.35	[1.18, 1.56]	18.29***
LS/CMI	1.01	[0.98, 1.04]	0.71
Comparison 2 (*n* = 142)			
BARR-2002R	1.13	[0.94, 1.35]	1.74
SIR-R1	1.06	[1.02, 1.10]	8.17**
Comparison 3 (*n* = 258)			
BARR-2002R	1.50	[1.32, 1.70]	38.65***
Static-2002R	0.95	[0.87, 1.04]	1.14
Violent (including sexual) recidivism			
Comparison 1 (*n* = 185)			
BARR-2002R	1.47	[1.19, 1.82]	12.38***
LS/CMI	1.00	[0.96, 1.04]	0.01
Comparison 2 (*n* = 141)			
BARR-2002R	1.09	[0.86, 1.38]	0.50
SIR-R1	1.07	[1.02, 1.12]	6.88**
Comparison 3 (*n* = 258)			
BARR-2002R	1.66	[1.38, 1.99]	29.23***
Static-2002R	0.89	[0.78, 1.01]	3.46

Note. Analyses conducted separately for each comparison. SIR scores were reversed scored, so that higher SIR-R1 scores represented higher recidivism risk. BARR-2002R = Brief Assessment of Recidivism Risk–2002R; CI = confidence interval; LS/CMI = Level of Service/Case Management Inventory; SIR-R1 = Statistical Information on Recidivism–Revised 1.
**$p < .01$. ***$p < .001$.

AUC = .70, 95% CI = [.63, .77], *n* = 258; and nonsexual nonviolent recidivism, AUC = .77, 95% CI = [.72, .83], *n* = 258. The BARR-2002R had a moderate to small effect size for the prediction of sexual recidivism, AUC = .60, 95% CI = [.49, .70], *n* = 258. In comparison, the LS/CMI total scores significantly predicted general recidivism, AUC = .68, 95% CI = [.60, .76], *n* = 185, but not any violent (including sexual) recidivism, AUC = .59, 95% CI = [.50, .67], *p* = .062, *n* = 185. For the SIR-R1, the effect sizes were similar to those observed for the BARR-2002R (general recidivism AUC = .75, 95% CI = [.67, .84], *n* = 140; and any violent (including sexual) recidivism AUC = .71, 95% CI = [.61, .82], *n* = 185). For the sample of offenders who had been scored on all three scales (*n* = 117), the predictive accuracy of the BARR-2002R was not significantly different from that of the LS/CMI or SIR-R1 for general and any violent (including sexual) recidivism, based on the Delong tests.[1]

Table 8 presents the incremental analyses of the scales. Scales with an asterisk predicted recidivism after controlling for the other scale (i.e., added above and beyond the second scale). As can be seen in Table 8, the BARR-2002R added incrementally to the LS/CMI total scores and Static-2002R for the prediction of general and violent (including
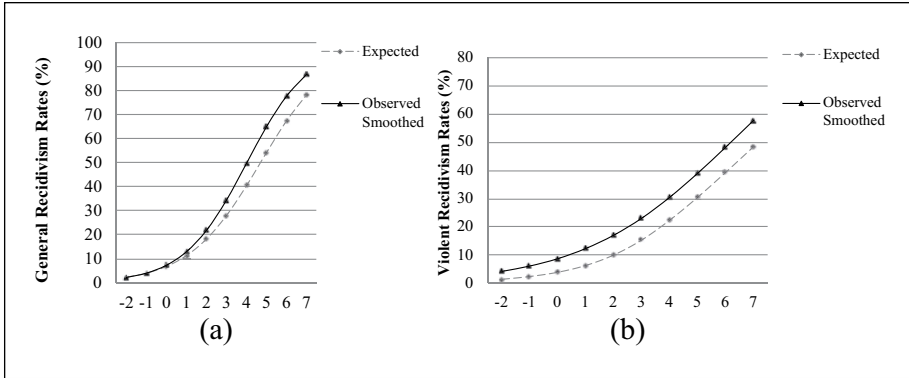
Figure 1. Observed and expected recidivism rates for general and violent recidivism. The observed recidivism rates were generated using the following logistic regression parameters: (a) for general recidivism, B_0 (centered on 2) = −1.28 (SE = 0.408) and B_1 = 0.632 (SE = 0.128); and (b) for violent recidivism, B_0 (centered on 2) = −1.63 (SE = 0.424) and B_1 = 0.441 (SE = 0.118). The expected values are the fitted, meta-analytic averages for Canadian samples (Table 5).

sexual) recidivism. The LS/CMI and Static-2002R no longer predicted general and violent (including sexual) recidivism after controlling for the BARR-2002R. In contrast, the BARR-2002R did not add incrementally to the SIR-R1 and no longer predicted general and violent (including sexual) recidivism after controlling for the SIR-R1.

Comparisons with the development samples. The parameters used to compute predicted recidivism rates of the BARR-2002R were aggregated using fixed-effect meta-analysis from the three Canadian samples (n = 1,112; see Table 5). Neither of the logistic regression parameters (i.e., discrimination [B_1] and calibration [B_0]) in the replication sample were significantly different from the meta-analytic averages in the development sample (see Figure 1). Specifically, the 5-year general recidivism rate associated with a BARR-2002R score of 2 was 19.9% in Study 2 compared with an expected value of 18.2%, $Q_{between}$ = 0.022, df = 1, p = .881; updated meta-analytical average = 18.5%. For any violent (including sexual) recidivism, the fitted value in Study 2 was 16.4%, which was not significantly different from the expected value of 12.3%, $Q_{between}$ = 0.200, df = 1, p = .655; updated meta-analytical average = 13.1%. The B_1 (discrimination) values were also similar for general recidivism, B_1 of 1.88 versus expected value of 1.75; $Q_{between}$ = 0.033, df = 1, p = .856; updated meta-analytical average = 1.78, and any violent recidivism, B_1 of 1.55 versus 1.53; $Q_{between}$ = 0.002, df = 1, p = .964; updated meta-analytical average = 1.54.

E / O indices were used to examine the fit (calibration) between the predicted 5-year general and any violent (including sexual) recidivism rates derived in Study 1 (see Appendix B) and the observed recidivism rates in Study 2. The observed number of general recidivists (n = 59) was not significantly different from the expected number of recidivists from the predicted 5-year estimates (n = 50), E / O = 0.85, 95% CI = [0.66,

1.10], $N = 104$. Similarly, the observed number of any violent (including sexual) recidivists ($n = 42$) was not significantly different than the number expected generated from the absolute recidivism norms ($n = 33$), $E / O = 0.78$, 95% CI = [0.58, 1.07], $N = 104$.

Discussion

Static, criminal history variables predict recidivism because they are markers for risk-relevant propensities, such as antisocial orientation and atypical sexual interests (Mann et al., 2010). Identifying the constructs responsible for recidivism risk is an important step forward for forensic risk assessment and has profound utility for prevention and intervention efforts. The Static-99 was originally intended to predict both sexual and any violent (including sexual) recidivism (see A. Harris et al., 2003). Although the base rates were obviously different, the scale was assumed to have similar relationships to both outcomes. However, research on the latent constructs implicit in the Static-99/R and Static-2002/R items suggested that the prediction of violent and general recidivism could be improved by using only a subset of items, namely, those related to age and general criminality. The purpose of the current pair of studies was to develop and validate a risk scale for the prediction of general recidivism using items of the Static-2002R, and to further refine the construct validity of a measure of general criminality.

Although items related to sexual criminality reliably predict sexual recidivism, they have limited association with nonsexual recidivism (Hanson & Morton-Bourgon, 2009). Consequently, removing the sexual criminality items should simplify the scoring and improve the prediction of largely nonsexual outcomes. The current results confirmed the feasibility of this approach. Specifically, a risk scale based only on the Static-2002R age item and general criminality items (the BARR-2002R) predicted any violent (including sex) and general recidivism better than Static-99R and Static-2002R total scores. In an independent sample, the BARR-2002R correlated with other risk assessment tools designed to predict general recidivism and the PCL-R. In addition, the general criminality scale was found to be related to other measures of antisocial traits (e.g., PCL-R) and criminality (e.g., LS/CMI Criminal History subscale). The BARR-2002R was also found to predict general and violent recidivism just as well as other, more complicated measures specifically designed for these outcomes (i.e., LS/CMI, SIR-R1). Importantly, Study 2 replicated the Study 1 finding that dropping the sexual criminality items from Static-2002R improved the prediction of general and violent recidivism (less is more). We also found that the BARR-2002R added incrementally over LS/CMI scores for general and violent recidivism outcomes. In contrast, the SIR-R1 added incrementally over the BARR-2002R, rendering the latter no longer significant. As in previous research (Hanson & Morton-Bourgon, 2009), the SIR-R1 was a robust predictor of general recidivism in this sample of high-risk sexual offenders. Readers should be cautioned, however, that these incremental analyses had low statistical power and are likely to be unstable. Samples in the thousands are required to identify reliable incremental effects of highly correlated risk scales (Tosteson, Buzas, Demidenko, & Karagas, 2003). With sufficiently large samples, incremental effects of even highly correlated measures should be expected (see Babchishin et al., 2012b, for discussion).

Another limitation of the scale comparisons was that different scoring procedures were used for the different measures. Research has often found that, compared with research studies, the scoring of risk scales in routine practice has lower reliability and weaker prediction effects (Boccaccini, Murrie, Caperton, & Hawes, 2009; Boccaccini et al., 2012). Such concerns, however, were somewhat mitigated in Study 2 because there was no clear pattern of superiority for the scales scored for research over the scores extracted from clinical files. Although not all findings were significant, the direction of the effects was that the BARR-2002R (researcher scored) did better than the Static-2002R total scores (researcher scored) and LS/CMI (researcher scored), but not the SIR-R1 (clinician scored).

The current results advance our understanding of risk assessment for sexual recidivism by providing a simple, practical method of identifying the source of the risk. Previous factor analytic studies have found that the items on the General Criminality subscale of Static-2002R consistently load on a common factor with moderate internal consistency (Boughner, 2010; Brouillette-Alarie, 2013; Ennis et al., 2011; Langton et al., 2007). Such findings have been interpreted as indicating a latent dimension related to antisocial orientation or psychopathy. The current study supports the validity of such an interpretation by finding moderate to strong correlations with other measures of the similar constructs. For example, the correlation between the general criminality factor and the PCL-R was .53, and .59 for LS/CMI total scores. Ennis, Buro, and Jung (2014) found that Static-2002R can classify sex offenders into meaningfully distinct groups differing on treatment need, including general criminality. Although not a comprehensive measure of antisociality, separately considering general criminality from the total Static-2002R score could support inferences concerning appropriate options for supervision and treatment (e.g., because this offender is high on general criminality, he is likely at risk for treatment dropout and noncompliance with community supervision).

Implications for Practice

We recommend evaluators use the BARR-2002R for predicting violent and general recidivism among sex offenders rather than the Static-99R or Static-2002R total scores. In addition, the BARR-2002R can be useful for screening for the psychological dimension of antisocial orientation. Sex offenders who score high on the BARR-2002R could also benefit from a more detailed risk assessment to identify risk-relevant propensities and inform treatment and supervision needs.

Modern actuarial risk scales can provide several types of risk-relevant information, including nominal risk categories (e.g., "low," "moderate," "high") and quantitative indicators of risk (e.g., estimated recidivism rates). The BARR-2002R provides percentiles, expected recidivism rates, and risk ratios (see static99.org) to aid in risk communication and interpretation of risk scores. Interpretation of risk assessments has found to be influenced by how the results are communicated (Varela, Boccaccini, Cuervo, Murrie, & Clark, 2014) and experts disagree on what "low," "moderate," and "high" risk actually represent, irrespective of experience with risk assessments (e.g., Hilton, Carter, Harris, & Sharpe, 2008; Slovic, Monahan, & MacGregor, 2000). As

such, providing nonarbitrary numerical definitions has the potential of improving the fidelity of risk communication (e.g., "high risk" means that the offender is in the top 15% in terms of risk for sexual recidivism). Given that decision makers and evaluators prefer communicating risk in nominal terms (e.g., Grann & Pallvik, 2002), we propose that BARR-2002R scores less than 1 represent low risk, 1 to 4 represent moderate risk, and scores of 5 or higher represent high risk of recidivism. These categories are based on using plus or minus one standard deviation to define typical scores (i.e., scores between the 16th and 84th percentiles are "moderate," scores in the bottom 16% are "low," and scores above the 84th percentile are "high").

Implications for Research

It would be expected that general offenders would score higher than sex offenders on general criminality and, hence, revised norms may be needed for nonsex offenders (e.g., percentile ranks). As such, future research should examine the generalizability of the BARR-2002R to a routine sample of nonsex offenders. The BARR-2002R and general criminality measure would need to be modified to be used for nonsex offenders. Specifically, the item time-free before sexual index would need to be recoded to time-free before *any* index. As well, researchers should consider whether the current time threshold (i.e., 3 years) equally applies to sexual and nonsexual recidivism. A shorter threshold is probably optimal for nonsexual recidivism because the criminal justice system's processing rate of nonsexual convictions is more rapid than the rate of sexual convictions. It is commonly accepted that validation studies require information about both discrimination and calibration (Moons, Royston, Vergouwe, Grobbee, & Altman, 2009). Researchers validating the BARR-2002R and other risk scales should also use both indices as these provide different information. For researchers interested in constructing local norms, we recommend that revisions to the norms be based on at least 100 recidivists from a clearly defined sample (Vergouwe, Steyerberg, Eijkemans, & Habbema, 2005), and that the reliability of both the item scoring and the recidivism information be verified.

In comparison with the general criminality factor, which is reasonably well-defined in theory and in evidence, the constructs assessed by the static sexual criminality items are less clear. A recent factor analysis has found that the sexual criminality items are assessing at least two constructs, one related to pedophilia/persistence and another describing young, unattached offenders (Brouillette-Alarie, 2013; Brouillette-Alarie et al., in press). Unfortunately, it is hard to establish the constructs being assessed by these sexual criminality items given that the latent constructs are typically assessed by too few items. Instead, more fundamental research on the construct validity of risk scales would (a) identify important risk-relevant propensities (e.g., pedophilia, sexual preoccupation, hostility toward women) and (b) use a large pool of possible indicators of these underlying constructs. Such research would be necessary to identify the risk-relevant propensities assessed by risk scales and provide important information for prevention and intervention efforts.

Conclusion

Given that sex offenders are more likely to reoffend with a nonsexual crime than a sexual crime (Hanson & Bussière, 1998), risk assessments of sex offenders should also include the assessment of risk for general and violent recidivism. Even when the focus is exclusively on the risk for sexual recidivism, it is useful to understand the source of recidivism risk. The current article presented an efficient method for assessing the risk of general recidivism by using only the Age at Release item and the General Criminality subscale from Static-2002R. Although more comprehensive measures of general criminality would be preferred for complete assessments (e.g., SIR-R1; VRAG-R), the scale developed in this study, the BARR-2002R, had excellent properties for a screening tool. One advantage of the BARR-2002R is that it can be easily computed from variables already routinely collected and from commonly available criminal history information. Instead of using Static-99R or Static-2002R total scores, we recommend evaluators use the BARR-2002R for predicting violent and general recidivism among sex offenders.

Appendix A

Table A1. BARR-2002R.

Age
1. Age at release
    18-34.9 = 2
    35-39.9 = 1
    40-59.9 = 0
    60 or above = −2
General Criminality
2. Any prior involvement with the criminal justice system
    No = 0
    Yes = 1
3. Prior sentencing occasions for anything:
    0-2 prior sentencing occasions for anything = 0
    3-13 prior sentencing occasions = 1
    14 or more prior sentencing occasions = 2
4. Any community supervision violation:
    No = 0
    Yes = 1
5. Years free prior to index sex offense:
    More than 36 months free prior to committing the sexual offense that resulted in the index conviction *and* more than 48 months free prior to index conviction = 0
    Less than 36 months free prior to committing the sexual offense that resulted in the index conviction *or* less than 48 months free prior to conviction for index sex offense = 1
6. Any prior nonsexual violence sentencing occasion:
    No = 0
    Yes = 1

Note. Coding rules for items and norms for BARR-2002R are available on www.static99.org. BARR-2002R = Brief Assessment of Recidivism Risk–2002R.

Appendix B

Recidivism Estimates

Table B1. Five-Year Observed and Estimated General Recidivism Rates for BARR-2002R.

Score	Fixed follow-up		Logistic regression estimates		
	Recidivists/ total	Observed recidivism rate (%)	Predicted recidivism rate (%)	95% CI	
−2	1/64	1.6	2.3	1.5	3.7
−1	1/15	6.7	4.0	2.7	5.8
0	8/124	6.5	6.8	5.0	9.2
1	13/129	10.1	11.3	9.0	14.2
2	39/195	20.0	18.2	15.4	21.4
3	28/117	23.9	28.0	25.1	31.1
4	59/138	42.8	40.4	37.3	43.6
5	52/112	46.4	54.2	50.5	57.9
6	86/119	72.3	67.4	63.0	71.6
7	72/95	75.8	78.3	73.6	82.4
8	3/4	75.0	—	—	—
Total	362/1,112	32.6			

Note. Recidivism estimates based on routine Canadian samples ($N = 1,112$, $n_{recidivists} = 362$, $k = 3$) and a weighted B_1 of 0.5575 ($SE = 0.0380$), a B_0 of −1.5021 ($SE = 0.1014$), and a median correlation of the estimates of −.754. Recidivism estimates are not presented for a score of 8 ($n = 4$). BARR-2002R = Brief Assessment of Recidivism Risk–2002R; CI = confidence interval.

Table B2. Five-Year Observed and Estimated Violent (Including Sexual) Recidivism Rates for BARR-2002R.

Score	Fixed follow-up		Logistic regression estimates		
	Recidivists/total	Observed recidivism rate (%)	Predicted recidivism rate (%)	95% CI	
−2	1/64	1.6	2.5	1.5	4.0
−1	1/15	6.7	3.8	2.5	5.7
0	7/124	5.6	5.6	4.0	7.9
1	8/129	6.2	8.4	6.4	10.9
2	23/195	11.8	12.3	10.1	15.0
3	20/118	16.9	17.8	15.4	20.4
4	40/138	29.0	24.9	22.4	27.6
5	31/112	27.7	33.7	30.5	37.0
6	43/119	36.1	43.8	39.3	48.4
7	47/96	49.0	54.5	48.4	60.5
8	3/4	75.0	—	—	—
Total	224/1,114	20.1			

Note. Recidivism estimates based on routine Canadian samples ($n = 1,114$, $n_{recidivists} = 224$, $k = 3$) and a weighted B_1 of 0.4282 ($SE = 0.0391$), a B_0 of −1.9604 ($SE = 0.1160$), and a median correlation of the estimates of −.797. Recidivism estimates are not presented for a score of 8 ($n = 4$). BARR-2002R = Brief Assessment of Recidivism Risk–2002R; CI = confidence interval.

Authors' Note

Acknowledgments

Declaration of Conflicting Interests

Funding

Note

1.   Of note, the Delong tests were based on a restricted sample size. Offenders had to have had a score on all three scales to be included in the analyses ($n = 117$).

References

References marked with an asterisk indicate studies included in the Study 1 meta-analysis.

Allison, P. D. (1984). *Event history analysis: Regression for longitudinal event data*. Thousand Oaks, CA: Sage.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

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

Andrews, D. A., & Bonta, J. (2010). *The psychology of criminal conduct* (5th ed.). New Providence, NJ: Matthew Bender.

Andrews, D. A., Bonta, J., & Wormith, J. S. (2004). *The Level of Service/Case Management Inventory (LS/CMI): User's manual*. Toronto, Ontario, Canada: Multi-Health Systems.

Babchishin, K. M., Hanson, R. K., & Blais, J. (2013). *User guide for the Brief Assessment for Recidivism Risk–2002R (BARR-2002R)*. Retrieved from http://www.static99.org/pdfdocs/Static2002-BARR-2002REvaluatorsHandbook-2013-12-03.pdf

Babchishin, K. M., Hanson, R. K., & Helmus, L. (2012a). Communicating risk for sex offenders: Risk ratios for Static-2002R. *Sexual Offender Treatment*, *12*(2), 1-12.

Babchishin, K. M., Hanson, R. K., & Helmus, L. (2012b). Even highly correlated measures can add incrementally to predicting recidivism among sex offenders. *Assessment*, *19*, 442-461. doi:10.1177/1073191112458312

Barbaree, H. E., Langton, C. M., & Peacock, E. J. (2006a). Different actuarial risk measures produce different risk rankings for sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, *18*, 423-440. doi:10.1007/s11194-006-9029-9

Barbaree, H. E., Langton, C. M., & Peacock, E. J. (2006b). The factor structure of static actu-arial items: Its relation to prediction. *Sexual Abuse: A Journal of Research and Treatment*, *18*, 207-226. doi:10.1007/s11194-006-9011-6

Beitchman, J. H., Zucker, K. J., Hood, J. E., DaCosta, G. A., Akman, D., & Cassavia, E. (1992). A review of the long-term effects of child sexual abuse. *Child Abuse & Neglect*, *16*, 101-118. doi:10.1016/0145-2134(92)90011-F

*Bengtson, S. (2008). Is newer better? A cross-validation of the Static-2002 and the Risk Matrix 2000 in a Danish sample of sexual offenders. *Psychology, Crime & Law*, *14*, 85-106. doi:10.1080/10683160701483104

*Bigras, J. (2007). *La prédiction de la récidive chez les délinquants sexuels* [Prediction of recid-ivism among sex offenders] (Doctoral dissertation). Retrieved from Dissertation Abstracts International. (UMI No. NR30941)

Blais, J., & Bonta, J. (2014). Tracking and managing high risk offenders: A Canadian initiative. *Law and Human Behavior*. Advance online publication. doi:10.1037/lhb0000109

Boccaccini, M. T., Murrie, D. C., Caperton, J. D., & Hawes, S. W. (2009). Field validity of the STATIC-99 and MnSOST-R among sex offenders evaluated for civil commitment as sexually violent predators. *Psychology, Public Policy, and Law*, *15*, 278-314. doi:10.1037/a0017232

Boccaccini, M. T., Murrie, D. C., Mercado, C., Quesada, S., Hawes, S., Rice, A. K., & Jeglic, E. L. (2012). Implications of Static-99 field reliability findings for score use and reporting. *Criminal Justice and Behavior*, *39*, 42-58. doi:10.1177/0093854811427131

*Boer, A. (2003). *Evaluating the Static-99 and Static-2002 risk scales using Canadian sexual offenders* (Unpublished master's thesis). University of Leicester, UK.

Bonta, J., Blais, J., & Wilson, H. A. (2014). A theoretically informed meta-analysis of the risk for general and violent recidivism for mentally disordered offenders. *Aggression and Violent Behavior*, *19*, 278-287. doi:10.1016/j.avb.2014.04.014

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: John Wiley.

Boughner, K. L. (2010). *Dimensions of the Static-2002: A statistical and theoretical perspective* (Doctoral dissertation). Retrieved from ProQuest, UMI Dissertation Publishing. (3404472)

Brouillette-Alarie, S. (2013, November). *Are there psychological constructs in the Static-99R and Static-2002R?* Paper presented at the 32nd Annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers, Chicago, IL.

Brouillette-Alarie, S., Babchishin, K. M., Hanson, R. K., & Helmus, L. M. (in press). Latent construct of the Static-99R and Static-2002R: A 3-factor solution. *Assessment*.

Brouillette-Alarie, S., Proulx, J., & Benbouriche, M. (2013). Mieux saisir les outils actuariels: Une analyse factorielle de la Statique-99R [Better understanding of actuarial tools: A factor analysis of Static-99R]. *Criminologie*, *46*, 199-219. doi:10.7202/1015299ar

Campbell, M. A., French, S., & Gendreau, P. (2009). The prediction of violence in adult offend-ers: A meta-analytic comparison of instruments and methods of assessment. *Criminal Justice and Behavior*, *36*, 567-590. doi:10.1177/0093854809333610

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155-159. doi:10.1037/0033-2909.112.1.155

Delong, E. R., Delong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics*, *44*, 837-845.

Ennis, L., Buro, K., & Jung, S. (2014). Identifying male sexual offenders subtypes using clus-ter analysis and the Static-2002R. *Sexual Abuse: A Journal of Research and Treatment*. Advance online publication. doi:10.1177/10790632114527481

Ennis, L., Choy, A., Jung, S., Corabian, G., & Hook, T. (2011, October). *Psychometric properties of the Static-2002*. Paper presented at the Annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers, Toronto, Ontario, Canada.

Gendreau, P., Goggin, C., & Smith, P. (2002). Is the PCL-R really the "unparalleled" measure of offender risk? A lesson in knowledge cumulation. *Criminal Justice and Behavior*, *29*, 397-426. doi:10.1177/0093854802029004004

Grann, M., & Pallvik, A. (2002). An empirical investigation of written risk communication in forensic psychiatric evaluations. *Psychology, Crime & Law*, *8*, 113-130. doi:10.1080/1068316020840181

Guay, J. P., Ruscio, J., Knight, R. A., & Hare, R. D. (2007). A taxometric analysis of the latent structure of psychopathy: Evidence for dimensionality. *Journal of Abnormal Psychology*, *116*, 701-716. doi:10.1037/0021-843X.116.4.701

*Haag, A. M. (2005). [Recidivism data from 198 offenders detained until their warrant expiry date. From: Do psychological interventions impact on actuarial measures: An analysis of the predictive validity of the Static-99 and Static-2002 on a re-conviction measure of sexual recidivism]. Retrieved from Dissertation Abstracts International. (NR05662). Unpublished raw data.

Hanson, R. K. (2009). The psychological assessment of risk for crime and violence. *Canadian Psychology/Psychologie canadienne*, *20*, 172-182. doi:10.1037/a0015726

Hanson, R. K., & Broom, I. (2005). The utility of cumulative meta-analysis: Application to programs for reducing sexual violence. *Sexual Abuse: A Journal of Research and Treatment*, *17*, 357-373. doi:10.1177/107906320501700402

Hanson, R. K., & Bussière, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology*, *66*, 348-362. doi:10.1037/0022-006X.66.2.348

*Hanson, R. K., Harris, A. J. R., Scott, T., & Helmus, L. (2007). *Assessing the risk of sexual offenders on community supervision: The Dynamic Supervision Project* (Corrections Research User Report No. 2007-05). Ottawa, Ontario: Public Safety Canada.

Hanson, R. K., Helmus, L., & Thornton, D. (2010). Predicting recidivism amongst sexual offenders: a multi-site study of static-2002. *Law and human behavior*, *34*, 198. doi:10.1007/s10979-009-9180-1

Hanson, R. K., Lloyd, C., Helmus, L., & Thornton, D. (2012). Developing non-arbitrary metrics for risk communication: Percentile ranks for the Static-99/R and Static-2002/R sexual offender risk tools. *International Journal of Forensic Mental Health*, *11*, 9-23. doi:10.1080/14999013.2012.667511

Hanson, R. K., & Morton-Bourgon, K. E. (2005). The characteristics of persistent sexual offenders: A meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology*, *73*, 1154-1163. doi:10.1037/0022-006X.73.6.1154

Hanson, R. K., & Morton-Bourgon, K. E. (2009). The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis of 118 prediction studies. *Psychological Assessment*, *21*, 1-21. doi:10.1037/a0014421

Hanson, R. K., & Thornton, D. (2000). Improving risk assessments for sex offenders: A comparison of three actuarial scales. *Law and Human Behavior*, *24*, 119-136.

Hanson, R. K., & Thornton, D. (2003). *Notes on the development of Static-2002* (User Report 2003-01). Ottawa, Ontario: Department of the Solicitor General of Canada.

Hare, R. D. (2003). *Psychopathy Checklist–Revised* (2nd ed.). Toronto, Ontario, Canada: Multi-Health Systems.

*Harkins, L., & Beech, A. R. (2007). *Examining the effectiveness of sexual offender treatment using risk band analysis*. Unpublished manuscript.

Harris, A., Phenix, A., Hanson, R. K., & Thornton, D. (2003). *Static-99 coding rules: Revised 2003*. Ottawa, Ontario: Department of the Solicitor General of Canada. Available from www.static99.org

Harris, G. T., Rice, M. E., & Quinsey, V. L. (1993). Violent recidivism of mentally disordered offenders: The development of a statistical prediction instrument. *Criminal Justice and Behavior*, *20*, 315-335. doi:10.1177/0093854893020004001

Hawes, S. W., Boccaccini, M. T., & Murrie, D. C. (2013). Psychopathy and the combination of psychopathy and sexual deviance as predictors of sexual recidivism: Meta-analytic findings using the Psychopathy Checklist–Revised. *Psychological Assessment*, *25*, 233-243. doi:10.1037/a0030391

Helmus, L., Thornton, D., Hanson, R. K., & Babchishin, K. M. (2012). Improving the predictive accuracy of Static-99 and Static-2002 with older sex offenders: Revised age weights. *Sexual Abuse: A Journal of Research and Treatment*, *24*, 64-101. doi:10.1177/1079063211409951

Higgins, J., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, *327*, 557-560. doi:10.1136/bmj.327.7414.557

Hilton, N. Z., Carter, A., Harris, G. T., & Sharpe, A. J. B. (2008). Does using nonumerical terms to describe risk aid violence risk communication? Clinician agreement and decision making. *Journal of Interpersonal Violence*, *23*, 171-188. doi:10.1177/0886260507309337

Janka, C., Gallasch-Nemitz, F., & Dahle, K.-P. (2011). Zur Altersabhängigkeit von Risikovariablen bei Sexualdelinquenz [Age dependency of risk variables for sexual offending]. *Forensische Psychiatrie, Psychologie, Kriminologie*, *5*, 37-44.

*Knight, R. A., & Thornton, D. (2007ab). *Evaluating and improving risk assessment schemes for sexual recidivism: A long-term follow-up of convicted sexual offenders* (Document No. 217618). U.S. Department of Justice.

Langton, C. M., Barbaree, H. E., Hansen, K. T., Harkins, L., & Peacock, E. J. (2007). Reliability and validity of the Static-2002 among adult sexual offenders with reference to treatment status. *Criminal Justice and Behavior*, *34*, 616-640. doi:10.1177/0093854806296851

*Lehmann, R. J. B., Hanson, R. K., Babchishin, K., Gallasch-Nemitz, F., Biedermann, J., & Dahle, K.-P. (2013). Interpreting multiple risk scales for sex offenders: Evidence for averaging. *Psychological Assessment*, *25*, 1019-1024. doi:10.1037/a0033098

Mann, R. E., Hanson, R. K., & Thornton, D. (2010). Assessing risk for sexual recidivism: Some proposals on the nature of psychologically meaningful risk factors. *Sexual Abuse: A Journal of Research and Treatment*, *22*, 191-217. doi:10.1177/1079063210366039

Moons, K. G., Royston, P., Vergouwe, Y., Grobbee, D. E., & Altman, D. G. (2009). Prognosis and prognostic research: What, why, and how? *British Medical Journal*, *338*, 1317-1320. doi:10.1136/bmj.b375

Nafekh, M., & Motiuk, L. L. (2002). *The Statistical Information on Recidivism, Revised 1 (SIR-R1) Scale: A psychometric examination*. Ottawa, Ontario: Correctional Service of Canada. Retrieved from http://publications.gc.ca/collections/collection_2010/scc-csc/PS83-3-126-eng.pdf

National Policing Improvement Agency. (2010). *Guidance on protecting the public: Managing sexual offenders and violent offenders* (2nd ed.). London, England. Retrieved from http://www.acpo.police.uk/mwg-internal/de5fs23hu73ds/progress?id=osBo7Gp5+6

Nuffield, J. (1982). *Parole decision-making in Canada: Research towards decision guidelines*. Ottawa, Ontario: Solicitor General of Canada.

Part XXIV. *Dangerous offenders and long term offenders*. Criminal Code of Canada, R.S.C 1985, c.46, s.752. Retrieved from http://canada.justice.gc.ca

Patrick, C. J. (Ed.). (2005). *Handbook of psychopathy*. New York, NY: Guilford Press.

Patrick, C. J., Fowles, D. C., & Krueger, R. F. (2009). Triarchic conceptualization of psychopathy: Developmental origins of disinhibition, boldness, and meanness. *Development and Psychopathology*, *21*, 913-938. doi:10.1017/S0954579409000492

Pintea, S., & Moldovan, R. (2009). The receiver-operating characteristic (ROC) analysis: Fundamentals and applications in clinical psychology. *Journal of Cognitive and Behavioral Psychotherapies*, *9*(1), 49-66.

Prentky, R. A., Lee, A. F., Knight, R. A., & Cerce, D. (1997). Recidivism rates among child molesters and rapists: A methodological analysis. *Law and Human Behavior*, *21*, 635-659. doi:10.1023/A:1024860714738

Public Safety Canada. (2011). *2011-2012 evaluation of the National Flagging System program*. Retrieved from http://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/vltn-ntnl-flggng-sstm-2011-12/index-eng.aspx

Rice, M. E., & Harris, G. T. (1995). Violent recidivism: Assessing predictive validity. *Journal of Consulting and Clinical Psychology*, *63*, 737-748.

Rice, M. E., Harris, G. T., & Lang, C. (2013). Validation of and revision to the VRAG and SORAG: The Violence Risk Appraisal Guide–Revised (VRAG-R). *Psychological Assessment*, *25*, 951-965. doi:10.1037/a0032878

Roberts, C. F., Doren, D. M., & Thornton, D. (2002). Dimensions associated with assessments of sex offender recidivism risk. *Criminal Justice and Behavior*, *29*, 569-589. doi:10.1177/009385402236733

Roberts, R., O'Connor, T., Dunn, J., & Golding, J. (2004). The effects of child sexual abuse in later family life; mental health, parenting and adjustment of offspring. *Child Abuse & Neglect*, *28*, 525-545. doi:10.1016/j.chiabu.2003.07.006

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., & Müller, M. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, *12*, Article 77. doi:10.1186/1471-2105-12-77

Rockhill, B., Byrne, C., Rosner, B., Louie, M. M., & Colditz, G. (2003). Breast cancer risk prediction with a log-incidence model: Evaluation of accuracy. *Journal of Clinical Epidemiology*, *56*, 856-861. doi:10.1016/S0895-4356(03)00124-0

Schulze, R. (2007). Current methods for meta-analysis: Approaches, issues, and developments. *Zeitschrift für Psychologie/Journal of Psychology*, *215*, 90-103. doi:10.1027/0044-3409.215.2.90

Slovic, P., Monahan, J., & MacGregor, D. G. (2000). Violence risk assessment and risk communication: The effects of using actual cases, providing instruction, and employing probability versus frequency formats. *Law and Human Behavior*, *24*, 271-296. doi:10.1023/A:1005595519944

Social Work Inspection Agency. (2009). *Multi-agency inspection: Assessing and managing offenders who present a high risk of serious harm* (Report No. 9781905501892). Edinburgh, Scotland: Social Work Inspection Agency, HM Inspectorate of Constabulary for Scotland, and HM Inspectorate of Prisons. Retrieved from http://www.scotland.gov.uk/Resource/Doc/275852/0082871.pdf

Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest*, *1*, 1-26. doi:10.1111/1529-1006.001

Tosteson, T. D., Buzas, J. S., Demidenko, E., & Karagas, M. (2003). Power and sample size calculations for generalized regression models with covariate measurement error. *Statistics in Medicine*, *22*, 1069-1082. doi:10.1002/sim.1388

Varela, J. G., Boccaccini, M. T., Cuervo, V. A., Murrie, D. C., & Clark, J. W. (2014). Same score, different message: Perceptions of offender risk depend on Static-99R risk communication format. *Law and Human Behavior*, *38*, 418-427. doi:10.1037/lhb0000073

Vergouwe, Y., Steyerberg, E. W., Eijkemans, M. J. C., & Habbema, J. D. F. (2005). Substantial effective sample sizes are required for external validation studies of predictive logistic regression models. *Journal of Clinical Epidemiology*, *58*, 475-483. doi:10.1016/j.jclinepi.2004.06.017

Wormith, J. S., Hogg, S., & Guzzo, L. (2012). The predictive validity of a general risk/needs assessment inventory on sexual offender recidivism and an exploration of the professional override. *Criminal Justice and Behavior*, *39*, 1511-1538. doi:10.1177/0093854812455741

Yang, M., Wong, S. C., & Coid, J. (2010). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin*, *136*, 740-767. doi:10.1037/a0020473