CANADIAN PSYCHOLOGICAL ASSOCIATION
SOCIÉTÉ CANADIENNE DE PSYCHOLOGIE

Canadian Psychology / Psychologie canadienne

0708-5591/20/$12.00

2020, Vol. 61, No. 1, 42–57
http://dx.doi.org/10.1037/cap0000182

Predictive Accuracy of the Static-99R and Static-2002R Risk Tools for Identifying Indigenous and White Individuals at High Risk for Sexual Recidivism in Canada

Seung C. Lee
Public Safety Canada, Ottawa, Ontario, Canada, and Carleton University

R. Karl Hanson and Julie Blais
Carleton University

In Canada, the issue of cultural bias in risk assessment tools is an ongoing concern, particularly for Indigenous peoples who are overrepresented in the criminal justice system. To date, there has been relatively little research on structured risk assessment tools for Indigenous peoples. This study examined the predictive accuracy of two sexual recidivism risk tools (Static-99R, Static-2002R) for White individuals (n = 1,560) and persons of Indigenous heritage (n = 653) from five independent Canadian samples. The results indicated that Static-99R predicted sexual recidivism with similar accuracy for both Indigenous and White individuals. Static-2002R predicted sexual recidivism only for White individuals. Indicators of general criminality (e.g., total prior offences) predicted sexual recidivism for both racial groups. In contrast, many of the sex crime–specific risk factors that have demonstrated validity in previous, predominantly White samples (e.g., male victim, non-contact sexual offence), only predicted sexual recidivism for the White study group and did not predict sexual recidivism for the Indigenous study group. Consistent with other studies, White individuals showed more indicators of paraphilic interests whereas Indigenous individuals in these samples displayed higher general criminality. The findings suggest that the treatment for individuals at risk for sexual offending may benefit from an increased focus on the distinctive risk-relevant characteristics of White and Indigenous peoples in the criminal justice system.

Public Significance Statement
The prediction tools used in the criminal justice system should not be culturally biased, particularly not in ways that increase the severity of sanctions for already disadvantaged groups (e.g., Indigenous peoples of Canada). This study found that the Static-99R sexual recidivism prediction tool demonstrated similar accuracy for Indigenous and White individuals, whereas another prediction tool (Static-2002R) predicted sexual recidivism only for White individuals. These findings reinforce the need to evaluate the potential for cultural bias in assessment measures used for applied decision-making.

Keywords: cultural bias, Indigenous peoples, sexual offending, Static-99R, Static-2002R

Supplemental materials: http://dx.doi.org/10.1037/cap0000182.supp

The purpose of the Canadian correctional system is to administer sentences, promote public safety, and assist in the rehabilitation of offenders (Corrections Act, 1988; Corrections and Conditional Release Act, 1992, 2012; Ministry of Correctional Services Act, 1990). Forensic risk assessments, commonly conducted by psychologists, are widely used to achieve these goals, routinely

This article was published Online First July 18, 2019.

Seung C. Lee, Public Safety Canada, Ottawa, Ontario, Canada, and Department of Psychology, Carleton University; R. Karl Hanson and Julie Blais, Department of Psychology, Carleton University.

We thank Patti McDonald and Leticia Gutierrez for their helpful feedback on an early version of this article. We also thank the following researchers for permitting us to use their data: Jim Bonta, Andy Haag, and Mark Olver. R. Karl Hanson is a co-author and certified Static-99R and Static-2002R trainer. The copyright for Static-99R and Static-2002R is held by the Government of Canada, and none of the authors receives royalties from this measure. This study was conducted by the lead author as part of his duties as an employee of Public Safety Canada; consequently, the copyright is held by the Government of Canada. This research project received research ethics approval from Carleton University's Research Ethics Boards (CUREB)-B. The opinions, findings, conclusions, and recommendations expressed are those of the authors and do not necessarily reflect those of the institutions that supported this research.

Correspondence concerning this article should be addressed to Seung C. Lee, Department of Psychology, Carleton University, 1125 Colonel By Drive, Ottawa, ON K1S 5B6, Canada. E-mail: seungc.lee@carleton.ca

informing decisions concerning who can be incarcerated indefinitely (e.g., Dangerous Offender hearings; ss. 752 and 753 of the Criminal Code of Canada [1985]) and who can be released early. Risk tools also inform the amount of rehabilitation services provided to offenders. Most Canadian correctional systems follow the Risk, Need, and Responsivity model of offender rehabilitation (RNR) in which the most intensive services are provided to the highest risk offenders (Andrews & Bonta, 2010; Andrews, Bonta, & Hoge, 1990).

An ongoing concern is the extent to which risk measures, developed and validated using predominantly White samples, can be applied to other ethnic and racialized minority groups (i.e., cultural bias). In Canada, the issue of cultural bias has mostly concerned individuals of Indigenous heritage, who are overrepresented at all stages of the criminal justice system (Haag, Boyes, Cheng, MacNeil, & Wirove, 2016). This issue has received renewed attention with the *Ewert v. Canada* (2015, 2018 case where the Supreme Court of Canada ruled that several commonly used risk assessment scales (e.g., Static-99; Hanson & Thornton, 2000; Psychopathy Checklist-Revised [PCL-R], Hare, 2003) had insufficient evidence on cross-cultural validity to justify their use with Indigenous peoples with a history of crime (for an overview of the Ewert case, see a special issue published in the *Journal of Threat Assessment and Management* [2016; Vol. 3, No. 2]). Given this context, the purpose of the current study was to examine the potential for cultural bias in the predictive validity of two widely adopted scales for individuals with a history of sexual crimes (i.e., Static-99R [Hanson & Thornton, 2000; Helmus, Thornton, Hanson, & Babchishin, 2012] and Static-2002R [Hanson & Thornton, 2003; Helmus, Thornton, et al., 2012]), one of which was specifically named in the Ewert case.

Indigenous Overrepresentation in Canadian Corrections

Whereas Indigenous peoples (First Nations, Métis, and Inuit) account for 4% of the total Canadian adult population (Statistics Canada, 2015), they account for more than 20% of the federal inmate population (Public Safety Canada, 2015) and 23% (444/1,898) of individuals serving a federal sentence for a sexual offence (MacDonald, 2014). Indigenous peoples are more likely to be placed in higher security prisons (Public Safety Canada, 2015), less likely to be granted statutory release (i.e., released after 2/3 of their sentence; Public Safety Canada, 2015), and more likely to have their parole revoked (Office of the Correctional Investigator, 2015).

There are several proposed reasons for the overrepresentation of Indigenous peoples in Canadian corrections. First, Indigenous peoples are more likely to be arrested and convicted than non-Indigenous peoples (e.g., Bonta, Laprairie, & Wallace-Capretta, 1997; Quann & Trevethan, 2000) and more likely to reoffend (Gutierrez, Wilson, Rugge, & Bonta, 2013; Rojas & Gretton, 2007; Sioui & Thibault, 2002). Indigenous individuals in the criminal justice system are also rated higher on the established risk factors for general and violent recidivism than non-Indigenous individuals (i.e., Central Eight risk/needs factors; Andrews & Bonta, 2010). For example, Indigenous individuals in the criminal justice system have lengthier criminal histories (e.g., Holsinger, Lowenkamp, & Latessa, 2006; Perley-Robertson, Helmus, & Forth, 2018; Shepherd, Adams, McEntyre, & Walker, 2014), higher rates of substance abuse (e.g., Ellerby & MacPherson, 2002; Shepherd et al., 2014), more problems with family/marital relationships (e.g., Shepherd et al., 2014; Trevethan, Moore, & Rastin, 2002), and less education/employment (Holsinger, Lowenkamp, & Latessa, 2003; Shepherd et al., 2014; Trevethan et al., 2002) than non-Indigenous individuals in the criminal justice system.

The presence of these risk factors for crime is not surprising given other social and economic indicators. Indigenous peoples have lower median incomes ($25,526 vs. $34,604), lower employment rates (52% vs. 60.5%), lower education levels (62% with a high school diploma vs. 80%; Statistics Canada, 2018), and poorer health outcomes (e.g., mental health issues and short life expectancy; Statistics Canada, 2005, 2015). Indigenous peoples also report higher rates of childhood maltreatment (neglect and sexual/physical abuse) and intimate partner violence (Boyce, 2016; Scrim, 2010; Sinha et al., 2011). They are also more likely than non-Indigenous peoples to figure as victims in official crime statistics (Boyce, 2016; Scrim, 2010; Sinha et al., 2011).

These indicators of adversity and social disadvantage need to be understood in the context of Canada's history of colonization, and the devastating effects of racist social policies toward Indigenous peoples (Truth and Reconciliation Commission of Canada, 2015). As the Royal Commission on Aboriginal Peoples (RCAP) makes clear, many of the most pressing problems in Indigenous communities are predictable consequences of the dominant social class' efforts to eliminate Indigenous culture (i.e., residential schools, the *Indian Act*, and enfranchisement, wherein the Canadian government would strip Status Indians of their status if they obtained a university degree, served in the armed forces, or if a Status Indian woman married a non-Indigenous or nonstatus person). Whether it is poverty, substance abuse, low levels of formal education, or alienation and isolation, the criminogenic factors that contribute to higher rates of crime among Indigenous peoples in Canada are rooted in some 500 years of Indigenous-settler relations. To quote the RCAP (1996):

> People who endure these disruptions may feel adrift—disoriented and unsure of how to get along in the sometimes hostile non-Aboriginal world. If their aboriginality has been devalued or ridiculed, they may have lost pride and self-esteem and be unable to build these qualities in their children. If they have been damaged in heart and soul, they may turn to alcohol, violence, crime or other forms of antisocial behaviour. (p. 36)

The reserve system has further marginalized Indigenous communities and fostered environments conducive to crime for the following reasons: (a) the reserve system creates large groups of disadvantaged people with limited resources; (b) these communities tend to exist separate from mainstream society creating the potential for culture clash (e.g., differing views of crime and justice); and (c) individuals in this environment are more likely to experience early and sustained abuse (Laprairie, 1996). Although such conditions do not describe all Indigenous communities, there is no question that the reserve system placed non-Indigenous peoples in control of most of Canada's natural resources—natural resources for which Indigenous peoples continue to be dependent upon.

Another factor that may lead to Indigenous overrepresentation is the effect of systemic discrimination and bias in how Indigenous

individuals are treated at all levels of the criminal justice system (*R v. Gladue*, 1999; Rudin, 2009; Truth and Reconciliation Commission of Canada, 2015). For example, there is evidence for the differential treatment of Indigenous peoples in policing (i.e., higher likelihood of arrest), sentencing (i.e., longer sentences), and rehabilitation (i.e., reduced access to culturally appropriate programming; LaPrairie, 2002; Mann, 2009; Rudin, 2009). Underlying this discrimination is the false assumption that Indigenous peoples, just because of their race, are more likely to commit a crime (i.e., higher risk) than non-Indigenous individuals. We turn now to a discussion of bias in assessing this level of risk.

Risk Assessment and Cultural Bias

Psychologists have the responsibility to use assessment methods that are supported by evidence and appropriate to the particular cultural and social contexts of the individuals being assessed (Canadian Psychological Association, 2017, p. 20). Considerable research has demonstrated that structured approaches to risk assessment, particularly those based on statistical methods (i.e., actuarial), provide a more accurate estimate for the likelihood of recidivism than unstructured judgments (Ægisdóttir et al., 2006; Grove, Zald, Lebow, Snitz, & Nelson, 2000; Hanson & Morton-Bourgon, 2009). Although recidivism risk tools can inform placement of latent psychological dimensions (e.g., antisociality), they are primarily evaluated based on predictive accuracy (Gutierrez, Helmus, & Hanson, 2016). Predictive accuracy has two major features: (a) discrimination, or the extent to which recidivists are different from nonrecidivists, and (b) calibration, how closely the predicted recidivism rates match the observed recidivism rates in replication studies (Helmus & Babchishin, 2017). A prediction tool is biased when either the discrimination or the calibration varies across groups (Reynolds & Suzuki, 2013).

There are several possible reasons for the cultural bias in risk tools (Geisinger & McCormick, 2013; see Haag et al. [2016] for a more thorough review). The items could mean different things to different cultural groups (e.g., the term *unrelated victim* could mean something different to someone from a collectivistic vs. an individualistic culture). As well, the *constructs* measured by a scale may mean different things to individuals with different cultural backgrounds (Geisinger & McCormick, 2013). For example, negative attitudes toward authority may be an indicator of narcissistic entitlement for affluent White individuals whereas it could be a natural reaction to colonial oppression for Indigenous parents whose children were abducted for indoctrination in residential schools. It is also possible that there are culturally unique constructs, which only have meaning for individuals of that culture. Such constructs are inherently difficult to identify and document; however, there is no reason to suppose that the constructs currently articulated in Western social science fully capture the lived experience of all peoples at all times (see the emic/etic distinction in Berry [1969]).

Given the ongoing concerns of Indigenous overrepresentation in the criminal justice system (e.g., Mann, 2009; Office of the Auditor General of Canada, 2016), the Ewert case has reinforced the need for more and better research on recidivism risk tools for individuals of Indigenous heritage. Mr. Ewert, a Métis, serving two concurrent life sentences for sexually motivated homicide and assault, challenged the use of psychological and actuarial risk assessment instruments (e.g., Static-99) by the Correctional Service of Canada (CSC). The claim was that Mr. Ewert's rights under s. 24 (1) of the *Corrections and Conditional Release Act* (i.e., the use of accurate and up-to-date information) and ss. 7 (i.e., right to life, liberty, and security of the person) and 15 (i.e., equality before and under the law) of the *Canadian Charter of Rights and Freedoms* (1982) were violated given that the instruments were not adequately validated for use with Indigenous peoples (*Ewert v. Canada, 2015*).

The 2015 trial resulted in a full victory for Mr. Ewert; however, this decision was completely overturned on appeal. The case, therefore, proceeded to the Supreme Court of Canada where it was determined that, although Mr. Ewert's Charter rights were *not* infringed upon, the court made a declaration that CSC had failed in its duties under s. 24 (1). More specifically:

> The CSC had long been aware of concerns regarding the possibility of these tools exhibiting cultural bias yet took no action to confirm their validity and continued to use them in respect of Indigenous offenders, despite the fact that research would have been feasible. In doing so, the CSC did not meet the legislated standard set out in s. 24 (1). (*Ewert v. Canada. 2018*, pp. 7–8)

To our knowledge, this is the only Supreme Court of Canada decision that considered the amount of research evidence needed before psychologists could use a specific assessment measure.

Risk Assessment With Indigenous Individuals With a History of Sexual Crimes

Research on samples of predominantly White people have identified two broad categories of sexual recidivism risk factors: indicators of general antisociality (e.g., impulsiveness, supervision failures, criminal history) and sex crime–specific factors (e.g., sexual preoccupation, emotional congruence with children; Hanson & Morton-Bourgon, 2005; Mann, Hanson, & Thornton, 2010; Whitaker et al., 2008). When comparing the presence of these risk factors, Indigenous individuals with a history of sexual crime tend to be younger, score higher on general antisociality, and score lower on sexual deviancy compared with non-Indigenous individuals (Babchishin, Blais, & Helmus, 2012; Ellerby & MacPherson, 2002; Helmus, Babchishin, & Blais, 2012).

Few studies have examined the predictive validity of risk assessment scales with Indigenous peoples with a history of sexual offending. Olver and his colleagues (2018) found that the Violence Risk Scale - Sex Offender version (VRS-SO; Wong, Olver, Nicholaichuk, & Gordon, 2003) significantly predicted sexual recidivism for both Indigenous and non-Indigenous individuals with a history of sexual crimes, and that treatment change was significantly related to reductions in risk for both groups. In contrast, Helmus, Babchishin, et al. (2012) found that the STABLE-2007 (Fernandez, Harris, Hanson, & Sparks, 2014) did not significantly predict sexual recidivism for Indigenous individuals with a history of sexual crimes. In their study, the general criminality items performed poorly, whereas sexual deviancy items performed equally well for Indigenous and non-Indigenous individuals.

Static-99R and Static-2002R are the most commonly used actuarial risk tools for adult males with a history of sexual crimes in Canada (Bourgon, Mugford, Hanson, & Coligado, 2018; Hill &

Demetrioff, 2019) and the United States. (Kelley, Ambroziak, Thornton, & Barahal, 2018). Static-99R contains 10 static (i.e., unchangeable) items (e.g., prior sex offences), and the total score (−3 to 12) is used to estimate the likelihood of sexual offending. Static-2002R was developed to improve the conceptual coherence of the Static-99R scale and, thus, to measure psychological constructs (e.g., deviant sexual interests, general criminality). Static-2002R also contains only static items (eight of 14 items from Static-99R), and the total score (−2 to 13) is used to estimate the likelihood of sexual offending.

Although primarily intended as prediction tools, the Static-99R and Static-2002R items can inform the assessment of psychologically meaningful constructs. Brouillette-Alarie, Babchishin, Hanson, and Helmus (2016) identified three major factors in the STATIC items: (a) persistence/paraphilia (e.g., rate of sexual offending, male victim); (b) youthful stranger aggression (e.g., age at release, nonsexual violence at index, unrelated victim); and (c) general criminality (e.g., breach of conditional release; prior nonsexual violence). More specifically, the first construct, persistence/paraphilia, is associated with deviant sexual interests, but without the intent to harm victims (e.g., pedophilia, voyeurism, and exhibitionism). The second construct, youthful stranger aggression, is related to the intent to sexually harm the victim (e.g., sexual sadism, hostility toward women). The construct of general criminality is related to antisocial traits (e.g., impulsivity, lack of remorse, rule violations; Brouillette-Alarie, Proulx, & Hanson, 2018).

In an earlier meta-analysis, Babchishin, Blais, and Helmus (2012) examined the predictive validity of Static-99R and Static-2002R among 1,588 individuals with a history of sexual crimes (Indigenous, $n = 319$; non-Indigenous, $n = 1{,}269$). Static-99R showed moderate accuracy in discriminating between recidivists and nonrecidivists for both groups (sexual recidivism; area under the curve [AUC] values of .71 for Indigenous and .74 for non-Indigenous individuals). Static-2002R, however, did not significantly predict sexual recidivism for Indigenous individuals with a history of sexual crimes (AUC of .61 vs. AUC of .76 for non-Indigenous individuals). This study did not evaluate the calibration of either Static-99R or Static-2002R (i.e., they did not consider whether the recidivism rates associated with each score were different for the different racial groups).

The Current Study

The current study provided an updated review of the predictive validity of both the Static-99R and Static-2002R for Indigenous and White individuals with a history of sexual crimes. Previous research has found that the recidivism rates associated with STATIC[1] scores are different for routine, relatively unselected samples compared with samples preselected to be high risk (Hanson, Thornton, Helmus, & Babchishin, 2016). Consequently, the samples in the current study only included individuals who were preselected to be above average risk for sexual recidivism. Another reason for focussing on high-risk samples is that they speak directly to the policy and practice concerns raised in Ewert.

Based on previous research, we expected Indigenous individuals to score higher on the total scores for both the Static-99R and Static-2002R, largely because of elevated scores on general antisociality. We also expected Indigenous individuals to score lower than White individuals on sex crime specific risk factors. Regarding predictive validity, we expected Static-99R to show moderate discrimination for the Indigenous peoples, but that discrimination would be lower than that observed for the White individuals. Given the similarity between Static-99R and Static-2002R, we expected similar results for both risk tools. Although previous studies with Indigenous peoples found that Static-2002R was less accurate than Static-99R, this finding lacks theoretical explanation and may be a sampling anomaly that could disappear with future research. We made no hypotheses regarding calibration analyses because no previous studies examined calibration of either tool for Indigenous peoples.

Method

Sample

The current study included an Indigenous study group ($n = 653$) and a White study group ($n = 1{,}560$) from five independent Canadian samples (Bonta & Yessine, 2005; Brankley, Lee, Babchishin, Hanson, & Harris, 2017; Haag, 2005; Lee, Mularczyk, Babchishin, Blais, & Bonta, 2018; Olver et al., 2018). The Indigenous study group was a mixed group of First Nations, Métis, and Inuit individuals. The White study group was individuals of European descent.

All individuals were adult males who had been convicted of sexual crimes and were classified by an administrative process as having a high level of risk and/or needs (e.g., referred to the National Flagging System, detained until the end of a federal sentence, or referred to a high-intensity treatment program). Although the samples were from all Canadian regions, most of the Indigenous study group were from the Prairies (see Supplemental Table S.1). The current samples partially overlap with the samples from Babchishin, Blais, et al. (2012; i.e., Bonta & Yessine [2005] and Haag [2005]).

Bonta and Yessine (2005). The original sample included three subgroups: (a) individuals flagged as potential Dangerous Offenders (considered for indeterminate sentences) by the National Flagging System, (b) individuals designated as Dangerous Offenders, and (c) individuals who committed a violent reoffence after being detained until their Warrant Expiry Date (WED). Only the first group (referred to the National Flagging System) was included in the current study because Static-99 scores were only available for this group. These individuals were released into the community between 1992 and 2004, and followed until April 2005. Sexual recidivism was defined as a conviction for a sexual offence (excluding prostitution offences, indecent phone calls, and possession of child pornography) after the release into the community. The sexual recidivism information was obtained from the Canadian Police Information Centre (CPIC).

Haag (2005). The original study sample included all male Canadian federal inmates serving a sentence for a sex offence whose WED was in 1995. For the current study, only individuals detained until WED and released into the community (about 25% of the total sample) were included. Sexual recidivism was defined

[1] STATIC refers to the following set of risk tools: Static-99, Static-99R, Static-2002, and Static-2002R.

as a conviction for a further sexual offence within seven years after release, and the sexual recidivism information was obtained from the CPIC.

Olver et al. (2018). This study included adult males who served federal sentences and attended sex crime specific treatment programs (e.g., Clearwater programs, National Sex Offender Program) between 1983 and 2008. Sexual recidivism was defined as a conviction for a sexually motivated offence (including child pornography offences). The recidivism information was obtained from the CPIC.

Brankley et al. (2017). This study was an extended follow-up to the Hanson and Harris (1998, 2000) study. Adult males who had served part of their sentences in the community (e.g., probation, parole) were selected from all Canadian federal and provincial correctional systems (except Prince Edward Island). All individuals had been convicted of a sexual offence involving physical contact with the victim(s). Sexual recidivism was defined as any subsequent offences that were considered sexually motivated (including offence, charge, or conviction) and their recidivism information was obtained from official (CPIC) and unofficial sources (news articles).

Lee et al. (2018). This study comprised individuals who had a history of sexual crimes from Blais and Bonta (2015), with updated recidivism information. Participants were placed into one of three mutually exclusive groups: (a) Flagged Offenders, (b) Dangerous Offenders, and (c) Long-Term Offenders. File information for all three groups was gathered from the National Flagging System coordinators across Canada. Sexual recidivism was defined as any subsequent offences that were considered sexually motivated (including offence, charge, or conviction) and their recidivism information was obtained from official (CPIC) and unofficial sources (news articles).

In accordance with the Tri-Council Policy Statement 2 (TCPS2, Article 9.17; Canadian Institutes of Health Research, 2010), we solicited experts' reviews and opinions from individuals of Indigenous heritage who were engaged professionally with the issue of Indigenous overrepresentation in Canada's criminal justice system (Leticia Gutierrez and Patti McDonald). Their contributions helped shape the interpretations and narratives of the research findings. This research project received research ethics approval from Carleton University's Research Ethics Boards (CUREB)-B.

Measures

Static-99R (Hanson & Thornton, 2000; Helmus, Thornton, et al., 2012). Static-99R is a 10-item empirical actuarial risk tool designed to assess the likelihood of sexual recidivism among adult males with a history of sexual offending. Static-99R is identical to Static-99 except for revised age weights. The total score (ranging from −3 to 12) is calculated by summing all items and can be used to place individuals in one of five risk categories: Level I – very low risk (scores of −3 to −2), Level II – below average risk (scores of −1 to 0), Level III – average risk (scores of 1 to 3), Level IVa – above average risk (scores of 4 to 5), and Level IVb – well above average risk (scores of 6 or higher; Hanson, Babchishin, Helmus, Thornton, & Phenix, 2017). When necessary, Static-99R scores were computed from Static-99 scores by using the individual's date of birth to calculate the updated age item.

Static-2002R (Hanson & Thornton, 2003; Helmus, Thornton, et al., 2012). Static-2002R is also an empirical actuarial risk assessment tool for an adult male with a history of sexual offending. The scale has 14 items grouped into five main subscales: Age at Release (one item), Persistence of Sex Offending (three items), Sexual Deviance (three items), Relationship to Victims (two items), and General Criminality (five items). The Static-2002R items are identical to the original Static-2002 except for updated age weights. The total score for Static-2002R can range from −2 to 13 and can be used to place individuals in one of five risk categories: Level I – very low risk (scores of −2 to −1), Level II – below average risk (scores of 0 to 1), Level III – average risk (scores of 2 to 4), Level IVa – above average risk (scores of 5 to 6), and Level IVb – well above average risk (scores of 7 or higher; Hanson et al., 2017). When necessary, Static-2002R scores were computed from Static-2002 scores by using the individual's date of birth to calculate the updated age item.

Psychological constructs of Static-99/R and Static-2002R. The items from Static-99R and Static-2002R were used as indicators of three psychological constructs that had been identified in previous research: (a) persistence/paraphilia, related to deviant sexual interests, but without the intent to harm their victims (e.g., pedophilia, voyeurism, and exhibitionism), (b) youthful stranger aggression, related to the intent to sexually harm the victim (e.g., sexual sadism, hostility toward women), and (c) general criminality, related to the antisocial trait (e.g., impulsivity, lack of remorse, rule violations; Brouillette-Alarie et al., 2016; Brouillette-Alarie et al., 2018).

Plan of Analysis

Aggregation of findings. AUC values and logistic regression parameters across the studies were aggregated using fixed-effect meta-analyses (Borenstein, Hedges, Higgins, & Rothstein, 2009; Hanson & Broom, 2005). Fixed-effect analyses have the advantage of providing an estimate of between-study variability (Q) that can be used to compare the variability across the studies (Borenstein et al., 2009). The Q statistics is distributed as a chi-square, with $k - 1$ degree of freedom (k = the number of studies).

Comparing the risk-relevant characteristics. The AUC statistic was used as an effect size statistic for comparing the frequencies of risk relevant characteristics between Indigenous and White study groups (Ruscio, 2008; Swets, Dawes, & Monahan, 2000). The AUC can vary between 0 and 1, with .50 indicating no difference between the groups. AUCs above .50 indicate that the Indigenous study group has higher levels of the risk relevant characteristic compared with the White study group. As a rough heuristic, an AUC of .56 corresponds to a small effect, .64 a moderate effect, and .71 a large effect (Rice & Harris, 2005). In contrast, AUCs below .50 indicate that the White study group has higher levels of the risk relevant characteristic compared with the Indigenous study group (.44 for a small effect, .33 for a moderate effect, and .29 for a large effect). An AUC value is statistically significant if the 95% confidence interval does not include .50.

When a risk factor was a binary variable, the odds ratio (instead of AUC) was used as the effect size indicator, with 0.5 added to each cell to stabilize the variance (Fleiss & Berlin, 2009). An odds

ratio is defined as $p/(1 - p)$, where p is the raw proportion of the sample with the characteristic. Odds ratios above 1 indicate that the Indigenous study group has higher levels of the risk relevant characteristic than to the White study group. No association is indicated when the 95% confidence interval of the odds ratio contains 1 (i.e., the odds are equal for both groups).

Predictive validity of Static-99R and Static-2002R. Assessing the predictive accuracy of a risk assessment instrument requires considering discrimination (how large are the differences in the scale's scores between those who reoffend and those who do not reoffend) as well as calibration (how well the estimated recidivism probability from the scale's norms corresponds with the observed recidivism probability of new samples). For discrimination, we used two statistical methods: (a) AUCs from receiver operating characteristic (ROC) analysis (Swets et al., 2000), and (b) odds ratios from logistic regression (Hosmer, Lemeshow, & Sturdivant, 2013). For calibration, we used a fixed-effect meta-analysis to compare logistic regression parameters (i.e., base rates; Borenstein et al., 2009; Hanson & Broom, 2005), as well as visual inspection of calibration plots.

Area under the curve. For prediction tools, AUC values can be interpreted as the probability that a randomly selected recidivist would have a more deviant score than a randomly selected nonrecidivist. To compare the unpaired AUC values, bootstrapping (n = 10,000) was used to obtain the standard error (SE) of the differences between AUC1 and AUC2 of the two (original) ROC curves with a *pROC* package for the statistical software R (Robin et al., 2011, 2014). The 95% confidence intervals of differences between AUC1 and AUC2 were calculated as follows:

$$95\%\ \mathrm{CI} = (AUC1 - AUC2) \pm 1.96 * SE$$

The difference of AUC values is statistically significant if the 95% confidence interval does not include 0.

Odds ratios. Odds ratios indicate the change in relative risk associated with one unit change in Static-99/R or Static-2002R scores. For example, Static-99R scores for the 5-year estimates are associated with a consistent relative risk increase of approximately 1.28 (Hanson et al., 2016) in the high-risk/high-needs samples. This indicates that the odds of recidivism increase 1.28 times as each Static-99R score increases. The primary advantage of odds ratios is that they are less influenced by a restriction of range than AUCs (Hanson, 2008).

Comparing logistic regression parameters. Calibration was examined by comparing the extent to which logistic regression parameters, specifically the intercept values, differed from the logistic regression parameters for the high-risk/high-needs sample norms (Hanson et al., 2016). *B0* coefficients were centered to represent the expected recidivism rate for individuals with a history of sexual crimes in the middle of the risk distribution. Specifically, the $B0_2$ represents the expected recidivism rate for a Static-99R score of 2 (p_2) in logit units ($\ln[p_2/\{1 - p_2\}]$) and the $B0_3$ represents the expected recidivism rate for a Static-2002R score of 3 (p_3) in logit units ($\ln[p_3/\{1 - p_3\}]$).

In addition, the recidivism rates in the current study were compared with the norms asserted by the test developers for high-risk/high-needs samples (Hanson et al., 2016; Phenix, Helmus, & Hanson, 2016), using the Q statistic from a fixed-effect meta-analysis (Borenstein et al., 2009; Hanson & Broom, 2005).

The differences between the parameters of the Indigenous and White study groups were calculated for each parameter by subtracting the b_i of the White samples from b_i of the Indigenous samples. The variance for the b_i *difference* was calculated according to the standard formula for the variance of a difference score (e.g., Ley, 1972):

$$Var\ (b_i\ diff) = s^2{}_x + s^2y - 2r_{xy}s_xs_y$$

where s_x is the standard deviation of the d_i from the aggregated Indigenous study group, s_x is the standard deviation of the d_i from the aggregated White study group, and r_{xy} is the correlation coefficient between the effect sizes in each sample for Indigenous and White individuals.

Results

The Indigenous study group was, on average, approximately 4 years younger than the White study group (36.1 vs. 40.6 years; AUC = .61, 95% CI [.57, .65]). The average Static-99R and Static-2002R total scores for both groups were in the Above Average range (Risk Level IVa, see Table 1), and were approximately a half-point higher for the Indigenous study group than the White study group (4.8 vs. 4.3 for Static-99R, AUC = .57, 95% CI [.53, .60]; 6.1 vs. 5.7 for Static-2002R, AUC = .55, 95% CI [.51, .59]).

Comparing Risk Relevant Characteristics

The Indigenous study group showed less persistence/paraphilia than the White study group (AUC = .39, 95% CI [.34, .43]). Compared with the White study group, the Indigenous study group

Table 1
Age at Release, Static-99R Scores, and Static-2002R Scores for Indigenous and White Study Groups

	Indigenous				White			
Study	n	Age at release M (SD)	Static-99R M (SD)	Static-2002R M (SD)	n	Age at release M (SD)	Static-99R M (SD)	Static-2002R M (SD)
Bonta & Yessine (2005)	37	36.8 (7.8)	5.1 (2.1)	—	118	40.4 (10.1)	5.0 (2.2)	—
Haag (2005)	50	33.7 (8.7)	4.2 (1.9)	5.8 (1.9)	141	38.6 (10.1)	3.9 (2.4)	5.5 (2.5)
Olver et al. (2018)	425	—	4.6 (2.4)	—	739	—	3.2 (2.9)	—
Brankley, Lee, Babchishin, Hanson, and Harris (2017)	36	35.5 (8.4)	4.7 (2.0)	6.3 (2.1)	346	40.4 (11.6)	4.2 (2.4)	5.9 (2.5)
Lee, Hanson, Calkins, and Jeglic (2018)	105	38.3 (9.2)	5.4 (2.1)	6.4 (2.1)	216	43.9 (12.3)	4.8 (2.7)	5.8 (2.8)
Average (fixed-effect)	653	36.1 (4.2)	4.8 (0.9)	6.1 (1.2)	1,560	40.6 (5.5)	4.3 (1.1)	5.7 (1.5)

was significantly less likely to have a history of noncontact sex offences (odds ratio = 0.45), male victims (odds ratio = 0.22), unrelated young victims (odds ratio = 0.27), and prior sex offences (AUC = 0.45; see Supplemental Table S.2).

In contrast, the Indigenous study group scored higher than the White study group on the items addressing youthful stranger aggression (AUC = .60, 95% CI [.55, .64]). The difference was attributable to two items (age at release and index nonsexual violence). Specifically, the Indigenous study group was significantly younger than the White study group (AUC of .59) and more likely to have been convicted of a violent index offence (odds ratio = 1.61; See Supplemental Table S.2).

As expected, the Indigenous study group scored higher on general criminality than the White study group (AUC = .72, 95% CI [.68, .76]). All seven items associated with general criminality were significantly higher for the Indigenous study group than the White study group (e.g., history of community supervision violation, prior nonsexual violent crimes, and any prior criminal involvement; see Supplemental Table S.2).

Predictive Accuracy of Three Psychological Structures and Items

Persistence/paraphilia predicted sexual recidivism for the White study group (AUC of .65), but not for the Indigenous study group (AUC of .58). The difference between these AUC values was, however, not statistically significant (see Table 2). Only two items (prior sex offences and unrelated young victim) significantly predicted sexual recidivism rates for the Indigenous study group in the expected direction. For one item (noncontact sex offences), the direction of the effect was reversed for the Indigenous study group (i.e., Indigenous individuals with a history of noncontact sex offences were significantly less likely to commit another sexual crime in the future than those without; Table 2).

Youthful stranger aggression predicted sexual recidivism for the White study group (AUC of .64), but not for the Indigenous study group (AUC of .53), which was significantly lower than that of the White study group. Specifically, the item of age at release for the Indigenous study group did not predict sexual recidivism and was significantly lower than the value for the White study group (AUCs of .51 vs. .62). In other words, sexual recidivism risk for the Indigenous study group was unrelated to age at release. Only one item (index nonsexual violence) predicted sexual recidivism for the Indigenous study group (AUC of .58) whereas it did not predict for the White study group (AUC of .51). The difference was, however, not statistically significant (see Table 2).

Overall, general criminality significantly and similarly predicted sexual recidivism for both the Indigenous and White study groups (AUC of .59 and 61). The only item that was significantly less predictive for the Indigenous study group than the White study group was any prior criminal involvement (AUC of .50 vs. .55, see Table 2).

Table 2
Predictive Accuracy of Static-99R and Static-2002R Items for Sexual Recidivism for Indigenous and White Study Groups

Items	Indigenous AUC	Indigenous *n/N*	White AUC	White *n/N*	Diff.	95% CI
Persistence/Paraphilia						
Prior sex offences (Static-99R; Item #5)	**.608 [.528, .688]**	59/219	**.626 [.582, .670]**	187/759	−.018	[−.108, .072]
Prior sentencing occasions for sexual offences (Static-2002R; Item #2)	.571 [.482, .660]	50/182	**.628 [.582, .674]**	171/641	−.057	[−.159, .045]
Rate of sexual offending (Static-2002R; Item #4)	.573 [.496, .650]	50/182	**.606 [.564, .647]**	171/641	−.033	[−.120, .054]
Noncontact sex offences (Static-99R; Item #7)	**.474 [.448, .499]**	59/219	.522 [.494, .551]	187/759	**−.048**	**[−.086, −.010]**
Male victim (Static-99R; Item #10)	.495 [.449, .541]	59/219	.538 [.498, .579]	187/759	−.043	[−.103, .017]
Two+ victims < 12 year-old, one unrelated (Static-2002R;Item #7)	**.572 [.502, .641]**	50/182	**.573 [.530, .617]**	171/641	−.001	[−.050, .048]
Persistence/Paraphilia	.581 [.489, .673]	50/182	**.645 [.597, .693]**	171/641	−.127	[−.333, .079]
Youthful stranger aggression						
Age at release (Static-99R; Item #1)	.514 [.436, .592]	59/219	**.616 [.574, .659]**	186/757	**−.102**	**[−.189, −.015]**
Juvenile sexual arrest prior to age 18 (Static-2002R; Item #3)	.502 [.458, .546]	50/182	**.534 [.507, .561]**	171/641	−.032	[−.084, .020]
Ever lived with a lover for +2 years (Static-99R; Item #2)	.470 [.396, .543]	59/218	.531 [.490, .571]	186/753	−.061	[−.143, .021]
Index nonsexual violence – any convictions (Static-99R; Item #3)	**.580 [.506, .654]**	59/218	.512 [.476, .549]	187/759	.068	[−.013, .149]
Unrelated victim (Static-99R; Item #8)	.516 [.466, .566]	59/219	**.537 [.516, .558]**	187/759	−.021	[−.074, .032]
Stranger victim (Static-99R; Item #9)	.513 [.438, .588]	59/219	**.577 [.536, .618]**	187/758	−.064	[−.148, .020]
Youthful stranger aggression	.529 [.439, .620]	50/182	**.644 [.597, .690]**	170/636	**−.115**	**[−.217, −.013]**
General criminality						
Any community supervision violation (Static-2002R; Item #12)	**.572 [.542, .602]**	50/182	**.616 [.582, .649]**	171/641	−.044	[−.089, .001]
Years free prior to index sex offence (Static-2002R; Item #13)	.517 [.459, .576]	50/182	**.579 [.539, .619]**	171/641	−.062	[−.133, .009]
Prior criminal involvement (Static-2002R; Item #10)	.500 [.461, .539]	50/182	**.548 [.522, .573]**	171/641	**−.048**	**[−.095, −.001]**
Prior sentencing occasions for anything (Static-2002R; Item #11)	.578 [.499, .578]	50/182	**.592 [.549, .635]**	171/641	−.014	[−.105, .077]
Prior sentencing dates (Static-99R; Item #6)	**.568 [.509, .626]**	59/219	**.566 [.525, .607]**	187/759	.002	[−.096, .100]
Prior nonsexual violence – any convictions (Static-99R; Item #4)	**.563 [.501, .625]**	59/219	**.541 [.500, .583]**	187/758	.022	[−.051, .095]
Prior nonsexual violence sentencing occasion (Static-2002R; Item #14)	**.545 [.475, .616]**	50/182	.515 [.472, .558]	171/641	.030	[−.051, .111]
General criminality	**.589 [.501, .677]**	50/182	**.613 [.568, .658]**	171/641	−.024	[−.122, .074]

Note. Numbers in bold indicate statistical significance (i.e., $p < .05$).

Table 3
Sexual Recidivism Rates and AUC Values for Static-99R for Indigenous and White Study Groups

	Indigenous			White		
	Sexual recidivism			Sexual recidivism		
Study	%	*n/N*	AUC	%	*n/N*	AUC
Fixed 5-year						
Bonta & Yessine (2005)	19.2	5/26	.60 [.25, .96]	14.4	13/90	**.71 [.58, .85]**
Haag (2005)	28.0	14/50	.60 [.42, .77]	17.7	25/141	**.70 [.59, .81]**
Olver et al. (2018)	14.3	57/400	**.62 [.55, .70]**	9.6	64/665	**.71 [.65, .77]**
Brankley et al. (2017)	11.4	4/35	.62 [.27, .97]	11.4	35/308	**.70 [.61, .80]**
Lee et al. (2018)	19.3	17/88	.56 [.41, .72]	15.8	27/171	**.69 [.59, .79]**
Average (fixed-effect)	16.2	97/599	**.61 [.55, .67]**^{ns}	11.9	164/1,375	**.70 [.66, .74]**^{ns}
Fixed 10-year						
Olver et al. (2018)	23.4	64/274	**.67 [.59, .74]**	16.7	72/432	**.71 [.65, .77]**
Brankley et al. (2017)	20.0	7/35	.60 [.34, .87]	21.7	65/299	**.74 [.67, .80]**
Lee et al. (2018)	28.8	19/66	**.66 [.51, .81]**	23.1	27/117	**.70 [.59, .82]**
Average (fixed-effect)	24.0	90/375	**.66 [.60, .72]**^{ns}	19.3	164/848	**.72 [.68, .76]**^{ns}

Note. *ns* indicates no significant variability across studies (*Q* statistics). Numbers in bold indicate statistical significance (i.e., $p < .05$).

Predictive Validity Static-99R and Static-2002R

Overall, sexual recidivism rates for the Indigenous study group were higher than the rates of the White study group (see Table 3). For the Indigenous study group, 16.2% sexually reoffended within a fixed 5-year follow-up period and 24.0% sexually reoffended within a fixed 10-year follow-up period. For the White study group, 11.9% sexually reoffended within a fixed 5-year follow-up period and 19.3% sexually reoffended within a fixed 10-year follow-up period.

Static-99R

Discrimination. Using the fixed 5-year follow-up, Static-99R was able to discriminate recidivists from nonrecidivists for both Indigenous and White study groups. Whereas the White study group had a moderate AUC value of .70, the AUC value for the Indigenous study group was smaller (AUC = .61). With the fixed 10-year follow-up, the AUC value of the Indigenous study group increased to a moderate effect size (AUC of .66) whereas the AUC value of the White study group at 10 years was .72 (see Table 3).

Before comparing predictive accuracies across racial groups, we first examined the degree of consistency within each racial group. For the Indigenous study group, both the 5-year (five studies) and 10-year sexual recidivism rates (three studies) at a Static-99R score of 2 ($B0_2$) were very similar (both $p > .50$, see Table S.3). The relative risk ($B1$) of Static-99R across the studies was also very similar for both the fixed 5-year follow-up (five studies) as well as the fixed 10-year follow-up (three studies; both $p > .70$, see Table S.3).

For the White study group, both the 5-year (five studies) and 10-year sexual recidivism rates (three studies) at a Static-99R score of 2 ($B0_2$) across the studies were very similar (both $p > .60$, see Table S.4). The relative risk ($B1$) of Static-99R across the studies was also very similar with a fixed 5-year follow-up (five studies) as well as with a fixed 10-year follow-up (three studies; both $p > .60$; Table S.4).

For the Indigenous study group, the 5-year sexual recidivism rates at a Static-99R score of 2 were significantly higher than that of the White study group (10.2% vs. 5.7%; Table 4; Figure 1). The 10-year sexual recidivism rates at a Static-99R score of 2 for the Indigenous study group were, however, similar to that of the White study group (11.9% vs. 9.7%; Table 4; Figure 2). The discrimination (change in relative risk) of Static-99R total scores for the White study group and Indigenous study group was equivalent with both 5-year ($e^{[.190]} = 1.21$ vs. $e^{[.306]} = 1.36$) as well as 10-year follow-up periods ($e^{[.271]} = 1.31$ vs. $e^{[.346]} = 1.41$; Table 4 and Figures 1 and 2).

Calibration. The sexual recidivism rates of the Indigenous study group were similar to the high-risk/high-needs norms for Static-99R for both 5-year and 10-year follow-up periods. Both the base rate (score of 2) and the discrimination parameter were not significantly different from the norms asserted for this measure (Hanson et al., 2016, see Table 4; Figure 2).

The base rate for the White study group, however, was significantly lower than the norms for both 5-year (11.3% vs. 5.7%) and 10-year (19.2% vs. 9.7%) follow-up periods (see Table 4). The discrimination parameter for the White study group was not significantly different from that expected from the norms for either follow-up period (see Table 4 and Figure 2).

Static-2002R

The sample size of the Indigenous study group was reduced to 173 individuals (3 studies) with a 5-year follow-up period and 101 individuals (2 studies) with a 10-year follow-up period.

Discrimination. Static-2002R discriminated recidivists from nonrecidivists for the White study group at both 5-year (AUC = .69) and 10-year (AUC = .72) follow-up periods. In contrast, the AUC values were not statistically significant at either follow-up period for the Indigenous study group (AUC of .56 at 5 years, .61 at 10 years; see Table 5).

Using logistic regression, the base rates and relative risk parameters for Static-2002R were reasonably consistent across studies within each racial group (all $p > .20$; Table S.5 and Table S.6).

Table 4
Comparison of Logistic Regression Parameters for Static-99R Predicting Sexual Recidivism With Norms (From Hanson et al., 2016)

Measure	HRHN norms (meta-average)	Indigenous	White	$B_{\text{Indigenous-White}}$	95% CI
Fixed 5-year					
Base rate					
B0$_2$ (*SD*)	−2.06 (.153) (11.3%)	−2.18 (.212) (10.2%)	−2.80 (.151) (5.7%)	**.623**	**[.179, 1.066]**
Q_Δ (*df* = 1)		.18	**11.60*****		
Relative risk					
B1 (*SD*)	.250 (.042)	.190 (.055)	.306 (.038)	−.116	[−.245, .013]
Q_Δ (*df* = 1)		.755	.971		
Fixed 10-year					
Base rate					
B0$_2$ (*SD*)	−1.44 (.186) (19.2%)	−2.00 (.246) (11.9%)	−2.23 (.164) (9.7%)	.232	[−.307, .770]
Q_Δ (*df* = 1)		3.64	**9.99****		
Relative risk					
B1 (*SD*)	.231 (.056)	.271 (.063)	.346 (.042)	−.075	[−.222, .071]
Q_Δ (*df* = 1)		.28	2.72		

Note. $r_{(\text{between Indigenous and White})}$ = .260 for *B0*$_2$ and .208 for *B1*. Numbers in bold indicate statistical significance (i.e., $p < .05$).
$^{**}p < .01$. $^{***}p < .001$.

The summary statistics for Static-2002R, however, were unstable because of the small number of studies in the meta-analysis. For example, the average base rate (*B0*$_3$) for the fixed 5-year follow-up for the Indigenous study group (16.4%) was higher than for the fixed 10-year follow-up (14.3%). This is only possible because different studies were included in each analysis.

The relative risk parameter (*B1*) of Static-2002R was statistically significant (greater than zero) for the White study group for both follow-up periods (Table S.6) but was not significant for either follow-up periods for the Indigenous study group (i.e., poor discrimination; Table S.5). The discrimination parameter was higher for the White study group than the Indigenous study group for the 5 year follow-up period (odds ratios = 1.10 vs. 1.34; Table 6 and Figure 3).

Calibration. The 5-year sexual recidivism rates for the Indigenous study group were significantly higher than for the White study group (16.4% vs. 6.0%; Table 6 and Figure 3).

For the Indigenous study group, neither the base rate nor the discrimination parameter was statistically different from the Static-2002R high-risk/high-needs norms; nevertheless, visual inspection of the calibration plot suggested meaningful discrepancies between expected and observed recidivism rates, particularly at the highest risk levels (i.e., overestimation; see Figure 3).

The sexual recidivism base rate for the White study group was significantly lower than the Static-2002R high-risk/high-needs norms (6.0% vs. 13.2%). The discrimination parameter, however, was not significantly different from the expected value (odds ratios $e^{[.218]} = 1.24$ [norms] vs. $e^{[.289]} = 1.34$ [White individuals]; Table 6 and Figure 3).

Discussion

The Ewert decision (2018) has renewed the call for research on the extent to which widely utilized risk assessment tools, such as

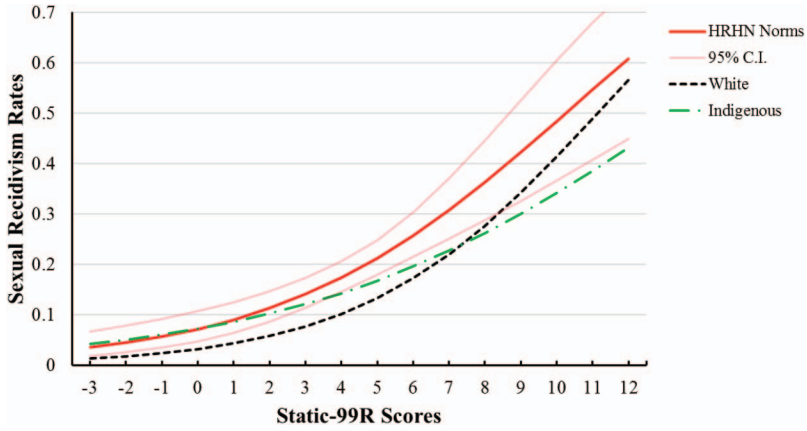

Figure 1. Estimated sexual recidivism rates according to Static-99R scores for Indigenous and White study groups with the High Risk/High Need Norms with a fixed 5-year follow-up. See the online article for the color version of this figure.

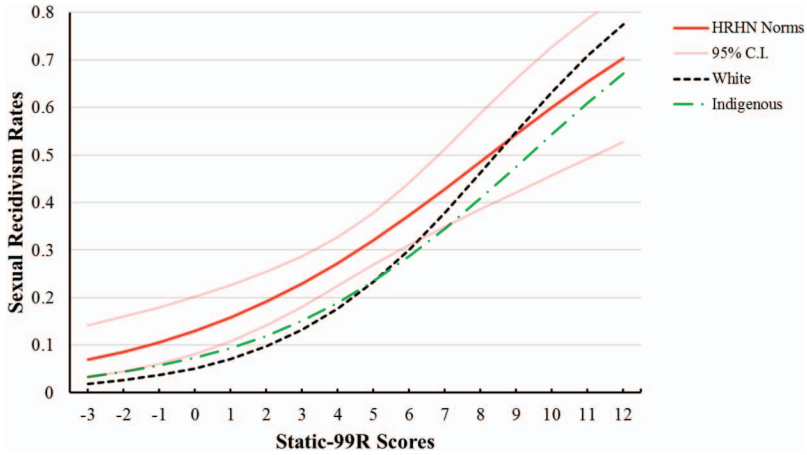

Figure 2. Estimated sexual recidivism rates according to Static-99R scores for Indigenous and White study groups with the High Risk/High Need Norms with a fixed 10-year follow-up. See the online article for the color version of this figure.

the Static-99R and Static-2002R, are valid for use with individuals of Indigenous heritage in the criminal justice system. The current study found that the Static-99R sexual recidivism risk tool showed similar predictive accuracy for both White and Indigenous study groups. The Static-2002R, however, did not.

Differences in discrimination between Static-99R and Static-2002R for Indigenous peoples was surprising given that the two scales contain many of the same items and are highly correlated (*r*s ranging from .70 to .92; Babchishin, Hanson, & Helmus, 2012). These findings may reflect chance variation because the sample size for the Static-2002R studies was small (e.g., the 10 year estimates were based on only two samples with only 101 Indigenous individuals, of which 26 were sexual recidivists). Babchishin, Blais, and Helmus's (2012) similar finding of relatively low discrimination for the Static-2002R compared with the Static-99R cannot be interpreted as an independent replication because of the substantial overlap in the samples included in both meta-analyses. Nevertheless, the current results support extreme caution when using Static-2002R with individuals of Indigenous heritage.

We also identified racial differences in the three major psychological constructs underlying the Static-99R and Static-2002R risk tools (Brouillette-Alarie et al., 2016). The Indigenous study group had significantly fewer indicators of persistence/paraphilia (e.g., less prior sex offences, noncontact offences, male victims, and young victims), and significantly more indicators of youthful stranger aggression (e.g., younger at release, more nonsexual violence at index) and general criminality (e.g., community supervision violations, prior criminal involvement, prior nonsexual violence, etc.) compared with the White study group. This pattern of results is consistent with previous research finding that individuals of Indigenous heritage in the criminal justice system are assessed as having more risk factors for general crime than non-Indigenous individuals (e.g., Babchishin, Blais, et al., 2012; Helmus, Babchishin, et al., 2012; Olver et al., 2018; Perley-Robertson et al., 2018; Shepherd et al., 2014).

As expected, all three factors were associated with sexual recidivism for the White study group; only the general criminality factor was associated with sexual recidivism for the Indigenous

Table 5
Sexual Recidivism Rates and AUC Values for Static-2002R for Indigenous and White Study Groups

	Indigenous			White		
	Sexual recidivism			Sexual recidivism		
Study	%	*n/N*	AUC	%	*n/N*	AUC
Fixed 5-year						
Haag (2005)	28.0	14/50	.56 [.36, .75]	17.7	25/141	**.69 [.59, .80]**
Brankley et al. (2017)	11.4	4/35	.57 [.27, .87]	11.4	35/308	**.70 [.61, .79]**
Lee et al. (2018)	19.3	17/88	.56 [.42, .70]	15.8	27/171	**.69 [.59, .79]**
Average (fixed-effect)	20.2	35/173	.56 [.45, .67]^{ns}	14.0	87/620	**.69 [.64, .75]**^{ns}
Fixed 10-year						
Brankley et al. (2017)	20.0	7/35	.64 [.40, .88]	21.7	65/299	**.74 [.67, .80]**
Lee et al. (2018)	28.8	19/66	.60 [.45, .75]	23.1	27/117	**.69 [.58, .81]**
Average (fixed-effect)	25.7	26/101	.61 [.49, .74]^{ns}	22.1	92/416	**.72 [.67, .78]**^{ns}

Note. *ns* indicates no significant variability across studies (*Q* statistics). Numbers in bold indicate statistical significance (i.e., $p < .05$).

Table 6
Comparison of Logistic Regression Parameters for Static-2002R Predicting Sexual Recidivism With Meta-Average (From Hanson et al., 2016)

Measure	Meta-average (HRHN Norms)	Indigenous	White	$B_{\text{Indigenous-White}}$	95% CI
Fixed 5-year					
Base rate					
B0₃ (SD)	−1.88 (.153) (13.2%)	−1.63 (.386) (16.4%)	−2.76 (.239) (6.0%)	**1.135**	**[.402, 1.868]**
Q_Δ (*df* = 1)		.36	**9.75****		
Relative risk					
B1 (SD)	.218 (.042)	.099 (.097)	.289 (.051)	**−1.190**	**[−.379, −.001]**
Q_Δ (*df* = 1)		1.24	1.15		
Fixed 10-year					
Base rate					
B0₃ (SD)	—	−1.78 (.509) (14.3%)	−2.43 (.257) (8.1%)	.639	[−.303, 1.581]
Q_Δ (*df* = 1)		—	—		
Relative risk					
B1 (SD)	—	.208 (.117)	.338 (.056)	−.130	[−.356, .095]
Q_Δ (*df* = 1)		—	—		

Note. $r_{(\text{between Indigenous and White})}$ = .359 for $B0_2$ and .280 for *B1*. Numbers in bold indicate statistical significance (i.e., $p < .05$).
** $p < .01$.

study group. For the persistence/paraphilia construct, the difference in discrimination could be attributed to two items: (a) noncontact sex offences and (b) any male victim. For White individuals, those items are well-established indicators of general sexual deviancy and paraphilic interests (e.g., pedophilia, exhibitionism, voyeurism). Although all explanations are highly speculative, it may be that exhibitionism, which is the strongest driver of the increased risk for noncontact offending among White individuals, is a culturally specific disorder of White males (like serial killing). Noncontact offending was rare among the Indigenous individuals in the current study, and it is similarly rare for African Americans (Lee, Hanson, Calkins, & Jeglic, 2018).

For the Indigenous study group, the lack of relationship between sexual recidivism and age was unexpected. Age (a key element of the youthful stranger aggression factor) is one of the most robust predictors of criminal recidivism, including sexual recidivism (Barbaree & Blanchard, 2008; Hanson, 2002). Although there are no data on the topic yet, it is possible that some aspects of Indigenous society result in different life course trajectories than those commonly observed for non-Indigenous individuals in conflict with the law. Consequently, sexual offending in middle age may mean something different from youthful offending for Indigenous individuals compared with White individuals.

Overall, the differences in the presence and predictive validity of risk relevant characteristics between Indigenous and White peoples suggests that there may be distinctive offence patterns or pathways that lead to sexual offending for Indigenous peoples. For example, Brouillette-Alarie and Proulx's (2018) etiological model (developed for non-Indigenous peoples in contact with the criminal justice system) includes two broad paths to sexual offending, one characterised by social isolation and sexual deviance (e.g., sexual interest in children) and the other characterised by antisociality (e.g., impulsivity), physical and psychological victimization, and sexual promiscuity. Whereas the first pathway is primarily related to sexual recidivism, especially sexual crimes against children, the second pathway is related to nonsexual violence and

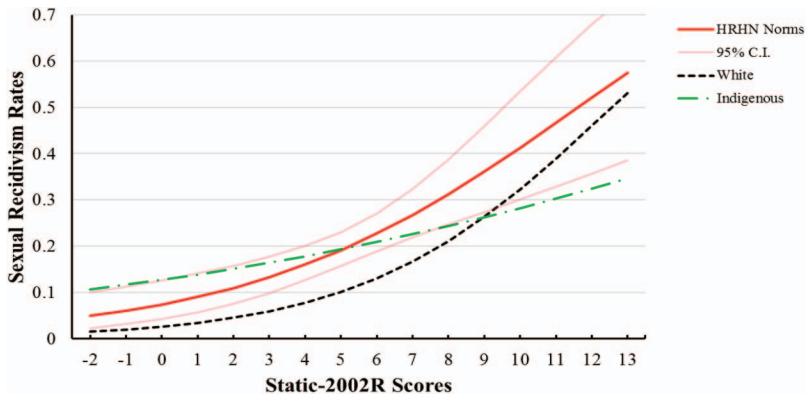

Figure 3. Estimated sexual recidivism rates according to Static-2002R scores for Indigenous and White study groups with the High Risk/High Need Norms with a fixed 5-year follow-up. See the online article for the color version of this figure.

sexual offending against adult women. In the current study, Indigenous individuals with a history of sexual offending more closely resembled the antisocial pathway; however, the pattern of results was not a clear fit to either pathway, indicating the potential for other distinctive pathways to sexual offending for Indigenous peoples that have yet to be identified.

In terms of broader implications, the current findings raise questions about how much similarity is needed among risk assessment tools to justify their use for different racial or ethnic groups. Currently, the scientific and professional communities do not have clear conventions or standards to evaluate the extent to which observed cultural variation in assessment tools is tolerable. One approach to setting a limit for cultural bias is to estimate the likelihood that the observed differences are primarily due to chance variation. From this perspective, the a priori expectation is that there are no real cultural differences and that any observed differences are the result of sampling (random or biased) from the same population. For many individuals concerned with Indigenous overrepresentation, the null hypothesis is a very questionable assumption. Instead, some degree of racial differences would be expected, which, from a Bayesian perspective, would lower the amount of evidence required to conclude that real differences were present. The primary problem with relying solely on statistically significant differences, however, is that it is possible to detect differences that are too small to matter.

Another approach to evaluating cultural bias is whether the evidence supports the same inferences for the same test score. From this perspective, the primary concern is the consistency of the decisions that would follow from the test scores. If there are real cultural differences in test performance, these differences are tolerable if the decision maker would nonetheless be justified in making the same decision regardless of the race/ethnicity of the case at hand. For example, parole would be equally denied to a White inmate whose risk was estimated at 30% and to an Indigenous inmate whose risk was estimated at 20% if the acceptable risk threshold was five recidivists of 100 parole releases. In contrast, the cultural bias in the test result would be problematic if the decision threshold was 25 of 100.

The above example suggests that it would be difficult to have a single standard for tolerable cultural bias: the amount and nature of tolerable bias would depend on the decisions that are required. Consequently, evaluators need to consider carefully the extent to which the specific inferences they make from a test score are justified based on the available evidence, and whether any observed racial/cultural differences would influence the decision-at-hand.

For the Static-99R and Static-2002R, the tools' developers provide several different types of information associated with test scores (Hanson et al., 2017; Phenix et al., 2016). The meaning of this information has been organized into standardized risk levels (i.e., Level I, II, III, IVa, and IVb), which are intended to convey quantitative information (percentile ranks, risk ratios, recidivism rates) as well as psychological characteristics (e.g., severity of criminogenic needs, developmental history, treatment needs). The current study focused only on certain quantitative information provided by these risk tools (i.e., risk ratios, recidivism rates). The validity of these risk tools for assessing other characteristics across ethnic groups remains unexplored.

Limitations

The current study combined First Nations, Métis, and Inuit people into a single group. Each of these peoples has their own history, language(s) and culture(s), and may differ on risk-relevant characteristics. Furthermore, individuals living off-reserve might show more similar risk relevant characteristics with White individuals than those living on reserve (e.g., 66% of the First Nations and almost all Métis and Inuit are living off reserve; Statistics Canada, 2018). Those variables were, however, not available in the current study.

In addition, all individuals with a history of sexual crimes in the current study were already identified as high-risk (e.g., from federal prisons, National Flagging System, sex offender treatment programs); consequently, the findings may not generalise to more routine/complete samples of Indigenous and White individuals.

Finally, despite having larger samples of Indigenous individuals than previous STATIC studies, the numbers were still well below the sample size required for stable logistic regression estimates (i.e., 100 recidivists and 100 nonrecidivists; Vergouwe, Steyerberg, Eijkemans, & Habbema, 2005).

Implications for Policy and Practice

Static-99R was one of the impugned measures identified in *Ewert v. Canada* (2015, 2018). Consistent with previous findings (e.g., Babchishin, Blais, et al., 2012), the current study supports the use of Static-99R for Indigenous peoples in the criminal justice system. The available research does not support the use of Static-2002R; however, the findings for Static-2002R were based on a small sample, and the negative conclusion could change as research accumulates.

Given that Indigenous individuals with a history of sexual crimes scored high on general criminality (e.g., criminal history, nonsexual violence) and low on sexual criminality (e.g., paraphilia), rehabilitation programs for Indigenous peoples who commit sexual crimes may benefit from an increased focus on general criminality (e.g., antisocial attitudes and behaviours). In addition, treatment programming will likely be most effective when it takes into consideration the cultural values or norms of Indigenous peoples (e.g., spirituality) as well as sociodemographic characteristics that influence their response to treatment (e.g., low education and socioeconomic status, systemic oppression, distrust of criminal justice system; Responsivity-principle; Gutierrez, Chadwick, & Wanamaker, 2018).

Conclusion

In *Ewert v. Canada* (2018), the Supreme Court of Canada ruled that CSC had not done its due diligence in verifying that commonly used risk assessment tools were valid for use for Indigenous peoples. Because of the rules of evidence, the Supreme Court justices hearing the Ewert case could only consider the scientific research deemed credible by Judge Phelan in the original, 2015 Federal Court hearing. The only expert relied on in that case, Dr. Stephen Hart, asserted in his 2012 report that the validity of the risk tools for Indigenous peoples was unknown because of insufficient research on the topic. Consequently, Chief Justice McLachlin and colleagues were unaware of any subsequent research,

including the 2012 meta-analysis of Babchishin and colleagues on the predictive accuracy of the Static-99R and Static-2002R for Indigenous peoples.

Would the courts have made a different decision if they had had access to the results of the current study? The answer is far from certain. As previously stated, we lack clear criteria by which to determine whether any observed amount of cultural bias is tolerable. In the Ewert case, we side with the minority opinion of Judges Rowe and Côté who argued that it is not the place of the courts to determine which risk tools psychologists should use. Decisions concerning the scientific credentials of psychological tests or criminological assessment tools are rightly made by professional communities of practice.

Based on the accumulated evidence and the current results, we believe that there is sufficient research to justify the use of Static-99R, but not Static-2002R, for assessing the likelihood of sexual recidivism among individuals with Indigenous heritage. Our position accepts that some cultural bias in Static-99R is likely, but interprets the available evidence as indicating that Static-99R scores will most often support the same general inferences concerning relative risk and absolute sexual recidivism rates regardless of whether the individual being assessed is White or Indigenous. Nevertheless, given that risk assessment tools are used at every point of contact with the criminal justice system (e.g., policing, sentencing, rehabilitation programming, parole), continued efforts to validate risk assessment scales with Indigenous peoples remains a research priority.

Résumé

Au Canada, la question de la subjectivité culturelle des outils d'évaluation du risque est une préoccupation constante, particulièrement pour ce qui est des peuples autochtones, qui sont surreprésentés dans le système de justice pénale. À ce jour, relativement peu de travaux de recherche ont été menés sur les outils d'évaluation du risque s'adressant particulièrement aux peuples autochtones. Cette étude s'est penchée sur l'exactitude prédictive de deux outils d'évaluation du risque de récidive des délinquants sexuels (statique-99R, statique-2002R) : l'un visant des individus blancs ($n = 1560$), l'autre des personnes de patrimoine autochtone ($n = 653$) prélevées de cinq échantillons indépendants canadiens. Les résultats indiquent que l'outil statique-99R prévoyait la récidive sexuelle avec à peu près la même précision pour les individus blancs et les individus autochtones. L'échelle statique-2002R prédisait le risque de récidive des délinquants sexuels pour les individus blancs seulement. Les indicateurs de criminalité générale (p. ex. le nombre total d'infractions antérieures) prédisaient la récidive des délinquants sexuels chez les deux groupes raciaux. En revanche, bon nombre des facteurs de risque propres aux crimes sexuels dont la validité a été démontrée dans d'autres échantillons composés principalement d'individus blancs (p. ex. les victimes de sexe masculin, les délits sexuels sans contact) prédisaient le risque de récidive sexuelle uniquement pour le groupe d'étude des individus blancs, et non pour celui des individus autochtones. Les individus blancs présentaient plus d'indicateurs de perversions sexuelles, tandis que les individus autochtones de ces échantillons affichaient un taux de criminalité générale plus élevé. Ces constatations donnent à penser qu'il serait bénéfique, pour le traitement des personnes à risque de délinquance sexuelle, d'accorder une plus grande attention aux caractéristiques d'importance sur le plan des risques en lien avec les individus blancs et les individus autochtones aux prises avec le système de justice pénale.

Mots-clés : subjectivité culturelle, peuples autochtones, délinquance sexuelle, Statique-99R, Statique-2002R.

References

Ægisdóttir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., Cook, R. S., . . . Rush, J. D. (2006). The meta-analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *The Counseling Psychologist, 34,* 341–382. http://dx.doi.org/10.1177/0011000005285875

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

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

Babchishin, K. M., Blais, J., & Helmus, L. (2012). Do static risk factors predict differently for Aboriginal sex offenders? A multi-site comparison using the original and revised Static-99 and Static-2002 scales. *Canadian Journal of Criminology and Criminal Justice, 54,* 1–43. http://dx.doi.org/10.3138/cjccj.2010.E.40

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

Barbaree, H. E., & Blanchard, R. (2008). Sexual deviance over the lifespan: Reduction in deviant sexual behavior in the aging sex offender. In D. R. Laws & W. T. O'Donohue (Eds.), *Sexual deviance: Theory, assessment, and treatment* (pp. 37–60). New York, USA: Guilford Press.

Berry, J. W. (1969). On cross-cultural comparability. *International Journal of Psychology, 4,* 119–128. http://dx.doi.org/10.1080/00207596908247261

Blais, J., & Bonta, J. (2015). Tracking and managing high risk offenders: A Canadian initiative. *Law and Human Behavior, 39,* 253–265. http://dx.doi.org/10.1037/lhb0000109

Bonta, J., LaPrairie, C., & Wallace-Capretta, S. (1997). Risk prediction and re-offending: Aboriginal and non-aboriginal offenders. *Canadian Journal of Criminology, 39,* 127–144.

Bonta, J., & Yessine, A. K. (2005). [Recidivism data for 155 released sexual offenders from the offenders identified in The National Flagging System: Identifying and responding to high risk, violent offenders (User Report 2005–04). Ottawa, ON: Public Safety and Emergency Preparedness Canada]. Unpublished raw data.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. West Sussex, UK: Wiley, Ltd. http://dx.doi.org/10.1002/9780470743386

Bourgon, G., Mugford, R., Hanson, R. K., & Coligado, M. (2018). Offender risk assessment practices vary across Canada. *Canadian Journal of Criminology and Criminal Justice, 60,* 167–205. http://dx.doi.org/10.3138/cjccj.2016-0024

Boyce, J. (2016). Victimization of Aboriginal people in Canada, 2014. Retrieved from https://www150.statcan.gc.ca/n1/en/pub/85-002-x/2016001/article/14631-eng.pdf?st=4WKRP8AE

Brankley, A., Lee, S. C., Babchishin, K. M., Hanson, R. K., & Harris, J. R. (2017). [Updated recidivism information for 409 individuals from the Dynamic Predictors Project (1998)]. Unpublished raw data.

Brouillette-Alarie, S., Babchishin, K. M., Hanson, R. K., & Helmus, L. M. (2016). Latent Constructs of the Static-99R and Static-2002R: A three-factor solution. *Assessment, 23,* 96–111. http://dx.doi.org/10.1177/1073191114568114

Brouillette-Alarie, S., & Proulx, J. (2018). The etiology of risk in sexual offenders: A preliminary model. *Sexual Abuse*. Advance online publication. http://dx.doi.org/10.1177/1079063218759325

Brouillette-Alarie, S., Proulx, J., & Hanson, R. K. (2018). Three central dimensions of sexual recidivism risk: Understanding the latent constructs of Static-99R and Static-2002R. *Sexual Abuse, 30,* 676–704. http://dx.doi.org/10.1177/1079063217691965

Canadian Charter of Rights and Freedoms, Part I of the Constitution Act, 1982, being Schedule B to the Canada Act 1982 (U. K.), 1982, c 11.

Canadian Institutes of Health Research. Natural Sciences and Engineering Research Council of Canada, & Social Sciences and Humanities Research Council of Canada. (2010). *Tri-council policy statement: Ethical conduct for research involving humans*. Ottawa, ON: Secretariat on Responsible Conduct of Research.

Canadian Psychological Association. (2017). *Canadian code of ethics for psychologists* (4th ed.). Ottawa, ON: Author. Retrieved from https://www.cpa.ca/docs/File/Ethics/CPA_Code_2017_4thEd.pdf

Corrections Act, R.S.N.W.T. 1988, c.C-22.

Corrections and Conditional Release Act (S. C. 1992, c. 20).

Corrections and Conditional Release Act (S. C. 2012, c. 1).

Criminal Code. RSC (1985). c.C-46.

Ellerby, L. A., & MacPherson, P. (2002). *Exploring the profiles of Aboriginal sexual offenders: Contrasting Aboriginal and non-Aboriginal sexual offenders to determine unique client Characteristics and potential implications for sex offender assessment and treatment strategies* (research report no. R-122). Ottawa, ON: Correctional Service of Canada. Retrieved from http://www.csc-scc.gc.ca/research/092/r122_e.pdf

Ewert v Canada, 2015 FC 1093.

Ewert v Canada, 2018 SCC 30.

Fernandez, Y., Harris, A. J. R., Hanson, R. K., & Sparks, J. (2014). *STABLE-2007 coding manual: Revised 2014*. Ottawa, ON: Public Safety Canada.

Fleiss, J. L., & Berlin, J. A. (2009). Effect sizes for dichotomous data. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 237–253). New York, USA: Russell Sage Foundation.

Geisinger, K. F., & McCormick, C. (2013). Testing and assessment in cross-cultural psychology. In J. R. Graham, J. A. Naglieri, & I. B. Weiner (Eds.), *Handbook of psychology: Vol. 10. Assessment psychology* (2nd ed., pp. 114–139). Hoboken, USA: Wiley.

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment, 12,* 19–30. http://dx.doi.org/10.1037/1040-3590.12.1.19

Gutierrez, L., Chadwick, N., & Wanamaker, K. A. (2018). Culturally relevant programming versus the status quo: A meta-analytic review of the effectiveness of treatment of indigenous offenders. *Canadian Journal of Criminology and Criminal Justice, 60,* 321–353. http://dx.doi.org/10.3138/cjccj.2017-0020.r2

Gutierrez, L., Helmus, L. M., & Hanson, R. K. (2016). What we know and don't know about risk assessment with offenders of indigenous heritage. *Journal of Threat Assessment and Management, 3,* 97–106. http://dx.doi.org/10.1037/tam0000064

Gutierrez, L., Wilson, H. A., Rugge, T., & Bonta, J. (2013). The prediction of recidivism with aboriginal offenders: A theoretically informed meta-analysis. *Canadian Journal of Criminology and Criminal Justice, 55,* 55–99. http://dx.doi.org/10.3138/cjccj.2011.E.51

Haag, A. M. (2005). [Recidivism data from 191 offenders detained until their warrant expiry date]. Unpublished raw data. 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*. Dissertations Abstracts International, *66,* 4531B. (UMI No. NR05662).

Haag, A. M., Boyes, A., Cheng, J., MacNeil, A., & Wirove, R. (2016). An introduction to the issues of cross-cultural assessment inspired by *Ewert v. Canada Journal of Threat Assessment and Management, 3,* 65–75. http://dx.doi.org/10.1037/tam0000067

Hanson, R. K. (2002). Recidivism and age: Follow-up data from 4,673 sexual offenders. *Journal of Interpersonal Violence, 17,* 1046–1062. http://dx.doi.org/10.1177/08862605-0201710-02

Hanson, R. K. (2008). What statistics should we use to report predictive accuracy? *Crime Scene, 15,* 15–17. Retrieved from http://www.cpa.ca/cpasite/UserFiles/Documents/Criminal%20Justice/Crime%20Scene%202008-04.pdf

Hanson, R. K., Babchishin, K. M., Helmus, L. M., Thornton, D., & Phenix, A. (2017). Communicating the results of criterion referenced prediction measures: Risk categories for the Static-99R and Static-2002R sexual offender risk assessment tools. *Psychological Assessment, 29,* 582–597. http://dx.doi.org/10.1037/pas0000371

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

Hanson, R. K., & Harris, A. J. (1998). *Dynamic predictors of sexual recidivism* (Corrections Research User Report No. 1998–01). Ottawa, ON: Department of the Solicitor General of Canada. Retrieved from https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/dnmc-prdctrs-sxl/dnmc-prdctrs-sxl-eng.pdf

Hanson, R. K., & Harris, A. J. (2000). Where should we intervene? Dynamic predictors of sex offense recidivism. *Criminal Justice and Behavior, 27,* 6–35. http://dx.doi.org/10.1177/0093854800027001002

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. http://dx.doi.org/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. http://dx.doi.org/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. http://dx.doi.org/10.1023/A:1005482921333

Hanson, R. K., & Thornton, D. (2003). *Notes on the development of Static-2002*. (Corrections Research User Report No. 2003–01). Ottawa, ON: Department of the Solicitor General of Canada. Retrieved from https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/nts-dvlpmnt-sttc/nts-dvlpmnt-sttc-eng.pdf

Hanson, R. K., Thornton, D., Helmus, L. M., & Babchishin, K. M. (2016). What sexual recidivism rates are associated with Static-99R and Static-2002R scores? *Sexual Abuse, 28,* 218–252. http://dx.doi.org/10.1177/1079063215574710

Hare, R. D. (2003). *Manual for the Psychopathy Checklist – Revised* (2nd ed.). Toronto, ON: Multi-Health Systems.

Helmus, L. M., & Babchishin, K. M. (2017). Primer on risk assessment and the statistics used to evaluate its accuracy. *Criminal Justice and Behavior, 44,* 8–25. http://dx.doi.org/10.1177/0093854816678898

Helmus, L., Babchishin, K. M., & Blais, J. (2012). Predictive accuracy of dynamic risk factors for aboriginal and non-aboriginal sex offenders: An exploratory comparison using STABLE-2007. *International Journal of Offender Therapy and Comparative Criminology, 56,* 856–876. http://dx.doi.org/10.1177/0306624X11414693

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, 24,* 64–101. http://dx.doi.org/10.1177/1079063211409951

Hill, D., & Demetrioff, S. (2019). Clinical-forensic psychology in Canada: A survey of practitioner characteristics, attitudes, and psychological assessment practices. *Canadian Psychology/Psychologie canadienne, 60,* 55–63. http://dx.doi.org/10.1037/cap0000152

Holsinger, A. M., Lowenkamp, C. T., & Latessa, E. J. (2003). Ethnicity, gender, and the Level of Service Inventory-Revised. *Journal of Criminal Justice, 31,* 309–320. http://dx.doi.org/10.1016/S0047-2352(03)00025-4

Holsinger, A. M., Lowenkamp, C. T., & Latessa, E. J. (2006). Exploring the predictive validity of the Level of Service Inventory-Revised with Native American offenders. *Journal of Criminal Justice, 34,* 331–337. http://dx.doi.org/10.1016/j.jcrimjus.2006.03.009

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). New York, USA: Wiley. http://dx.doi.org/10.1002/9781118548387

Kelley, S. M., Ambroziak, G., Thornton, D., & Barahal, R. M. (2018). How do professionals assess sexual recidivism risk? An updated survey of practices. *Sexual Abuse*. Advance online publication. http://dx.doi.org/10.1177/1079063218800474

Laprairie, C. (1996). *Examining aboriginal corrections in Canada*. (Catalogue. No. JS5-1/14-1996E). Ottawa, ON: Ministry of the Solicitor General. Retrieved from https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/xmnng-brgnl-crrctns/xmnng-brgnl-crrctns-eng.pdf

LaPrairie, C. (2002). Aboriginal over-representation in the criminal justice system: A tale of nine cities. *Canadian Journal of Criminology, 44,* 181–208.

Lee, S. C., Hanson, R. K., Calkins, C., & Jeglic, E. (2018). Paraphilia and antisociality: Motivations for sexual offending may differ for American Whites and Blacks. Manuscript submitted for publication.

Lee, S. C., Mularczyk, K., Babchishin, K. M., Blais, J., & Bonta, J. (2018). [Updated recidivism information for 360 individuals from the National Flagging System Project (2015)]. Unpublished raw data.

Ley, P. (1972). *Quantitative aspects of psychological assessment: An introduction*. London, UK: Duckworth.

MacDonald, S. F. (2014). *Profile of Aboriginal men offenders: Custody and supervision snapshots* (Report No. 2014 N° R-321). Retrieved from http://publications.gc.ca/site/eng/9.810320/publication.html

Mann, M. (2009). *Good Intentions, Disappointing Results: A Progress Report on Federal Aboriginal Corrections*. Ottawa, ON: Office of the Correctional Investigator. Retrieved from http://www.oci-bec.gc.ca/cnt/rpt/pdf/oth-aut/oth-aut20091113-eng.pdf

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, 22,* 191–217. http://dx.doi.org/10.1177/1079063210366039

Ministry of Correctional Services Act, R. S. O. 1990, c. M. 22.

Office of the Auditor General of Canada. (2016). *Preparing Indigenous offenders for release*. Ottawa, ON: Author. Retrieved from http://www.oag-bvg.gc.ca/internet/English/parl_oag_201611_03_e_41832.html

Office of the Correctional Investigator. (2015). *Annual report of the office of the correctional investigator, 2014–2015* (Catalogue no. PS100E-PDF). Retrieved from http://www.oci-bec.gc.ca/cnt/rpt/pdf/annrpt/annrpt20142015-eng.pdf

Olver, M. E., Sowden, J. N., Kingston, D. A., Nicholaichuk, T. P., Gordon, A., Beggs Christofferson, S. M., & Wong, S. C. P. (2018). Predictive accuracy of Violence Risk Scale – Sexual Offender Version Risk and change scores in treated Canadian Aboriginal and non-Aboriginal sexual offenders. *Sexual Abuse, 30,* 254–275. http://dx.doi.org/10.1177/1079063216649594

Perley-Robertson, B., Helmus, L. M., & Forth, A. (2018). Predictive accuracy of static risk factors for Canadian indigenous offenders compared to non-indigenous offenders: Implications for risk assessment scales. *Psychology, Crime & Law*. Advance online publication. http://dx.doi.org/10.1080/1068316X.2018.1519827

Phenix, A., Helmus, L., & Hanson, R. K. (2016). *Static-99R & Static-2002R evaluators' workbook*. Retrieved from http://www.static99.org/

Public Safety Canada. (2015). *Corrections and conditional release statistical overview* (Catalogue No. PS1-3E-PDF). Retrieved from https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/ccrso-2015/ccrso-2015-en.pdf

Quann, N. L., & Trevethan, S. (2000). *Police-reported Aboriginal crime in Saskatchewan* (Catalogue no. 85F0031XIE). Ottawa, ON: Statistics Canada. Retrieved from http://www.publications.gc.ca/site/eng/9.685779/publication.html

R v Gladue, 1999 1 S.C.R. 688.

Reynolds, C. R., & Suzuki, L. A. (2013). Bias in psychological assessment: An empirical review and recommendations. In J. R. Graham, J. A. Naglieri, & I. B. Weiner (Eds.), *Handbook of psychology: Vol. 10. Assessment psychology* (pp. 82–113). Hoboken, USA: Wiley.

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's *d,* and *r*. *Law and Human Behavior, 29,* 615–620. http://dx.doi.org/10.1007/s10979-005-6832-7

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,* 77. http://dx.doi.org/10.1186/1471-2105-12-77

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J., & Müller, M. (2014). *Package 'pROC'*. Retrieved from https://cran.r-project.org/web/packages/pROC/pROC.pdf

Rojas, E. Y., & Gretton, H. M. (2007). Background, offence characteristics, and criminal outcomes of Aboriginal youth who sexually offend: A closer look at Aboriginal youth intervention needs. *Sexual Abuse, 19,* 257–283. http://dx.doi.org/10.1177/107906320701900306

Royal Commission on Aboriginal Peoples. (1996). *Highlights from the Report of the Royal Commission on Aboriginal People* (Catalogue no. Z1-1991/1-6E).

Rudin, J. (2009). *Aboriginal peoples and the criminal justice system*. Retrieved from http://www.archives.gov.on.ca/en/e_records/ipperwash/policy_part/research/pdf/Rudin.pdf

Ruscio, J. (2008). A probability-based measure of effect size: Robustness to base rates and other factors. *Psychological Methods, 13,* 19–30. http://dx.doi.org/10.1037/1082-989X.13.1.19

Scrim, K. (2010). Aboriginal victimization in Canada: A summary of the literature. *Victims of Crime Research Digest, 3*. Retrieved from http://www.justice.gc.ca/eng/rp-pr/cj-jp/victim/rd3-rr3/p3.html

Shepherd, S. M., Adams, Y., McEntyre, E., & Walker, R. (2014). Violence risk assessment in Australian aboriginal offender populations: A review of the literature. *Psychology, Public Policy, and Law, 20,* 281–293. http://dx.doi.org/10.1037/law0000017

Sinha, V., Trocmé, N., Fallon, B., MacLaurin, B., Fast, E., Prokop, S. T., . . . Richard, K. (2011). *Kiskisik Awasisak: Remember the children. Understanding the overrepresentation of First Nations children in the child welfare system*. Ontario, ON: Assembly of First Nations. Retrieved from http://cwrp.ca/publications/2280

Sioui, R., & Thibault, J. (2002). Examining reintegration potential for Aboriginal offenders. *Forum on Corrections Research, 14,* 49–51.

Statistics Canada. (2005). *Projections of the Aboriginal populations, Canada, provinces and territories- 2001 to 2017* (Catalogue no. 91–547-XIE). Retrieved from https://www150.statcan.gc.ca/n1/pub/91-547-x/91-547-x2005001-eng.pdf

Statistics Canada. (2015). *Aboriginal statistics at a glance* (Catalogue no. 89–645-X2015001). Retrieved from https://www150.statcan.gc.ca/n1/en/pub/89-645-x/89-645-x2015001-eng.pdf?st=_GMvlN8d

Statistics Canada. (2018). *Aboriginal Population Profile. 2016 Census* (Catalogue no. 98–510-X2016001). Retrieved from http://www12.statcan.gc.ca/census-recensement/2016/dp-pd/abpopprof/details/page.cfm?Lang=E&Geo1=PR&Code1=01&Data=Count&SearchText=Canada&SearchType=Begins&B1=All&SEX_ID=1&AGE_ID=1&RESGEO_ID=1

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

Trevethan, S., Moore, J.-P., & Rastin, C. J. (2002). A profile of Aboriginal offenders in federal facilities and serving time in the community. *Forum*

on Corrections Research, 14, 17–19. Retrieved from http://www.csc-scc.gc.ca/text/pblct/forum/e143/143f_e.pdf

Truth and Reconciliation Commission of Canada. (2015). *Honouring the truth, reconciling for the future*. Winnipeg, MB: Author.

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

Whitaker, D. J., Le, B., Karl Hanson, R., Baker, C. K., McMahon, P. M., Ryan, G., . . . Rice, D. D. (2008). Risk factors for the perpetration of child sexual abuse: A review and meta-analysis. *Child Abuse & Neglect, 32,* 529–548. http://dx.doi.org/10.1016/j.chiabu.2007.08.005

Wong, S., Olver, M. E., Nicholaichuk, T. P., & Gordon, A. (2003). *The Violence Risk Scale: Sexual Offender version (VRS-SO)*. Saskatoon, SK: Regional Psychiatric Centre and the University of Saskatchewan.

Received January 8, 2019
Revision received April 25, 2019
Accepted April 28, 2019 ■