

Too Risky to Use, or Too Risky Not to? Lessons Learned from Over 30 Years of Research on Forensic Risk  
Assessment with Indigenous Persons

Mark E. Olver

Keira C. Stockdale

L. Maaike Helmus

Phil Woods

Jordan Termeer

Jessica Prince

In Press: *Psychological Bulletin* ©American Psychological Association, 2023. This paper is not the copy of record and may not exactly replicate the authoritative document published in the APA journal.

Author note:

Mark E. Olver, Department of Psychology and Health Studies, University of Saskatchewan; Keira C. Stockdale, Department of Psychology and Health Studies, University of Saskatchewan; L. Maaike Helmus, Department of Criminology, Simon Fraser University; Phil Woods, College of Nursing, University of Saskatchewan; Jordan Termeer, Department of Psychology and Health Studies, University of Saskatchewan; Jessica Prince Department of Psychology and Health Studies, University of Saskatchewan.

This study was funded through a Memorandum of Understanding with the Correctional Service of Canada. We all thank Dr. Leslie Anne Keown for her support, feedback, and comments on this body of work and the foundational report for this manuscript. L. Maaike Helmus also thanks Ashley Kyne for her invaluable assistance locating key literary works on Indigenous peoples and having shaped her thinking

on culture and risk assessment issues. Finally, Mark E. Olver and Keira C. Stockdale dedicate this work to their daughter, Mirabel, who arrived into this world in the very midst of the writing, analyses, and revisions of the present work.

The views expressed in this document are those of the authors and do not necessarily represent those of the University of Saskatchewan, Simon Fraser University, Saskatoon Police Service, or Correctional Service of Canada. In the spirit of recommendations for non-Indigenous researchers to acknowledge their privilege and challenge their biases we respectfully acknowledge that we completed this work in Canada on traditional lands of Treaty 6 Territory and the Homeland of the Metis as well as the traditional and unceded territory of the Coast Salish Peoples. We recognize these nations whose sovereignty has not been ceded.

Mark E. Olver is co-author of the Violence Risk Scale-Sexual Offense version, Keira C. Stockdale is co-author of the Violence Risk Scale-Youth Version, and L. Maaiké Helmus is co-author of the Static-99R. Each receives occasional remuneration for consultation and training services on the respective tools. Please direct correspondence to Mark. E. Olver, Department of Psychology and Health Studies, University of Saskatchewan, 9 Campus Drive, Arts 154, Saskatoon, SK, Canada, S7N 5A5, [mark.olver@usask.ca](mailto:mark.olver@usask.ca)

### Abstract

Indigenous peoples are overrepresented in correctional systems internationally, reflecting a history of systemic racism and colonial oppression, and the practice of risk assessment with this population has been a focus of legal and sociopolitical controversy. We conducted a systematic review and meta-analysis of the risk assessment literature comparing Indigenous and non-Indigenous (White majority) groups. We retrieved 91 studies featuring 22 risk tools and 15 risk/need/cultural domains ( $N = 59,693$ , Indigenous;  $N = 237,729$ , non-Indigenous/White), and 4 documents identifying culturally relevant factors. Most measures demonstrated moderate predictive validity but often had significant ethnoracial differences, particularly for static measures. The Service Planning Instrument/Youth Assessment Screening Inventory, Level of Service Inventory youth variants, Psychopathy Checklist-Revised and Youth Version, and the Violence Risk Scale and its Sexual Offense version had the strongest predictive validity and least ethnoracial discrepancy. The Static Factors Assessment and Dynamic Factors Identification and Analysis-Revised had the weakest predictive validity. For Indigenous persons, the strongest individual predictors of recidivism were low education/employment, substance abuse, antisocial pattern, and poor community functioning, while mitigating factors that predicted decreased recidivism were measures of risk-change (i.e., from culturally integrated programs combining mainstream and traditional healing approaches), cultural engagement/connectedness, and protective factors. In practice, static measures need to be supplemented with dynamic ones, and assessors should select measures with at least moderate predictive validity and ideally the least ethnoracial bias. These conclusions are tempered by the quantity and quality of the literature coupled with the circumstance that some study authors have co-authored tools in this review.

*Key words:* risk assessment; Indigenous; culture; recidivism; meta-analysis

**Public Significance Statement:** Most risk assessment measures can predict reoffending with a reasonable (i.e., moderate) level of accuracy for Indigenous people undergoing risk assessments in the

justice system; however, these measures often have better accuracy when used with non-Indigenous (and primarily White) clients, particularly static measures that are heavily based on criminal history. The “best” potential measures tend to be dynamic tools, as these have less prediction bias, can drive services, measure reductions in risk (e.g., with culturally integrated treatment programs), and inform release decisions and community supervision strategies. Of note, other unmeasured or non-risk relevant factors (e.g., over-policing, aggressive prosecution) can inflate recidivism estimates for Indigenous peoples and decrease predictive validity of risk measures, which should be considered when conducting risk assessments. Careful, critical examination of current risk measures and further research is necessary to increase fair, humane, and ethical use of these measures with Indigenous clients in forensic, correctional, and legal settings.

**Too Risky to Use, or Too Risky Not to? Lessons Learned from Over 30 Years of Research on Forensic Risk Assessment with Indigenous Persons**

Over the past 30 years, there have been significant advancements in the development and validation of structured approaches to assess and manage risk for recidivism with forensic and correctional populations. During this time period, research has established the superiority of structured assessment approaches over unstructured clinical judgment in terms of improvements in predictive accuracy, reliability, objectivity, fairness, accountability, and utility (Bonta et al., 1998; Dawes et al., 1989; Grove et al., 2000; Hanson & Morton-Bourgon, 2009; Vincent & Viljoen, 2020). The use of structured assessment approaches is especially important for evaluating forensic risk given the social, institutional, human, fiscal, and public health costs of recidivism and the human rights consequences of overpredicting low base rate phenomena (i.e., false positives) such as unnecessary detention. Further, the accurate identification and classification of individuals in terms of their level of risk and criminogenic need informs correctional intervention and recidivism prevention to decrease societal victimization, promote public safety, and enhance reintegration and wellbeing among those convicted of offenses. These important advances and functions notwithstanding, a critical focus of discussion and debate has centered around the applications of forensic risk assessment instruments in correctional and psycholegal contexts with ethnocultural minorities in general, and Indigenous peoples in particular (Gutierrez et al., 2016; Hart, 2016; Olver, 2016; Rugge, 2006; Tamatea, 2017). The use of forensic risk assessment measures with justice involved Indigenous persons is made all the more salient given the overrepresentation of Indigenous peoples and many ethnocultural minorities in the criminal justice system and their collective experiences of systemic racism, marginalization, and intergenerational trauma (Haskell et al., 2009; Public Safety Canada, 2019; Rugge, 2006; Truth and Reconciliation Commission [TRC] of Canada, 2015).

The present study is a systematic review and meta-analysis examining the predictive properties

of static and dynamic forensic risk assessment tools and their constituent risk-need domains for Indigenous justice-involved adults and youth. We also review cultural-specific approaches to risk assessment. Current findings are synthesized with clinical and contextual considerations to inform recommendations for risk assessment practices with Indigenous persons. We included literature from around the globe but much of it focuses primarily on the Canadian landscape as most contemporary risk assessment tools have been developed in Canada and have been adopted in some form in other countries worldwide. As such, most of the available research conducted to date has examined Canadian correctional clientele and predominantly men. That said, the currency of the psychometric and assessment issues explored within these lines of research may extend to other correctional clientele (e.g., ethnocultural minorities, different contexts and jurisdictions) that share important historical, contextual, and methodological parallels (e.g., United States, Australia, New Zealand).

It is recognized that a literature review of empirical studies represents a limited way of knowing or understanding. Despite the purported objectivity of quantitative scientific research, it is still a reflection of colonial values (Anderson & Kukutai, 2016), which may influence our interpretation of data. Research conducted by White researchers has historically been used to control, exploit, and alienate Indigenous communities and devalue their knowledge (Smith, 2012; Victor, 2007). Although these research methods have often clashed with Indigenous ways of knowing, there are many researchers who believe Western and Indigenous methodologies can be integrated in a way that dismantles colonialism (Asselin & Basile, 2018; Peltier, 2018; Smith, 2012) and advances social justice (Datta, 2018).

This article presents our synthesis and interpretation of a diverse base of literature, which is influenced by our own identities and cultures, which have been shaped by historical privilege. We also recognize that our perspective and contribution to this complicated literature is one piece of the important puzzle of understanding the role of risk assessment within a colonial system. Alternative methodologies and collaborations with Indigenous communities are needed to complement this

research. The present work on risk assessment with Indigenous persons is part of an ongoing collaboration with the Correctional Service Canada<sup>1</sup> (CSC), which incorporates traditional approaches to knowledge keeping as well as consultation with Indigenous peoples and communities. We extend our respect to the traditional custodians of ancestral lands and we do not attempt to represent the knowledge or stories of Indigenous peoples. We do begin by introducing historical, clinical, forensic, and legal issues and controversies in risk assessment with Indigenous persons in order to situate this literature and its human impact.

### **Criminal Justice and Historical Contexts of Forensic Risk Assessment with Indigenous Peoples**

Correctional and forensic mental health systems serve an important function to rehabilitate, reduce and manage risk, and reintegrate people who come in contact with the criminal justice system; all with the goal of preventing their return and reducing societal crime and victimization. The risk, need, responsivity (RNR) model is the predominant model guiding assessment, intervention, and case management efforts with justice involved persons in correctional systems around the world (Andrews et al., 1990; Andrews & Bonta, 1994-2010, Bonta & Andrews, 2007; Bonta & Andrews, 2017; for a review and critique, see Polaschek, 2012). The risk principle states that individuals can be accurately classified according to risk level and that service intensity should be matched to level of risk (i.e., higher risk entails more services). The need principle stipulates that dynamic risk factors, also known as criminogenic needs (i.e., “crime causing” variables, such as antisocial attitudes, criminal peers, or substance-use concerns), involved in the onset and continuation of offending should be prioritized for treatment services. As such, changes in criminogenic needs from treatment or other change agents (e.g., aging, health status) should be associated with changes (reductions) in recidivism; further, programs that primarily target factors not empirically related to recidivism should not expect a reduction in

---

<sup>1</sup> CSC administers all custodial sentences of two or more years in Canada.

recidivism. Finally, the responsivity principle states that services should be adapted to the unique characteristics of clientele (e.g., culture, learning style, motivation) within the context of cognitive behavioral methods of change. Implementation of RNR leans heavily on the use of risk assessment measures in order to identify: 1. the level of risk with respect to a given outcome (e.g., sexual, violent, general recidivism) to inform treatment intensity; 2. where to intervene to help prioritize services as well as potential changes in risk upon reassessment; and 3. how to intervene in light of other considerations that may have arisen during the assessment process (Olver & Wong, 2019).

As risk measures are employed in correctional and forensic mental health systems worldwide, invariably they will also be employed with ethnocultural minorities. Concerns have been raised that these measures have tended to be developed on White majority samples (Shepherd & Lewis-Fernandez, 2016; Shepherd & Willis Esqueda, 2017). Although diverse groups have representation within the construction and validation samples of such measures, it is essential to establish the psychometric properties of these measures with diverse groups to ensure that their use aids their intended risk management and reintegration applications. Otherwise, should such tools contain psychometric bias or promote unfair and inequitable decisions (e.g., access to programming, conditional release), the use of mainstream risk measures has the potential to perpetuate problems endemic to the overincarceration of certain ethnocultural minority groups.

### ***Overrepresentation of Indigenous Persons in Correctional Settings***

Several ethnocultural minorities are overrepresented in correctional and forensic mental health systems throughout the world including in the United States with Black, Hispanic, and North American Indigenous persons (Nellis, 2016; Travis et al., 2014), New Zealand Māori (Anyia, 2015), Indigenous Australians (Day et al., 2018; Krieg, 2006), the United Kingdom's British Black citizens (UK Ministry of Justice, 2013), and Canada's Indigenous (First Nations, Métis, and Inuit) youth and adults (Department

of Justice, 2017; Public Safety Canada, 2019).<sup>2</sup> Indigenous overincarceration specifically has occurred on an international scale, and particularly in the countries of Canada, New Zealand, Australia, and the United States (US), all of which have common ground in their colonial origins and policies and practices in criminal justice, corrections, and policing that have contributed to a mounting social and health crisis (Cunneen, 2009).

In Canada, Indigenous youth and adults are overrepresented by about 5-8-fold in their respective justice systems. Although Indigenous persons represent only about 4% of the general Canadian population, according to Public Safety Canada's Corrections and Conditional Release Statistical Overview (2019), in 2018-19 Indigenous persons accounted for 25.5% of the total offender population and 29.5% of the custody population (28.8% of men, 41.5% of women), 18.7% of the community correctional population, and 35.5% of the dangerous offender (DO) population; Canada's Indigenous population in custody has also increased by 47.4% over the past 10 years. In New Zealand, the Māori peoples represent about 17% of the country's population (Government of New Zealand, 2017) but for several years they have steadily comprised at least 50% of the total prison population (New Zealand Department of Corrections, 2021). Australia's Aboriginal and Torres Strait Islander peoples, in turn, comprise 30% of the prison population, an increase of 8% from 2020 to 2021 (Australian Bureau of Statistics, 2021), despite representing just over 3% of the Australian population (Australian Government, 2021).

In the US context, incarceration has been increasing more than any other country in the world (Travis et al., 2014). The US incarcerates more of its citizens per capita than any other country in the world and this disproportionately affects ethnocultural minorities (Travis et al., 2014). Black people are the most heavily incarcerated group in the United States prison system with a rate of incarceration that

---

<sup>2</sup> Some ethnocultural groups, such as East Asian individuals, by contrast, tend to be underrepresented (Brankley & Lee, 2016; Hu & Esthappan, 2017).

is approximately 5 times higher than White persons, in addition to serving longer prison terms on average for the same offenses (United States Sentencing Commission, 2017). According to the National Congress of American Indians (NCAI, 2020), American Indians and Alaskan Natives represent just under 1% of the general American population, but they have had some of the fastest rising counts in US prisons and jails and are heavily overrepresented in certain states, such as North and South Dakota, Montana, Arizona, and New Mexico, where the incarceration rates are up to 7 times higher than for White persons (United States Sentencing Commission, 2021). Further, unique to the US context are “Indian Country Jails,” facilities located on tribal lands in certain states which also incarcerate and house Indigenous persons; although the Indigenous population in these facilities (as well as the number of facilities) had been steadily increasing up to 2019, the population declined by 30% in 2020 during the COVID-19 pandemic, owing to a pandemic-related reduction in admissions and concordant increase in expedited releases (Minton, 2021).

### ***Social and Health Impacts and Criminal Justice Implications of Colonialism and Racial Trauma***

Common to Indigenous peoples in Canada, US, New Zealand, and Australia is a devastating collective history of colonization by White settlers and the resultant forced assimilation and policies and practices bent on the destruction of Indigenous cultures through the loss of language, traditions, ceremonies, spirituality, and family systems (Reyhner & Singh, 2010; Spring, 1994). This includes the expropriation of Indigenous lands and resources and segregating Indigenous peoples to often remote and less desirable would-be reserves, as well as a residential school systems, which tore families apart and uprooted Indigenous children, who would then be subjected to educative efforts at “deculturalization” (Spring, 1994) and horrific abuses some of which resulted in death (Sterritt & Dickson, 2021) by the hands of the very people tasked with their care (Reyhner & Singh, 2010; TRC, 2015). The impacts of collective racial trauma among Indigenous peoples are widespread and substantial. Today, social, health, and economic inequality are prevalent.

Indigenous peoples are at higher risk for violent victimization (Brzozowski et al., 2006; Perrault, 2011; Scrim, 2010), are most likely to be victims of interracial violence (National Institute of Justice, 2016), and to be killed by police than any other ethnocultural group (Males, 2014). Indigenous persons have higher rates of trauma (Skewes & Bloom, 2019), substance misuse and dependence (Skewes & Bloom, 2019), and morbidity and mortality from medical concerns (e.g., hypertension, type II diabetes) (Indian Health Service, 2019; Scrim, 2010). Many Indigenous communities in Canada and the US are “food deserts” (i.e., areas lacking access to affordable, quality, fresh food) and Indigenous persons in the US face food insecurity at four times the national United States average (Partnership with Native Americans, 2017). According to the 2018 United States Census, Native Americans have the highest rate of poverty in the United States at 25.4% (Poverty USA, n.d.), and from 2015 to 2019, the median household income for Indigenous persons was \$43,825 vs. the national median of \$62,843 (Guzman, 2020). Indigenous peoples also have lower educational attainment (Postsecondary National Policy Institute, 2021) and higher rates of unemployment (Maxim et al., 2022).

Colonization has long enacted systemic racism and inequality and the criminal legal system has been one of many societal mechanisms to perpetuate it. Almost exactly 100 years ago, Haudenosaunee Chief Deskaheh traveled to the League of Nations to complain about arbitrary prosecution and incarceration of Indigenous peoples in Canada (Deskaheh, 1923). Today, people in correctional settings continue to be drawn most heavily from marginalized segments of society, and in an “era of high incarceration rates, prison admission and return have become commonplace in minority neighborhoods characterized by high levels of crime, poverty, family instability, poor health, and residential segregation” (Travis et al., p. 5, 2014). Concerns about over-policing and over-prosecuting Indigenous communities have also been raised in the US, Canada, Australia, and New Zealand (Fergusson et al., 2003; Nettlebeck & Smandych, 2010; Perry, 2006).

In correctional settings, Indigenous persons score higher on established risk factors, such as

Andrews and Bonta's (2010) Central Eight, which include criminal history, antisocial attitudes, antisocial peers/associates, antisocial personality pattern, employment/education, family/marital, leisure/recreation, and substance abuse (drug/alcohol). Indigenous persons also score higher on risk tools comprised of these factors (Wormith et al., 2015), given that they have higher rates of substance abuse (Cram & MacDonald, 2019; Shepherd et al., 2014; Stewart et al., 2017), lower education and higher unemployment (Shepherd et al., 2014; Stewart et al., 2017), criminogenic peer ties (Sittner & Hautala, 2016; Stewart et al., 2017; Wormith et al., 2015), adverse family/marital circumstances (Shepherd et al., 2014; Stewart et al., 2017; Wormith et al., 2015), and lengthier criminal histories (Perley-Robertson et al., 2019), relative to non-Indigenous persons. In turn, justice involved Indigenous persons are more likely to be classified as high risk (Helmus & Forrester, 2014a,b; Ruge, 2006; Stewart et al., 2017; Wormith et al., 2015), to be assigned to higher security levels on intake (Barnum & Gobeil, 2012b; Gobeil, 2011), and have higher rates of recidivism (Stewart & Wilton, 2019). Thus, risk assessment becomes highly contentious, and may be cynically viewed as yet another mechanism of Indigenous oppression.

Some legal developments in Canada have been established to mitigate the problem of Indigenous overincarceration (e.g., *R. v. Ipeelee*, 2012; *R. v. Gladue*, 1999). For instance, the Gladue provision (*R. v. Gladue*, 1999) introduced by the Supreme Court of Canada (SCC) (CCC Sec. 718.2 [e]) stipulates that the court must take into consideration: a. the unique background or contextual factors that contributed to an Indigenous person coming into contact with the law, and b. appropriate sentencing options or sanctions (e.g., restorative, rehabilitative) in light of the individual's Indigenous heritage. For instance, an individual's community and case worker may attest to how that person's experience with racism, foster care, and abuse against the backdrop of colonialism contributed to a dislocation from their culture and life concerns with substance misuse, education gaps and unemployment, mistrust of authority, and relationship instability that contributed to the current crime

(e.g., thefts, assaults). The presiding judge may adjust the sentence or decide on alternatives to incarceration (e.g., probation and a community treatment order, attendance at a healing lodge) reflecting the individual's circumstances to promote desistance, healing, and community reintegration. Although formal justice avenues such as invoking Gladue principles and issuing a Gladue report are a means of formally responding to Indigenous heritage in legal decision making, critics contend that the provision is inconsistent in its frequency of use and the nature of its application, and that it has done little to reduce Indigenous overrepresentation in criminal justice contexts (Jackson, 2016).

Canadian legal precedents such as Gladue introduce a conundrum to risk assessment with Indigenous persons in Canada and beyond. For instance, the so-called "Gladue factors" may converge with risk factors associated with recidivism such as low income, unemployment, education gaps, substance abuse, and community fragmentation (Jackson, 2016). Do risk assessment tools unfairly penalize the personal and social circumstances of Indigenous persons in general by virtue of their item content, thereby perpetuating White colonial oppression? Or do the historical and lifetime circumstances of justice involved Indigenous persons create personal and environmental circumstances that create fertile ground for the development of such social and psychological factors that increase risk for crime and violence? Or can it be both?

### ***Ethnocultural Criticisms of Risk Assessment Tools***

There has been extensive discussion that conventional risk tools may potentially disadvantage Indigenous persons by virtue of including items or indicators on which they will be overrepresented (Clark, 2019; Shepherd & Anthony, 2018; Shepherd & Willis Esqueda, 2017), but also, through failure to consider variables unique to the history, circumstances, lived experience, and cultural heritage of Indigenous peoples (Boer et al., 2003; Gutierrez, 2018; Shepherd, 2016; Tamatea, 2017). Existing risk tools do not have formal item content with considerations unique specifically to Indigenous peoples. In this vein, there have been three sets of criticisms launched at these measures: 1. that the item content

reflects the lived experiences, perceptions, and world views of White people, especially the test developers; 2. that predictive validity magnitudes are often lower for Indigenous than for White individuals; and 3. it has not been sufficiently established whether mainstream instruments assess risk in the same manner (i.e., in structure and content) across ethnocultural groups (Hart, 2016; Shepherd, 2016; Shepherd & Lewis-Fernandez, 2016; Tamatea, 2017). As such, there have been calls for consultations with Indigenous communities about item content in existing risk assessment tools; otherwise, “there is a risk that the process of risk assessment could be perceived as yet another way that members of the dominant culture impose their values on Indigenous people” (Hanson, 2020, par. 3).

A largely untested assumption has also been made that the psychometric properties (i.e., validity and reliability) of such tools extend to Indigenous persons; a recent Canadian Supreme Court decision (*Ewert v. Canada*, 2018; see also *Ewert v. Canada*, 2015; *Canada v. Ewert*, 2016) has explicitly required this assumption be tested. *Ewert v. Canada* (2015) called into question the cultural and psychometric appropriateness of risk assessment tools. Jeffrey Ewert, a Métis man who had served 30 years of two concurrent life sentences for the second degree murder and attempted murder of two teenage females over two separate sexually motivated incidents, sued Correctional Service Canada (CSC, Canada’s federal correctional department) on the grounds that he had been caused harm through the use of risk assessment tools and related structured forensic assessment measures, given that these measures were developed on largely White Euro-Canadian samples and were culturally inappropriate in light of his Métis heritage (see Haag et al., 2016). The impugned tools were the Psychopathy Checklist-Revised (PCL-R), Static-99 (Hanson & Thornton, 1999/2000), Violence Risk Scale-Sexual Offense version (VRS-SO; Wong et al., 2003-2017), Violence Risk Appraisal Guide (VRAG; Harris et al., 1993), and the Sex Offender Risk Appraisal Guide (SORAG; Quinsey et al., 1998).

Several critical psychometric issues were identified in *Ewert v. Canada* (2015). First, there is an

issue of generalizability from a dominant aggregate group, to a smaller diverse sample of Indigenous persons, such that generalization of psychometric properties cannot be assumed. Second, actuarial tools came under fire, and specifically, that quantitative measures that yielded numeric estimates of recidivism probability were potentially problematic and prone to misuse; the discussion focused specifically on static actuarial measures and the problems that can arise from a lack of professional discretion around their use. Third, the lack of psychometric research specifically with Indigenous persons was noted. Among the myriad psychometric issues highlighted (see Hart, 2016), the one that seemed to feature most prominently was predictive invariance, that is, to what extent a risk instrument predicts a targeted outcome (i.e., recidivism) with an acceptable and comparable degree of accuracy for Indigenous and non-Indigenous persons.

Although the Plaintiff won, the presiding judge ruling that Mr. Ewert's Charter rights had been violated and admonishing CSC about using such tools absent proper validation, the matter was overturned on appeal the following year (*Canada v. Ewert, 2016*), before being heard in the Supreme Court of Canada (SCC; *Ewert v. Canada, 2018*). Although the SCC found that Mr. Ewert's Charter rights had not been infringed upon, the Court did rule that "the CSC breached its obligation under s. 24(1) of the *CCRA* [Corrections and Conditional Release Act] to take all reasonable steps to ensure that any information about an offender that it uses is as accurate as possible" (p. 208). What the matter helped accomplish was to set a clear expectation that cross-cultural validity research of risk assessment with justice involved Indigenous persons is critical and to set higher standards for this research. The implications are far reaching and extend to risk assessment research and practice with Indigenous persons and other ethnocultural minorities in correctional systems elsewhere.

### **Psychometric Considerations and Extant Findings in Risk Assessment with Indigenous Persons**

Several important psychometric research and practice considerations emerged from the aforementioned legal developments in tandem with the state of the literature. One set of key issues

concerns the form and content of these tools, as well as their relative strengths and weaknesses, to which we now turn.

### ***Risk Tool Classification and Terminology***

***Risk Tool Classification.*** Risk tools can be classified into four basic approaches, with variations therein (Hanson & Morton-Bourgon, 2009); Table 1 provides a summary of the form, structure, and content for several of these measures. Unstructured clinical judgment involves an informal and subjective appraisal of recidivism probability, usually through an interview and possibly a review of collateral documents. The accuracy of such approaches was poor and usually erred on the side of overpredicting or overclassifying risk (i.e., false positives, predicting someone will reoffend or behave dangerously when they do not) (Grove & Meehl, 1996; Grove et al., 2000; Meehl, 1954; Mossman, 1994). The remaining approaches (empirical actuarial, structured professional judgment, mechanical) involve imposing greater structure and have some element of standardization in their administration (drawing from file review and interview) and scoring. Although these approaches tend to have better predictive validity and interrater reliability, invariably, assessment processes that rely on observers or interviewers are vulnerable to the biases and cultural misunderstandings of the observer.

Empirical actuarial measures consist of items that are numerically scored and summed to generate risk scores, which are then arranged into risk bands (traditionally some variant on low, medium, and high) and recidivism probabilities over a defined follow-up. Items can be selected from theory and research or be statistically driven and the tools themselves may comprise items that are static (i.e., historical, generally unchanging factors such as criminal history or age at first offense) and/or dynamic (i.e., potentially changeable social, psychological, or environmental factors such as antisocial attitudes or criminal peers). Relevant examples of static empirical actuarial tools include the Statistical Information on Recidivism (SIR; Nuffield, 1982; Nafekh & Motiuk, 2002) scale, Violence Risk Appraisal Guide Revised (VRAG-R; Rice et al., 2013), and the Static-99R (Helmus et al., 2012; see also Hanson &

Thornton, 1999/2000). Examples of dynamic empirical actuarial tools include the Level of Service Inventory-Revised (LSI-R; Andrews & Bonta, 1995), and its variants, referred to as the Level of Service (LS) scales, the STABLE-2007 to assess sexual violence risk (Hanson et al., 2007), or the Violence Risk Scale (Wong & Gordon, 1999-2003) to assess general violence risk.

Structured professional judgment (SPJ) measures involve rating the items and integrating the information they convey into a summary risk rating (SRR), usually of low, medium, or high. In SPJ, the items are not summed to generate risk scores, and these are not linked to recidivism probabilities. Prominent SPJ examples include the Historical Clinical Risk Management-20 Version 2 (HCR-20; Webster et al., 1997) and 3 (Douglas et al., 2013) for general violence, the Sexual Violence Risk-20 (SVR-20; Boer et al., 1997) for sexual violence, and the Spousal Assault Risk Assessment Version 2 (SARA; Kropp et al., 1995) and 3 (Kropp & Hart, 2015) for intimate partner violence (IPV). Finally, mechanical tools involve numerically summing items on risk tools to generate a numeric risk score, however, the score is not linked to recidivism probabilities (e.g., violence within the next 3 years). This would characterize summing the items from SPJ tools to yield a numeric score, typically done to examine the measure's psychometric properties, although this differs from how the SPJ tool is used in clinical practice.

***Generations of Risk Assessment.*** The different risk measures intersect with Bonta's (1996; see also Andrews et al., 2006) generational risk assessment framework, which outlines the nature and application of risk tools in a risk assessment chronology. The first generation is synonymous with unstructured clinical judgment, while static empirical actuarial tools characterize the second generation. Static empirical actuarial tools are limited in their capacity to inform treatment and they cannot assess change meaningfully; yet their objectivity, reliability, and predictive accuracy for recidivism is far superior to unstructured clinical judgment (Grove et al., 2000; Hanson & Morton-Bourgon, 2009; Mossman, 1994). By contrast, dynamic measures, whether they be actuarial or SPJ, comprise what Bonta (1996) has termed the third and fourth generations. As dynamic measures, they can be used to

drive treatment planning, evaluate changes in risk, and to guide future management efforts. Third generation instruments are theoretically informed in some capacity (e.g., general personality and cognitive social learning theory of criminal behavior, or GPCSL; Andrews & Bonta, 1994-2010; Bonta & Andrews, 2017), they can inform service intensity and foci, and they have the capacity to explicitly assess changes in risk through treatment or other change agents. Measures such as the LSI-R and HCR-20 (and concordant SPJ measures) arguably characterize the third generation (although Andrews et al., 2006 labeled SPJ measures a variant on the first generation). Fourth generation instruments go a step further and can be used from the point of intake to case closure (Andrews et al., 2006), and specifically, as dynamic tools, they can drive the nature and foci of services, as well as assess change (Campbell et al., 2009). Examples of fourth generation tools include the Level of Service/Case Management Inventory (LS/CMI; Andrews et al., 2004), VRS and its sexual offense specific counterpart, the Violence Risk Scale-Sexual Offense version (VRS-SO; Wong et al., 2003-2017), and we would also suggest the HCR-20 Version 3, in light of its distinction between the prevalence vs. risk relevance of items to aid case formulation and risk management.

The aforementioned discussion regarding the classification and generational framework of risk tools refers to scales specifically designed to predict reoffending, which are criterion-referenced measures (Helmus & Babchishin, 2017) and they have their own unique psychometric literature (Hanson, 2022). Additionally, other tools (i.e., norm-referenced measures) are frequently used in risk assessment external to this generational or typological framework. Hare's Psychopathy Checklist-Revised (PCL-R; Hare, 1991, 2003), and its variants, such as the Screening Version (PCL: SV; Hart et al., 1995) and Youth Version (PCL: YV; Forth et al., 2003), are clinical rating scales that measure a psychological construct with substantial criminal justice relevance and which are frequently used in forensic risk assessment. A large literature supports the predictive properties of the PCL scales for institutional behavior problems (Olver, Stockdale, et al., 2020), future violence (Campbell et al., 2009;

Yang et al., 2010), general recidivism (Leistico et al., 2008), and failure on conditional release (Hart et al., 1988), among other outcomes. Hare (2003) warns that the PCL-R is not intended to be used as a standalone tool, but in conjunction with other measures as part of a balanced comprehensive risk assessment.

### ***Discrimination and Calibration***

A further critical issue is evaluating the threshold of scientific and clinical acceptability of effect size magnitudes for the prediction of different recidivism outcomes from extant findings, and the amount of disparity that is permissible in predictive accuracy for Indigenous and non-Indigenous groups. Two prediction concepts here are essential: discrimination and calibration.

***Discrimination (Relative Risk).*** Discrimination or relative risk refers to the accuracy with which recidivists can be differentiated from non-recidivists on the basis of test (risk) scores; a tool with good properties of discrimination will be able to accurately assign recidivists higher risk scores than non-recidivists well beyond chance level (e.g., flip of a coin). Common prediction metrics used to examine the discrimination properties of risk tools include receiver operator characteristic analyses (ROC) that generate an area under the curve (AUC) statistic (ranging from 0 to 1.0), point biserial ( $r_{pb}$ ) or phi ( $\phi$ ) correlation (ranging from -1.0 to + 1.0), standardized mean difference (Cohen's  $d$  or equivalent), and risk ratios from logistic regression or Cox regression survival analysis (which tend not to have a defined range). Applying the conventions from Cohen (1992), Rice and Harris (2005) provide the following effect size language guidelines for the interpretation of prediction metrics: small/low AUC = .556,  $r_{pb}$  = .10, and  $d$  = 0.20; medium/moderate AUC = .636,  $r_{pb}$  = .24,  $d$  = 0.50, and large/high AUC = .714,  $r_{pb}$  = .37,  $d$  = 0.80. The prediction metrics have their relative strengths and weaknesses.

For instance, AUCs are ranked ordered statistics that are relatively unaffected by fluctuating base rates, however, they will be attenuated by low variance in scores (i.e., range restriction; see Howard, 2017). AUCs can be interpreted as the percent probability that a randomly selected recidivist

will have a higher score on a given risk measure than a randomly selected nonrecidivist. For instance, an AUC of .50 would indicate no more accuracy than the flip of a coin; an AUC of 1.0 would indicate perfect predictive accuracy (i.e., all recidivists can be accurately identified on the basis of a given risk score), and values below .50 would indicate negative predictive accuracy. So, a medium AUC of .65 would translate into a 65% probability that a randomly selected recidivist would have a higher risk score than a randomly selected nonrecidivist. Further,  $d$  represents the size of the difference in test scores in standard deviation units between recidivists and non-recidivists (e.g., if a risk tool has  $d = 0.50$  for predicting future violence and an  $SD = 8$ , then violent recidivists would score 4 points (i.e.,  $0.50 \times 8$ ) higher on average than non-recidivists, a medium effect). As with AUCs,  $d$  value magnitudes are not impacted by extreme base rates, and rather their computation and standard error directly incorporates base rate information.

Finally,  $r_{pb}$  is a correlation between continuous test scores and binary recidivism ( $\phi$  would equal the correlation between dichotomous risk categories and binary recidivism). Rosenthal's (1991) binomial effect size display is often used to interpret correlations in risk assessment research (as opposed to squaring the correlation to obtain an estimate of variance accounted for). For instance, a correlation of .20 could be interpreted as representing a 20% difference in recidivism base rates between groups scoring high vs. low on a given measure. For instance, if the base rate was 25%, the higher scoring group would have a 35% base rate of recidivism, and the low group, a 15% base rate; this is an interpretive heuristic that should be taken cautiously, however, given that dichotomizing risk scores into extreme groups and using this to make binary predictions is not how these measures are (or should be) employed in assessment practice, given that doing so reduces variance, oversimplifies communication of the individual's overall risk level, and may undermine other important risk management functions (e.g., treatment planning, graduated release) (Helmus & Babchishin, 2017). Moreover, although correlations are a common metric that are readily understood, they will be attenuated by loss of variance as well as

extreme base rates (i.e., very high or very low) (Babchishin & Helmus, 2016). For instance, an  $r_{pb}$  of .10 with a 50% recidivism base rate would equate to a  $d = 0.20$ ; however, with a recidivism base rate of 10%, an  $r_{pb}$  of .10 would equate to a  $d = 0.34$  (Rice & Harris, 2005).

There are essential questions to be considered when evaluating the adequacy of the predictive properties of these measures for Indigenous persons. First, what level of predictive accuracy (discrimination) is required for use of a tool with justice involved Indigenous persons? Do the effect sizes have to be large, or will medium/moderate suffice? Second, what degree of disparity is acceptable for Indigenous and non-Indigenous persons? If certain risk tools yield larger effect size magnitudes for non-Indigenous persons than Indigenous persons, when does this become a concern? If the disparities are small and not significant, or still fall within an effect size category (e.g., upper end of medium vs. lower end of medium), is this of concern? Results of meta-analyses of the discrimination properties of risk assessment tools demonstrate that there will be effect size variability and the effect size magnitudes for a given instrument can range from small to large in any single study depending on moderators such as sample, setting, definition of the recidivism criterion variable, base rate, length of follow-up, information quality, rater training, scoring accuracy, and amount of variability in test scores (re: range restriction) (Helmus et al., 2021; Yang et al., 2010).

Moreover, across instrument and outcome, different risk instruments (including the PCL-R) all generally yield broadly moderate or medium aggregate effects in the prediction of violent recidivism (Campbell et al., 2009; Olver et al., 2009; Singh et al., 2011; Yang et al., 2010), general recidivism (Gendreau et al., 1996; Gendreau et al., 2002; Leistico et al., 2008; Olver et al., 2009, 2014; Wilson & Gutierrez, 2014), sexual recidivism (Hanson & Morton-Bourgon, 2009), institutional violence and general misconducts (Campbell et al., 2009; Guy et al., 2005; Olver et al., 2020), and IPV recidivism (Hanson et al., 2007; Helmus & Bourgon, 2011; van der Put et al., 2019). Given that different risk measures will have item content specific to a target population (e.g., sexual offending, violence) to drive service planning

and risk management, any given tool should be used for its intended purpose.

**Calibration (Absolute Risk).** Calibration or absolute risk concerns the rates of recidivism associated with risk scores or groups of scores. These are unique to actuarial risk scales given that they are the only method that include recidivism probability estimates (Hanson & Morton-Bourgon, 2009). Calibration can be evaluated via logistic regression in which continuous risk scores are used to predict a given outcome, and the resulting regression model contains a constant ( $B_0$ , the log odds of recidivism for those with a score of 0 on all predictors) and one or more predictor variable regression coefficients ( $B_i$ ) for a risk score, which represent the increase in the logged odds of recidivism between adjacent scores on the predictor. Logistic regression also generates an odds ratio which can be interpreted in a similar manner. The information from logistic regression can then be used to model trajectories of recidivism at specific scores (e.g., every possible LSI-R score) through the logistic function. This has an important advantage given that actual observed rates of recidivism tend to be unstable and will fluctuate considerably at extreme and infrequent scores. Logistic regression, by estimating the rates of recidivism predicted with a given score, can generate accurate projections of recidivism for all scores. Logistic regression has been used to examine the calibration properties of risk measures for Indigenous persons, specifically, by comparing the estimated trajectories of recidivism predicted by risk scores and then comparing the curves between Indigenous and non-Indigenous persons, including the Static-99R and Static-2002R (Lee et al., 2020), VRS-SO (Olver, Sowden, et al., 2018), LS/CMI (Wilson & Gutierrez, 2014), and PCL-R (Olver, Neumann et al., 2018) as will be discussed in further detail.

Calibration can also be examined via the expected/observed (E/O) Index, which represents the extent to which recidivism rates associated with risk scores expected from a normative sample correspond to those observed in a different sample or setting (Hanson, 2017). E/O index values above 1.0 represent overestimation of projected recidivism rates from the normative sample, while values under 1.0 represent an underestimate of recidivism rates; values of 1.0 represent perfect calibration.

The E/O index is significant when the 95% CI does not overlap with 1.0, indicating concerns with calibration between expected and observed recidivism rates. Although research has been more limited using the E/O index to compare recidivism rates between Indigenous and non-Indigenous persons as a function of risk scores, it has been employed with the Ontario Domestic Assault Risk Assessment (Hegel et al., 2021) and the LSI-Saskatchewan Youth Edition (Prince et al., 2021).

### ***Predictive Properties of Risk Tools with Indigenous Persons: Results from Previous Meta-Analyses***

With these prediction concepts in forensic risk assessment defined, a further central issue is what prior reviews have demonstrated about the predictive properties of extant measures with Indigenous persons. In this line of research, predictive invariance (Hart, 2016) or predictive equivalence can be examined through formal comparison of prediction effect sizes for a given risk measure and outcome between ethnocultural groups. Predictive equivalence or invariance could be demonstrated by non-significant ethnoracial differences in effect sizes in addition to both effect sizes being within the same class or magnitude, either at the individual study level or through meta-analysis. Increasing the sample size and length of follow-up, or (in the case of meta-analysis) number of independent investigations, can add rigor. This stands in contrast to other traditional forms of measurement invariance testing, such as through multigroup confirmatory factor analysis and item response modelling, which examines invariance in the structural properties of an instrument across groups while imposing increasing constraints such as equivalence in factor loadings, item thresholds, or residual item variances (Putnick & Bornstein, 2016). While traditional forms of invariance testing have formalized options of adding constraints to test the equivalence of measurement properties between groups (e.g., metric, scalar, strict), meta-analysis is limited by the nature and number of moderators that can be imposed or other design features to increase rigor (e.g., limiting to within study comparisons). Additionally, the 2014 Standards for Educational and Psychological Testing from the Joint Committee of the American Educational Research Association, the American Psychological Association, and the

National Council on Measurement in Education emphasize that the types of reliability and validity evidence that are needed to justify the use of a tool will depend on the purpose of using the tool (Joint Committee, 2014). Given that the purpose of risk assessment tools is to assess the likelihood of a future outcome, validation efforts (including cross-cultural comparisons) should primarily be based on predictive accuracy, as opposed to structural equivalence (Hanson, 2022; Helmus & Babchishin, 2017).

The literature to date has featured meta-analyses of the discrimination properties of two sets of risk tools for use with Indigenous persons; the Level of Service (LS) Inventory measures, a family of dynamic general risk-need tools, and the Static-99R and Static-2002R, which are two static empirical actuarial sexual offense specific risk assessment measures. Both sets of measures are described in detail in the Method and summarized in Table 1. Three previous meta-analyses have examined the predictive properties of the LS measures as a function of Indigenous and non-Indigenous ancestry (see aggregate summary table of all studies [supplementary Tables S1 and S2] and Table 4 for greater detail). First, Olver et al. (2009) conducted a meta-analysis of youth variants of the LSI, the SAVRY and youth variants of the PCL-R. Sufficient  $N$  and  $k$  were available to meta-analyze studies that examined YLS/CMI and LSI-Sk among Indigenous and non-Indigenous groups. Across five studies (all from Canadian samples), Olver et al. (2009) found the youth variants of the LSI to have significant and equivalent moderate in magnitude predictive effects for general recidivism among Indigenous ( $r_w = .35$ ,  $d = 0.75$ ,  $n = 860$ ) and non-Indigenous youth ( $r_w = .32$ ,  $d = 0.68$ ,  $n = 462$ ). Second, Olver et al. (2014) subsequently conducted a large-scale meta-analysis of the LS measures as a function of moderators such as gender and ethnocultural heritage. Among Indigenous samples specifically (Canadian, Australian, and US), the LS total score had moderate predictive accuracy for general recidivism ( $r_w = .29-.30$ ,  $d = 0.63-.66$ ,  $k = 13$ ,  $n = 5,354$ ), which was not significantly different from the White sample effect size ( $r_w = .29-.32$ ,  $d = 0.63-.70$ ,  $k = 24$ ,  $N = 40,989$ ).

Third, Wilson and Gutierrez (2014) conducted a meta-analysis of the LS measures across 16

samples ( $N = 64,322$ ) to draw comparisons between Indigenous (Canada, Australia, and US) and non-Indigenous subgroups in predictive validity for general and violent recidivism. In addition, the authors only included studies that made direct comparisons between Indigenous and non-Indigenous groups within the same study. As with Olver et al. (2014), LS total scores in general evinced moderate predictive validity for general recidivism overall in the Indigenous samples, with a comparable  $d = 0.62$  ( $k = 13$ ,  $n = 21,508$ ); the LS measures in general demonstrated moderate effects for violent recidivism ( $d = 0.49$ ). Wilson and Gutierrez (2014) did not report the ES for the non-Indigenous samples, but formally tested the magnitude of difference in effect sizes in the prediction of outcome; effect sizes were significantly higher for non-Indigenous samples for both general and violent (fixed effects models only) recidivism.

The Static-99R and 2002R have also undergone quantitative reviews of their predictive efficacy among Indigenous and non-Indigenous men with a conviction history of sex offenses (see supplementary Table S1 and Tables 12 and 13). Babchishin et al. (2012) conducted the first meta-analysis of the Static-99R and Static-2002R measures comparing predictive validity for Canadian Indigenous and non-Indigenous samples of men convicted for sexual offenses. The study received considerable discussion in Canadian legal proceedings on use of these tools with Indigenous men convicted for sex offenses and has been critiqued owing to the small number of available studies. Babchishin et al. (2012) found that Static-99R evinced high predictive validity for sexual recidivism across both Indigenous ( $AUC_w = .713$ ,  $n = 319$ ,  $k = 5$ ) and non-Indigenous ( $AUC_w = .736$ ,  $n = 1,269$ ,  $k = 5$ ) groups. By contrast, Static-2002R had small non-significant predictive effects for sexual recidivism with Indigenous men ( $AUC_w = .608$ ,  $n = 209$ ,  $k = 3$ ) but large prediction effects for non-Indigenous men ( $AUC_w = .759$ ,  $n = 955$ ,  $k = 3$ ).

Lee et al. (2020) conducted an expanded and updated meta-analysis of the predictive properties of Static-99R and 2002R among samples preselected for being unusually high risk for sexual recidivism. The meta-analysis partially overlapped with Babchishin et al. (2012), incorporating 2 of the smallest 5

samples from the previous meta-analysis, combining a small treated sexual offending sample with the much larger combined sample from Olver, Sowden, et al. (2018), and adding 2 new samples. Static-99R and Static-2002R data were included for 5-year ( $k_s = 5$  and  $3$ , respectively) and 10-year ( $k_s = 3$  and  $2$ , respectively) sexual recidivism in Canadian Indigenous and White adult male samples. Among Indigenous men, Static-99R had a small effect for 5-year sexual recidivism ( $AUC_w = .61$ ,  $n = 599$ ) and moderate effect for 10-year sexual recidivism ( $AUC_w = .66$ ,  $n = 375$ ), while for the White samples it had a strong moderate effect for 5-year sexual recidivism ( $AUC_w = .70$ ,  $n = 1,375$ ), and a large effect for 10-year sexual recidivism ( $AUC_w = .72$ ,  $n = 848$ ). Differences in discrimination for Static-99R were not statistically significant between groups. As with Babchishin et al. (2012), predictive validity was much poorer for Static-2002R with Indigenous men for both 5-year ( $AUC_w = .56$ ,  $n = 173$ ) and 10-year ( $AUC_w = .61$ ,  $n = 101$ ) sexual recidivism, but considerably stronger for White samples in the prediction of both 5-year ( $AUC_w = .69$ ,  $n = 620$ ) and 10-year ( $AUC_w = .72$ ,  $n = 416$ ) outcomes. The difference in discrimination effect size magnitudes was significant for 5-year, but not 10-year, outcomes.

Finally, concurrent to the present study, Ahmed et al. (2023) conducted a meta-analysis of the Static-99R's predictive validity across multiple ethnocultural groups of persons with histories of sex crimes. A small to medium  $AUC_w = .635$  ( $k = 10$ ,  $n = 2,043$ ) was found for Indigenous samples in the prediction of sexual recidivism, and a medium  $AUC_w = .682$  was obtained for White or non-Indigenous samples ( $k = 9$ ,  $n = 13,756$ ). Within study comparisons demonstrated that Indigenous samples scored significantly higher on the Static-99R than White or non-Indigenous groups ( $d = .407$ ), however, the measure's predictive accuracy for sexual recidivism was significantly lower for the Indigenous groups ( $AUC_{diff} = -.0669$ ). Unlike Babchishin et al. (2012) and Lee et al. (2020), this meta-analysis, larger in scope, found significant differences in AUC prediction magnitudes between Indigenous and non-Indigenous groups.

We delve into much greater detail on the issue of ethnoracial differences in predictive accuracy

of risk tools later, but briefly, there are two points worth emphasizing. First, pragmatically, Static-99R scores predict the targeted outcome of sexual recidivism; however, its lower predictive accuracy means that Static-99R scores reflect more error in assessing risk for sexual violence for Indigenous persons. The heavy weighting of criminal history variables (on which Indigenous persons are overrepresented) mean that higher Static-99R scores also partly reflect the culmination of systemic and historical factors that contribute to Indigenous overrepresentation in the criminal legal system (Ahmed et al., 2023). Second, the results speak to the necessity of augmenting static tools with measures that are dynamic, changeable, and which include a greater breadth of risk and need areas. The addition of dynamic tools may strengthen predictive accuracy but doing so is also more meaningful in the ability to inform treatment, recidivism prevention, and successful reintegration; all of which necessitates formal meta-analytic examination of the predictive properties of a range of tools.

### **Current Study**

Indigenous persons are overincarcerated worldwide and overrepresentation has either remained stable in certain jurisdictions (New Zealand) or only continued to increase in others (Canada, US, Australia). Risk assessment measures exist to prioritize and drive treatment and risk management services to inform release and reintegration decisions; their use is critical within correctional and forensic mental health systems. Invariably, Indigenous persons within these systems will have one or more of these measures administered to them which will be used to make critical decisions about sentencing, security classification, programming, conditional release, and community reintegration. It is critical to understand the psychometric properties of these tools, including their strengths and shortcomings, with a vulnerable population that has been historically and repeatedly oppressed, victimized, and marginalized in order to evaluate the utility of these measures and to assist in decision making regarding their potential use with Indigenous persons.

Existing quantitative reviews to date have focused on a narrow selection of risk tools—the LS

and Static-99R/2002R measures—however, given the nature and volume of general and specific measures used, a comprehensive review is vital. There have been legal calls for further research, for instance, in the Supreme Court of Canada where they write, “what is required, at a minimum, is that if the CSC wishes to continue to use the impugned tools, it must conduct research into whether and to what extent they are subject to cross-cultural variance when applied to Indigenous offenders” (*Ewert v. Canada, 2018*, p. 199). Similar calls have been advanced for cross cultural validation research of psychological assessment measures with Indigenous peoples by other groups, including scholars (Gutierrez et al., 2016; Hart, 2016; Shepherd & Lewis-Fernandez, 2016) and professional bodies (e.g., Canadian Psychological Association; CPA, 2018a).

There is also a further need to identify and better understand the primary risk-need domains of crime and violence with Indigenous persons, which has implications for the role and relevance of the item content and underlying latent constructs of risk assessment tools (Gutierrez et al., 2016; Shepherd & Lewis-Fernandez, 2016). As part of this process, there is an urgent need to examine strengths, protective factors, resiliencies, other risk mitigating agents (e.g., treatment change, see CPA, 2018b), and culturally relevant factors, all of which can inform assessment, intervention, and risk management. Taken together, a systematic review and meta-analysis of the empirical research literature examining the predictive properties of risk assessment and protection measures, risk-need domains, and potentially culturally relevant risk factors with Indigenous persons, could promote further understanding of current approaches and controversies. This is viewed as an important first step in a larger collaboration seeking to promote culturally safe and responsive assessment practices.

## **Method**

### **Selection of Studies**

We conducted a systematic review and meta-analysis of the risk assessment and recidivism prediction research literatures with justice involved Indigenous persons. Our search followed the

Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Moher et al., 2009) guidelines (see Figure 1) in locating, reviewing, and excluding sources obtained from the search. A coordinated search was conducted of major scholarly repositories, both national and international, including PsycINFO, Web of Science, Google Scholar, iPortal (Indigenous studies portal research tool), Criminal Justice Abstracts, and Proquest Dissertations and Theses selecting different combinations of search terms from the following list: *Indigenous, Aboriginal\*, Native American, American Indian, First Nations, Métis, Torres Strait, Sami, Maori, offen\*, risk, risk assessment, risk prediction, predictive validity, protective factor, culture, recid\*, Canad\**.<sup>3</sup> Additional sources included documents culled from meta-analyses, research syntheses or review papers (including their reference sections), and individual empirical works examining the predictive properties of risk factors, risk assessment measures, and related structured assessment measures with Indigenous persons.

The following inclusion criteria were employed: 1. the focus needed to be Indigenous persons; 2. studies needed to be empirical (i.e., data driven); 3. the predictive association of risk factors or a risk measure with recidivism was examined in some form; 4. predictive validity findings were presented separately for Indigenous and non-Indigenous (usually White) persons; 5. the study could be a quantitative single study, multi-sample study, or meta-analysis; 6. individual single studies of risk assessment measures that had been included in a meta-analysis drawing Indigenous/non-Indigenous comparisons were extracted and reaggreated with additional works obtained from the search through updated meta-analyses.

In total, 91 studies from 88 documents were retrieved that met the above inclusion criteria. Four additional sources included in the qualitative synthesis that identified pertinent culturally relevant

---

<sup>3</sup> The search included all international sources, however, the term "Canad\*" was also included as a term for some searches due to the fact that the major risk assessment measures were developed in Canada and imported to international jurisdictions, as well as given the volume of Canadian research and overrepresentation of Canadian Indigenous persons in correctional settings.

risk factors but did not examine associations with outcome were also obtained. Other pertinent works are cited and reviewed although these may not be empirical in nature or they may provide other helpful information and insights for the purpose of the review. Of the 91 studies, 57 studies were refereed journal articles (12 featured secondary analysis of unpublished data), 17 were government reports, 11 were unpublished graduate theses or dissertations, 4 came from unpublished data or manuscripts under review, and 2 were conference presentations; the relatively high volume of unpublished sources, in part, reflected the type of outlet for certain groups of tools. For instance, research on jurisdiction specific measures such as the Dynamic Factors Identification and Analysis-Revised (DFIA-R) and the Statistical Information on Recidivism (SIR) scale, are more likely to be disseminated as government reports. The four additional documents included in the qualitative synthesis were one each of a government report (Boer et al., 2003), doctoral thesis (Gutierrez, 2018), undergraduate thesis (Kyne, 2022), and refereed journal article (Shepherd & Willis Esqueda, 2017).

The mean and median years of the studies were 2012 and 2013, respectively; as seen in Table 2, there has been a substantial uptick in risk assessment research in Indigenous samples within the past 20 years. Most studies ( $k = 48$ ) were field prospective (i.e., ratings completed in field as routine service delivery and individuals followed up in real time), followed by archival retrospective ( $k = 23$ ) (i.e., ratings completed from past institutional file blind to outcome on individuals currently released), mixed archival/field ( $k = 7$ ) (i.e., hybrid sample of archivally coded and clinically rated cases), meta-analyses ( $k = 7$ ), and face-to-face prospective ( $k = 5$ ) (i.e., ratings completed on individuals face-to-face in a research context and followed up in real time). Finally, most studies ( $k = 69$ ) featured Canadian samples, followed by the US ( $k = 10$ ), Australian ( $k = 8$ ), and multi-country, although the latter were all meta-analyses conducted by Canadian researchers.

### **Measures**

Twenty-two risk assessment (or related structured assessment) measures and 15 risk factor

predictor domains were reviewed, focusing primarily on research that has examined their predictive properties for various recidivism outcomes with Indigenous and non-Indigenous (predominantly White) correctional and forensic populations. The risk measures are organized from general in scope and purpose (general risk-need) to more specialized (e.g., sexual offense, violent, or IPV specific). Each measure is reviewed briefly as follows; as noted previously, Table 1 provides a summary of all measures, including representative indices of reliability.

### ***Level of Service (LS) Inventory Measures and their Youth Variants***

**LS Adult Variants.** The Level of Service Inventory and its variants refer to a family of general risk-need measures (herein referred to as the LS scales) with item content broadly organized into the Central Eight risk-need domains, consisting of one static/historical domain (Criminal History), and seven dynamic criminogenic need areas of Antisocial/Procriminal Attitudes, Antisocial Peers/Associates/Companions, Antisocial Personality Pattern, Employment/Education, Family/Marital, Substance Abuse (or Drug/Alcohol), and Leisure/Recreation. Additional domains may be covered by some LS variants such as Financial, Accommodations, and Personal/Emotional. Items are rated in a binary manner and summed to generate risk-need totals for each domain, as well as an overall total risk score, that is in turn, arranged into five risk bands with different recidivism probabilities; most typically reincarceration within one year of release. As such, the LS measures are an empirical-actuarial set of tools with dynamic items. The LS measures are founded on a general personality and cognitive social learning (GPCSL) theory of criminal behavior and, per the RNR model, are designed explicitly to link risk assessment with treatment planning, service delivery, and case management (i.e., to drive services to prevent recidivism). The variants of the LS measures include additional sections for summarizing primary areas of risk-need, overall risk level, other specific areas of risk, strengths, and intervention and case management considerations. Several variants of the LS measures have been developed which include the Level of Service Inventory-Revised (LSI-R; Andrews & Bonta, 1995), Level of Service Inventory-Ontario Revision

(LSI-OR; Andrews et al., 1996) and its commercially available counterpart, the Level of Service/Case Management Inventory (LS/CMI; Andrews et al., 2004), the LSI: Screening Version (LSI: SV) and Self-Report (LSI: SR; Motiuk et al., 1992).

**LS Youth Variants.** LS youth variants have been developed with the Youth Level of Service/Case Management Inventory (YLS/CMI; Hoge & Andrews, 2003), and Level of Service Inventory: Saskatchewan Youth Edition (LSI-Sk; Andrews et al., 2000) having been the most thoroughly examined in comparative evaluations of Indigenous and non-Indigenous youth. The measures are youth adaptations of the LS/CMI and LSI-OR, respectively, and structured around the Central Eight with developmental emphases to reflect family circumstances and parenting (in lieu of Family/Marital), education (in Education/Employment), and antisocial pattern (e.g., verbal and physical aggression).

#### ***Offender Risk Assessment Management System (ORAMS)***

The ORAMS (Government of Saskatchewan, 1997) is a 15-item actuarial general risk-need measure developed by the Government of Saskatchewan to assess general recidivism risk in men and women under provincial corrections jurisdiction. The items are differentially scaled and comprise a collection of static (e.g., age, criminal history), and dynamic (e.g., family/marital, emotional stability, employment) indicators that broadly converge with the Central Eight. Given its close convergence with the LS measures, the ORAMS results are presented in tandem with this group of measures.

#### ***Arizona Risk Need Assessment (ARNA) Instrument***

The ARNA (LeCroy et al., 1998) is a 10-item general risk-need actuarial tool for justice involved youth developed by a research team in collaboration with the juvenile justice system in Arizona. Items are binary and reflect a combination of offense history and single indicators reflective of Central Eight domains (e.g., peer delinquency, substance use). With total scores ranging from 0 to 10, youth are grouped into low, medium, and high levels of risk-need (Schwalbe, 2009).

#### ***Statistical Information on Recidivism (SIR) scale***

The Statistical Information on Recidivism (SIR) scale is a 15-item static actuarial recidivism prediction tool to guide release decision making for the CSC. Items are a combination of criminal history, sentencing, and demographic variables that are statistically weighted to optimize prediction of general recidivism, although SIR scores also predict violent and sexual recidivism to a lesser degree (Bonta et al., 1996; Hanson & Morton-Bourgon, 2009). Possible scores range from -30 to +27 and are organized into five risk levels representing the likelihood that the individual will not be returned to federal custody within three years of release, ranging from Very Good (4 in 5) to Poor (1 in 3). The SIR scale is administered to every non-Indigenous man admitted to Canadian federal custody upon intake; it is not administered to women or Indigenous men. The SIR scale has undergone minor revision over the years from the original General Statistical Information on Recidivism (or GSIR; Nuffield, 1982), to the SIR-R1 and its parallel form, the SIR-Proxy, developed by Nafekh and Motiuk (2002) for research purposes with federally sentenced Indigenous men and women. The SIR-Proxy has analogous and isomorphic item content to the SIR-R1, and Nafekh and Motiuk (2002) found that it correlated at .90 for raw scores and .85 for risk groupings, and predicted slightly better than the SIR-R1 for non-Indigenous men (AUC = .752 vs. .745 for general recidivism, respectively, and .726 vs. .708 for violence, respectively).

### ***Static Factors Analysis (SFA) and Dynamic Factors Identification and Analysis-Revised (DFIA-R)***

**SFA.** The Static Factors Analysis (SFA; Motiuk, 1993) and Dynamic Factors Identification and Analysis-Revised (DFIAR-R; Brown & Motiuk, 2005) refer to a pair of CSC-developed SPJ-based risk assessment tools. The measures are intended to inform correctional planning and service delivery and are administered to all men and women admitted to Canadian federal correctional custody. The SFA consists of 137 static or historical risk indicators scored in a binary manner, arranged into three subscales: Criminal History Record (CHR, 38 indicators reflecting density of criminal history), Offence Severity Record (OSR, 71 indicators reflecting harms to victims/society from offending behavior), and (for select inmates), the Sex Offence Severity Checklist (28 indicators). Based on indicators present

across the domains, SPJ is used to assign a static summary risk rating (SRR) of low, moderate, or high.

**DFIA-R.** The DFIA-R comprises 107 dynamic risk indicators organized into eight need domains analogous to the Central Eight: Employment, Marital/Family, Associates, Substance Abuse, Community Functioning, Personal/Emotional, Attitudes, and Responsivity. Each need domain is rated on a five-point ordinal scale of asset, no need, some need, moderate need, and considerable need (except Substance Abuse and Personal/Emotional, which do not have an “asset” option). Operationalized guidelines are followed to integrate domain ratings into an SPJ summary risk-need rating of Low, Medium, or High.

***Service Planning Instrument (SPIn) and Youth Assessment and Screening Instrument (YASI)***

**SPIn.** The Service Planning Instrument (SPIn; Orbis Partners, 2003) is a general risk-need tool intended to assess risk and identify treatment needs and strengths for men and women in correctional settings. The SPIn is comprised of 11 static (e.g., criminal history) and dynamic (e.g., substance abuse, employment, attitudes, social-cognitive skills, mental health) risk-need domains that converge with the Central Eight, and it can be administered in a 35-item pre-screen or 90-item full assessment format. Items are scored on a 6-point Likert scale and summed to generate total scores that can be arranged into low, medium, or high risk bands. The pre-screen is typically initially administered at the front end of service delivery to determine if a fuller assessment is needed, and thus, psychometric research has featured the former.

**YASI.** The Youth Assessment and Screening Instrument (YASI, Orbis Partners, 2000) is a counterpart of the SPIn intended for risk assessment and service planning with justice involved youth. The measure has the same platform and covers similar item content across nine risk and strength domains. The YASI consists of a 34-item Pre-Screen variant administered initially to ascertain if a full assessment (comprising 90 items across 10 domains) is warranted, with psychometric research again tending to feature the former.

***Violence Risk Scale (VRS), Sexual Offense Version (VRS-SO), and Youth Version (VRS-YV)***

**VRS and VRS-SO.** The Violence Risk Scale (VRS; Wong & Gordon, 1999-2003) and the Violence Risk Scale-Sexual Offense version (VRS-SO; Wong et al., 2003-2017) refer to a pair of dynamic risk assessment and treatment planning tools designed to assess risk for general violence or sexual violence in adult populations, respectively, identify targets for intervention, evaluate readiness to change, and assess changes in risk from treatment or other change agents. The VRS measures have been identified as fourth generation tools (Campbell et al., 2009). The tools employ the same platform for evaluating risk and change, although the item content differs in some important respects to reflect the populations for which the tools were developed. Most items on the tools are dynamic, with a smaller subset that are static (VRS, 6 static, 20 dynamic; VRS-SO, 7 static, 17 dynamic). Each item is rated on a four-point (0, 1, 2, 3) ordinal scale with higher scores representing increased risk for violent or sexual recidivism. Static items represent criminal history and personal demographic variables, while the dynamic items represent potentially changeable aspects of social, psychological, and interpersonal functioning or environmental circumstances linked to risk. Total scores on the VRS (range 0-81) and VRS-SO (range 0-72) are assigned to one of five risk levels ranging from Level I or very low risk (VRS, 0-20, VRS-SO, 0-15) to well above average risk, Level IVb for the VRS-SO (range 51-72) or Level V for the VRS (range 65-81) (Olver, Mundt et al., 2018, 2022). Unique to the VRS measures is that they include a mechanism for assessing treatment readiness and change via a modified application of the stages of change model (Prochaska et al., 1992), which documents the cognitive, behavioral, and experiential processes as the individual addresses problem areas. While reductions in risk are credited with point deductions, deteriorations can be similarly captured through an increase in score; change ratings are then summed to yield a change score. Of note, the tools were developed on the Canadian Prairies in a maximum security correctional mental health facility (the Regional Psychiatric Centre, or RPC) that provided high intensity treatment and stabilization services to high risk, high need men and women. Due to their overrepresentation, especially in the Prairie Region, Indigenous men comprised anywhere from one third to one half of the

development and validation samples for the tools.

**VRS-YV.** The Violence Risk Scale-Youth Version (VRS-YV; Wong et al., 2004-2011) is designed to integrate violence risk assessment and intervention in youth and is a developmental adaptation of the VRS with 23 items (4 static/stable, 19 dynamic). It employs a the same four-point item scaling, structured change evaluation platform, and actuarial format as the parent measure. Total scores (range 0 -69) are numerically arranged into low (0-35), medium (36-49), and high (50+) levels of violence risk and the intensity of risk management services. Developmental considerations include parent/child relations, family stress, school attitude/performance, community stability, and exposure to antisocial role models/parenting. The VRS-YV was developed (Stockdale et al., 2014) and revalidated (Lovatt et al., 2022) on Indigenous majority samples of youth with convictions for violence or other serious offenses.

#### ***STABLE-2007***

STABLE-2007 (Hanson et al., 2007) is an empirical actuarial sexual recidivism risk tool with 13 stable dynamic items, and intended to be used in tandem with a static risk tool (e.g., Static-99R, Static-2002R; Brankley et al., 2017). STABLE-2007 is intended to identify targets for sexual offense risk reduction and management and to capture changes in risk with reassessment. The 13 items can be arranged into five domains: Significant Social Influences, Intimacy Deficits, General Self-Regulation, Sexual Self-Regulation, and Cooperation with Community Supervision. Items are scored on a three-point (0, 1, 2) ordinal scale and total scores are assigned to need levels of low (0-4), medium (5-11), or high (12-26).

#### ***Static-99R and Static-2002R***

Static-99R (Hanson & Thornton, 2000; Helmus et al., 2012) and Static-2002R (Hanson & Thornton, 2003; Helmus et al., 2012) are both static empirical actuarial sexual recidivism risk tools with statistically weighted items assessing sexual and nonsexual criminal history, victim characteristics, and perpetrator demographics. Static-99R has 10 items (total score range -3 to 12) while Static 2002R has 14

items (total score range -2 to 14) with expanded item content and different weighting in some domains. Total scores on each tool are numerically assigned to one of five risk levels ranging from Level I (very low risk, Static-99R range -3, -2, Static-2002R range -2, -1) to Level IVb (well above average risk, Static-99R range 6-12, Static-2002R range 7-14). With over 50 replications and more than a dozen translations (Helmus et al., 2022), Static-99R is the most frequently used sexual recidivism risk assessment tool in the field (Kelley et al., 2020; Neal & Grisso, 2014).

### ***Ontario Domestic Assault Risk Assessment (ODARA)***

The Ontario Domestic Assault Risk Assessment (ODARA; Hilton et al., 2004) is a 13-item static empirical actuarial risk tool originally developed to appraise risk for future male-to-female intimate partner violence (IPV). It is frequently used by police first responders to assess IPV risk on domestic violence calls. It contains 13 binary items that capture elements of criminal and specifically IPV history as well as perpetrator and victim characteristics and circumstances. Possible scores range from 0-13 and are arranged into 7 risk bins based on scores of 0, 1, 2, 3, 4, 5 or 6, and 7+.

### ***Spousal Assault Risk Assessment guide (SARA)***

The Spousal Assault Risk Assessment (SARA; Kropp et al., 1995; Kropp & Hart, 2015) is a clinician administered SPJ risk assessment for future IPV. The second edition of the SARA (Kropp et al., 1995) has marshalled the most psychometric research and is a 20-item tool with a combination of static and dynamic items organized into four sections: criminal history, psychosocial adjustment, spousal assault history, and current offense. The SARA third edition (SARA-V3; Kropp & Hart, 2015) contains similar item content and is arranged into three sections: Nature of IPV, Perpetrator Factors, and Victim Vulnerability Factors. Clinical applications of the SARA involve using SPJ methods to arrange ratings into low, medium, and high IPV risk bands, although most predictive validity research has examined the tool in a mechanical manner by summing the ratings to generate a total score (Helmus & Bourgon, 2011).

### ***Structured Assessment of Violence Risk in Youth (SAVRY)***

The Structured Assessment of Violence Risk in Youth (SAVRY; Borum et al., 2006), is a 30-item SPJ-based violence risk assessment tool for youth. It consists of 24 risk items arranged into three sections: Historical (10 largely static items such as history of violence, stability of upbringing), Social/Contextual (6 dynamic items such as coping, peer influences, and parental relations), and Individual/Clinical (8 dynamic items such as anger problems, attitudes, substance use, school problems). There is a fourth Protect section for protective factors (6 items that mitigate risk such as social supports, prosocial attitudes, positive attachments). The 24 risk items are rated on a 3-point (0, 1, 2) ordinal scale corresponding to what degree the item is present, while the Protective domain is scored in a binary (0, 1) manner. As an SPJ tool, SAVRY item ratings inform a summary risk rating (SRR) of low, medium, or high violence risk using professional judgement. Psychometric research has examined the tool both in terms of SRR ratings (i.e., the SPJ approach) as well as numeric scores generated from summing the items and thus employing a “mechanical” approach.

***Structured Assessment of Protective Factors (SAPROF) and the SAPROF-Youth Version (SAPROF-YV)***

The Structured Assessment of Protective Factors (SAPROF; de Vogel et al., 2009) is a 17-item clinician rated measure of protective factors arranged into three sections of internal (5 items, e.g., self-control, coping), motivational (7 items, e.g., life goals, motivation for treatment), and external (5 items, e.g., social network, living circumstances). The SAPROF-Youth Version (SAPROF-YV; de Vries Robbé et al., 2015) is a 16-item youth adapted variant of its adult counterpart arranged into four subsections of resilience (4 items, e.g., coping, self-control), motivational (6 items, e.g., for treatment, school/work), relational (3 items, e.g., parents/guardians), and external (3 items, e.g., professional care) domains. Each item is scored on a three-point scale (0 absent, 1 possibly/partially present, 2 present), with higher item, subscale, and total scores representing a greater number and density of protective (and hence, risk mitigating) factors. As SPJ measures, SAPROF and SAPROF-YV item ratings can be arranged into an SRR although the psychometric properties of the tools have been evaluated in a mechanical manner (e.g.,

Koh et al., 2020; Kleeven et al., 2020). The SAPROF and SAPROF-YV are each intended to be used with a risk assessment tool (e.g., HCR-SO, VRS for the SAPROF; SAVRY, VRS-YV for the SAPROF-YV), to identify protection enhancing/risk mitigating considerations for recidivism prevention.

***Psychopathy Checklist-Revised (PCL-R) and its Youth Version (PCL: YV)***

**PCL-R.** The Psychopathy Checklist-Revised (PCL-R; Hare, 1991, 2003) is not a risk tool but assesses a personality disorder construct that is overrepresented in criminal justice populations, has forensic and psycholegal relevance, and is associated with varied antisocial behavioral outcomes. As such, it is heavily used in forensic risk assessment in tandem with other measures (e.g., Blais & Forth, 2014; Neal & Grisso, 2014). The PCL-R is a 20-item symptom construct rating scale designed to assess the features of psychopathy. Items are rated on a three-point ordinal scale of 0 (absent), 1 (possibly, partially present), or 2 (yes, present), from comprehensive file and interview information. Possible scores range from 0-40 with higher scores representing a greater density of psychopathic traits; the summation of item ratings would represent a mechanical application of the instrument for risk assessment purposes (e.g., percentile ranks). Traditionally, PCL-R total scores of 30 have been used to characterize psychopathy, although with a standard error of measurement around 3-4, modified cutoffs have been employed (e.g., scores of 25 for research purposes on file-based ratings). The PCL-R item content can be arranged into two factors (termed Factor 1 and 2) or four facets. Factor 1 is subdivided into two facets dubbed Interpersonal (e.g., superficial, grandiose, deceitful, manipulative) and Affective (e.g., callous, shallow affect, lack of remorse), while Factor 2 is arranged into two facets labeled Lifestyle (e.g., parasitic, lacks goals, impulsivity, irresponsibility) and Antisocial (e.g., poor behavioral controls, early behavior problems, juvenile delinquency, criminal versatility).

**PCL: YV.** The Psychopathy Checklist: Youth Version (PCL: YV; Forth et al., 2003) is also a 20-item symptom construct rating scale developed to assess features of psychopathy, but in youth populations; it is a developmental adaptation of the PCL-R with the same item scaling (0, 1, 2), factor structure

(correlated four factor model), and score range as the parent measure. Although the PCL: YV is also not a risk tool, the juvenile psychopathy construct has substantial criminal justice relevance and the PCL: YV is frequently used in risk assessment contexts with youth. Given the dynamism of youth and problems of labeling, there is no cut score on the PCL: YV to characterize juvenile psychopathy; rather, PCL: YV scores are employed in a mechanical manner and the measure is typically used alongside other measures of psychological functioning (Forth et al., 2003; Olver & Stockdale, 2010).

### ***Risk Factor Predictor Domains***

Fifteen risk-need domains, representing static or dynamic correlates of crime and violence for correctional and forensic populations were identified from our search and review of the literature. These domains tend to comprise the item content, or constellations of items, that are operationalized and scored on risk assessment and related structured forensic assessment measures, briefly described as follows:

**Criminal history (general).** This includes formal prior offense history of crime and violence, captured either through offense history counts or as operationalized through individual risk scale items and/or their composites (e.g., SFA, Static-99R).

**Antisocial personality pattern/Psychopathy.** Antisocial personality disorder (ASPD) and psychopathy are not one and the same (Hare & Neumann, 2010), however, the two have tended to be examined in tandem through quantitative reviews of the risk factor literature. This includes operationalizations such as the PCL-R or other indicators of a chronic antisocial lifestyle, persistent pattern of irresponsibility, and generalized trouble such as formal ASPD diagnosis, the antisocial personality pattern domain on the LS measures, or other item constellations capturing the personality and behavioral attributes of ASPD/psychopathy (e.g., VRS-SO, STABLE-2007).

**Education/employment.** This is captured through variables reflecting unemployment or low school achievement either measured directly or through education/employment risk-need domains on

measures such as the LS scales and the DFIA-R.

**Substance abuse.** Represents a problematic history of substance use either measured directly (e.g., diagnosis), through drug and alcohol risk-need domains on measures such as the LS scales and DFIA-R, or individual item indicators on select risk assessment measures (e.g., VRS-SO).

**Companions/peers/associates.** Presence of a network of friends or acquaintances involved in crime and violence and who otherwise serve as negative social influences, often captured on risk-need domains on measures such as the LS scales and DFIA-R, or individual item indicators on select risk assessment measures (e.g., STABLE-2007).

**Procriminal attitudes.** Represents the presence of values, attitudes or thinking patterns that legitimize crime in general or specific categories of antisocial behavior (e.g., rape myths, child sexual abuse distortions), as measured through general criminal attitude domains on the LS measures and the DFIA-R or item specific content on specialized measures (e.g., VRS-SO, STABLE 2000/2007).

**Community functioning.** Represents the adequacy of structures, supports, supervision, and resources in the community (e.g., accommodations, social supports, leisure, employment), including the willingness of the individual to utilize same, to promote successful reintegration and recidivism prevention. This tends to be captured by the community functioning domain of the DFIA-R as well as relevant item content from select measures (e.g., VRS-SO, STABLE-2007).

**Leisure/recreation.** Represents poor use of leisure time and existence of much unstructured time as most typically captured by the same named risk-need domain on the LS measures.

**Family/marital.** This domain captures family, marital, and/or broader relationship problems such as criminal spouse, family dysfunction, or a lack of emotionally intimate or stable relationships, as most typically measured by individual items or relevant constellations on select risk measures (e.g., VRS-SO, Static-99R/STABLE-2007 composite) or the family/marital domain on the LS measures and DFIA-R.

**Sexual deviance.** This is a primary risk-need domain for sexual offense recidivism and represents

a pattern of atypical or illegal sexual interests (e.g., paraphilias), sexual preoccupation and compulsivity, and sexually criminalized behavior (e.g., sexual assault, accessing child sexually explicit materials). Sexual deviance has operationalizations in the sexual violence risk assessment field that can be static (e.g., sexual offense history count, Static-99R item content) or stable dynamic (e.g., item composites from the VRS-SO and STABLE-2007).

**General risk-need.** Also known as antisocial orientation (Hanson & Morton-Bourgon, 2005), this is a constellation of usually dynamic general risk-need variables, consistent with the Central Eight, that reflect a broad propensity for rule violation as captured on risk measures through factor analysis (e.g., VRS-SO) or rational selection of items (e.g., Static-99R/STABLE-2007 composites).

**Personal/emotional.** This domain has some overlap with emotional distress as listed below, but is distinct as a specific risk-need domain reflecting serious mental health concerns (e.g., psychosis) or chronic distress that is present on certain variants of the LS measures (e.g., LSI-R) as well as the DFIA-R.

**Emotional distress.** Represents the presence of emotional concerns (e.g., anxiety, depression) such as formalized diagnoses or other measured indicators (e.g., self-report) of emotional distress.

**Prior victimization.** Represents a history of physical, sexual, verbal, and/or emotional abuse/neglect experienced in one's upbringing and/or lifetime.

**Cultural engagement/connectedness.** Prosocial connectedness to one's cultural heritage such as language and involvement in cultural customs, traditions, rituals, values, and practices. This domain is reviewed within the final section of the Results addressing culturally relevant risk factors for Indigenous persons along with the distinct but concordant concept of *cultural identity*, examined in the same study.

## Procedure

The results are organized into three sections: 1. risk assessment tools and related standardized forensic measures (e.g., PCL-R), arranged by tool and relevant predictive validity research (discrimination and calibration) for adults and youth; 2. risk factors or risk factor constellations (e.g.,

latent constructs or subscales of risk measures); and 3. culturally relevant risk factors as identified from sources obtained through the search. Updated meta-analyses of a single instrument or groups of instruments were conducted to aggregate effect sizes into a single metric to aid communication and comparisons between Indigenous and non-Indigenous groups. Individual studies were often reviewed in greater individual detail than would typically be done in a meta-analysis, so as to provide information about the sample and their nuances, consistent with the purpose of a systematic review. Supplementary Table S1 provides a detailed summary of all studies including sample, location, gender, ethnocultural group, *M* and *SD* across measures, recidivism base rates and criteria, and prediction effect sizes for all measures and criteria. Table S2 is an abridged version of Table S1 that provides a listing of all studies with their design and sample characteristics, minus numeric data, and serves as a much condensed and easier to navigate version for readers. Finally, Table 2 proper in the main document provides an aggregate descriptive summary of means and frequencies of the study descriptions of the information presented in the two supplemental tables (S1 and S2).

**Effect size coding.** Effect sizes were reported in their original metric, usually *d*, *r<sub>pb</sub>*, AUC (or the analogous Harrell's *C*), and then reported in a summary table converted to Cohen's *d* or standardized mean difference (i.e., between recidivists and nonrecidivists) using the tables and/or formulae from Rice and Harris (2005) for continuity among risk measures and their findings. When raw data were available (i.e., *M* and *SD* for recidivists and non-recidivists) as indicated for a given measure, *d* and corresponding SE were directly computed. For other metrics (e.g., Hazard ratios), the formula and source are cited where employed. Prediction metrics were reported for general, violent, and sexual recidivism, or IPV recidivism (for IPV instruments only). Although individual scale components were reviewed (static, dynamic, total, subscale scores), for the purposes of summary and aggregation, only the total score was reported, as the most comprehensive measure for the tool. When multiple effect sizes were reported within a category (e.g., 3- vs. 5- vs. 10-year recidivism), the largest and most inclusive N was chosen,

which was typically 3-year fixed follow-ups for general and violent recidivism and 5-year fixed follow-ups for sexual recidivism. If different operationalizations for an outcome (e.g., charges vs. convictions) were present within a given category, the effect size was averaged.

**Data aggregation and planned analyses.** A series of focused meta-analyses were conducted for each individual instrument in the prediction of a given recidivism outcome when the number of studies permitted (typically  $k \geq 3$ ). In some instances, when  $k = 2$  and effect size variability was low (e.g., ODARA) meta-analysis was used to generate an aggregate effect for illustrative and comparative purposes with other measures. Meta-analytic aggregations were most typically conducted using Comprehensive Meta-Analysis 3.0 (Borenstein et al., 1998-2014) using  $d$  and the SE, computed from the variance of  $d$  based on the number of recidivists and non-recidivists for a given effect. For our updated meta-analysis of the Static-99R and Static-2002R, meta-analysis was conducted examining AUCs and SEs, given how the findings were reported in the original sources, using MedCalc version 19.0.6 (MedCalc Software bvba, Ostend, Belgium, 2019) and the final effects were reported both in AUC units and their  $d$  equivalents. We chose to use  $d$  and AUC both owing to their practical interpretability and given that their magnitudes are not impacted by extreme base rates.

While fixed effects analyses are weighted by the inverse of the variance for the effect (which is primarily determined by sample size), this enables large samples to potentially disproportionately influence an aggregated effect size. Random effects analyses, however, also include a constant representing unexplained variation across studies; relatively greater weight is then given to smaller studies compared to the fixed-effects model, with the random-effects model more closely approximating the unweighted mean as variability across studies increases. Although the random-effects model is often conceptually preferable due to greater generalizability, it is unstable when the number of studies is small ( $< 30$ , Schulze, 2007), which was the case for nearly all analyses. In contrast, however, the greater the variability across studies (and in our analyses, moderate to large variability was

not uncommon), the more unrealistically narrow the confidence intervals from fixed-effect analyses are. Consequently, we reported both fixed- and random-effects analyses in the tables; however, for parsimony, in the main text of the Results, we report fixed effects given the instability of random effects for small  $k$ . Variability in findings across studies was reported using Cochran's  $Q$  statistic and the  $I^2$  effect size statistic (Borenstein et al., 2009). As a rough heuristic,  $I^2$  values of 25%, 50%, and 75% can be considered low, moderate, and high variability, respectively (Higgins et al., 2003). Significance testing between effect size magnitudes for Indigenous and non-Indigenous persons was conducted through the  $Q$  statistic between groups for Cohen's  $d$  values via CMA, and through the Delong test for independent AUCs via MedCalc.

Given the potential for publication bias, we conducted Orwin's fail safe  $N$  (Orwin, 1983) using CMA for the primary sets of findings on the aggregate tools for the three recidivism outcomes, calculating the number of studies with a mean  $d = 0.00$  that would be required to reduce the overall prediction effect size below the threshold of a small effect (i.e.,  $d = 0.20$ ) for each ethnocultural group. The FSN results are reported in two aggregate comparison summary tables of substantive predictive validity findings for Indigenous and non-Indigenous groups (Tables 11 and 12). We chose Orwin's fail safe  $N$  over tests of asymmetry of publication bias (e.g., funnel plot, trim and fill), given that these measures of require a minimum  $k$  (e.g., 10 studies per effect) to yield accurate results (Ioannidis & Trikalinos, 2007), which is greater than most of the individual effect sizes generated in this study.

As noted, planned analyses focused on the total score as the most comprehensive and frequently used measure for a given instrument (e.g., risk classification, treatment prioritization, release planning). For space considerations, we did not examine the subscales of every instrument, but rather prioritized the main risk-need domains identified in the extant literature, and which would increment and not duplicate existing efforts (e.g., Gutierrez et al., 2013). Given the importance of dynamic and protective factors in capturing change, rehabilitation benefit, and strengths/resiliencies that could aid

release and reintegration for Indigenous peoples, we separately examined these components within the risk tools. For risk change score associations which tended to be inverse (i.e., changes in risk associated with decreased recidivism), the corresponding AUC was reported on a positive scale (i.e.,  $> .50$ ) for continuity with other measures (e.g.,  $d = -0.50$  would translate into an AUC = .64 as opposed to .36). In addition, when the volume of studies permitted, we conducted moderator analyses for gender, developmental group, and country of study. Finally, given that Indigenous persons historically scored higher on risk measures and evidenced higher rates of recidivism, we also conduct meta-analyses of group differences in scores (Table 3) on these measures and computed unweighted recidivism rates across measures and outcomes (see supplementary Table S3).

### **Transparency and Openness**

The present study is a quantitative and narrative review of the forensic risk assessment literature with justice involved Indigenous persons. Consultation and feedback were provided by the Correctional Service of Canada who commissioned the original work as part of a larger interdisciplinary project examining forensic risk tools with Indigenous persons, which included a full legal analysis and Indigenous community engagement. Unpublished effect sizes reported for the purpose of meta-analytic aggregation were obtained from studies that had received prior agency, ethical, and/or judicial approval. The present study was not pre-registered. Per Simmons et al.'s (2012) 21-word solution, "We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study" (par. 6), which included conducting our search and reporting substantive findings in accordance with PRISMA guidelines. All requisite data to conduct the meta-analyses (effect size, base rates, cell n, AUC SE) are provided in supplementary Table S1 and Tables 4a through 13. These data (including SE values for  $d$ ), as well as datafiles for study descriptives are available online through the Open Science Framework and can be accessed at <https://osf.io/48v9d/> (Olver et al., 2023).

Finally, the authorship team acknowledges that three of its members are co-authors on 6 of the

22 assessment measures (and related user products, such as interpretive manuals and training materials) featured in this meta-analysis and systematic review. Specifically, MEO is co-author of the VRS-SO as well as user products related to the VRS and VRS-YV; KCS a co-author of the VRS-YV; and LMH is co-author of Static-99R, Static-2002R, and STABLE 2007, including user products for these measures (the materials for these latter scales are typically released via a Canadian non-profit organization whose mission is to support evidence-based risk assessment practices). While the authors occasionally receive financial compensation for training and consultation services with their respective tools, the remuneration is comparatively modest and does not serve as a substantive incentive for this research, especially when compared to the primary intellectual-scholarly, human rights-welfare, and clinical-legal motives driving this work. Additionally, the authors do not receive any royalties for the use of these tools or their products, with the exception of the VRS-SO rating manual; however, all other VRS-SO user products such as score sheets, the recidivism calculator, User's Workbook, and norms are all available free from [www.psynergy.ca](http://www.psynergy.ca). The frequency of trainings vary by instrument, many (or most) are delivered by approved trained trainers or the other instrument co-authors for their own profit, and many such trainings are conducted by the authors pro bono or heavily discounted, especially if delivered to local smaller agencies with few resources who rely heavily on these tools. The authors consider any perceived conflicts of interest seriously, both in terms of a perceived financial interest and a vested interest in the performance of these tools. In all, our authorship team endeavored to conduct our search, analyses, and reporting and interpretation of findings with the highest level of rigor and integrity and to facilitate replication by the readership.

## Results

### **Forensic Risk Assessment with Indigenous Adults and Youth: A Review of Measures and Findings**

We provide a detailed review of the risk assessment literature with Indigenous correctional populations arranged by instrument and outcome (i.e., recidivism type). In this review, we prioritize the

predictive properties in terms of discrimination (relative risk, e.g., ROC analyses, Cohen's  $d$ ) and, where available, calibration (absolute risk, e.g., logistic regression, E/O index) for adult and youth populations, followed by interim conclusions on the predictive properties of a given measures in Indigenous samples.

Table 2 provides an aggregate summary of the 91 studies included in the meta-analysis and systematic review including demographic variables, risk measures, and outcome measures. Across all risk assessment studies and documents retrieved, but excluding previous meta-analyses (due to overlapping samples), the grand  $N = 59,693$  for Indigenous persons, and  $N = 237,729$  for non-Indigenous persons across non-overlapping samples. Indigenous samples were significantly younger on average overall ( $d = 0.31, k = 33, p < .001, I^2 = 80.10$ ), and within youth ( $d = 0.29, k = 6, p < .001, I^2 = 0.00$ ) and adult ( $d = 0.31, k = 27, p < .001, I^2 = 83.67$ ) developmental subgroups. Most samples came from custody ( $k = 42$ ) or a combination of custody and community settings ( $k = 36$ ), whereas a minority ( $k = 9$ ) were community-based samples only. Further, the majority of studies featured adult ( $k = 54$ ) vs. youth ( $k = 21$ ) unique samples. Finally, a minority of the study samples (13%,  $k = 11$ ) were exclusively female, whereas the remainder were either male ( $k = 40$ ), or mixed male and female ( $k = 26$ ) samples.

Table 2 reports the M and SD for Indigenous and non-Indigenous (White majority) groups on the forensic assessment measures, while formal comparisons through meta-analysis are reported in Table 3. As seen in Table 3, Indigenous persons scored about one half standard deviation ( $d = 0.57$ ) higher than non-Indigenous persons in fixed effects meta-analysis across the entirety of measures representing  $k = 90$  sets of comparisons and 212,724 cases available to make such comparisons; the magnitude of difference was identical ( $d = 0.57$ ) for random effects. There was substantial variability in the magnitude of group differences. Indigenous persons also had higher rates of recidivism across non-overlapping samples; over a mean 4.0-year follow-up ( $SD = 3.7, k = 63$ ), the unweighted recidivism base rates for Indigenous and non-Indigenous samples for general recidivism was 51.5% ( $k = 70$ ) vs. 38.9% ( $k = 65$ ), respectively, violent recidivism 36.6% ( $k = 38$ ) vs. 23.6% ( $k = 37$ ), respectively, sexual recidivism 13.6% vs.

9.6% ( $k = 15$ ) respectively, and IPV recidivism, 31.6% vs. 15.8% ( $k = 3$ ), respectively (see supplementary Table S3 for breakdown by instrument).

***Level of Service (LS) Inventory and Other General Risk-Need Measures: Adult and Youth Variants***

**Discrimination properties.** LS scores on adult variants were  $d = .67$  higher for Indigenous adults and  $d = .39$  higher for Indigenous youth (Table 3). Unweighted rates of violent and general recidivism, respectively, were 15.0% vs. 40.0% for Indigenous persons and 14.5% vs. 33.3% for non-Indigenous persons (supplementary Table S3). For youth, unweighted violent and general recidivism rates, respectively, were 50.6% vs. 61.8% for Indigenous youth and 33.6% vs. 49.9% for non-Indigenous youth. Table 4 presents the discrimination findings for each of the 22 individual studies that have examined the predictive validity of the LS measure among Indigenous and non-Indigenous samples; all studies included direct within study comparisons between the two groups. Table 5 presents the results of past meta-analyses and our updated meta-analysis, which to date is the most comprehensive examination of the LS measures' comparative predictive properties for Indigenous and non-Indigenous samples. The number of studies in this meta-analysis also permitted examination of Indigenous and non-Indigenous group differences as a function of gender, development, and geographic region. Nuances for coding and analytic decisions are detailed in the Table 4 note.

Across  $k = 22$  studies with unique effects, LS total score evinced moderate in magnitude effects for the prediction of general recidivism among Indigenous samples ( $N = 23,044$ ,  $d = 0.53$ ,  $AUC = .65$ ) and moderate to large effects in non-Indigenous samples ( $N = 92,274$ ,  $d = 0.69$ ,  $AUC = .69$ ); effect sizes were significantly higher for non-Indigenous samples. This pattern was observed across male, female, and adult samples, with LS measures demonstrating moderate predictive effects among Indigenous subgroups, and often significantly higher moderate to large effects for non-Indigenous (predominantly White) groups. There were no significant differences in prediction magnitudes between ethnocultural subgroups for youth variants of the LS measures, however; there was consistent moderate in magnitude

effects across Indigenous and non-Indigenous samples in 11 studies and effect size heterogeneity was considerably lower than the other subgroups. Indigenous and non-Indigenous group differences were also observed in Canadian, Australian, and United States (US) samples, with predictive accuracy highest across ethnocultural subgroups within the Canadian samples, followed by the Australian and US samples. Finally, a smaller subset of studies ( $k = 6$ ) examined violence as an outcome. These yielded predominantly moderate in magnitude prediction effects across all studies and within female and Canadian Indigenous and non-Indigenous subgroups. With the exception of large fixed effects for some non-Indigenous moderator subgroups, none of the random effects models for the prediction of violence were significantly different between Indigenous and non-Indigenous groups.

Two additional instruments not included in the updated LS meta-analysis but worth brief mention are reported in Tables 1, 2, supplementary Table S1, and later in Tables 16 and 17 with all other risk measures. First, O'Byrne (2002) examined Saskatchewan's Offender Risk Assessment and Management System (ORAMS) Primary Risk Assessment (PRA) tool, a general risk-need measure with Central Eight item content similar to the LS measures, in a large provincial correctional sample of men and women. Similar to the aforementioned findings, over an average 3-year follow-up, PRA field ratings had moderate to large prediction effects for general recidivism among Indigenous ( $n = 1,560$ ,  $r_{pb} = .33$ ,  $AUC = .70$ ) and non-Indigenous ( $n = 667$ ,  $r_{pb} = .41$ ,  $AUC = .74$ ) groups; the difference in AUC magnitudes was not significant ( $z = -1.41$ ,  $p = .157$ ). Second, Schwalbe (2009) reported the results of an ethnocultural cross validation of the Arizona Risk/Needs Assessment (ARNA) instrument based on a large sample (Indigenous  $n = 1,576$ , White  $n = 12,945$ ) of youth on probation from 3 consecutive fiscal cohorts. Schwalbe also reported results from the original validation study (LeCroy et al., 1998, Indigenous  $n = 448$ , White  $n = 5,225$ ) which did not compute prediction statistics, but reported recidivism frequencies. The two reports in tandem were sufficient to generate a meta-analytic effect reported in Table 17, which yielded medium effects for Indigenous ( $d = 0.60$ ,  $AUC = .66$ ) and White ( $d = 0.72$ ,  $AUC = .70$ ) youth

in the prediction of general recidivism; the fixed effect was significantly larger for White youth. There was significant effect size heterogeneity owing to a substantially larger effect observed for both ethnocultural groups in the inaugural validation.

**Calibration properties.** At least two studies have examined the calibration properties of the LS measures in Indigenous and non-Indigenous samples, one adult and one youth. In addition to their meta-analysis of the LS measures, in a further prong to their investigation, Wilson and Gutierrez (2014) examined the calibration properties of LS/CMI scores in a large sample of provincially sentenced Indigenous ( $n = 1,692$ ) and non-Indigenous ( $n = 24,758$ ) sample of men and women from Wormith and Hogg (2012). Using logistic regression modelling, they found that LS/CMI scores and Indigenous ancestry each uniquely predicted general recidivism. Further, LS/CMI scores predicted comparable rates of 1-year recidivism at moderate to high scores for Indigenous and non-Indigenous groups; however, predicted recidivism estimates generated solely for Indigenous persons in this sample were notably higher at lower scores compared to estimates generated for the non-Indigenous persons (i.e., suggesting that the LS/CMI underestimates the actual recidivism rates for low risk Indigenous persons). In other words, developing Indigenous-specific recidivism estimates would result in low-scoring Indigenous people being classified at higher absolute risk of reoffending compared to a non-Indigenous person with the same score. Moreover, Prince et al. (2021) examined the calibration properties of the LSI-SK comparing 18-month general recidivism rates for Indigenous and non-Indigenous court adjudicated youth at each risk level via the E/O index. Although Indigenous youth had higher rates of recidivism than non-Indigenous youth overall, E/O index values at each risk level were not significantly different from 1.0 for Indigenous and non-Indigenous youth (E/O range = 0.74 to 1.04), demonstrating good correspondence of observed recidivism rates between ethnocultural groups at each LSI-SK risk level.

**Conclusions.** The LS measures are the most heavily researched risk measures with Indigenous persons in correctional settings across international jurisdictions. The available literature, including

results from an updated meta-analysis, demonstrate consistent moderate prediction effects for violent and general recidivism for male, female, youth and adult Indigenous subgroups; effect sizes are moderate to large and significantly higher for non-Indigenous persons, the exception being youth samples and for violence as an outcome. There appears to be some variability among specific LS variants (i.e., stronger prediction for LSI-OR or LS/CMI; lower for LSI-R). Part of this may also be attributable to country as a moderator, given that past research has shown Canadian studies (where the LSI-OR is more likely to be used) to have larger effects (Olver et al., 2014), followed by Australian, and US studies (where the LSI-R is more frequently used) to have smaller effects; the results of the updated meta-analysis reaffirm this. Despite recent research inroads, most LS evaluations have not separately examined the predictive properties among male and female Indigenous persons owing to insufficient power, and further such research is needed. Further, findings for the ORAMS and ARNA revealed medium effects that contribute to the broader pattern of extant findings for general risk-need tools. Finally, two LS studies, one each on a youth and an adult sample, demonstrated generally good calibration between ethnocultural groups in terms of rates of recidivism associated with risk scores; the exception was higher than expected rates of recidivism for low scoring Indigenous adults, suggesting that factors aside from general risk-need may be contributing to observed ethnoracial differences in rates of recidivism.

### ***Statistical Information on Recidivism (SIR) Scale***

As noted previously, the SIR scale is administered to all non-Indigenous men sentenced within Canada's federal correctional department (Correctional Service Canada; CSC), but not women or Indigenous persons. At least seven studies, however, have examined the predictive properties of variations on the SIR scale with Indigenous persons (see Table 6, all of which are from CSC samples; Barnum & Gobeil, 2012; Bonta & Rugge, 2004 as cited in Rugge, 2006; Hann & Harman, 1989, 1993; Nafekh & Motiuk, 2002; Wong & Gordon, 2006; Wormith & Olver, 2002). Several of these involve

examination of the original GSIR as developed by Nuffield (1982), while others employ slight adaptations such as the SIR-R1 or parallel forms such as the SIR-Proxy. Hann and Harman (1989, 1993) conducted the inaugural studies and employed smaller federal samples ( $Ns = 49$  and  $269$ ) of Indigenous men, finding that GSIR scores and risk bands had large prediction effects for general recidivism ( $AUCs = .71-.73$ ). Similar results with moderate to large effects have been found in subsequent examinations of the SIR scale with Indigenous and non-Indigenous men, including men who had attended a high intensity violence reduction program (Wormith & Olver, 2002), a multisite Prairie Region sample of violent offenders (Wong & Gordon, 2006), and a large-scale recidivism study of released CSC cohorts (Bonta & Rugge, 2004; as cited in Rugge, 2006). The largest, and most representative studies (Barum & Gobeil, 2012; Nafekh & Motiuk, 2002) examined the SIR-Proxy with Indigenous persons, contrasted with the SIR-R1. Both investigations found significant predictive validity for the SIR variants within ethnocultural subgroups, although generally larger effects for non-Indigenous men and women. Notably, in a large-scale revalidation of the SIR-R1 and SIR-Proxy, Barnum and Gobeil (2012) had sufficient sample size to examine the measures among Indigenous and non-Indigenous women. SIR-Proxy scores had moderate predictive accuracy for general and violent recidivism among Indigenous women ( $AUCs = .65$  and  $.70$ , respectively,  $n = 246$ ) and high predictive accuracy for these outcomes among non-Indigenous women ( $AUCs = .74$  and  $.82$ , respectively,  $n = 666$ ).

We conducted a meta-analysis of the 7 studies and  $k = 8$  samples examining SIR variants among Indigenous and non-Indigenous men and women (Tables 6 and 7). The SIR had a moderate effect in the prediction of general recidivism for Indigenous persons ( $n = 5,448$ ,  $d = .64$ ,  $AUC = .68$ ) but large effect for non-Indigenous persons ( $n = 24,810$ ,  $d = 0.83$ ,  $AUCs = .72$ ), which was significantly different. Across 6 samples, SIR scores had a small effect for violence among Indigenous persons ( $d = 0.42$ ,  $AUCs = .62$ ) and a moderate effect for non-Indigenous persons ( $d = 0.61$ ,  $AUCs = .67$ ) with the differences in ES magnitude significant across both models.

**Conclusions.** In short, research supports the predictive properties of the SIR scale variants (GSIR, SIR-R1, SIR-Proxy) among Indigenous men and women, with moderate predictive accuracy for general recidivism and small effects for violence. This stands in contrast to significantly higher and large prediction effects for general recidivism among non-Indigenous men and women and moderate prediction effects for future violence. Although Indigenous persons score on average  $d = 0.50$  SD higher than non-Indigenous persons on the SIR scale variants, the tool has lower discrimination magnitudes for recidivism outcomes with this population.

***Static Factors Analysis (SFA) and Dynamic Factors Identification and Analysis-Revised (DFIA-R)***

While the SIR scale is not employed with women or Indigenous persons, the SFA and DFIA-R are examples of other jurisdiction-specific instruments that are administered to all Canadian federally sentenced persons on intake; unlike the SIR, which is an actuarial measure, these employ SPJ to generate risk and need ratings. These each have a limited number of evaluations, although these are large in scale as the ratings can be easily culled from CSC's electronic filing system. The results are reported in supplementary Table S1 and Tables 16 and 17 with all other measures.

**Static Factors Analysis (SFA).** Helmus and Forrester (2014b; see also Perley-Robertson et al., 2019) conducted the lone predictive validity examination of the SFA ratings among separate groups of Indigenous men ( $n = 1,500$ ) and women ( $n = 149$ ) and non-Indigenous men ( $n = 6,884$ ) and women ( $n = 377$ ) for three outcomes: revocation without a new offense within 8 months, any new offense within 5 years, and any violent offense within 5 years. For Indigenous men, SFA summary risk ratings (SRR) had significant but small subthreshold effects for predicting the three outcomes (AUCs = .529-.544), Criminal History Record (CHR) had small to moderate effects (AUCs = .605-.654), and Offense Severity Record (OSR) had subthreshold and largely non-significant effects (AUCs = .471 - .545). AUC magnitudes for the SFA ratings were significantly larger for non-Indigenous men, with small effects for the SRR (AUCs = .598-.623), moderate to large effects for CHR (.677-.718), and subthreshold to moderate effects for OSR

(AUCs = .544-.641). AUCs for women tended to higher than for men, across ethnoracial groups; for Indigenous women the AUCs for SFA SRRs ranged from small to medium (AUCs = .562-.684), and for non-Indigenous women, the AUCs ranged from small to large (AUCs = .576 - .748) (Helmus & Forrester 2014b). The SFA subcomponents had similar themes, with both the CHR and OSR ratings having predictive validity ranging from small to medium for Indigenous women (AUCs = .594-.762) and small to large for non-Indigenous women (AUCs = .572 - .830). Across both female ethnoracial groups, the SFA and its two subscales predicted revocations with an offense (violent or general) over the 5 year follow up better than 8-month revocations alone, although AUCs were higher for non-Indigenous women. A meta-analysis of male and female prediction effects within ethnoracial groups demonstrated significantly larger effects for Indigenous persons (fixed effects) across both revocation offense outcomes (see Table 17).

**Dynamic Factors Identification and Analysis-Revised (DFIA-R).** The DFIA-R, by contrast, has seen more extensive validation with Indigenous persons. In a sample of 16,244 federally sentenced men and women (15,479 men, 765 women, 2,593 Indigenous), Brown and Motiuk (2005) examined the psychometric properties of the original DFIA, in the process of developing the DFIA-R, and found that its SRR had significant predictive accuracy for returns to federal custody within 3 years of release (AUC equivalents reported as converted from phi correlation of chi square): moderate effects for men (AUC = .65) and women (AUC = .69) were reported for the aggregate sample, and a small effect was reported for the Indigenous subsample (AUC = .57). A pair of updated investigations of the DFIA-R have extended previous findings. Stewart et al. (2017) examined the predictive validity of DFIA-R summary ratings and domain need ratings (discussed in next section) for revocations post-release in a sample of 18,752 non-Indigenous men and 5,856 Indigenous men. DFIA-R summary ratings had small significant predictive accuracy (Harrell's C, a time dependent ROC generated from Cox regression analogous to AUC) for revocations with or without an offense for both Indigenous (C = .55 and .56, respectively) and non-

Indigenous ( $C = .61$  and  $.63$ , respectively) men, with notably lower accuracy for Indigenous individuals. Meta-analysis was used to aggregate the effects across the two studies and to compare the effect sizes ( $d$ ); this was a conservative test, given that Brown and Motiuk (2005) presented the necessary information to compute comparison effects for the aggregate sample, but not non-Indigenous persons. As seen in Table 17, the DFIA-R had a medium effect in the prediction of revocations for non-Indigenous persons across the two studies ( $d = 0.51$ ), but a significantly lower, small effect for Indigenous men and women ( $d = 0.22$ ) in the prediction of this outcome.

Given that the DFIA-R overall risk rating (and the need ratings therein) are generated by parole officers employing an SPJ approach, it is possible that incorporation of numeric ratings to generate overall need ratings and an overall DFIA-R rating (more in keeping with an actuarial approach), could influence predictive validity, particularly since staff may be conservative when it comes to classifying people overall as low risk. Wilton et al. (2017) subsequently examined the predictive validity of calculated DFIA-R ratings overall and the need domains in construction and validation samples with respect to revocations (per Stewart et al., 2017). Across both testing and validation samples, the actuarial (i.e., calculated) DFIA-R ratings had improved accuracy for both Indigenous ( $C = .62$ ,  $n = 2,321$  and  $2,319$ ) and non-Indigenous ( $C = .68$ ,  $n = 7,773$  and  $7,714$ ) men, though still with somewhat lower accuracy for the Indigenous sample.  $C$  magnitudes were identical across testing and validation samples and weighting the indicators prior to inclusion in the overall DFIA-R risk category did not strengthen prediction beyond simple summation alone.

**Conclusions.** In all, despite its use for over 20 years, the SFA has been validated only once. Although statistically significant, AUCs were small (or worse) for Indigenous and non-Indigenous men, and although effect size magnitudes were larger for both women samples, overall, effect sizes were significantly greater for non-Indigenous persons as demonstrated through meta-analysis. Across two large independent cohorts of federal sentenced Indigenous and non-Indigenous men, the DFIA/DFIA-R

has demonstrated significant predictive validity for readmissions to federal custody, with the results of meta-analysis demonstrating a small effect for Indigenous men and a significantly greater medium prediction effect for non-Indigenous men. There is evidence to suggest that prediction improves when actuarial/mechanical methods are used to generate the overall DFIA-R rating as opposed to SPJ alone.

### ***Service Planning Instrument (SPIn) and Youth Assessment and Screening Instrument (YASI)***

Three studies have examined the predictive properties of either the SPIn or YASI Pre-Screen scores with Canadian Indigenous and non-Indigenous samples, two of which were based in Alberta, Canada (Jones et al., 2014, 2016) and a smaller study in the United States (Matz et al., 2021). The results are reported in Tables 2 and 10 in detail. The results from these three validations are particularly impressive and constitute large effects in the prediction of violent and general recidivism across Indigenous and non-Indigenous groups; the one exception was a moderate effect in Matz et al.'s (2021) small youth Indigenous subsample. A meta-analysis of the three studies, however, affirmed these large prediction magnitudes across ethnocultural groups ( $d = 0.98-1.00$ ,  $AUC = .76$ ; Table 16/17). Although separate AUCs for the strength domains of the SPIn and YASI in the prediction of decreased recidivism were not reported by ethnocultural group in the original sources, the results of hierarchical logistic regressions demonstrated SPIn strength scores to incrementally predict decreased recidivism, controlling for risk score in both Indigenous and non-Indigenous subgroups (Jones et al., 2014).

**Conclusions.** Indigenous youth and adults do score higher on the SPIn and YASI ( $d = 0.61$ ), however, results of meta-analysis of the three studies on Indigenous samples to date demonstrate robust large effects in the prediction of general recidivism across ethnocultural groups.

### ***Violence Risk Scale (VRS) and its Sexual Offense (VRS-SO) and Youth (VRS-YV) Versions***

**Discrimination properties.** The discrimination properties of the VRS-SO, VRS, and VRS-YV for recidivism outcomes have been examined across a series of Canadian samples of Indigenous and non-Indigenous justice involved men and male and female youth. These investigations pooled individual

samples into a combined sample for analysis, but we also used meta-analysis to separately aggregate individual effects into a global effect size for comparison purposes. Results for the three tools are described as follows.

***Violence Risk Scale-Sexual Offense version (VRS-SO).*** Olver, Sowden et al. (2018) examined the predictive properties of VRS-SO risk and change scores across three Canadian correctional (CSC) samples of 1,093 treated men (Indigenous  $n = 393$ , non-Indigenous  $n = 670$ ) convicted for sexual offenses who had been rated on the VRS-SO pre and posttreatment and followed up in the community an average 12.3 years post-release. A combined sample was utilized, following calls to employ large, representative samples, that would be adequately powered to conduct stratified Indigenous and non-Indigenous psychometric analyses (Hart, 2016). For Indigenous men, broadly moderate effects (AUCs = .62-.70) were observed for VRS-SO static, dynamic, and total scores in the prediction of 5- and 10-year sexual recidivism; prediction magnitudes were moderate to large for non-Indigenous men (AUCs = .69-.75), although the difference in AUC values were not significant between the groups. Similar prediction effects were observed for violent and general recidivism (Table 8). VRS-SO change ratings were significantly associated with decreased sexual and violent recidivism for both Indigenous ( $e^B = .906-.912$ ,  $ps < .05$ ) and non-Indigenous ( $e^B = .900-.915$ ,  $ps < .01$ ) groups, controlling for baseline risk (pretreatment total score).

The potential nonetheless may exist that Indigenous and non-Indigenous group differences occurring at the individual sample level, are disguised when the samples are combined in aggregate; however, meta-analysis should be able to disentangle possible differences, if they exist. The results of meta-analysis in Table 9 for the three VRS measures focus on the total score (i.e., static + dynamic) and residualized dynamic change scores (i.e., controlling for pretreatment score), for parsimony. We also examined the overall meta-analytic effect for all three measures across  $k = 9$  samples. Scrutiny of Table 9 shows that fixed and random effects analyses generally yielded identical findings across measures,

effect size heterogeneity was very small ( $Q$  and  $I^2$ ), and none of the differences in effect size magnitudes between Indigenous and non-Indigenous persons were significant; all of which supports the stability of the generated effects as well as the viability of combining samples into a large aggregate. For instance, for the VRS-SO, the meta-analytically generated effects were substantively the same for both Indigenous and non-Indigenous subgroups, compared to when they were examined in aggregate; moderate effects in the prediction of sexual, violent, and general recidivism for Indigenous men ( $d = 0.61-0.66$ ,  $AUC = .67-.68$ ), and large effects in the prediction of the three outcomes for non-Indigenous men ( $d = 0.72-0.98$ ,  $AUCs = .71-.79$ ). VRS-SO change scores had significant small in magnitude effects for associations with decreased sexual and violent (but not general) recidivism for both ethnocultural subgroups; these effect sizes were within the same class of magnitude (i.e., small) and not significantly different.

**Violence Risk Scale (VRS).** A preliminary examination of the predictive properties of the VRS has been conducted by Olver, Cortoni et al. (2021) with five pooled Canadian samples of Indigenous ( $n = 406$ ) and non-Indigenous ( $n = 549$ ) men adjudicated for violent offenses. Four of the samples were federal Canadian corrections admissions; three were samples from high intensity violence treatment programs (Coupland & Olver, 2020; Higgs et al., 2020; Lewis et al., 2013) while one was from the Prairie Region Violent Offender Survey (Wong & Gordon, 2006). The fifth sample was a small predominantly White, predominantly male, forensic inpatient sample (Hogan & Olver, 2019). While the five samples could be combined to examine Indigenous and non-Indigenous group differences in aggregate, only four of the samples had a sufficient  $n$  to examine such differences at the individual sample level to permit meta-analysis (Table 8). First, the aggregate multisite sample was followed up in the community an average of 4.9 years post release ( $SD = 3.5$ ). For both Indigenous and non-Indigenous subgroups, large prediction effects were observed for VRS total scores in the prediction of violent ( $AUC = .75$  and  $.78$ , respectively) and general ( $AUCs = .75$  and  $.77$ , respectively) recidivism. Further, the results of Cox regression survival analysis demonstrated changes in risk, controlling for VRS pretreatment score, were

significantly associated with decreased violent recidivism within the Indigenous subsample ( $e^B$  VRS change = 0.919,  $p < .001$ ), but not for the non-Indigenous (predominantly White) subsample ( $e^B$  VRS change = 0.980,  $p = .489$ ). That is, the risk reduction effect as measured by the VRS, was stronger for Indigenous men.

As seen in Table 9, the results of meta-analysis of VRS scores were consistent with those in the aggregate sample. Although the meta-analytic average of VRS total score prediction effects were lower than those observed in the aggregate sample (likely due to a loss of variance from a large combined sample), this was uniform across ethnocultural groups. VRS scores had moderate in magnitude effects for the prediction of violent and general recidivism for both Indigenous ( $d = 0.55$  and  $0.52$ , AUC =  $.65$  and  $.64$ , respectively) and non-Indigenous ( $d = 0.55$  and  $0.60$ , AUC =  $.65$  and  $.66$ , respectively) men. Further, VRS change scores had moderate effects for decreased violent and general recidivism for Indigenous persons ( $d = -0.56$  and  $-0.53$ , AUCs =  $.65$ , respectively) and small effects for non-Indigenous men ( $d = -0.29$  and  $-0.31$ , AUC =  $.58$  and  $.59$ , respectively).

**Violence Risk Scale-Youth Version (VRS-YV).** Two studies have examined the youth version of the tool, both of which had a majority sample of Indigenous youth and featured independent samples of boys and girls referred for special assessment and/or intervention services at a youth forensic mental health program (Lovatt et al., 2022; Stockdale et al., 2014). As seen in Table 9, VRS-YV total scores had moderate effects in the prediction of violent and general recidivism for Indigenous youth ( $d = 0.63$ , AUC =  $.67$ , and  $d = 0.54$ , AUC =  $.65$ , respectively) and moderate to large effects for non-Indigenous ( $d = 0.80$ , AUC =  $.71$ , and  $d = 0.74$ , AUC =  $.70$ , respectively) youth; differences in effect size magnitude were not significant between groups for either outcome.

**Calibration properties.** Two studies of VRS measures have examined their calibration properties for Indigenous persons. First, Olver, Sowden et al. (2018) conducted VRS-SO calibration analyses of 5 and 10-year sexual and violent recidivism through logistic regression. They found that group differences

in base rates of sexual recidivism could be explained by individual differences in static and dynamic risk. Further, VRS-SO scores generated slightly higher estimates of sexual recidivism for non-Indigenous men at very high scores, but slightly higher predicted estimates for Indigenous men at low scores. For general violence, however, calibration analyses demonstrated much unexplained variability in rates of general violent recidivism remained after controlling for baseline risk and change, with higher rates of violent recidivism observed for Indigenous men irrespective of VRS-SO score (i.e., a given score does not translate into the same violent recidivism estimate for Indigenous and non-Indigenous persons). In all, the VRS-SO demonstrated acceptable calibration for sexual recidivism risk for both Indigenous and non-Indigenous men, but it under detected violent recidivism risk for Indigenous men.

Second, Lovatt et al. (2021) conducted calibration analyses using logistic regression to examine trajectories of 5-year violent recidivism among Indigenous and non-Indigenous youth as a function of VRS-YV total score. Projected rates of violent recidivism increased as a function of VRS-YV score for each group, with good calibration at moderate to high scores. VRS-YV scores also under classified low scoring Indigenous youth, who recidivated violently at levels higher than would be expected by their scores, consistent with calibration evaluations of adult tools such as the LS/CMI (Wilson & Gutierrez, 2014).

**Conclusions.** The results from seven independent samples support the predictive properties of the VRS and VRS-SO with Indigenous men convicted, respectively, for violent or sexual offenses. The importance of evaluating dynamic risk and change was underscored, as changes in risk from evidence informed interventions was associated with decreased recidivism for Indigenous men. Change effects were equivalent across ethnocultural groups, even though predictive validity of the risk scores was often slightly higher for non-Indigenous men; none of the meta-analytic effects were significantly different between groups. Two additional studies broadly supported the discrimination and calibration properties of the VRS-YV, particularly with respect to future violence, among Indigenous youth. These results underscore the importance of using dynamic measures that can capture changes in risk for Indigenous

persons, such as from correctional and cultural programming or other change agents, to inform release planning, community supervision, and reintegration efforts.

### ***Structured Assessment of Violence Risk in Youth (SAVRY)***

Table 10 presents the results from four studies that have examined the predictive efficacy of the SAVRY among Indigenous and non-Indigenous youth, three Canadian (Lovatt et al., 2021; Meyers & Schmidt, 2008; Muir et al., 2020), one Australian (Shepherd et al., 2014). Each of these evaluations examined the SAVRY as a mechanical tool (i.e., summing the items to generate risk scores), while comparatively fewer examined the SPJ-based summary risk ratings. Shepherd et al.'s (2014) Australian validation, also compared Indigenous, English-speaking background (ESB), and culturally and linguistically diverse (CALD) youth; of note, the SAVRY had comparably poor predictive properties among CALD youth. More recent validations of the SAVRY have employed larger samples, with Muir et al. (2020) having conducted the largest scale examination of the SAVRY examining community youth worker field ratings with sufficient sample size to examine male and female Indigenous and non-Indigenous effects separately.

We conducted a meta-analysis of SAVRY studies across  $k = 5$  effects which demonstrated moderate and substantively equivalent effects for numeric scores in the prediction of violent and general recidivism for Indigenous youth. We aggregated effects sizes for both the SAVRY total score and Protect domain. As seen in Table 11, the effect size magnitudes of SAVRY total scores for violence (the targeted criterion of this measure) were moderate in magnitude and not significantly different across Indigenous ( $d = .65$ ,  $AUC = .67$ ) and non-Indigenous ( $d = .71$ ,  $AUC = .69$ ) youth groups. In the prediction of general recidivism, SAVRY total score effect sizes were large and significantly higher for non-Indigenous (predominantly White) youth samples but moderate for Indigenous youth samples. Further, the SAVRY Protect domain had small to moderate associations with decreased general recidivism and small significant associations with decreased violence within Indigenous youth samples. In non-

Indigenous youth samples, SAVRY Protect domain effect sizes were large for decreased general recidivism and moderate for decreased violence. Effect size heterogeneity was generally low and nonsignificant for both outcomes within ethnocultural groups, with the exception of medium heterogeneity of the Protect domain effects for non-Indigenous youth.

**Conclusions.** Results from meta-analysis demonstrate that SAVRY total risk scores, when summed mechanically, have moderate and largely equivalent predictive accuracy across ethnocultural groups, but significantly different, and large effects, in the prediction of general recidivism for non-Indigenous youth. Protective factor scores (also summed mechanically, consistently have small effects for associations with decreased recidivism for Indigenous youth, but medium to large effects for non-Indigenous youth. These findings are tempered by our finding, however, that Indigenous youth score higher on SAVRY total (risk) scores, but significantly lower on SAVRY protective factors. In all, the available research supports the predictive properties of the SAVRY with Indigenous youth for violent and general recidivism, and the potential risk mitigating properties of its Protect domain.

### ***STABLE-2007***

There have been at least three examinations of STABLE-2007 with Indigenous men. Two evaluations featured STABLE-2007 ratings conducted in the field by community supervision officers, and as such, these are both prospective designs as well as illustrations of the STABLE-2007's field validity (Helmus et al., 2012; 2021). The other evaluation was an archival retrospective examination of the tool on an incarcerated sample of men treated for sexual offending and followed up nearly 10 years in the community post release (Sowden & Olver, 2017) (see Tables 2 and 16/17). Helmus et al. (2012) conducted the first of these evaluations on the same Canadian sample used to develop the STABLE-2007 (Indigenous,  $n = 88$ ; non-Indigenous,  $n = 509$ ). For Indigenous men, STABLE-2007 total scores had weak nonsignificant effects for sexual and violent recidivism (AUCs = .53) and small significant effect for general recidivism (AUC = .58). By contrast, among non-Indigenous men, STABLE-2007 scores had

moderate to large effects for sexual (AUC = .70), violent (AUC = .71), and general (AUC = .72) recidivism. The second evaluation of the STABLE-2007 with Indigenous persons (Helmus et al., 2021) featured field ratings completed on a large Canadian sample of 4,511 men in British Columbia who were being supervised in the community for sexual offending. STABLE-2007 showed the strongest predictive validity for sexual recidivism with a moderate effect for the White subgroup ( $C = .69$ ), followed by other ethnocultural minorities ( $C = .67$ ,  $n = 774$ ; reported by Lee et al., 2016), and the lowest predictive accuracy, with a small effect for the Indigenous subgroup (AUC = .59). STABLE-2007 scores also had slightly higher prediction magnitudes for violent ( $C = .61$ ) and general ( $C = .64$ ) recidivism with Indigenous men in contrast to moderate effects ( $C = .66$  and  $.68$ , respectively) for White men. The predictive validity findings from these two field investigations were generally consistent with Sowden and Olver (2017), whose study featured a higher risk, higher need incarcerated sample, with a roughly equal proportion of Indigenous ( $n = 87$ ) and non-Indigenous ( $n = 90$ ) men. Aggregating the three studies through meta-analysis (Table 17) yielded small effects in the Indigenous samples ( $n = 1,096$ ) for the prediction of sexual ( $d = 0.26$ , AUC/ $C = .57$ ) and violent ( $d = 0.40$ , AUC = .61) recidivism and moderate effect for general recidivism ( $d = 0.51$ , AUC = .64). Among non-Indigenous persons ( $n = 3,216$ ), predictive effects were consistently moderate for the three outcomes ( $d = 0.61-0.72$ , AUC = .68-.69) and significantly higher than the observed effects for Indigenous persons. Effect size heterogeneity was smallest for sexual recidivism across ethnocultural groups and more variable for the other outcomes.

**Conclusions.** Meta-analytic results aggregating findings from two field studies and an archival study on an incarcerated treated sample indicate that the STABLE-2007 does significantly predict recidivism for Indigenous men, but it has better discrimination properties for sexual recidivism with White persons and other ethnocultural minorities (medium effects) than it does for Indigenous men with sexual offense conviction histories (small effects). The discrimination prediction magnitudes were closer to parity between ethnocultural groups for higher base rate outcomes, particularly general

recidivism, although assessing risk for any recidivism and general violence is not the primary purpose of the tool. Additional research is warranted evaluating STABLE 2007's discrimination properties for sexual recidivism before firm conclusions can be made regarding its clinical utility for Indigenous men with conviction histories of sexual offending.

### ***Static-99R and Static-2002R***

**Discrimination properties.** Table 12 reports the prediction effects sizes for Static-99R and 2002R for 5 and 10-year sexual recidivism in Indigenous and non-Indigenous (White) samples; quantitative reviews have tended to focus on sexual offense recidivism, given that this is the targeted criterion of these two measures. Most individual Canadian and United States samples ( $k = 11$ ) employ White persons as the non-Indigenous comparison group, while the Australian samples employ a broad non-Indigenous (although predominantly White) aggregate. Prior meta-analyses are reported in Table 13, and an updated meta-analysis conducted from the present review is presented immediately above it.

We extend the work of previous meta-analyses through conducting updated meta-analysis of the two measures for sexual recidivism (Static-99R,  $k = 12$ , Static-2002R,  $k = 4$ ) using fixed 5 and 10-year follow-ups, per Lee et al. (2020). Although Smallbone and Rawlings (2013) had a short and variable follow-up period ( $M = 29$  months, range 15-53 months), we included it with the 5-year effects computation given that this was the closest timeframe (NB: excluding it did not change the 5-year effect); further, Helmus et al. (2021) had a mean follow-up close to five years and used time dependent ROCs (Harrel's C). As seen in Table 13, fixed and random effects analyses generated substantively similar findings across ethnocultural groups and measures; again, fixed effects results are reported for parsimony. The 5-year effect for Static-99R with Indigenous samples was  $AUC = .63$  ( $d = 0.45$ ,  $n = 1,996$ ), and for White/non-Indigenous samples,  $AUC = .70$  ( $d = 0.74$ ,  $n = 6,191$ ); the differences in AUCs (both models) was statistically significant,  $z = 3.20$ ,  $p = .001$ . Moderate 10-year effects were observed for Static 99-R with Indigenous samples ( $AUC = .65$ ,  $d = 0.58$ ,  $n = 513$ ), and a large effect for White/non-Indigenous

samples, ( $AUC = .71$ ,  $d = 0.79$ ,  $n = 1,547$ ); the difference in AUCs was not significant. For the Static-2002R, significant small in magnitude effects were observed for Indigenous men with respect to 5-year ( $AUC = .63$ ,  $d = 0.45$ ,  $n = 281$ ) and 10-year ( $AUC = .63$ ,  $d = 0.47$ ,  $n = 162$ ) sexual recidivism. For White/non-Indigenous men, Static-2002R scores had a moderate effect in the prediction of 5-year sexual recidivism ( $AUC = .69$ ,  $d = 0.72$ ,  $n = 708$ ) and large effect for 10-year sexual recidivism ( $AUC = .74$ ,  $d = 0.93$ ,  $n = 460$ ). Static-2002R AUC magnitudes were significantly higher for White persons in the prediction of 10-year sexual recidivism ( $z = 2.26$ ,  $p = .024$ ) but not 5-year sexual recidivism.

**Calibration properties.** Lee et al. (2020) conducted calibration analyses examining logistic regression generated rates of 5 and 10-year sexual recidivism. Static-99R scores predicted comparable trajectories of recidivism for both White and Indigenous men over these timepoints. Although non-significant, somewhat higher rates of sexual recidivism were estimated for White men at very high scores, while somewhat higher rates were predicted for Indigenous men at lower scores. Any differences in projected rates of sexual recidivism between the ethnocultural groups tended to decrease in magnitude as length of follow-up increased. As with the discrimination analyses, however, Static-2002R displayed poor calibration for Indigenous men and did not significantly predict 5- or 10-year sexual recidivism. Among the White subsample, the recidivism rates controlling for Static-2002R were significantly lower than those predicted by Phenix et al.'s (2016) high risk/need norms.

**Conclusions.** Results of four meta-analyses, including our updated meta-analysis, support the predictive validity of Static-99R for sexual recidivism among Indigenous men convicted for sexual offenses, with accuracy improving over longer follow-ups, and prediction magnitudes being small to moderate; effect sizes are consistently higher and medium to large for White/non-Indigenous samples. Results of an updated meta-analysis including US, Australian, and additional Canadian samples extended the findings of, and were highly consistent with, Lee et al. (2020), and most recently, Helmus et al. (2023). (As an aside, in both Australian samples, Static-99R demonstrated moderate predictive validity

for violence with Indigenous men as noted in supplementary Table S1 while yielding small or subthreshold effects for sexual recidivism.) The inclusion of a fourth Static-2002R sample strengthened the predictive prospects of this tool for 5- and 10-year sexual recidivism with Indigenous men; however, the research remains much more limited for this measure and thus caution remains warranted for the use of Static-2002R with this population.

### ***Ontario Domestic Assault Risk Assessment (ODARA)***

**Discrimination properties.** As reported in the aggregate summary supplementary Table S1 and with all other measures in Tables 16/17, two studies, both based in Saskatchewan, Canada, have examined the predictive properties of the ODARA with Indigenous persons convicted of IPV offenses (Table 16). First, in an unpublished doctoral dissertation, Buchanan (2009) examined the predictive validity of ODARA scores on an urban Saskatchewan sample of 53 Indigenous and 41 non-Indigenous men convicted for IPV who were under Ministry of Corrections Public Safety and Policing (CPSP) jurisdiction. ODARA scores demonstrated moderate predictive accuracy (AUC = .64) for Indigenous men and a large predictive effect (AUC = .77) for non-Indigenous men with respect to this outcome. Hegel et al. (2021) subsequently scored the ODARA from police files for 300 men convicted for IPV randomly selected from six Royal Canadian Mounted Police (RCMP) detachment regions in Northern Saskatchewan, which has the highest rates of IPV in the country. Most of the sample ( $n = 278$ , 92.7%) identified as Indigenous ancestry. In the large Indigenous subsample, ODARA scores significantly predicted IPV, violent, and nonviolent recidivism with small effects (AUCs = .62-.63) and general recidivism with a moderate effect (AUC = .65); in the much smaller non-Indigenous subsample ( $n = 22$ ), ODARA scores yielded small effects for IPV recidivism but moderate to large effects for other outcomes. Aggregating the two studies through meta-analysis yielded small to moderate effect in the Indigenous samples for IPV recidivism ( $d = 0.43$ , AUC = .62) and moderate effect for non-Indigenous persons ( $d = 0.56$ , AUC = .66) (Table 17).

**Calibration properties.** Logistic regression analyses of the Hegel et al. (2021) sample demonstrated that rates of IPV recidivism increased with successive increases in ODARA score and risk level, in correspondence with actual, observed rates, for all but the extreme lowest scores (which had very small cell  $n$ ). Formal E/O index analyses demonstrated IPV recidivism rates for the ODARA risk bands from the predominantly White normative sample (Hilton et al., 2004) had close correspondence to those in Hegel et al.'s (2021) predominantly Indigenous sample at low and moderate scores (bottom five risk bands); however, Hilton et al.'s (2004) risk bands significantly overpredicted IPV recidivism at the top two risk bands. Further, E/O index comparisons to Olver and Jung's (2017) sample from a neighboring Canadian Prairie province using ODARA risk bands collapsed at lower risk levels and a fixed 3-year follow-up, demonstrated higher rates of IPV recidivism in Hegel et al. (2021) irrespective of risk level. The results suggest the ODARA may be better calibrated for low to moderate risk levels than it is for the highest risk levels in an Indigenous IPV sample.

**Conclusions.** Taken together, the results from two Canadian studies support the discrimination properties of the ODARA for future IPV with Indigenous men; effect sizes were small for Indigenous men and moderate for non-Indigenous men. These prediction magnitudes were slightly lower than the overall meta-analytic average for the ODARA ( $AUC = .69$ ,  $k = 22$ , van der Put et al., 2019) in published studies. Perhaps owing to its heavy weighting of offense history and certain demographic and historic clinical indicators (e.g., substance abuse), Indigenous men scored considerably higher across these two samples ( $d = 0.65$ ) and were more likely to score in the higher risk tiers, increasing the potential for false positives. Although the tool does predict its targeted outcomes and has the potential to aid IPV prevention, the tool should be used in conjunction with a dynamic risk measure (e.g., SARA, VRS, LS measure) to better target intervention planning and risk management efforts for Indigenous men with a history of IPV offenses.

***Spousal Assault Risk Assessment guide (SARA)***

There are two studies, both unpublished featuring Canadian samples, examining the predictive accuracy of this SPJ measure with Indigenous men with a history of domestic violence (Ahmed, 2022; Hogan & Corabian, 2023). Both studies examine the predictive validity of SARA scores computed mechanically, as well as SPJ summary risk ratings (Table 16). In an unpublished MA thesis, Ahmed (2022) examined the SARA Version 2 among Indigenous ( $n = 2,949$ ) and White ( $n = 6,105$ ) men on community supervision for IPV in British Columbia Corrections. Three recidivism outcomes were examined: domestic violence, any violence, and any criminal recidivism. For domestic violence recidivism, SARA summary risk ratings (SRR) had significant and similar predictive accuracy for Indigenous ( $AUC = .593$ ) and White men ( $AUC = .617$ ), although the effects were small for both groups. Effects for violent and any criminal recidivism were also small for both Indigenous ( $AUC = .574$  for both outcomes) and White men ( $AUC = .606$  and  $.602$ , respectively), but with significantly lower effect sizes for Indigenous men.

Hogan and Corabian (2023) examined SARA Version 3 scores and SPJ ratings in the prediction of any violent recidivism in a sample of 82 Indigenous men and 108 non-Indigenous (White majority) men who were referred by law enforcement agencies throughout Alberta for violence risk assessment. The SARA had subthreshold effects for Indigenous men ( $AUCs < .56$ ) and subthreshold to small effects for SARA SPJ and mechanical ratings for White men ( $AUCs = .54$  and  $.59$ , respectively). A meta-analysis of the two studies (Table 17) generated small effects for both SPJ and mechanical ratings among Indigenous men in the prediction of general violence ( $d = 0.27$  and  $0.33$ , respectively) and small to medium effects, respectively, for non-Indigenous men ( $d = 0.37$  and  $0.50$ , respectively); SARA mechanical scores had significantly stronger predictive validity for general violence in non-Indigenous samples compared to Indigenous ones.

**Conclusions.** Although statistically significant, AUCs and aggregated  $d$  values were small for both Indigenous and non-Indigenous persons, and they were significantly smaller for Indigenous men for predicting violent and any recidivism in the Ahmed (2022) thesis, and in predicting violence across the

two samples through meta-analysis. The analyses suggested that prediction improves slightly when actuarial/mechanical methods are used to generate a total score as opposed to SPJ alone; the two studies, carried out in neighboring Canadian provinces with prospectively administered SARAs had considerable continuity in findings, despite large differences in N. On the basis of these two studies, the SARA is minimally defensible to use, but has very modest accuracy for both ethnocultural groups.

### ***Structured Assessment of Protective Factors (SAPROF) and its Youth Version (SAPROF-YV)***

The SAPROF and its youth version (SAPROF-YV) were each examined in a pair of Canadian studies. Olver and Riemer (2021) examined the predictive properties of SAPROF protection scores, and its intersection with PCL-R measured psychopathy, in a sample of treated men with sexual offense conviction histories. Secondary analyses of SAPROF scores (posttreatment), demonstrated that Indigenous men ( $n = 185$ ) had fewer protective factors than non-Indigenous men ( $n = 236$ ),  $d = -0.26$ ,  $p < .001$ . Total scores had medium effects in the prediction of decreased general and violent recidivism ( $d = 0.69$  and  $0.68$ , respectively) for Indigenous men, but among non-Indigenous men, small effects for general recidivism and a large effect for violence ( $d = 0.46$  and  $0.86$ , respectively) (Table 16). Lovatt et al. (2021) has conducted the only evaluation of the SAPROF-YV with Indigenous youth ( $n = 171$ , non-Indigenous  $n = 55$ ) to date (Table 16), and found that the relational and motivational domains, as well as the total score, had the most consistent predictive accuracy for decreased violent, nonviolent, and general recidivism (AUCs = .58-.66). The resiliency and external subscales yielded smaller effects by comparison and were seldom significant. For non-Indigenous youth, the same pattern was observed, with moderate to large effects for the motivational, relational, and total scores, and small effects for resilience and external scores. A meta-analysis of these two studies yielded medium effects for both Indigenous and non-Indigenous groups in terms of associations with decreased violent ( $d = 0.65$  and  $0.71$ , respectively) and general recidivism ( $d = 0.68$ , both groups).

**Conclusions.** These preliminary findings suggest promise for use of the SAPROF and its Youth

Version as a measure of protective factors with justice involved Indigenous youth and adults; specifically, higher levels of protective factors (e.g., positive coping, self-control, empathy, prosocial leisure, supportive relationships) are associated with decreased general and violent recidivism across ethnocultural groups with medium effects. Protective factors have equal if not greater import for marginalized groups although the challenge remains to what extent these measures can be meaningfully integrated with risk assessment measures to inform service planning; a point that we revisit in further detail in the Discussion.

***Psychopathy Checklist-Revised (PCL-R) and its Youth Version (PCL: YV)***

**Discrimination properties.** The predictive properties of the PCL-R and its youth version (PCL: YV) have been examined with Indigenous men, boys, and girls, in a series of adult ( $k = 5$ ) and youth ( $k = 5$ ) correctional samples (see Tables 14 and 15). PCL-R and PCL: YV scores were significantly higher ( $d = 0.22$  and  $0.15$ , respectively for Indigenous youth and adults, a small significant effect (Table 3). Unweighted rates of violent and general recidivism, respectively, were 45.6% vs. 65.1% for Indigenous men and 30.9% vs. 35.7% for non-Indigenous men. For youth, unweighted violent and general recidivism rates, respectively, were 45.2% vs. 72.6% for Indigenous youth and 32.6% vs. 54.8% for non-Indigenous youth (supplementary Table S3).

***Psychopathy Checklist-Revised (PCL-R).*** Olver et al. (2013) was the inaugural publication for the PCL-R on the topic and received a fair amount of psycho-legal attention (e.g., *Ewert v. Canada, 2015*). The study featured a sample of 766 men rated on the PCL-R via file review, with 435 individuals who were released (Indigenous  $n = 171$ , non-Indigenous  $n = 264$ ) and followed up 26 months in the community post-release. PCL-R Lifestyle, Antisocial, and total scores significantly predicted each outcome for Indigenous and non-Indigenous men; effects were broadly moderate for Lifestyle (AUCs = .64-.65 and .68-.74, respectively) and total scores (AUCs = .63-.66 and .64-.70, respectively), and large for Antisocial scores (AUCs = .71 and .73-.77, respectively), across ethnocultural groups. The

Interpersonal and Affective facets had similarly small to weak and non-significant predictive validity for each outcome across ancestral groups (AUCs = .51-.58 and .47-.57, respectively). Olver et al. (2013) was critiqued as a single, unreplicated study with a limited sample size for the prediction analyses within ethnocultural subgroups (*Ewert v. Canada, 2015*). In response, Olver et al. (2018) added three non-overlapping Canadian federal corrections (CSC) Prairie Region samples with PCL-R, recidivism, and ethnocultural heritage information nearly tripling the sample to 1,163 federally sentenced men (Indigenous  $n = 452$ , non-Indigenous  $n = 711$ ), followed up an average 11.7 years post release. PCL-R total, Factor, and facet scores significantly predicted general and violent recidivism with moderate to high accuracy across both groups (AUCs = .63-.72 and .67-.73, respectively). PCL-R total scores, specifically, had large effects (AUCs = .71-.73) for the two outcomes across ethnocultural groups.

We conducted a meta-analysis of the four studies included in Olver, Neumann et al. (2018) to examine whether the effects observed in the aggregate sample remained invariant when examined at the individual sample level; a fifth non-overlapping sample (Olver & Riemer, 2021) was added from a subsequent investigation, as noted previously, that examined the intersection of psychopathy and protective factors from a sexual violence treatment program. All  $d$  values were computed directly from the original data and aggregated (see Table 14 for individual samples). As seen in Table 15, fixed and random effects models yielded generally identical effects within and between ethnocultural groups with very low effect size heterogeneity. PCL-R scores had medium effects for general recidivism for Indigenous ( $d = 0.56$ , AUC = .65) and non-Indigenous men ( $d = 0.51$ , AUC = .64), as well as medium prediction effects for violence ( $d = 0.55$ , AUC = .65 and  $d = 0.63$ , AUC = .67, respectively). Meta-analytic effects were uniformly smaller for both ethnocultural groups than that reported in the combined aggregate sample of Olver, Neumann et al. (2018), likely owing to its greater range and variance to discriminate recidivists from non-recidivists through statistics such as  $d$  and AUC.

***Psychopathy Checklist: Youth Version (PCL: YV).*** At least five studies have examined the PCL: YV

with Indigenous youth, four Canadian (Lovatt, 2020; McCuish et al., 2017; Schmidt et al., 2006; Stockdale et al., 2010), and one Australian (Shepherd & Strand, 2016); largely the same research groups who have examined the LS youth variants and SAVRY with Indigenous youth have carried out these evaluations (see Table 14). As with the PCL-R, we conducted a meta-analysis of these five studies by converting AUCs or  $r_{pb}$  values to  $d$ , or directly computed (e.g., Lovatt, 2020), and then aggregated using the  $SE$  computed from the recidivism base rates for each study. As reported in Table 15, the results (Indigenous  $n = 483$ , non-Indigenous  $n = 580$ ) yielded broadly moderate effects in the prediction of general (Indigenous  $d = 0.61$ ,  $AUC = .67$ , non-Indigenous  $d = 0.54$ ,  $AUC = .64$ ) and violent (Indigenous  $d = 0.55$ ,  $AUC = .65$ , non-Indigenous  $d = 0.41$ ,  $AUC = .61$ ) recidivism. There was significant heterogeneity for the non-Indigenous subgroup and some improvement in prediction with the random effects model for violent recidivism; however, for Indigenous youth, fixed and random effects were highly consistent and the  $Q$  statistic was non-significant. For general recidivism, fixed and random effects yielded the same findings across ethnocultural groups and effect size heterogeneity was nil. Combining youth and adult samples across PCL variants ( $k = 9$ , Indigenous  $n = 932$ , non-Indigenous  $n = 1,291$ ) extended these findings. Equivalent moderate in magnitude effects were found for the prediction of violent and general recidivism across ethnocultural groups, with very little effect size heterogeneity for Indigenous samples and low to moderate heterogeneity for non-Indigenous samples.

**Calibration properties.** Olver Neumann et al. (2018, see also Olver et al., 2013) conducted a Cox regression survival analysis entering PCL-R total and Indigenous group membership as covariates and found that both variables uniquely predicted increased recidivism; that is, higher PCL-R scores only partly accounted for higher observed rates of recidivism for Indigenous men, indicating that variables unique to Indigenous group membership likely accounted for this association with outcome. Follow-up calibration analyses through logistic regression examining the rates of recidivism associated with PCL-R scores found that higher rates of violent and general recidivism were associated with increasing PCL-R

score across both groups. Indigenous men, however, continued to have higher rates of recidivism associated with any given PCL-R score than non-Indigenous men irrespective of outcome or follow-up.

**Conclusions.** The discrimination properties of PCL-R and PCL: YV scores for general and violent recidivism are generally from the same population of effect sizes (medium) for Indigenous and non-Indigenous justice involved groups, not significantly different in magnitude, and have minimal effect size heterogeneity across studies; the one exception is slightly lower predictive validity for the PCL: YV (small effects) for violent recidivism among White majority non-Indigenous youth groups. The results of calibration analyses demonstrate that Indigenous correctional populations have higher rates of recidivism associated with a given PCL-R score, indicating that factors unique to Indigenous group membership, in addition to individual differences in levels of psychopathy, account for observed patterns of base rate variation between ethnocultural groups.

#### ***Additional Risk Assessment Measures in Need of Psychometric Research with Indigenous Persons***

Three out of the five measures impugned in Canadian legal proceedings and their variants have been the focus of dedicated research efforts with Indigenous persons in adult male Canadian correctional samples (i.e., VRS-SO, Static-99R, and PCL-R). The VRAG and SORAG have since been supplanted by the Violence Risk Appraisal Guide-Revised (VRAG-R, Rice et al., 2013), which has important improvements over its predecessors (e.g., simplicity in scoring, removal of contentious or counterintuitive predictors, removal of clinical diagnoses) but awaits validation with Indigenous persons. Certain major SPJ tools (e.g., HCR-20, SVR-20, and RSVP) have been noticeably absent in Indigenous risk assessment research with adult populations, or have been under researched (e.g., two unpublished evaluations of the SARA).

Arguments have been advanced that SPJ tools, by virtue of being nonnumeric in nature will not generate psychometric bias, but like all measures, have some potential for rater bias or examiner bias (*Ewert v. Canada*, 2015; Venner et al., 2020). We assert that the potential for psychometric bias still

exists, or at least disparities in levels of risk and need identified and consequently, predictive accuracy. For instance, it is possible that Indigenous persons may be disproportionately assigned a summary risk rating of high compared to non-Indigenous persons, or that there is some difference in the accuracy of SRRs in classifying recidivists and non-recidivists. Both sets of findings were found in the Helmus and Forrester (2014a) examination of the SFA as well as to a lesser degree, Stewart et al.'s (2017) examination of the DFIA-R. In contrast, Ahmed (2022) found that Indigenous men were not more likely to be rated as high risk on the SARA, after controlling for SARA item total scores. Given the limited and conflicting findings, further research on possible biases in the professional judgement rating of Indigenous individuals is warranted.

***Interim Conclusions on the Predictive Properties of Structured Forensic Assessment Measures with Justice Involved Indigenous Adult and Youth Populations***

Two global summary tables have been provided for readers summarizing key meta-analytic effects for the preceding set of 22 tools. Table 18 provides a synopsis of general recidivism prediction effects by ethnocultural group, while Table 19 summarizes violent and sexual recidivism effects. Each table directly juxtaposes the computed weighted standardized mean difference ( $d$ ), both fixed and random effects, for Indigenous and non-Indigenous groups, along with the AUC equivalents per Rice and Harris (2005). Tables 18 and 19 also provide the summary label for the overall effect (e.g., small, medium, large), whether the difference is significant between groups (i.e., generally higher for non-Indigenous/White samples), followed by Orwin's fail safe N (FSN) of number of studies required with  $d = 0.00$  to reduce the observed effect below the threshold for a small prediction effect for Indigenous vs. non-Indigenous/White (NIW) groups. Which tools show the greatest vs. least ethnoracial discrepancy in predictive validity?

First, as seen in Tables 18 and 19, most risk assessment tools on average have medium predictive validity effect sizes (i.e.,  $d = 0.50-0.79$  or  $AUC = .64-.70$ ) for recidivism across ethnocultural

groups, and the level of predictive accuracy is high enough to accurately classify individuals in different risk levels of recidivism probability. Research has been generated from international Indigenous samples for the LS variants, Static-99R, and PCL measures, and from Canadian samples for the SPIn/YASI, VRS measures, STABLE-2007, Static-2002R, ODARA, and SARA as well as jurisdiction specific measures such as the SIR scale, SFA, and DFIA-R. The differences in discrimination prediction magnitudes for Indigenous and non-Indigenous persons on risk assessment tools vary according to measure and type of outcome, although on average, effect sizes are larger for non-Indigenous samples with the most pronounced differences appearing to be on static measures. In addition, the FSN, when computed, shows generally a larger number of studies required with zero effect, than the number of studies reported generating a given effect, in order to bring the effect size below the threshold of small. For instance, even though the LS measures have a large effect for non-Indigenous female persons and a small-medium (i.e., fixed-random, respectively) effect for Indigenous female persons, in the prediction of general recidivism based on 6 studies, respectively 18 and 9 studies with  $d = 0.00$  would be required to reduce the effect to  $d < 0.20$  (Table 18). This would give confidence in the robustness of effect sizes and argue against the presence of meaningful publication bias for the meta-analytic effects generated.

Second, purely static measures with large emphases on criminal history and demographics tend to have lower predictive validities and some of the greatest disparities between Indigenous and non-Indigenous groups. For instance, although the SFA is used routinely with Indigenous persons, and the SIR scale (or its variants) is not, paradoxically, the SIR has stronger predictive properties than SFA from the extant research and has been examined more extensively. The SFA and SIR scale have lower categories of prediction effect sizes for general recidivism for non-Indigenous persons and the difference is significant. In the prediction of violent and sexual recidivism (Table 19), the same trend emerges for the SIR scale as well as for the Static-99R and 2002R, with small effects observed most frequently for Indigenous samples and medium effects for non-Indigenous samples. These are robust effects and

differences in effect size magnitudes are not consistently significant; further, greater effects are observed with longer follow-ups. The ODARA (in the prediction of IPV recidivism) also follows this trend, however, this is based on only two studies, and the effect sizes are respectively in the upper bounds of small and lower bound of medium for Indigenous and non-Indigenous persons, respectively. All told, strictly static criminal history and demographics-based tools appear to demonstrate the greatest risk and presence of bias for Indigenous correctional populations. Static risk scores will in part reflect the cumulation of adverse historic, colonial, and systemic influences contributing to criminal history, which can serve to undermine release and reintegration possibilities and compound the problems of correctional overrepresentation. This does not mean static tools should not be used; only that they be done so with caution and preferably alongside a dynamic measure, to where we now turn.

Third, the dynamic measures tended to be closer to predictive parity between ethnocultural groups, both in terms of magnitude of effect size, but also whether differences in prediction magnitude were significant. General risk-need measures organized around the Central Eight robustly predicted general recidivism across ethnocultural groups with broadly medium effects, with some exceptions. DFIA-R had lower effects for Indigenous persons across two large samples, while the SPIn/YASI measures aggregated across youth and adult samples had large effects for both ethnocultural groups, as well as the largest effects overall; this is tempered by the fact that only 3 studies contributed to the effect, in contrast to the LS measures, which had a set of 22 within study comparisons between ethnocultural groups for the same outcome. The LS measures tended to have the same population of effect sizes overall, across males, and for developmental groups; the differences were often significant, however, in part due to the large aggregate sample (i.e., over 100,000 in some cases) and higher  $k$ . The LS measures with Indigenous women and girls had a unique pattern; significantly lower effects in the prediction of general recidivism, but medium effects and not significantly different prediction for future violence. In sum, from a pure prediction standpoint, the greatest potential for bias appeared to be for the DFIA-R,

the least for the SPIn/YASI, and a small (but variable) amount for the LS measures, depending on group and outcome. (The latter would extend to the ORAMS which had non-significant differences in effect size and strong prediction across ethnocultural groups, but medium effect for the Indigenous sample and a large effect for non-the Indigenous sample).

Fourth, the specialized dynamic measures had some variability in evidence for potential bias in the prediction of outcome, and sometimes this was more apparent when it came to the prediction of the targeted outcome (i.e., violence or sexual recidivism) in Table 19. The three VRS measures consistently generated medium effect sizes in the prediction of general, violent, and sexual recidivism in Indigenous samples, while the corresponding non-Indigenous group effect size was greater only for the VRS-SO and VRS-YV in the prediction of sexual or violent recidivism; none of the differences were significant. Further, change effect size associations with decreased recidivism were most frequently small for the VRS-SO across ethnocultural groups; however, on the VRS, change associations were actually larger for Indigenous groups (medium) than for non-Indigenous persons (small). On these bases, the evidence for bias would appear to be low for this group of measures (the order being, in the direction of decreasing bias, VRS, VRS-SO, and VRS-YV), particularly when factoring in change; there appeared to be a small advantage for the VRS proper in Indigenous groups. The STABLE 2007 had large disparities in prediction effects for sexual recidivism between ethnocultural groups (and the greatest evidence for potential bias), although the disparity decreased in magnitude for more general, non-targeted outcomes with the least evidence of bias in the prediction of general recidivism. And in its lone sample, the SARA had uniformly small effects in the prediction of general recidivism and future IPV, although owing to the size of the samples, non-Indigenous group effect sizes were significantly different, suggesting slight potential for bias. Finally, the PCL measures showed some of the lowest risk of bias in the prediction of general and violent recidivism, with generally identical medium range effects across ethnocultural groups in the prediction of these outcomes.

Sixth, the predictive accuracy of structured risk assessment tools with Indigenous justice involved youth appears broadly consistent with that observed for adult counterpart measures. Effect sizes vary somewhat but are generally comparable in magnitude to the adult literature (i.e., medium effects), with well-established tools such as youth variants of the LS and PCL measures. There also appears to be greater research support for use of SPJ measures with Indigenous youth compared to adult samples, specifically the SAVRY. SAVRY effect sizes were medium and equivalent in the prediction of violence across ethnocultural groups, suggesting low bias, however, they were significantly larger for non-Indigenous youth in the prediction of general recidivism, its non-targeted outcome, suggesting the presence of some bias. Further, sample sizes for justice involved Indigenous women and girls have often been fairly small and further research on diverse samples is required to support culturally sensitive, gender responsive, and developmentally appropriate assessment practices for this population. The disparities in prediction magnitudes for the youth tools between ethnocultural groups was often smaller than with adult samples; what may account for this? We believe this may reflect a combination of factors such as: 1. proportionately more dynamic items in some youth measures; 2. the possibility of treatment and reintegration services provided to youth as captured in instrument ratings; 3. less accumulated criminal history at an early developmental juncture, and hence colonial systemic bias, that may adversely influence item ratings and risk scores; 4. the inclusion of unique developmental considerations in youth item content (e.g., role and relevance of family and parenting, school and neighborhood); and 5. possible differences among raters, for instance, service providers immersed in a rehabilitative-oriented youth justice context with an eye toward preventing adult justice system contact.

In all, the measures that appeared to show the least ethnoracial discrepancies in predictive validity were the SPIn/YASI, LS youth variants, PCL-R and PCL: YV, and the VRS measures, including the VRS-SO and VRS-YV. Showing evidence for partial/possible discrepancies (e.g., different categories of effect size across majority of outcomes and/or significant differences despite effect sizes being within

the same category) would be the LS adult variants, ORAMS, SAVRY, SAPROF/YV, SARA, and ODARA. Finally, measures that showed the greatest evidence for Indigenous ethnoracial discrepancies (i.e., different categories of effect size and significant differences in predictive validity across majority of outcomes) were the SFA, SIR, DFIA-R, Static-99R/2002R, and STABLE 2007.

Wilson and Gutierrez (2014) raise four possibilities for such disparities in prediction magnitude. The first is the possibility of systemic racial discrimination in the justice system, for instance, Indigenous persons may be more likely to be arrested and charged and are less likely to enter not guilty pleas; as such, they are more likely to come before the justice system for reasons unrelated to their risk, which in turn will inflate recidivism rates and reduce the predictive accuracy of risk tools. Second, they propose that Indigenous justice involved persons have more risk factors and higher scores on average (e.g., owing to historical disadvantage and circumstances), and that the disparity in predictive accuracy could be reflecting that lower scoring individuals may have additional unmeasured risk factors. Third, it is possible that the same risk factors have different underlying mechanisms or indicators, such that they operate in a different way that may have differential associations with recidivism. For instance, substance abuse may reflect self-regulation problems and increases in risk for antisocial behavior, and/or it may also reflect an attempt to cope with trauma and adversity, which may have particular salience for Indigenous persons (e.g., intergenerational trauma). A fourth possibility is that key risk factors are missed on current tools, including cultural and historical variables such as cultural connectedness, instability of upbringing, or intergenerational trauma. It is also possible that the four explanations all offer a partial account for any observed disparities in prediction.

### **Risk Factors for Recidivism in Indigenous Forensic and Correctional Samples**

Risk assessment measures are ultimately collections of variables with demonstrated empirical or theoretical links to recidivism outcomes. In addition to examining the psychometric properties of risk tools, a growing body of literature has examined specific risk-need domains or psychological risk factors

that may have salience for Indigenous peoples. Minimally, if risk scales are being used for Indigenous peoples, it is important to establish that the risk factors within those tools are valid predictors for this subgroup. These lines of research include meta-analyses and individual studies examining single items or composite scores with Indigenous and non-Indigenous persons. As with the risk measures, these lines of research demonstrate that justice involved Indigenous persons tend to have a larger number of, and/or score higher on, major risk-need domains than non-Indigenous persons (Gutierrez et al., 2013; Huang et al., 2021; Lockwood et al., 2018; McCuish & Corrado, 2018; Prince et al., 2021; Wilson & Gutierrez, 2014). The critical question these lines of research address is what is it about these factors that help them predict, or that may account for prediction disparities? Are any risk-need domains or psychological risk factors more or less predictive or more or less salient for Indigenous persons? As such, this section examines the evidence for individual risk factors for recidivism with Indigenous persons, in addition to considering variables that may not be predictively or conceptually relevant.

The findings are summarized in Table 20 (see also Table 3 for additional study and sample details) and organized into 15 risk-need domains. To provide a summary effect size for a given risk-need domain and recidivism criterion, median values are presented taking the midpoint of ranked effect sizes ( $d$ ); to ensure that single studies (i.e., that were not included in the meta-analyses reported) do not receive undue weight compared to meta-analyses (e.g., Babchishin et al., 2012; Gutierrez et al., 2013; Lee et al., 2020; Wilson & Gutierrez, 2014), the unweighted mean  $d$  value was computed from the single studies and ranked accordingly with the other  $d$  values from meta-analytic effects. For these reasons, as well as for continuity with effect size metrics reported from the meta-analyses of the risk tools, we also reported the overall unweighted mean effect size. (NB: Helmus et al., 2012 and 2021 were treated meta-analytically by taking weighted averages of composite variables when both samples were involved.) The median and unweighted average  $d$  values are each presented for Indigenous and non-Indigenous samples at the bottom of each risk-need domain in Table 20.

***Criminal History (General)***

The extant literature has demonstrated a nonspecific history of prior offenses to be among the most robust predictors of criminal recidivism for Indigenous and non-Indigenous persons, with denser and longer criminal histories to be associated with increased probability of recidivism. Prediction of general and more specific types (e.g., sexual, violent) recidivism, have been demonstrated through meta-analysis (e.g., Babchishin et al., 2012; Gutierrez et al., 2013; Lee et al., 2020; Wilson & Gutierrez, 2014) and large-scale individual studies (e.g., Helmus & Forrester, 2014b; Perley-Robertson et al., 2019) with broadly moderate predictive effects. For instance, large scale meta-analyses of the Central Eight in general (Gutierrez et al., 2013) and from the LS measures in particular (Wilson & Gutierrez, 2014) have demonstrated general criminal history to be a robust moderate predictor of violent and general recidivism among Indigenous persons; both investigations have found a stronger effect in non-Indigenous samples.

Helmus and Forrester (2014b), as noted previously, found the SFA's CHR (Criminal History) subscale to be a consistent predictor of various indicators of recidivism among federally sentenced Indigenous and non-Indigenous men and women (Table 20); offense seriousness as measured by the OSR (Offence Severity) subscale, by contrast, had weaker associations. In the same dataset, Perley-Robertson et al. (2019) subsequently examined the predictive accuracy of the individual indicators of the SFA's CHR and OSR subscales on a large Indigenous ( $n = 1,500$ ) and non-Indigenous ( $n = 6,684$ ) men. While most CHR indicators positively predicted outcomes for both Indigenous and non-Indigenous men, substantially fewer of the OSR indicators positively significantly predicted the same outcomes. There tended to be a pattern of lower effect sizes for Indigenous peoples; several items had significantly lower accuracy for the Indigenous subsample, and few or no items had higher accuracy. Within the sexual offense literature, general criminal history has demonstrated links to recidivism, with effects being smaller for sexual recidivism. Lee et al.'s (2020) Static-99R/Static-2002R Indigenous meta-analysis

examined predictive validity effects for sexual recidivism for individual items and item composites. General Criminality had the strongest predictive accuracy at the item and composite level for Indigenous and White samples, the results of which extended findings from the Babchishin et al. (2012) meta-analysis.

### ***Antisocial Personality Pattern/Psychopathy.***

The Gutierrez et al. (2013) and Wilson and Gutierrez (2014) meta-analyses each found Antisocial Personality Pattern to be among the most consistent predictors of general and violent recidivism in Indigenous samples; the former of the two meta-analyses observed significantly higher effects for non-Indigenous samples, while the latter did not. As noted previously, Olver, Neumann et al. (2018) found PCL-R measured psychopathy (total score) had large in magnitude associations (reported in Table 14) with general and violent recidivism in large multisite correctional Indigenous and non-Indigenous samples. When the individual samples were re-examined through meta-analysis (Table 15), PCL-R scores had broadly moderate in magnitude associations with the two outcomes that were not significantly different between ethnocultural subgroups; a pattern that was also found for youth operationalizations of the tool and construct.

Finally, two large multisite samples of Indigenous and non-Indigenous men with a history of sexual offending examined the associations of psychopathy composite variables developed from selected items from sexual violence risk measures. For Indigenous men, the psychopathy composite developed from the VRS-SO (D4 criminal personality, D6 interpersonal aggression, D13 impulsivity, D14 compliance community supervision) had large associations with general and violent recidivism, and a moderate association with sexual recidivism (Olver, Kingston, & Sowden, 2021). The psychopathy composite developed from the STABLE-2007 (lack of concern for others, impulsive acts, poor problem solving, negative emotionality, and cooperation with community supervision) had significant small in magnitude associations with the three outcomes (Helmus et al., 2021).

### ***Employment/Education***

Meta-analytic reviews have demonstrated small to moderate associations between various operationalizations of the employment/education domain and general and violent recidivism (Gutierrez et al., 2013; Wilson & Gutierrez, 2014) in Indigenous samples. The significance of the difference in prediction magnitude has varied depending on whether fixed vs. random effects models were examined. Four independent examinations of the DFIA-R measures in released cohorts of Indigenous and non-Indigenous men and women have also investigated the predictive accuracy of its seven need domains for general recidivism. Two large samples of federally sentenced Indigenous men and women (Brown & Motiuk, 2005,  $n = 2,635$ ; Stewart et al., 2017,  $n = 3,785$ ), and one federal correctional sample of men (Gobeil, 2011,  $n = 733$ ) found the domain had small to large associations with returns to custody (with or without a new offense,  $d$  equivalents = 0.37 to 0.87), while a sample of Indigenous women observed an inverse association with this outcome (Barnum & Gobeil, 2012,  $n = 90$ ).

### ***Substance Abuse***

The presence of substance use problems linked to crime and violence has been one of the most frequently investigated and empirically established risk markers within Indigenous correctional populations. Both the Gutierrez et al. (2013) and Wilson and Gutierrez (2014) meta-analyses found significant small to moderate effects for substance abuse and general recidivism; effects were slightly smaller but significant in the prediction of future violence, as well as higher still for non-Indigenous samples. Each of the four DFIA-R investigations found significant associations between the measure's substance abuse domain and general recidivism across male and female Indigenous samples, with  $d$  equivalents (0.26 to 0.70) ranging from small to large (Barnum & Gobeil, 2012b; Brown & Motiuk, 2005; Gobeil, 2011; Stewart et al., 2017). Finally, in Olver, Kingston, and Sowden's (2021) treated multisite sexual offense sample, the VRS-SO's substance abuse item (D9; reflecting a link between substance abuse and sexual violence), had consistent moderate in magnitude associations with general, violent,

and sexual recidivism for Indigenous men ( $d = 0.44$  to  $0.53$ ).

### ***Companions/Peers/Associates***

Having a network of friends and acquaintances involved in crime and violence is a well-established risk factor for recidivism. The Gutierrez et al. (2013) and Wilson and Gutierrez (2014) meta-analyses found identical significant small in magnitude effects ( $d = 0.39$ ) for this domain and general recidivism in Indigenous samples, and a smaller significant effect for violence. The collection of DFIA-R studies from Canadian correctional samples listed in Table 20 (Barnum & Gobeil, 2012b; Brown & Motiuk, 2005; Gobeil, 2011; Stewart et al., 2017) similarly reported significant effects that ranged from small to large in the prediction of general recidivism for Indigenous men and women. Finally, Helmus et al.'s (2012, 2021) community based sexual offending samples found the Significant Social Influences item of the STABLE-2007 had significant small effects in the prediction of violent and general recidivism for Indigenous men, but did not predict sexual recidivism; the domain by contrast had significant small to moderate effects in the prediction of each outcome in White samples.

### ***Procriminal Attitudes***

Attitudes supportive of criminal behavior in general as well as specific forms of antisocial behavior (e.g., sexual offending; Helmus et al., 2013) is a well-established Central Eight predictor, with robust small effects obtained in the prediction of general and violent recidivism within Indigenous samples in the Gutierrez et al. (2013) and Wilson and Gutierrez meta-analyses. As seen in Table 20, when measured by the DFIA-R, the criminal attitudes domain had widely varied effects for general recidivism ranging from nil to moderate (Barnum & Gobeil, 2012b; Brown & Motiuk, 2005; Gobeil, 2011; Stewart et al., 2017). Attitudes composite variables from sexual violence risk tools (Helmus et al., 2012, 2021; Olver, Kingston & Sowden, 2021) had small and frequently subthreshold effects for general, violent, and sexual recidivism among Indigenous subsamples of men convicted for sexual offending. The effects were larger, by contrast for non-Indigenous/White samples in the prediction of these outcomes,

except for violent recidivism, which had larger effects for Indigenous samples.

### ***Community Functioning***

Stable community functioning in the form of established residence, positive prosocial supports and services, adequate supervision, and prosocial means of leisure and income are essential for successful reintegration. This domain was among the strongest predictors of general recidivism from the DFIA-R for Indigenous men and women with a moderate median effect (Barnum & Gobeil, 2012b; Brown & Motiuk, 2005; Gobeil, 2011; Stewart et al., 2017). It was similarly predictive of violent and sexual recidivism in Olver, Kingston, and Sowden's (2021) treated multisite sexual offending sample that examined a community functioning composite variable (D10 community support, D11 released to high risk situations, D14 compliance community supervision) from the VRS-SO. Allan and Dawson (2002; see also Allan et al., 2019 for a review) conducted an examination of potential predictors for sexual and violent recidivism among Western Australian Indigenous and non-Indigenous men convicted for one or both classes of offenses. In this study, they examined the incremental validity of item composites extracted from established risk scales and used stepwise procedures to identify the strongest variables and to optimize prediction (NB: the authors did not report the predictive validities of these tools or their subscales.) They identified three predictor variables with particular relevance to community functioning—poor coping, lack of goals, and unrealistic release plans—that provided the strongest prediction. The subgroups were small and AUCs unusually high for unspecified recidivism—AUCs exceeded .90 in subgroups (see supplementary Table S1) and an overall AUC = .84 was found 226 persons across combined Indigenous and non-Indigenous subsamples. Thus, the stability of the “three-predictor” model is unclear and it does not appear to have been replicated.

### ***Leisure/Recreation***

The Gutierrez et al. (2013) and Wilson and Gutierrez (2014) meta-analyses found virtually identical small significant effects for the leisure recreation domain and general recidivism in Indigenous

samples ( $d = 0.33$  and  $0.34$ , respectively); Wilson and Gutierrez (2014) further noted an association with future violence. There was no significant difference between Indigenous and non-Indigenous groups in effect size magnitudes on this domain in either quantitative review.

### ***Family/Marital***

The family/marital domain has been well researched in Indigenous correctional samples. As seen in Table 20, family and marital instability, as operationalized through the LS scales, DFIA-R, or other methods have tended to show a consistent small in magnitude association with general recidivism, and even smaller association with future violence, within Indigenous correctional samples (Barnum & Gobeil, 2012b; Brown & Motiuk, 2005; Gobeil, 2011; Gutierrez et al., 2013; Stewart et al., 2017; Wilson and Gutierrez, 2014). In the sexual violence risk assessment field, this domain has been captured through risk instrument item content reflecting intimacy concerns or relationship instability, with generally small effects observed across different recidivism outcomes in Indigenous samples, and slightly greater small in magnitude effects observed for non-Indigenous/White samples (Helmus et al, 2012, 2021; Olver, Kingston, & Sowden, 2021).

### ***Sexual Deviance***

Sexual deviance is a well-established risk-need domain for men with a history of sexual offending and features prominently in assessment measures and focused intervention programs for sexual violence (Hanson & Morton-Bourgon, 2005). Its relevance to Indigenous persons who have committed sex crimes is less established, and four lines of research reported in Table 20 have examined the predictive properties of the sexual deviance domain with this population; the four studies show a high degree of continuity. First, in their Static-99R/2002R meta-analysis, Lee et al. (2020) found that a Persistence/Paraphilia static composite variable had a small, significant effect for Indigenous men (AUC = .58) and moderate effect for the White subgroup (AUC = .65); the results were consistent with averaged AUCs for concordant items in Babchishin et al. (2012). Further, Olver, Kingston, and Sowden

(2021) examined the predictive and change properties of the VRS-SO's dynamic Sexual Deviance factor (D1 sexually deviant lifestyle, D2 sexual compulsivity, D3 offense planning, D12 sexual offending cycle, D16 deviant sexual preference) in their treated multisite sexual offending sample. The domain significantly predicted 5-year sexual recidivism with a moderate effect (AUC = .69) among Indigenous men who targeted child victims, however, it had a small non-significant effect for those men who had adult victims (AUC = .57). The aggregate effect for the prediction of sexual recidivism in the Indigenous sample as a whole as reported in Table 20 was small ( $d = 0.26$ , AUC = .57), while the effect for non-Indigenous/White men was greater although still small, overall ( $d = 0.40$ , AUC = .61). Of note, the authors found that pre-posttreatment changes in the sexual deviance domain were significantly associated with decreased sexual recidivism in both Indigenous ( $e^B = 0.68$ , 95%CI = 0.51, 0.91) and non-Indigenous ( $e^B = 0.63$ , 95%CI = 0.52, 0.76) groups, supporting the viability of this construct as a treatment and risk management priority across ethnocultural groups.

Finally, in their examination of Static-99R/STABLE-2007 items from a large community sexual offending sample, Helmus et al. (2021) created a sexual criminality composite—young age, non-contact sex offence, male victim, prior sex offences, emotional congruence with children, sex drive, sex as coping, and deviant sexual interest—based on the work of Brouillette-Alarie et al. (2016) and Brouillette-Alarie and Hanson (2015). The effect reported in Table 20 is based on the combined effect from this domain and the commensurate dynamic items from Helmus et al. (2012); a small, non-significant effect was found for Indigenous men in the prediction of sexual recidivism, while a significant association (closer to medium effect, AUC = .62,  $d = 0.42$ ) was found for this outcome for White men.

### **General Risk-Need**

A general risk-need domain, often referred to as antisocial orientation and which overlaps with the Central Eight has been examined in the sexual offense literature in the Olver, Kingston, and Sowden (2021) VRS-SO and Helmus et al. (2021) Static/STABLE study samples. The VRS-SO Criminality domain

comprising 6 dynamic variables (D4 criminal personality, D6 interpersonal aggression, D9 substance abuse, D10 community support, D13 impulsivity, D14 compliance with community supervision) was a ubiquitous predictor for Indigenous men, with large effects for violent and general recidivism; for non-Indigenous men, higher effects were observed for sexual recidivism, but equivalent effects for violent and general recidivism (see Table 20). The authors also found changes in the Criminality factor to be significantly associated with decreased recidivism across the three outcomes for both Indigenous and non-Indigenous groups. Moreover, Helmus et al. (2021) found a static/stable composite variable that they termed General Criminality (prior sentencing occasions, prior nonsexual violence, hostility towards women, lack of concern, impulsivity, poor problem-solving, cooperation with supervision, significant social influences, negative emotionality/hostility) had a small significant predictive effect for sexual recidivism and moderate effects for violent and general recidivism in their large Indigenous subsample. Among White persons, this general risk-need variable had moderate effects for sexual recidivism and large (and significantly higher) effects for violent and general recidivism.

### ***Personal/Emotional***

The personal-emotional domain reflects a pattern of psychological concerns linked to antisocial behavior, including mental health difficulties and poor self-regulation (e.g., problems with stress management, emotional concerns, impulsivity). The DFIA-R's operationalization straddles elements of the same-named domain of the LSI-R as well as the antisocial personality pattern domain from other LS variants. The Gutierrez et al. (2013) and Wilson and Gutierrez (2014) reviews did not explicitly examine this factor, and rather, Gutierrez et al. (2013) examined emotional concerns/distress more specifically (as reviewed next). The four DFIA-R investigations each found small effects that bore a fair amount of consistency in magnitude for general recidivism among Indigenous and non-Indigenous samples, with prediction slightly stronger for the latter (Table 20). This domain tended to be the least predictive of the DFIA-R's seven domains across both ethnocultural groups and across all investigations.

***Emotional Distress and Prior Victimization***

The Gutierrez et al. (2013) meta-analysis found that emotional concerns (e.g., depression, low self-esteem, distress) had small in magnitude, significant associations with general recidivism ( $d = 0.18$ ); however, it had weak, non-significant associations with future violence ( $d = 0.06$ ). This meta-analysis found that a history of victimization, also had significant small in magnitude associations with general recidivism ( $d = 0.23$ ) within Indigenous correctional and forensic samples; again, this domain did not predict future violence.

**Conclusions.** In conclusion, the extant literature examining the predictive properties of risk factor need domains with Indigenous persons broadly support the risk relevance of the Central Eight and related risk-need domains for Indigenous persons, particularly low Education/Employment, and concerns with Substance Abuse, Antisocial Personality Pattern, Community Functioning, and General Risk-Need. Most effects range from small to moderate in magnitude. Other domains had risk relevance but smaller effect sizes (e.g., procriminal attitudes, family/marital). Sexual deviance was a less consistent predictor of sexual recidivism for Indigenous men in sexual offending samples, although the available research would suggest that this is moderated by victim profile (e.g., men who preferentially offend against children). The results have implications for intervention planning and risk management with justice involved Indigenous persons.

**Culturally Relevant Risk Factors and Assessment Measures for Indigenous Persons**

As noted previously, much discussion has raised concerns that mainstream risk assessment tools may disadvantage Indigenous persons through the inclusion of items or indicators on which they will be overrepresented (Clark, 2019; Shepherd & Anthony, 2018; Shepherd & Willis Esqueda, 2017) and by failure to consider factors that reflect the historical circumstances, lived experience, and cultural heritage of Indigenous peoples (Boer et al., 2003; Gutierrez, 2018; Shepherd, 2016; Tamatea, 2017). Very few studies, however, have taken steps to identify or empirically examine culturally relevant risk

factors or assessment measures for Indigenous persons and even fewer examined links to recidivism. In this final section, we review and discuss culturally relevant risk factors, assessment measures, and contextual considerations for Indigenous persons identified from these works.

### ***Culturally Relevant Factors***

**The SAVRY Revisited.** Shepherd and Willis Esqueda (2017) conducted a “cultural peer review” of the SAVRY in a mailout survey to 30 United States Indigenous professionals from health, legal, and education sectors; none were forensic or correctional service providers nor had any been trained in risk assessment. The group was highly critical of the SAVRY’s item content and voiced concerns about negative item labels (e.g., poor school achievement as a criminalization of cognitive or learning challenges), an individualized focus of the instrument (to the neglect of community factors), absence of Indigenous cultural norms, and a lack of attention to Indigenous cultural context. Further, the group identified a number of examples from the instrument’s item content in which Indigenous cultural practices and circumstances may be unfairly pathologized or misinterpreted. For instance, negative attitudes toward authority and reading hostile undertones into benign interactions could be a product of adverse experiences with White authority and might be adaptive for Indigenous youth. Alternatively, a history of violence may be viewed differently if this occurred in self-defense or theft due to economic hardship (e.g., financial and food insecurity). Moreover, truancy and school absenteeism may result if a young person was a caretaker at home for siblings or elders and could not attend school. Similarly, living with a miscellany of extended family, having grandparents as caregivers, or multiple relocations between residences within the family unit may not reflect instability, but rather family norms (in contrast to the “traditional nuclear family”). With these and other examples, the same behavior may be interpreted differently, and it may or may not have any risk relevance.

Rather than abandoning the SAVRY, the group believed the SAVRY had pertinent risk markers and could be helpful; however, it could be strengthened through provision of a cultural addendum or

supplement and adaptation of some of the language and wording to reduce the potential for misinterpretation in order to increase relevance and fairness for use with Indigenous youth. The group was also enthusiastic about the SAVRY's consideration of protective factors—which demonstrated risk mitigation for Indigenous youth in the present study's meta-analysis of the Protect domain—as well as the importance of cultural engagement and connectedness to buffer risk, and by contrast, the perils of cultural alienation that may exacerbate risk (see also, Shepherd, 2015; Shepherd & Anthony, 2018).

**Yókw'tól.** The “Yókw'tól” (Boer et al., 2003) was developed in consultation with Indigenous Elders from CSC's Pacific Region's Mountain and Kwìkwèxwelhp Institutions, the Yukon territory, and other service providers, as a set of professional guidelines for violence risk assessment and management with Indigenous persons. Named by Elder Ritchie and formally endorsed in a ceremony at Kwìkwèxwelhp Healing Lodge, the measure is an SPJ tool comprised of 20 items, each operationalized in a rating manual and scored on a three-point scale (0, 1, 2), where higher ratings generally indicate greater presence of a potential problem relevant to risk. The Indigenous community was opposed to making this a numeric tool that would generate a score divisible into risk bands. There is no psychometric research on the Yókw'tól, and the items are a combination of seemingly static and dynamic factors that take into account the unique experiences and circumstances of Canadian Indigenous peoples as well as variations on other established risk markers. The items include those that are culturally specific (e.g., traditional teachings, ceremonies, and customs; relationship to heritage); that identify historical, generational, and contextual issues with legal relevance (e.g., historical/generational issues, child abuse, foster care history); and victim and community reconciliation considerations (e.g., victim support), given the potential for individuals released from custody and victims within small communities to interact, and to what extent provisions can be in place for the victim and community to heal. Other items appeared to have continuity with the Central Eight (e.g., family/marital relationships, drug/alcohol use, attitudes toward offending, supervision attitudes and

compliance), had empirical support and analogues on mainstream risk measures (e.g., impulsivity, lifestyle stability, risk management plan, support), or emphasized growth and recovery (e.g., gains, insights, behavioral changes; unique resiliency factors).

The language of the items was framed using Indigenous specific examples, even those consistent with the Central Eight or present in other forms on other risk measures. This language seemed to increase the relevancy for how a given common risk factor might be uniquely manifested with Indigenous persons (e.g., Shepherd & Willis Esqeda, 2017; Wilson & Gutierrez, 2014). Despite the considerable effort and Indigenous collaboration leading to the development of this Indigenous-specific structured risk assessment measure, the measure has yet to be validated or adopted into use for unknown reasons. As previous legal developments highlight (re: *Ewert v. Canada*, 2015; 2018), insufficient effort had been invested in validating existing risk tools for Indigenous peoples. Moreover, alternatives already available had also not been adequately explored or validated. As such, there is a pressing need for researchers to comprehensively examine the psychometric properties of culturally-informed measures such as the Yókw'tól.

**Culturally Relevant Factors Questionnaire.** Gutierrez (2018) conducted a series of group and individual interviews with an Indigenous Advisory Group of 51 professionals (29 women and 22 men, 37 of whom were Indigenous) who worked with Indigenous persons in correctional settings to identify culturally relevant factors. Thematic analysis of group narratives generated 12 key areas: 1. Cultural Background and Indigenous Identification; 2. Stability of Upbringing; 3. Residential School Experience (direct); 4. Family Experience of Residential School; 5. History of Exposure to Abuse; 6. Coping Strategies; 7. Extreme Poverty; 8. Preparation for and Experience of Discrimination; 9. Community; 10. Views on Justice; 11. Employment and Educational Issues; and 12. Substance Abuse. As with the Yókw'tól, the themes identified were a combination of culturally specific, potentially culturally salient, and common risk-need variables. These themes were used to generate a Culturally Relevant Factors Questionnaire

(CRFQ), which was then distributed to 36 Indigenous men with prior legal involvement under community supervision and living in First Nations Manitoba communities. Many of the themes resonated with the sample in their completion of the CRFQ, although there was also considerable variability in the endorsement of item areas. Due to the small sample size, associations with criminal behavior could not be examined, although the themes and the CRFQ may have promise and warrant further research with Indigenous correctional populations to inform possible applications.

Kyne (2022) used a modified and expanded version of the CRFQ to identify correlates of having a criminal record among 144 Indigenous participants residing in Canada. Participants were recruited from online research platforms via CloudResearch, and 47 (32.6%) self-reported having a charge or conviction for a criminal offence. This study was exploratory and temporal ordering of some of the risk factors and criminal charges/convictions is unknown (although childhood variables would obviously precede the criminal behavior). The strongest correlates of having a criminal record were related to substance abuse. Those who participated in traditional Indigenous activities or ceremonies within the last 12 months were less likely to have a criminal record, with a small effect that did not reach significance ( $AUC = .41$ ,  $p = .093$ ). Many childhood trauma variables (e.g., physical or sexual abuse) were not significantly related to having a criminal record, although effect sizes were small ( $AUCs = .57$  to  $.58$ ) and exceeded Mann et al.'s (2010) threshold ( $d = 0.15$  or  $AUC \approx .54$ ) to be a worthy treatment target. Having been sent to residential schools (or having family members sent to residential schools) were generally unrelated to crime ( $AUCs = .52$  to  $.55$ ). Having been placed in foster care or group homes, however, was significantly related to having a criminal record ( $AUC = .60$ ). History of suicide attempts in childhood and adulthood had small but meaningful relationships to crime ( $AUC = .59$  and  $.61$ , respectively, although only the latter was statistically significant). Consequently, although general mental health problems (e.g., anxiety/depression) are not considered criminogenic needs (Bonta & Andrews, 2017), this study suggests that they may be risk-relevant for Indigenous peoples.

**Cultural Identity, Engagement, and Connectedness.** Finally, at least two studies have examined cultural identity, engagement, or connectedness to decreased recidivism in Indigenous samples. First, Shepherd, Delgado et al. (2018), in a sample of 122 men and women from 11 Australian prisons in the state of Victoria, examined the role and relevance of Indigenous cultural identity and engagement to violent offending post release. Although cultural identity (i.e., the extent to which an Australian Indigenous person identified with their cultural heritage) was unrelated to outcome ( $e^B = 0.97$ ), as seen in Table 20, cultural engagement was significantly inversely associated with violent recidivism ( $e^B = 0.80$ ). That is, individuals who reported deeper commitment and immersion in their culture had lower rates of violent recidivism; a conservative conversion to  $d$  would make this the equivalent of a small effect ( $d = 0.17$ ). The finding is consistent with the notion of cultural connectedness to be a protective factor or resiliency (e.g., Boer et al., 2003; Gutierrez, 2018; Shepherd & Willis Esqueda, 2017) and its absence, a risk factor. Shepherd (2015) further notes that Indigenous social and emotional wellbeing “is conceptualized in a holistic framework that encompasses mental, physical, social, and spiritual aspects” (p. 1299). It is strengthened by cultural connectedness, while conversely, complex factors stemming from colonization (e.g., ongoing marginalization, disadvantage, discrimination) are associated with its disruption. Second, Hansen (2018) examined the predictive properties of the LSI-R and a measure of cultural connectedness developed by the author (the Cultural Connectedness Scale, CCS) in a US correctional sample of 166 Indigenous men and women, released to a Montana reserve through the Flathead Reservation Reentry Program. Directly akin to the notion of cultural engagement per Shepherd, Delgado et al. (2018), the CCS measured connectiveness, knowledge, access, participation and desire to learn Indigenous culture per the Flathead Reservation. The results of hierarchical logistic regression analyses demonstrated that LSI-R scores uniquely predicted increased recidivism ( $e^B = 1.110$ ), while CCS scores uniquely predicted decreased recidivism ( $e^B = 0.888$ ); that is, after controlling for baseline risk and need, cultural connectedness seemed to have a protective effect toward decreasing recidivism.

In the realm of general psychology, Fox et al. (2018) developed a self-report measure of cultural embeddedness among Māori people that may have promise as a protective factor for criminal behaviour. The 16-item measure assesses whānau (extended family) connectedness, connectedness to Māori peers, cultural awareness, and te reo Māori language attitudes and fluency. In a longitudinal study of 403 youth aged 9 to 15, their measure predicted adaptive coping, and adaptive coping mediated a relationship between cultural embeddedness and overall wellbeing. Whether cultural embeddedness predicts other outcomes such as reduced recidivism is a worthy area of exploration. Although the measure was developed in the Māori context and includes terminology in te reo Māori, the item content appears as though it could be adapted to Indigenous groups in other countries (or phrased more broadly), although psychometric validation of such adaptations would be necessary.

### ***Some Cultural Cautions and Conclusions***

There may be room to psychometrically or contentually augment existing risk measures, particularly for item domains or language that may reflect the historical and lived circumstances of Indigenous peoples. But there are very few potentially cultural-specific predictors and the research to date is scant; candidate variables await further validation. Caution may also be warranted about developing replacement culture specific tools or radically overhauling the form and content of established risk tools with ethnocultural indicators that might have a high level of face validity, but little or no predictive validity (Shepherd et al., 2018; Shepherd & Spivak; 2021). For one, variables that characterize a majority or possibly even all members of a group of people (e.g., historical/generational trauma), run the risk of becoming a proxy for race or culture, and will not be able to differentiate people at different levels of risk. Psychometrically speaking, the variable may be such a high base rate phenomenon that it almost becomes a constant, incapable of discriminating recidivists from nonrecidivists, given that the two would be equally likely to have the characteristic. As Shepherd and Spivak (2021) warn, “efforts to culturally amend instruments, could paradoxically reduce accuracy and

increase bias.” (p.1). If a forensic risk tool is culturally augmented, then predictive accuracy for relevant outcomes needs to be consistent with established, orthodox risk instruments; clients and agencies would be at a disservice if predictive validity was sacrificed for face validity. Shepherd and Spivak (2021) further caution that the inclusion of ethnoculturally specific variables that are unique to a group and their circumstances (e.g., cultural erosion, experiences of racism), runs the risk of: 1. criminalizing such experiences as they become a variable that is now a marker of elevated recidivism risk, and 2. potentially inflating the risk levels of individuals from that ethnocultural group since given that they more than likely possess the experience or characteristic.

If the aforementioned review tells us anything, it is that the content of any given assessment measure is of obvious importance, but this takes a backseat to the content of the very service providers who work with vulnerable and marginalized populations. Cultural sensitivity, safety, humility, and general responsiveness are critical in humane and effective risk assessment and correctional service delivery; it behooves service providers to not recapitulate deep seeded oppression and inequity by way of culturally insensitive practice. Cultural safety (i.e., self-reflexivity of cultural belief systems) and cultural humility (limitations of cultural awareness and competence) (Shepherd, 2019) in particular, are evolving professional competencies required across human service delivery, including risk assessment and associated client assessor interactions (Kois & Chauhan, 2016; Santiago Schmitt et al., 2021). We argue that strengthening such competencies in tandem with use of no less evidence informed and empirically based methods can contribute to effective, fair, and accurate risk assessment practices that can also be culturally responsive.

### **Discussion**

Forensic risk assessment has a contentious past with ethnocultural minority groups and Indigenous persons specifically. The first president of the American Psychological Association, G. Stanley Hall, is reported to have “described First Peoples as childlike, calling for ‘civilizing programs’ that

destroyed their culture, language, and spiritual life” (Deangelis & Andoh, 2022, p. 24). Such a characterization is emblematic of a long history of enacted colonialism, systemic racism, deculturation, and inequality experienced by Indigenous persons, perpetuated by criminal justice systems, and culminating in the overrepresentation of Indigenous peoples in correctional settings.

Research on forensic risk assessment measures with Indigenous peoples has intensified in the last 5 to 10 years, particularly since the *Ewert v. Canada* (2015, 2018) decision in Canada which has reverberated to international jurisdictions (e.g., Day et al., 2018; Shepherd & Lewis-Fernandez, 2016). Previous quantitative reviews have featured specific tools such as the LS measures (Olver et al., 2009, 2014; Wilson & Gutierrez, 2014) and Static-99R (Ahmed et al., 2023; Lee et al., 2020), focused on the Central Eight (Gutierrez et al., 2013), or were narrative literature reviews (e.g., Allen et al., 2019; Day et al., 2018; Olver, 2016; Olver & Stockdale, 2021; Rugge, 2006; Shepherd & Lewis-Fernandez, 2016). To our knowledge this is the first systematic review and meta-analysis of the youth and adult Indigenous risk assessment literature across multiple measures—22 measures in total across 91 studies and 84 published and unpublished source documents from three countries. Here we review key themes and contextualize our findings with implications for risk assessment research and practice with justice involved Indigenous persons.

### **Ethnocultural Group Disparities and the Threshold of Predictive Accuracy**

Most risk tools and related forensic measures were developed in Canada and imported to other countries to assess and manage risk with their justice involved populations. Although important research inroads have been made, applications of this body of work require replication and extension to different samples, settings, and jurisdictions. Empirical actuarial measures have been the most heavily examined approaches for use with Indigenous persons. Results of our meta-analysis demonstrated that Indigenous persons had higher rates of recidivism and on average scored higher on risk measures by approximately one half standard deviation with much variability. Score differences seemed particularly

evident on static measures (around 0.50 SD) that heavily weight criminal history as well as personal and victim demographics (e.g., ODARA, Static-99R/2002R); although that said, notable differences were observed for the LS measures ( $> 0.50$  SDs) and the youth version of the VRS. Score differences were frequently smaller in magnitude on certain dynamic and protection measures as well as risk change (e.g., from treatment, discussed later below).

Research on large Canadian samples of Indigenous and non-Indigenous men has been generated for the LS measures, VRS, VRS-SO, and VRS-YV, STABLE-2007, Static-99R/2002R, PCL-R, SARA, as well as the SIR and CSC's "in house" SFA and DFIA-R measures; the results were bolstered for a number of these measures through meta-analysis. Similar findings have been obtained for parallel youth measures with both male and female Indigenous youth. The findings have broadly supported the predictive (discrimination and calibration) properties of the tools for justice involved Indigenous youth and adults. Most risk assessment tools on average have moderate predictive validity (e.g.,  $d = 0.50-0.79$  or  $AUC = .64-.70$ ) for recidivism with Indigenous persons with some variability; the SPIn and YASI demonstrated large effects while the SFA and DFIA-R demonstrated small (or lower) effects.

The differences in discrimination prediction magnitudes for Indigenous and non-Indigenous persons on risk assessment tools varied according to measure and type of outcome, although several prediction effect sizes were larger for non-Indigenous samples. For instance, the effect sizes for the LS measures overall, and in adult samples specifically, pointed to large effects sizes for non-Indigenous persons and consistently medium effect sizes for Indigenous samples. A significant difference in prediction magnitudes was observed in female Indigenous samples and across international jurisdictions. By contrast, the differences in predictive validity of the LS measures for youth samples was not significant and effect sizes were consistently moderate in magnitude for Indigenous and non-Indigenous youth. Moreover, the Static-99R and 2002R predicted 5 and 10-year sexual recidivism across ethnocultural groups, with the results of our updated meta-analysis demonstrating small to moderate

predictive validity for Indigenous samples (AUC = .62-.66) and moderate to large prediction effects for non-Indigenous samples (AUC = .70-.77). Significant ethnoracial group differences were frequently seen in the magnitude of prediction effects for the two measures, however, consistent with a recent Static-99R meta-analysis (Ahmed et al., 2023).

In turn, the results from our review of the Central Eight need domains supported their risk relevance for Indigenous persons, particularly problems with education/employment, substance abuse, antisocial personality pattern (and psychopathy), as well as a community functioning domain. The findings suggest some continuity in the risk-relevance of the risk constructs captured by risk assessment tools, but also those that may have some salience for Indigenous persons (discussed in “Forensic Risk Assessment Considerations”). For instance, meta-analyses of PCL-measured psychopathy (which converges with antisocial personality pattern) demonstrated broadly moderate effects in the prediction of general and violent recidivism in youth and adult samples; effect size magnitudes were not significantly different between Indigenous and non-Indigenous groups. Rather, the PCL-R/YV was one of the measures with the least amount of evidence for possible ethnoracial bias in the prediction of general and violent recidivism. Although measures such as the PCL-R/YV are prone to misuse and require caution and proper training in their use, integrated within a comprehensive risk assessment (Olver, Stockdale, et al., 2020), the results demonstrate the risk relevance of the psychopathy construct for Indigenous and non-Indigenous persons in forensic assessment contexts.

With the above findings taken together, what is the bare minimum level of acceptable predictive accuracy? Mann et al. (2010) propose that a minimum Cohen’s *d* of 0.15 (AUC  $\approx$  .54) is sufficient to say a variable is a worthy risk factor or intervention target (i.e., worth assessing and directing services toward). That said, we would expect a scale (combining risk factors) should do better than a single risk factor alone. Viljoen et al.’s (2021) umbrella review of structured vs. unstructured judgment summarized different reviews and meta-analyses, with the median AUC for unstructured professional judgement

around .56, which also corresponds to the bottom end of a small effect size per Rice and Harris (2005) and Cohen (1992). As such, a strong argument can be made that any meaningful or significant improvement over  $AUC = .56$  ( $d = 0.20$ ) would be justifiable to use a given risk measure; most of the risk measures examined in this quantitative review exceeded this threshold. The evidence from our meta-analysis suggests that the potentially best measures to attain the strongest and least ethnoculturally biased predictive validity would appear to be the SPIn/YASI, LS youth variants, PCL-R and YV, and VRS/VRS-SO; the weakest measures would be the SFA and DFIA-R. Other measures have acceptable predictive validity yet also significant ethnoracial differences in effect size (e.g., Static-99R/2002R). The “best” measures, in turn, will be able to distinguish about 65-75% of potential recidivists from non-recidivists on the basis of their risk scores to be identified for risk management services to prevent future crime and violence.

This begs the question of what degree of prediction disparity is acceptable for Indigenous and non-Indigenous persons? This issue is even more challenging to address, as no single tool should be used in isolation and the use of one or more measures to drive service delivery matters more than any given prediction metric (Kroner, 2016). It also depends on the availability of empirically validated alternatives. For instance, a scale with moderate accuracy for Indigenous persons but with substantial prediction disparity may be preferable to a scale with low accuracy but no disparity, at least until improved options are available. We offer the following considerations when using any given tool or set of tools with Indigenous criminal justice populations. First, tools should have moderate predictive validity for its targeted outcome when used with Indigenous populations. A tool with moderate predictive validity will have value for risk classification, informing treatment intensity or dosage, identifying targets for intervention, informing release decisions, and guiding post-release risk management. This does not mean that tools with small effects should never be used, but that there needs to be clinical and empirical justification for doing so. For instance, small significant effects may be present across

ethnoracial groups (as seen with some IPV measures, such as the SARA), absolute difference in prediction magnitude between groups may be minimal or not significant, or a measure with a small effect may have the largest base of empirical support and prediction is improved when used with another measure (e.g., Static-99R and STABLE 2007).

Second, and relatedly, when selecting a tool for use, it is preferred that prediction metrics fall within the same general category of effect size (e.g., medium) across ethnoracial groups, and/or are not significantly different (e.g., in the case of a small vs. medium effect, or a medium vs. large effect). Finally, assessors should consider if there are other, central features of the tool that may offset potential prediction or classification disparities, with clinical and empirical evidence to suggest their use may have beneficial applications (rather than cause harm) with Indigenous persons. Such examples would include using tools that can identify treatment targets or providing valid appraisals of change to capture risk reduction. For instance, there may be evidence for changes across repeated assessments that can be reasonably attributed to treatment or risk management, and such changes are associated with reductions in recidivism (discussed in further detail on “The Need to be Dynamic”).

How do we proceed with tools that show significant ethnoracial group differences in prediction magnitudes for a targeted outcome, even if the measure significantly predicts that outcome? When should a tool not be used? Assessors need to consider the nature and magnitude of a prediction disparity; is it significant in a statistical and/or clinical sense, and if so, does use of the measure potentially disadvantage Indigenous persons? Is there a less biased or unbiased alternative that can be used? We submit that if the use of a given tool would appear to disadvantage an Indigenous client owing to the presence of ethnoracial psychometric bias, it should not be used. To do otherwise would constitute committing an act of social injustice in addition to contravening ethical standards in professional psychology practice in most jurisdictions. Examples of this scenario would be standalone use of tools for which there is strong evidence for small or subthreshold predictive accuracy, or wide

disparities in prediction magnitudes (e.g., small vs. large effect) between ethnoracial groups, without qualifying the limitations of the scale's applicability to Indigenous persons. We suggest this could also include singular use of one or more static tools in the absence of a validated dynamic tool with Indigenous persons, where it is reasonably feasible to use a dynamic tool. In such instances, when tools are used inappropriately, they can harm and hinder, especially for marginalized groups.

It then behooves assessors to find the best, and least biased alternative, in a given risk assessment scenario. As such, when selecting tools for use in risk assessments with justice involved Indigenous clientele, assessors should utilize measures with: 1. a stronger research base, including replication across Indigenous samples; 2. good prediction effect sizes (e.g., medium or higher, unless clinically and empirically justified to use a tool with a small effect); and 3. evidence for minimal prediction-centered bias in properties of discrimination and calibration. Unfortunately, in some situations these three considerations will not point to the same tool, or there may be no tools that reasonably satisfy all criteria. There are benefits to using multiple tools, with some kind of "averaging" approach being optimal (Lehmann et al., 2016). Consequently, in more comprehensive assessments or for particularly high stakes decisions, it may make sense to use a combination of tools and note the strengths and weaknesses of the research support for each when integrating the results. In any situation where no tool meets all three of the criteria above, evaluators should clearly outline the limitations of the best available tool and be particularly cognizant of ways in which the results of the tool could disproportionately impact Indigenous individuals. Note that this latter suggestion is the best of a set of suboptimal choices and should not be mistaken for an acceptable status quo for the future. It is an interim solution for decisions that must be made in the present, but clearly signals the ethical responsibility of governments to remediate this situation (i.e., to take clear and actionable steps towards developing the research base to satisfy the three criteria above).

Our review of the calibration properties (i.e., absolute risk) across the suite of existing measures

revealed a different, although related, set of issues regarding the potential for ethnoracial prediction bias. As expected, increasing scores on all the risk tools we surveyed were generally associated with higher observed and predicted rates of recidivism within a given recidivism category; however, we found the same score on risk tools can sometimes generate different recidivism estimates for Indigenous and non-Indigenous persons. This issue of calibration with tools such as the LS/CMI, Static-99R, VRS-SO, VRS-YV, and PCL-R seems to particularly affect the lower score bands, where logistic regression modelling demonstrates that Indigenous persons appraised as lower risk, have higher rates of recidivism at low scores than would be predicted by the tool, whereas non-Indigenous persons have rates consistent with low projections. E/O index analyses on LSI-SK and ODARA scores have yielded a similar pattern. For this reason, separate norms on such measures for Indigenous and non-Indigenous persons would likely be contraindicated, as the recidivism projections for Indigenous persons would be higher than those generated from a common set of risk norms.

As we previously note, per Wilson and Gutierrez (2014), differences in the discrimination and calibration properties of risk measures between ethnoracial groups may be attributable to several factors such as unmeasured risk factors, different ethnocultural manifestations of the same risk factor (e.g., substance abuse), or non-risk relevant reasons (see below) that are contributing uniquely to higher recidivism rates for Indigenous peoples than would be forecasted by a risk tool. For this reason, differences in predictive properties may not necessarily reflect deficiencies in a given risk tool inasmuch as they may reflect systemic, historical, contextual, or environmental factors not captured by the tool (and not necessarily tied to risk) that increase the potential for being charged and convicted for a new crime. Non-risk relevant reasons (over-policing, aggressive prosecution, systemic racism and disadvantage) are likely to play some role in disparate findings (Fergusson et al., 2003; Nettlebeck & Smandych, 2010; Perry, 2006) as risk assessment cannot be removed from the colonial and racist context of societies in which these tools are administered (Muir et al., 2023). Other explanations (e.g.,

unmeasured or insufficiently operationalized risk factors) may continue to play an important role, but there is insufficient research exploring these explanations.

So, what is the remedy? We do not yet have an ideal accepted practice. Ultimately, for instrument developers and researchers, this remains a work in progress given the inherent challenge of disentangling measurement issues from systemic differences in society. We believe the answer may partly lie in further research on the potential sources of variation or moderators of predictive accuracy of existing measures which in turn may inform the application or even refinement of these measures for use with Indigenous persons. A potential remedy, we believe, also lies within the application of culturally responsive, comprehensive, and contextualized assessment practices incorporating the responsible application of empirically substantiated measures, to which we speak in greater detail in the final two sections of this Discussion.

#### ***A Word on Structured Professional Judgment Measures***

With the exception of the SAVRY, SARA, and CSC's two inhouse intake risk assessment measures (SFA and DFIA-R), there is a noticeable absence of Indigenous risk assessment research with other prominent SPJ tools such as HCR-20, SVR-20, and RSVP. The development and item composition of these measures is similar to empirical actuarial tools, so it would be reasonable to anticipate that these measures could also have adequate discrimination and calibration properties, but until this is formally examined, it remains an untested proposition. Per the Supreme Court of Canada (*Ewert v. Canada*, 2018) ruling, research demonstrating the psychometric properties (validity and reliability) of measures intended to be used with Indigenous persons should be in place when employing any given instrument or approach with those persons. No measure should be exempt given that biases from myriad sources (e.g., examiner, psychometric) can influence the administration, scoring, and/or interpretation of results for any measure. Although the item composition tends to be similar between actuarial and SPJ measures, the structure is different, with the latter injecting greater professional judgement. The

increased discretion could be beneficial (e.g., cultural sensitivity in considering the risk factors and unique circumstances) or could exacerbate human biases and prejudices (or both). Helmus and Forrester's (2014a,b) results from the SFA suggest the latter effect may be more likely.

### **Forensic Risk Assessment Considerations with Justice Involved Indigenous Populations**

Our findings have implications for forensic risk assessment research and practice with justice involved Indigenous persons including: 1. the need to be dynamic, 2. the role and relevance of protective factors, and 3. critical cultural and diversity considerations for forensic practice with Indigenous peoples. We discuss each of these in turn, including outlining some recent guidance for risk assessment with Indigenous clients from the Society for the Advancement of Actuarial Risk Needs Assessment (SAARNA).

#### ***A Need to be Dynamic***

Our findings highlight a need to incorporate dynamic measures into service delivery with Indigenous persons. First, from a pure prediction standpoint, several dynamic measures evinced moderate or higher predictive accuracy for their targeted recidivism criteria within Indigenous samples, such as the LS and SPIn/YASI measures for general recidivism, the VRS, VRS-YV, and SAVRY for violence, and the VRS-SO for sexual recidivism. Second, smaller and frequently non-significant meta-analytic group differences were found between Indigenous and non-Indigenous men on specialized dynamic measures such as the VRS and VRS-SO dynamic scores ( $d = 0.46$  and  $0.16$ , respectively); there were no significant group meta-analytic differences on VRS or VRS-SO change scores ( $d = 0.06$  and  $-0.12$ , respectively), which reflect positive reductions in risk from intervention or other change agents. Third, dynamic measures can drive services—they can inform where to intervene, they can track and monitor changes in risk, and they can inform community reintegration and risk management efforts.

Fourth, and finally, the changes captured by dynamic measures have risk relevance. Justice involved persons can and do make changes for the better, and when completing evidence informed and

culturally responsive programming, their risk to reoffend is reduced. Our meta-analytic results of risk-change effect sizes from the VRS and VRS-SO with treated Canadian violent and sexual offending samples demonstrated that Indigenous men made substantively similar amounts of risk change as non-Indigenous persons. The VRS samples featured men who had attended high intensity violence reduction programs while the VRS-SO samples similarly had a high proportion of men attending high intensity sexual offense treatment or prison-based sexual offense programs of different intensity levels across Canada. The mainstay of these programs is that they all integrated Indigenous cultural programming (e.g., sweat lodge), staff (e.g., Elders), spiritual practices (e.g., smudging), with programming that is CBT-based, westernized, and criminogenically focused. Accordingly, such changes in risk were associated with decreased sexual and violent recidivism; for the VRS-SO, the change-recidivism associations were significant for Indigenous and non-Indigenous samples ( $d = -0.30$  and  $-0.37$ , respectively), and in the case of the VRS, the change associations were higher for Indigenous samples, particularly with respect to decreased violence ( $d = -0.56$  and  $-0.29$ , respectively).

Taken together, several factors may partly account for what would appear to be reduced bias for the trio of VRS measures in their predictive properties of risk and change scores with Indigenous persons. These measures are predominantly dynamic in nature and were developed and validated on samples with large Indigenous representation, within treatment settings that provided integrated mainstream correctional and culturally responsive intervention services. Measuring and incorporating risk-relevant change has always been a *raison d'être* of the VRS measures (Wong & Gordon, 2006), while the operationalization and contextualization of risk and change embedded within the instruments' item content would have been informed by clinical contact with, and the lived experiences of, Indigenous clients. Finally, methodologically, the raters of the measures, whether these be clinical program staff or student research assistants, in light of their educational, training, and practice contexts with Indigenous persons, may exercise greater cultural discretion in their ratings.

### ***The Role and Relevance of Protective Factors***

Capturing decreases in risk can equate to the development of prosocial skills and competencies (e.g., Coupland & Olver, 2020b), but there is also a need consider strengths, protective factors, and resiliencies that can mitigate risk and promote positive outcomes with all clientele. Miller (2006) convincingly argues that a singular narrowed focus on risk factors can fuel pessimism among service providers and create an orientation toward of trying to “predict the bad” rather than “encourage positive outcomes” (e.g., obtaining work, finishing school, supporting a family, maintaining a stable roof over one’s head). Further, Miller (2006) and others elsewhere (e.g., Laub & Lauritsen, 1994; Rogers, 2000) argue that focusing solely on risk factors may generate a negativity bias, contribute to unbalanced reports, and translate into longer periods of detention; all of which is contrary to client welfare and the best interests and missions of correctional systems and would seem to make protective factors particularly salient for Indigenous persons. Most typically, protective factors have been evaluated indirectly on measures such as the DFIA-R and LS measures which consider domains that are not problematic to be an “asset” or a “strength.” The generation of formalized operationalizations of strengths or protective factors moves away from the practice of assessing protective factors as an absence of a risk factor, and rather, as an indicator of positive prosocial competencies.

Protective factors may be included through informal means by way of identifying factors that mitigate risk (e.g., prosocial coping, cultural and family supports, spirituality, positive leisure) or via measures such as the SAPROF and its youth version, the SAPROF-YV; these too require psychometric research with Indigenous persons which, to date, is limited. Indeed, in our review, less than 5% of the studies culled formally examined protective factors. Our review showed that there is preliminary support for the role and relevance of protective factors for Indigenous youth and adults through respective use of the SAPROF-YV (Lovatt et al., 2022) and SAPROF proper (updated data analysis from Olver & Riemer, 2021), in addition to the Protect domain of the SAVRY (for youth). Although Indigenous

youth and adults had lower protective factor scores from these reviews, protective factors were associated with decreased violent and general recidivism across developmental groups. But there is no consensus on how protective factors may be harnessed in order to adjust appraisals of risk, or how the two may be integrated in a useful, systematic, and evidence informed manner. One approach is to examine recidivism trajectories at different levels of risk and protective factors such as through survival analysis or logistic regression (Lovatt et al., 2022; Olver & Riemer, 2021); the evidence, however, is inconsistent as to what extent protective factors may be incrementally predictive of outcome beyond changes in risk (e.g., from treatment) (Coupland & Olver, 2020; Lovatt et al., 2022). This is also influenced by the lack of clarity on the extent to which risk and protective factors overlap in underlying constructs (Helmus, 2018). Continued exploration of strengths and mitigating factors, including protective factors for Indigenous peoples, is worthwhile and may serve to strengthen correctional and reintegration plans.

***Culturally Relevant Factors, Cultural Competence, and Cultural Diversity: Implications for Practice***

Our review also identified several potential candidates for culturally relevant risk-need factors and protective factors, some which converged with the Central Eight or item domains captured on other risk measures to varying degrees (e.g., instability of upbringing). Other variables identified included salient cultural, historical, and contextual issues for Indigenous persons. Few of these culturally relevant variables had been empirically examined. An exception was the notion of cultural engagement or connectedness, associated with decreased recidivism (or cultural alienation, increased recidivism), while another was the concept of Indigenous social and emotional wellbeing (the lack of which would be a risk factor, per Gutierrez et al., 2013; Shepherd, 2015). For variables to be effective in accurately appraising risk and driving services, they cannot, however, serve as proxies for race, culture, or ethnicity (e.g., economic disadvantage, intergenerational trauma).

The literature also cautions about overgeneralizing among Indigenous peoples given their

substantial diversity, for instance, with over 630 First Nations communities in Canada representing 50 nations and 50 languages (Government of Canada, 2021) and more than 570 federally recognized North American Indigenous tribes in the United States (Bureau of Indian Affairs, n.d.). This includes making presumptions regarding an individual's desired personal connection to their culture, and its perceived relevance to assessment and intervention planning (Shepherd, 2019). Culture is also multifaceted and may include age, gender, national origin, sexual orientation, spirituality, as well as other aspects of sociocultural diversity and microcultures (Hays, 2001; Santiago Schmitt et al., 2021; Tamatea, 2017). Professional guidelines have been developed advocating for a layered ecological model (i.e., of nested, interacting social systems) for multicultural service delivery (American Psychological Association, 2017).

Further research is needed on incorporating Indigenous perspectives and ancestral ways of knowing, with Western knowledge and ways of knowing—what has been referred to as “two-eyed seeing (Canadian Psychological Association; CPA, 2018a)—to better connect forensic assessment and case formulation to effective and culturally responsive interventions (CPA, 2018b; Olver & Stockdale, 2021; *R. v. Gladue*, 1999). There are many ways to possibly achieve this. Consultation with Indigenous communities on assessment content as well as risk construct operationalization and communication (e.g., language, framing) is essential (Hanson, 2020). Moreover, the use of an interview guide could help increase sensitivity in lines of questioning and information solicited to complete item ratings (e.g., Yókw'tól, DSM-5 cultural formulation interview). We would also recommend using collaborative risk assessment practices (e.g., self-assessment of risk factors, collaborative safety planning) to empower clients towards positive risk management (Markham, 2020).

Trauma is inherent to correctional populations and services may be augmented by incorporating a trauma informed lens per the responsivity principle. People with complex trauma histories as identified through the assessment and intake process may also benefit from focused services that acknowledge, validate, and help to process trauma (Haskell & Randall, 2009; Levenson, 2014; Olver &

Stockdale, 2021). Although research examining mental health and risk is extensive with correctional and forensic populations in general (Bonta et al., 1998, 2014), this area of research with Indigenous persons is comparatively lacking. Kyne's (2022) preliminary research, however, suggests some trauma-related variables may have particular risk salience for Indigenous persons. It is also important to consider community variables that could impact successful reintegration which may include barriers to accessing services and supports (e.g., housing, employment) and other potentially destabilizing social and environmental factors (e.g., negative community response to release, financial and residential instability, food insecurity, health system discrimination) (Olver & Stockdale, 2021).

Although the forensic relevance of cultural variables may be unclear, culture in a general and humanitarian sense, matters. Perlin and McClain (2009) define cultural competence as an ability to "grasp, reason, and behave effectively when faced with culturally diverse situations, where assumptions, values, and traditions differ from those to which they are accustomed" (p. 259). Service providers cannot be expected to have intimate familiarity with all the different cultures of the clients they work with, nor should there be assumptions that an Indigenous person will identify with all aspects of their traditional heritage. Service providers (conducting assessment or intervention) are recommended, however, to be aware of the limits to their knowledge, their assumptions and biases, the strengths and limitations of their assessment methods, and to exercise respect and cultural humility in their service delivery activities (e.g., interviewing, information data gathering, risk item scoring, case formulation, recommendations) (CPA, 2018a; Olver & Stockdale, 2021; Rioja & Rosenfeld, 2018). Relatedly, the inclusion of family members, knowledge keepers, and cultural and community supports in a manner that is consistent with a correctional client's wishes throughout the assessment and case management process is, in our view, critical, and also consistent with the risk salience of the community functioning domain as demonstrated in our review. Service providers should endeavor to practice cultural allyship, that is, "to stand as advocates for those that have been oppressed by colonial systems" (CPA, 2018a, p.

6), and engage in self-reflection/education to better use their power and privilege to change inequitable systems (e.g., anti-racist accomplices and co-conspirators).

**Recent Guidance on Risk Assessment with Indigenous Clients.** Recently, the Canadian-based non-profit organization SAARNA (Society for the Advancement of Actuarial Risk Needs Assessment), which includes the developers of Static-99R and STABLE-2007, has released a set of guidelines for using these scales with Indigenous clients (SAARNA, 2023). These guidelines, however, could be applicable for all types of structured risk assessments with Indigenous populations. Reproduced with permission from SAARNA, these guidelines state:

Evaluators should consider the following recommendations should they aspire to increase their confidence in their assessments of Indigenous men:

1. Conduct the assessment in a setting that is perceived as culturally safe. Culturally safe environments are those that communicate understanding and respect. Consider the attitudes and behaviors (e.g., openness, tolerance, respective language) of all the individuals in the setting (e.g., security staff, management), not just the evaluator. The physical environment also can promote cultural safety (e.g., Indigenous art, inclusion of posters/flags in support of Indigenous communities, absence of symbols of colonialism).<sup>4</sup>
2. Ensure the evaluators have sufficient cultural competence. Evaluators are more likely to obtain useful information if they are familiar with the history and culture of the individual being assessed. This context generally includes the harms of colonization, current and historical racism, and the over-representation of Indigenous peoples in the criminal legal system.

---

<sup>4</sup> We acknowledge that the physical and psychological attributes of a correctional environment such as a prison, parole, or probation office inherently defies a feeling of physical or cultural safety, and as such, we cite these as guidelines to create a physical, psychological, and interpersonal space and setting that can contribute to cultural safety, and help mitigate negative impacts as much as reasonably possible.

Culturally competent evaluators are familiar and comfortable with culture-specific modes of communication, including differences in body language and expressions.

3. Seek input/assistance through consultation with experts in Indigenous culture (e.g., Elders, staff specially trained in working within Indigenous cultures). These consultations can assist in interpreting the cultural context of behavior, which may be important in assessing risk factors and crafting overall recommendations. (SAARNA, 2023).

### **Strengths and Limitations**

Our systematic review and meta-analysis has important study strengths and limitations with implications for future research. A key set of strengths was the size and scope of the review which included 22 risk assessment tools or related forensic assessment measures, across 91 studies, and nearly 300,000 Indigenous and non-Indigenous (majority White) justice involved youth and adults. The review included an examination of key moderators such as gender, developmental group, recidivism outcome, and geographic jurisdiction to draw conclusions about the predictive properties of these tools with Indigenous persons. The authorship team also had the means to conduct focused meta-analyses for 18 of the 22 tools, including measures of protective factors and risk change, which increased the level of depth, detail, and rigor of the analyses, and the trustworthiness of conclusions.

That said, perhaps the most significant limitation is that insufficient data were available to conduct analyses that reflect the diversity of Indigenous peoples in each of the different countries. For instance, and as noted above, Canada alone has hundreds of First Nations communities, in addition to the Métis and Inuit peoples, over its geographical expanse, that examining such a group so diverse in their language, customs, traditions, and belief systems in aggregate invariably has some artificiality. Future research would benefit from more nuanced exploration of Indigenous differences in forensic risk assessment that better reflects this diversity and better examines issues of intersectionality (e.g., justice involved Indigenous women and girls). A second limitation is that for space considerations we did not

examine Indigenous and non-Indigenous differences on all the individual subscales of the various measures through meta-analysis, but rather prioritized the total score given its comprehensiveness and frequency of use, and also the most frequently reported metric of a risk tool. In lieu of this, we reviewed some of the primary risk-need domains identified though and subsequent to, meta-analyses of this literature; we also examined dynamic and protective factors, specifically, given that these drive services to promote release and reintegration. Third, the literature was scant identifying culturally relevant variables; more research simply needs to be done, for instance, examining tools such as the Yókw'tól and the CRFQ, including their predictive properties and convergence with conventional risk tools. Finally, there are limits to the use of meta-analysis to examine measurement invariance, which are not limited to quantity and quality of available literature, but also the ability to systematically impose more specific and nuanced constraints beyond scrutiny of ethnoracial group differences in effect size magnitude and sources of effect size heterogeneity.

### **Conclusion**

Are risk measures “too risky to use” or is it “too risky not to use” them with Indigenous persons in legal and/or correctional contexts? Based on our meta-analytic data from 91 studies and 22 measures, our overall conclusion is that risk assessment guided by the use of structured evidence-based tools is more likely to yield accurate, informed, and useful decisions than unstructured judgment, for both clients and their communities. Although there are disparities in prediction magnitudes between Indigenous and non-Indigenous samples, depending on the tool and outcome, generally speaking, we concur with Gutierrez et al. (2016) that “abandoning their use is not defensible, unless they are replaced with a method empirically demonstrated to have superior accuracy” (p. 103) (see also Harris, 2009; 2011; Vincent & Viljoen, 2020). Risk assessment measures do more than inform the potential for an unwanted outcome; they are intended to drive the intensity, foci, and delivery of services, they can be used to appraise changes in risk, and can guide release planning and community supervision. All of this is

done in a spirit toward preventing recidivism, strengthening community safety, and promoting client wellbeing.

Importantly, however, even if current risk measures have moderate accuracy, inform a set of risk management and prevention functions, and are better than certain alternatives (e.g., unstructured clinical judgment), we do not wish to encourage complacency with the status quo. Given historical and systemic racism, including within the field of psychology which is the predominant background of many risk assessment tool developers, and the unique position of Indigenous peoples in imposed criminal legal systems where they are grossly overrepresented and where these tools are used, it is our responsibility to listen and learn from Indigenous experiences and methodologies and strive for better assessment practices. Ultimately, how a risk measure is used matters as much or more than what a given score conveys (Kroner, 2016; Olver & Stockdale, 2021). As any assessment measure has the potential for misuse, it is the responsibility of service providers, correctional agencies, and professional bodies to ensure the responsible application of forensic risk, and any other assessment measures. Such applications should be conducted in a culturally responsive and anti-racist manner to promote decision making that can maximize benefit and minimize harm for justice-involved persons, enhance community safety, and advocate for human rights and social justice. We hope this work will serve as a call to action to promote greater work in this important area, both in terms of research and clinical/correctional practices, in collaboration with Indigenous peoples and communities.

### References

#### References marked with an asterisk \* included in the systematic review and meta-analyses

- \*Ahmed, S. (2022). *Validating the Spousal Assault Risk Assessment-Version 2 (SARA-V2) guide with Indigenous men supervised by British Columbia Corrections*. Unpublished master's thesis, Simon Fraser University, Burnaby, BC, Canada.
- \*Ahmed, S., Lee, S. C., & Helmus, L. M. (2023). Predictive accuracy of Static-99R across different racial/ethnic groups: A meta-analysis. *Law and Human Behavior, 47*(1), 275-291.  
<https://doi.org/10.1037/lhb0000517>
- \*Allan, A., & Dawson, D. (2002). *Developing a Unique Risk of Violence Tool for Australian Indigenous Offenders*. Canberra, Australia: Criminology Research Council. (CRC 6/00-01)
- Allan, A., & Dawson, D. (2006). Prediction of the risk of male sexual reoffending in Australia. *Australian Psychologist, 41*(1), 60-68. <https://doi.org/10.1080/00050060500391886>
- Allan, A., Parry, C.L., Ferrante, A., Gillies, C., Griffiths, C.S., Morgan, F., Spiranovic, C., Smallbone, S., Tubex, H., & Wong, S.C.P. (2019). Assessing the risk of Australian Indigenous sexual offenders reoffending: A review of the research literature and court decision. *Psychiatry, Psychology, and Law, 26*(2), 274-294. <https://doi.org/10.1080/13218719.2018.1504242>
- American Psychological Association (2017). *Multicultural Guidelines: An Ecological Approach to Context, Identity, and Intersectionality*. Retrieved from: <http://www.apa.org/about/policy/multicultural-guidelines.pdf>
- Anderson, C., & Kukutai, A. (2016). Sources and methods in Indigenous Studies. In C. Andersen & J. M. O'Brien (Eds.) *Sources and methods in Indigenous studies* (pp. 41-48). Routledge.  
<https://doi.org/10.4324/9781315528854>
- Andrews, D. A. (1982). *The Level of Supervision Inventory: The first follow-up*. Toronto, Canada: Ontario Ministry of Correctional Services.

Andrews, D. A., & Bonta, J. (1994-2010). *The psychology of criminal conduct* (1<sup>st</sup> to 5<sup>th</sup> Ed.). Cincinnati, OH: Anderson Publishing.

Andrews, D. A., & Bonta, J. (1995). *Level of Service Inventory-Revised (LSI-R): An offender assessment system. User's guide*. Toronto, ON: Multi-Health Systems.

Andrews, D. A., & Bonta, J. (2010). Rehabilitating criminal justice policy and practice. *Psychology, Public Policy, and Law*, 16(1), 39-55. <https://doi.org/10.1037/a0018362>

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

Andrews, D. A., Bonta, J., & Wormith, J. S. (2004). *Level of Service/Case Management Inventory (LS/CMI): An offender assessment system. User's guide*. Toronto, ON: Multi-Health Systems.

Andrews, D. A., Bonta, J., & Wormith, J. S. (2006). The Recent Past and Near Future of Risk and/or Need Assessment. *Crime & Delinquency*, 52(1), 7-27. <https://doi.org/10.1177/0011128705281756>

Antoine, D. (2017). Pushing the academy: The need for decolonizing research. *Canadian Journal of Communication*, 42(1), 113-120. <https://doi.org/10.22230/cjc.2017v42n1a3091>

Asselin, H., & Basile, S. (2018). Concrete ways to decolonize research. *International Journal for Critical Geographies*, 17(3), 643-650. <https://www.acme-journal.org/index.php/acme/article/view/1707>

Australian Bureau of Statistics (2021). *Profile of Indigenous Australians*.

<https://www.aihw.gov.au/reports/australias-welfare/profile-of-indigenous-australians>

Australian Government (2021). *Prisoners in Australia*. <https://www.abs.gov.au/statistics/people/crime-and-justice/prisoners-australia/latest-release>

\*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), 1-43.

<https://doi.org/10.3138/cjccj.2010.E.40>

Babchishin, K. M., Hanson, R. K., & Blais, J. (2016). Less is more: Using Static-2002R subscales to predict violent and general recidivism among sexual offenders. *Sexual Abuse, 28*(3), 187–217.

<https://doi.org/10.1177/1079063215569544>

Babchishin, K. M., & Helmus, L. M. (2016). The influence of base rates on correlations: An evaluation of proposed alternative effect sizes with real-world data. *Behavior Research Methods, 48*(3), 1021–1031. <http://dx.doi.org/10.3758/s13428-015-0627-7>

\* Barnum, G., & Gobeil, R. (2012a). *Prediction of re-offence using the SIR-R1 and a proxy*. (Corrections Research Report No. R-281). Ottawa: Correctional Service of Canada.

\*Barnum, G., & Gobeil, R. (2012b). Revalidation of the Custody Rating Scale for Aboriginal and non-Aboriginal women offenders. (Corrections Research Report No. R-273). Ottawa: Correctional Service of Canada.

Blais, J., & Forth, A. E. (2014). Prosecution-retained versus court-appointed experts: Comparing and contrasting risk assessment reports in preventative detention hearings. *Law and Human Behavior, 38*(6), 531-543. <https://doi.org/10.1037/lhb0000082>

Boer, D., Couture, J., Geddes, C., & Ritchie, A. (2003). *Yókw'tól Risk Management Guide for Aboriginal Offenders Structured Guidelines for the Assessment of Risk Manageability for Aboriginal Violent Offenders (RESEARCH VERSION)*. Aboriginal Initiatives Branch, Pacific Region, Correctional Service of Canada.

Boer, D. P., Hart, S. D., Kropp, P. R., & Webster, C. D. (1997). *Manual for the Sexual Violence Risk-20: professional guidelines for assessing risk of sexual violence*. Vancouver, BC: Institute against Family Violence and the Mental Health, Law, and Policy Institute, Simon Fraser University.

\*Bonta, J. (1989). Native inmates: Institutional response, risk, and needs. *Canadian Journal of Criminology, 31*(1), 49–62. <https://doi.org/10.3138/cjcrim.31.1.49>

Bonta, J. (1996). Risk needs assessment and treatment. In A.T. Hartland (Ed.), *Choosing correctional*

*options that work: Defining the demand and evaluating the supply* (pp. 18-32). Thousand Oaks, CA: Sage.

Bonta, J., & Andrews, D. A. (2007). *Risk-need-responsivity model for offender assessment and rehabilitation* (User Report 2007-06). Public Safety Canada.

Bonta, J., & Andrews, D. A. (2017). *The psychology of conduct, sixth edition*. New York, NY: Routledge.

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*(3), 278–287. <https://doi.org/10.1016/j.avb.2014.04.014>

Bonta, J., Harman, W. G., Hann, R. G., & Cormier, R. B. (1996). The prediction of recidivism among federally sentenced offenders: A re-validation of the SIR scale. *Canadian Journal of Criminology, 38*(1), 61–79. <https://doi.org/10.3138/cjcrim.38.1.61>

\*Bonta, J., LaPrairie, C., & Wallace-Capretta, S. (1997). Risk prediction and re-offending: Aboriginal and non-aboriginal offenders. *Canadian Journal of Criminology, 39*(2), 127-144.

Bonta, J., Law, M., & Hanson, K. (1998). The predication of criminal and violent recidivism violent recidivism among mentally disordered offenders: a meta-analysis. *Psychological Bulletin, 123*, 123-142. <https://doi.org/10.1037/0033-2909.123.2.123>

Bonta, J., Ruge, T., & Dauvergne, M. (2003). The reconviction rate of federal offenders. (Corrections Research Report No. 2003-02). Ottawa: Department of the Solicitor General of Canada.

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

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (1998-2014). *Comprehensive meta-analysis* (Version 2). Englewood, NJ: Biostat

Borum, R., Bartel, P., & Forth, A. (2006). *Manual for the Structured Assessment of Violence Risk in Youth (SAVRY)*. Odessa FL: Psychological Assessment Resources.

Brankley, A. E., Helmus, L. M., & Hanson, R. K. (2017). *STABLE-2007 Evaluator Workbook – Revised 2017*. Ottawa, ON.

Brankley, A. E., & Lee, S. C. (2016, November). *The utility of Static-99R and STABLE-2007 across different ethnic groups: A prospective field study* [Paper presentation]. Association for the Treatment of Sexual Abusers 35<sup>th</sup> Annual Research and Treatment Conference, Orlando, FL, United States.

\*Brews, A. L. (2009). *The Level of Service Inventory and female offenders: Addressing issues of reliability and predictive validity*. Unpublished master's thesis, University of Saskatchewan, Saskatoon, Saskatchewan, Canada.

\*Brown, S. L., & Motiuk, L. L. (2005). *The Dynamic Factors Identification and Analysis (DFIA) component of the Offender Intake Assessment (OIA) Process: A meta-analytic, psychometric, and consultative review*. (Corrections Research Report No. R-164). Ottawa: Correctional Service of Canada.

Brouillette-Alarie, S., Babchishin, K. M., Hanson, R. K., & Helmus, L. M. (2016). Latent constructs of static risk scales for the prediction of sexual recidivism: A 3-factor solution. *Assessment, 23*(1), 96-111.  
<http://dx.doi.org/10.1177/1073191114568114>

Brouillette-Alarie, S., & Hanson, R. K. (2015). Comparaison de deux mesures d'évaluation du risque de récidive des délinquants sexuels [Comparison of two measures of recidivism risk assessment of sexual offenders]. *Canadian Journal of Behavioural Science, 47*(4), 292-304.  
<http://dx.doi.org/10.1037/cbs0000019>

Brzozowski, J. A., Taylor-Butts, A., & Johnson, S. (2006). Victimization and offending among the Aboriginal population in Canada. *Juristat, 26* (3). Ottawa: Statistics Canada, Canadian Centre for Justice Statistics

\*Buchanan, K. (2009). *Risk assessment and spousal violence: Predictive validity and cultural applicability*. [Unpublished doctoral dissertation]. University of Regina.

Bureau of Indian Affairs (n.d.). Bureau of Indian Affairs (home page). <https://www.bia.gov>

Campbell, M. A., French, S., & Gendreau, P. (2009). The prediction of violence in adult offenders: A meta-analytic comparison of instruments and methods of assessment. *Criminal Justice and Behavior*, 36(6), 567–590. <http://doi.org/10.1177/0093854809333610>

*Canada v. Ewert* (2016). FCA 203

Canadian Psychological Association (2018a). *Psychology's response to the Truth and Reconciliation*

*Commission of Canada's report*. Retrieved from:

[https://cpa.ca/docs/File/Task\\_Forces/TRC%20Task%20Force%20Report\\_FINAL.pdf](https://cpa.ca/docs/File/Task_Forces/TRC%20Task%20Force%20Report_FINAL.pdf)

Canadian Psychological Association (2018b). *Outcomes and progress monitoring in psychotherapy*.

Retrieved from

[https://cpa.ca/docs/File/Task\\_Forces/Treatment%20Progress%20and%20Outcome%20Monitoring%20Task%20Force%20Report\\_Final.pdf](https://cpa.ca/docs/File/Task_Forces/Treatment%20Progress%20and%20Outcome%20Monitoring%20Task%20Force%20Report_Final.pdf)

Clark, S. (2019). *Overrepresentation of Indigenous people in the Canadian criminal justice system: Causes and responses*. Research and Statistics Division, Department of Justice Canada.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159. <https://doi.org/10.1037/0033-2909.112.1.155>

Correctional Service Canada (2019). *Correctional programs for Indigenous offenders*. Retrieved from:

<https://www.csc-scc.gc.ca/002/002-0004-en.shtml#pim>

\*Coupland, R. B. A., & Olver, M. E. (2020). [Violence Risk Scale and recidivism information for 155

treated men from the Aggressive Behaviour Control Program.] Unpublished raw data. From:

Assessing protective factors in treated violent offenders: Associations with recidivism reduction and positive community outcomes. *Psychological Assessment*, 32(5), 493-508.

<https://doi.org/10.1037/pas0000807>

Cram, S., & MacDonald, S. F. (2019). *Substance use patterns of Indigenous and non-Indigenous women offenders*. (Research in Brief No. RIB-19-08) Ottawa: Correctional Service of Canada.

- Cunneen, C. (2009). Indigenous incarceration: The violence of colonial law and justice. In Scraton, P. and McCulloch, J. (Eds) *The Violence of Incarceration* (pp. 209-224). Routledge.
- Datta, R. (2018). Decolonizing both researcher and research and its effectiveness in Indigenous research. *Research Ethics*, 14(2), 1-24. <https://doi.org/10.1177/1747016117733296>
- Dawes, R. M., Faust, D., Meehl, P.E. (1989). Clinical versus actuarial judgment. *Science*, 243, 1668-1674. <https://doi.org/10.1126/science.2648573>
- Dawson, D. (1999). Risk of violence assessment: Aboriginal offenders and the assumption of homogeneity. Conference presentation.
- Day, A., Tamatea, A. J., Casey, S., & Geia, L. (2018). Assessing violence risk with Aboriginal and Torres Strait Islander offenders: considerations for forensic practice. *Psychiatry, Psychology, and Law*, 25(3), 452-464. <https://doi.org/10.1080/13218719.2018.1467804>
- Deangelis, T., & Andod, E. (2022, March). Confronting past wrongs and building an equitable future. *Monitor on Psychology*, 53(2), 20-26.
- Department of Justice Canada (2017). *Spotlight on Gladue: Challenges, experiences, and possibilities in Canada's criminal justice system*. Retrieved from <https://www.justice.gc.ca/eng/rp-pr/jr/gladue/p2.html>
- Deskaheh, C. (1923). *Chief Deskaheh tells why he is over here again*. [https://fisherpub.sjfc.edu/cgi/viewcontent.cgi?article=1008&context=georgepdecker\\_leagueofnations](https://fisherpub.sjfc.edu/cgi/viewcontent.cgi?article=1008&context=georgepdecker_leagueofnations)
- de Vogel, V., de Ruiter, C., Bouman, Y. H. A., & de Vries Robbe', M. (2009). SAPROF; Guidelines for the assessment of protective factors for violence risk. Utrecht, The Netherlands : Forum Educatief.
- De Vries Robbé, M., Geers, M., Stapel, M., Hilterman, E., & de Vogel, V. (2015). SAPROF youth version: Guidelines for the assessment of protective factors for violence risk in juveniles. Retrieved from: <http://www.saprof.com/saprof-youth-version/>

- Douglas, K. S, Hart, S. D., Webster, C.D., & Belfrage, H. (2011). *Historical Clinical Risk Management (Version 3): Professional guidelines for evaluating risk of violence [Draft 2.1]*. Vancouver, BC: Mental Health, Law, and Policy Institute, Simon Fraser University.
- Douglas, K. S. & Kropp, P. R. (2002). A prevention-based paradigm for violence risk assessment: Clinical and research applications. *Criminal Justice and Behavior*, 29, 617-658.
- Ewert v. Canada* (2015). FC 1093
- Ewert v. Canada* (2018). SCC 30
- Fergusson, D. M., Horwood, L. J., & Swain-Campbell, N. (2003). Ethnicity and criminal convictions: Results of a 21-year longitudinal study. *The Australian and New Zealand Journal of Criminology*, 36(3), 354-367. <https://doi.org/10.1375/acri.36.3.354>
- Forth, A. E., Kosson, D. S., & Hare, R. D. (2003). *Psychopathy Checklist: Youth Version (PCL: YV)*. Toronto, Ontario, Canada: Multi-Health Systems.
- Fox, R., Neha, T., & Jose, P. E. (2018). Tū Māori Mai: Māori cultural embeddedness improves adaptive coping and wellbeing for Māori adolescents. *New Zealand Journal of Psychology*, 47(2), 14-24.
- \*Gobeil, R. (2011). Use of the Custody Rating Scale with male offenders. (Corrections Research Report No. R-273). Ottawa: Correctional Service of Canada.
- \*Gossner, D., & Wormith, J. S. (2007). The prediction of recidivism among young offenders in Saskatchewan. *The Canadian Journal of Police & Security Services*, 5, 70-82.
- Government of Canada (1969). *Statement of the Government of Canada on Indian policy*. Department of Indian Affairs and Northern Development. Retrieved from [https://oneca.com/1969\\_White\\_Paper.pdf](https://oneca.com/1969_White_Paper.pdf)
- Government of Canada (2017). Indigenous peoples and communities. Retrieved from: <https://www.rcaanc-cirnac.gc.ca/eng/1100100013785/1529102490303>
- Government of Canada (2021). Indigenous peoples and communities. Retrieved from:

<https://www.rcaanc-cirnac.gc.ca/eng/1100100013785/1529102490303>

Government of New Zealand (2021). *Māori population estimates: At 30 June 2021*.

<https://www.stats.govt.nz/information-releases/maori-population-estimates-at-30-june-2021#:~:text=New%20Zealand%27s%20estimated%20M%C4%81ori%20ethnic,17.1%20percent%20of%20national%20population>

Government of Saskatchewan (1997). *Offender risk assessment management system manual*.

Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical-statistical controversy.

*Psychology, Public Policy, and Law*, 2(2), 293–323. <http://dx.doi.org/10.1037/1076-8971.2.2.293>

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(1), 19-30. <https://doi.org/10.1037/1040-3590.12.1.19>

Gutierrez, L. L. (2018). *Walls of Red Wing: An examination of culturally-informed sentencing, risk/need factors, and treatment for peoples of Indigenous heritage in Canada's criminal justice system*.

Unpublished dissertation, Carleton University, Ottawa, ON.

Gutierrez, L. L., Helmus, 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(2), 97-106. <https://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*, 50, 55-99. <https://doi.org/10.3138/cjccj.2011.E.51>

Guy, L. S., Edens, J. F., Anthony, C., & Douglas, K. S. (2005). Does psychopathy predict institutional misconduct among adults? A meta-analytic investigation. *Journal of Consulting and Clinical Psychology*, 73, 1056–1064. <http://dx.doi.org/10.1037/0022-006X.73.6.1056>

Guzman, G. (2020, December). Household income by race and Hispanic origin: 2005-2009 and 2015-2019. *American Survey Briefs*, ACSBR-007, U.S. Census Bureau, Washington, DC.

Haag, A. M., Boyes, A., Cheng, J., MacNeil, A., & Wirove, R. (2016a). An introduction to the issues of cross-cultural assessment inspired by *Ewert v. Canada*. *Journal of Threat Assessment and Management*, 3(2), 65-75. <https://doi.org/10.1037/tam0000067>

\*Hann, R. G., & Harman, W. G. (1989). *Release risk prediction: Testing the Nuffield scoring system for native and female inmates*. (Report No. 1989-04). Ottawa: Ministry of the Solicitor General of Canada.

\*Hann, R. G., & Harman, W. G. (1993). *Predicting release risk for Aboriginal penitentiary inmates*. (Report No. 1993-12). Ottawa: Ministry of the Solicitor General of Canada.

\*Hansen, C. D. (2018). *Risk and resiliency factors in predicting recidivism among Native Americans on a Montana reservation*. Unpublished master's thesis, University of Montana, Missoula, MT, USA.

Hanson, R. K. (2016, January 15). An affidavit. [http://www.static99.org/pdfdocs/Legal-Thorton-Hanson\\_Affidavits\(2016\).pdf](http://www.static99.org/pdfdocs/Legal-Thorton-Hanson_Affidavits(2016).pdf)

Hanson, R. K. (2017). Assessing the calibration of actuarial risk scales: A primer on the E/O index. *Criminal Justice and Behavior*, 44(1), 26-39. Doi:10.1177/0093854816683956

Hanson, R. K. (2020). *Affidavit: R. v. Monsieur Joe Kritik (Salowatseak)*. Cour du Quebec, Court File No. 635-01-014601-166

Hanson, R. K. (2022). *Prediction statistics for psychological assessment*. American Psychological Association.

Hanson, R. K., Bourgon, G., McGrath, R. J., Kroner, D., D'Amora, D. A., Thomas, S. S., & Tavarez, L. P. (2016). *A five-level risk and needs system: Maximizing assessment results in corrections through the development of a common language*. Justice Center Council of State Governments: Washington, DC.

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.

Hanson, R. K., Harris, A. J. R., Scott, T., & Helmus, L. (2007). *Assessing the risk of sex offenders on community supervision*. (User Report No. 2007-05). Ottawa: Public Safety and Emergency Preparedness Canada.

Hanson, R. K., Helmus, L., & Bourgon, G. (2007). *The validity of risk assessments for intimate partner violence: A meta-analysis* (Corrections Research User Report No. 2007-07). Ottawa, ON: Public Safety Canada.

\*Hanson, R. K., Helmus, L. M., & Harris, A. J. R. (2015). [Unpublished Indigenous-non-Indigenous comparisons on Static-99R scores.] Assessing the risk and needs of supervised sexual offenders: A prospective study using STABLE-2007, Static-99R, and Static-2002R. *Criminal Justice and Behavior*, 42(12), 1205-1224. <https://doi.org/10.1177/0093854815602094>

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), 1-21. <https://doi.org/10.1037/a0014421>

Hanson, R. K., & Thornton, D. (2000). Improving risk assessments for individuals convicted of sexual offences: A comparison of three actuarial scales. *Law and Human Behavior*, 24(1), 119-136. <https://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, Ontario: Department of the Solicitor General of Canada.

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: A Journal of Research and Treatment*, 28(3), 218-252. Doi:10.1177/1079063215574710

Hare, R. D. (1991). *Hare Psychopathy Checklist Revised*. Toronto, ON: Multi-Health Systems.

Hare, R. D. (2003). *Hare Psychopathy Checklist Revised* (2<sup>nd</sup> ed.). Toronto, ON: Multi-Health Systems.

- Harris, A. J. R., Cousineau, S., Pagé, C. A., Sonnichsen, P., & Varrette, S. (2009). *Feasibility of an Inuit specific violence risk assessment instrument*. (User Report R-206). Ottawa: Correctional Service of Canada.
- Harris, A. J. R., Cousineau, S., Pagé, C. A., Sonnichsen, P., & Varrette, S. (2011). Recidivism risk assessment for Aboriginal males: A brief review of scientific literature. *Aboriginal Policy Research Consortium International*
- 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*, 20(4), 315-335.  
<https://doi.org/10.1177/0093854893020004001>
- Hart, S. D. (2016). Culture and violence risk assessment: The case of Ewert v. Canada. *Journal of Threat Assessment and Management*, 3(2), 76-96. <https://doi.org/10.1037/tam0000068>
- Hart, S. D., Cox, D. N., & Hare, R. D. (1995). *The Hare Psychopathy Checklist: Screening Version (PCL: SV)*. Toronto, ON: Multi-Health Systems.
- Haskell L, & Randall M. (2009). Disrupted attachments: A social context complex trauma framework and the lives of Indigenous peoples in Canada. *Journal of Aboriginal Health*, 5(3), 48-99.  
<https://doi.org/10.3138/ijih.v5i3.28988>
- Hays, P. A. (2001). Addressing cultural complexities in practice: A framework for clinicians and counselors. Washington, DC. American Psychological Association.
- \*Hegel, J., Pelletier, K., & Olver, M. E. (2021). Predictive properties of the Ontario Domestic Assault Risk Assessment (ODARA) in a Northern Canadian Prairie sample. *Criminal Justice and Behavior*, 49(3), 411-431. <https://doi.org/10.1177/00938548211033631>
- Helmus, L. M. (2018). Sex offender risk assessment: Where are we and where are we going? *Current Psychiatry Reports*, 20: 46. <https://doi.org/10.1007/s11920-018-0909-8>
- Helmus, L. M., & Babchishin, K. M. (2017). Primer on risk assessment and the statistics used to evaluate

its accuracy. *Criminal Justice and Behavior*, 44(1), 8–25. <https://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(6), 856-876.

<https://doi.org/10.1177/0306624X11414693>

Helmus, L., & Bourgon, G. (2011). Taking stock of 15 years of research on the Spousal Assault Risk

Assessment Guide (SARA): A critical review. *International Journal of Forensic Mental Health*, 10(1),

64-75. <https://doi.org/10.1080/14999013.2010.551709>

\*Helmus, L., & Forrester, T. (2014a). *Construct validity of the Static Factors Assessment in the offender*

*intake assessment process* (Research Report R-309). Ottawa, ON: Correctional Service of Canada.

\*Helmus, L., & Forrester, T. (2014b). *Static Factors Assessment (SFA): Relationship to release and*

*recidivism outcomes* (Research Report No. R-339). Ottawa, ON: Correctional Service of Canada.

Helmus, L. M., Hanson, R. K., Murrie, D. C., & Zaborauckas, C. L. (2021). Field validity of Static-99R and

STABLE-2007 with 4,433 men serving sentences for sexual offences in British Columbia: New findings

and meta-analysis. *Psychological Assessment*, 33(7), 581–595. <https://doi.org/10.1037/pas0001010>

Helmus, L. M., Kelley, S. M., Frazier, A., Fernandez, Y. M., Lee, S. C., Rettenberger, M., & Boccaccini, M. T.

(2022). Static-99R: Strengths, limitations, predictive accuracy meta-analysis, and legal admissibility

review. *Psychology, Public Policy, and Law*, 28(3), 307–331. <https://doi.org/10.1037/law0000351>

\*Helmus, L. M., Lee, S. C., & Zaborauckas, C. L. (2021). *Do Static-99R and STABLE-2007 work with*

*Indigenous people charged or convicted of sexual offences? A prospective field validity study.*

Unpublished manuscript.

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(1), 64-101. <https://doi.org/10.1177/1079063211409951>

- \*Higgs, T., Olver, M. E., Nunes, K., & Cortoni, F. (2020). [Recidivism information for 127 treated men from the Violence Prevention Program.] Unpublished raw data. From: Risk relevance of psychometric assessment and evaluator ratings of dynamic risk factors in high-risk violent offenders. *Legal and Criminological Psychology*, 25(2), 219-236. <https://doi.org/10.1111/lcrp.12173>
- Higgins, J., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, 327, 557-560. <https://doi.org/10.1136/bmj.327.7414.557>
- Hilton, N. Z., Harris, G. T., Rice, M. E., Lang, C., Cormier, C. A., & Lines, K. J. (2004). A brief actuarial assessment for the prediction of wife assault recidivism: The Ontario Domestic Assault Risk Assessment. *Psychological Assessment*, 16, 267–275.
- \*Hogan, N. R., & Corabian, G. (2023). Race and intimate partner violence risk assessment: Structured professional judgement ratings among Canadian Indigenous and non-Indigenous individuals. In N. R. Hogan (Chair), *In pursuit of parity: Investigating violence risk assessment among Indigenous persons*. Symposium presented at the Fifth North American Correctional and Criminal Justice Psychology Conference, Toronto, ON.
- \*Hogg, S. M. (2011). *The Level of Service Inventory (Ontario Revision) scale validation for gender and ethnicity: Addressing reliability and predictive validity*. Unpublished master's thesis, University of Saskatchewan, Saskatoon, Saskatchewan, Canada.
- \*Holsinger, A. M., Lowenkamp, C. T., & Latessa, E. J. (2006). Exploring the validity of the Level of Service Inventory-Revised with Native American offenders. *Journal of Criminal Justice*, 34(3), 331-337. Doi: 10.1016/j.jcrimjus.2006.03.009
- Howard, P. D. (2017). The effect of sample heterogeneity and risk categorization on Area Under the Curve (AUC) predictive validity metrics. *Criminal Justice and Behavior*, 44(1), 103-120. <https://doi.org/10.1177/0093854816678899>
- \*Hsu, C., Caputi, P., & Byrne, M. K. (2010). Level of Service Inventory–Revised: Assessing the risk and

need characteristics of Australian Indigenous offenders. *Psychiatry, Psychology and Law*, 17(3), 355–367. <https://doi.org/10.1080/13218710903089261>

\*Hsu, C., Caputi, P., & Byrne, M., K. (2011). The Level of Service Inventory-Revised and Australian offenders: Factor structure, sensitivity, and specificity. *Criminal Justice and Behavior*, 38(6), 60-618. <https://doi.org/10.1177/0093854811402583>

Hu, C., & Esthappan, S. (2017, May 23). Asian Americans and Pacific Islanders, a missing minority in criminal justice data. *Urban Wire: Crime and Justice*. <https://www.urban.org/urban-wire/130anad-americans-and-pacific-islanders-missing-minority-criminal-justice-data>

\*Huang, S., Peterson-Badali, M., Jang, E., & Skilling, T. A. (2021). IRT-based differential items functioning analysis of the Youth Level of Service/Case Management Inventory across Indigenous and non-Indigenous youth. *Criminal Justice and Behavior*, 48(4), 502-517. <https://doi.org/10.1177/0093854820968877>

Indian Health Service (2019, October). Disparities. Retrieved from: <https://www.ihs.gov/newsroom/factsheets/disparities/>

Jackson, N. (2016). The substantive application of Gladue in Dangerous Offender proceedings: Reassessing risk and rehabilitation for Aboriginal offenders. *Canadian Criminal Law Review*, 20, 77-91.

Joint Committee on Standards for Educational and Psychological Testing of the American Educational Research Association, American Psychological Association and National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.

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

- \*Jones, N. J., Brown, S. L., Robinson, D., & Frey, D. (2016). Validity of the youth assessment and screening instrument: A juvenile justice tool incorporating risks, needs, and strengths. *Law and Human Behavior, 40*(2), 182–194. <https://doi.org/10.1037/lhb0000170>
- \*Jung, S., & Rawana, E. P. (1999). Risk and need assessment of juvenile offenders. *Criminal Justice and Behavior, 26*(1), 69-89.
- Kelley, S. M., Ambroziak, G., Thornton, D., & Barahal, R. M. (2020). How do professionals assess sexual recidivism risk? An updated survey of practices. *Sexual Abuse, 32*(1), 3-29.  
<https://doi.org/10.1177/1079063218800474>
- Kleeven, A. T. H., de Vries Robbé, M., Mulder, E. A., & Popma, A. (2020). Risk assessment in juvenile and young adult offenders: Predictive validity of the SAVRY and SAPROF-YV. *Assessment, 29*(2), 181-197.  
<https://doi.org/10.1177/1073191120959740>
- Koh, L., Day, A., Klettke, B., Daffern, M., & Chu, C. (2022). The predictive validity of three youth violence assessment instruments: The SAVRY, VRS-YV, and SAPROF-YV. *International Journal of Offender Therapy and Comparative Criminology, 63*(2-3), 168-185.  
<https://doi.org/10.1177/0306624X20970887>
- Kois, L., & Chauhan, P. (2016). Forensic evaluators' self-reported engagement in culturally competent practices. *International Journal of Forensic Mental Health, 15*(4), 312-322.  
<https://doi.org/10.1080/14999013.2016.1228089>
- Kraemer, H. C., Kazdin, A. E., Offord, D. R., Kessler, R. C., Jensen, P. S., & Kupfer, D. J. (1997). Coming to terms with the terms of risk. *Archives of General Psychiatry, 54*(4), 337-343.  
<https://doi.org/10.1001/archpsyc.1997.01830160065009>
- Krieg, A. S. (2006). Aboriginal incarceration: Health and social impacts. *The Medical Journal of Australia, 184*, 534-536.
- Kroner, D.G. (2016). The *Ewert v. Canada* judgment: Moving forward. *Journal of Threat Assessment and*

*Management*, 3(2), 122-127. <http://dx.doi.org/10.1037/tam0000069>

Kropp, P. R., & Hart, S. D. (2015). *The Spousal Assault Risk Assessment Guide Version 3 (SARA-V3)*.

Vancouver, Canada: ProActive ReSolutions Inc.

Kropp, P. R., Hart, S. D., Webster, C. W., & Eaves, D. (1995). Manual for the Spousal Assault Risk

Assessment Guide, 2<sup>nd</sup> ed. Vancouver, BC: British Columbia Institute on Family Violence.

Krysiak, J., & LeCroy, C. W. (2002). The empirical validation of an instrument to predict risk of recidivism

among juvenile offenders. *Research on Social Work Practice*, 12(1), 71–

81. <https://doi.org/10.1177/104973150201200106>

Kyne, A. (2022). *A culturally-informed examination of risk factors for crime among Indigenous Peoples*.

Simon Fraser University, Vancouver, Canada. Unpublished honours thesis.

Laub, J. H., & Lauritsen, J. L. (1994). The precursors of criminal offending across the life course. *Federal*

*Probation*, 58, 51–57.

Lederman, J. (1999). Trauma and healing in Aboriginal families and communities. *Native Social Work*

*Journal*, 2(1), 59–90.

\*Lee, S. C., Hanson, R. K., & Blais, J. (2020). 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.

*Canadian Psychology/Psychologie 132canadienne*, 61(1), 42–57. <https://doi.org/10.1037/cap0000182>

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(3), 1019-1024. <https://doi.org/10.1037/a0033098>

Leistico, A. M. R., Salekin, R. T., DeCoster, J., & Rogers, R. (2008). A large-scale meta-analysis relating the hare measures of psychopathy to antisocial conduct. *Law and Human Behavior*, 32(1), 28–45.

<https://doi.org/10.1007/s10979-007-9096-6>

Levenson J. (2014). Incorporating trauma-informed care into evidence-based sex offender treatment.

*Journal of Sexual Aggression*, 20(9)-22. <https://doi.org/10.1080/13552600.2013.861523>

\*Lewis, K., Olver, M., & Wong, S. (2013). [Violence Risk Scale and recidivism information for 152 treated men from the Aggressive Behaviour Control Program.] Unpublished raw data. From: The Violence Risk Scale: Predictive validity and linking treatment changes with recidivism in a sample of high-risk offenders with psychopathic traits. *Assessment*, 20(2), 150-164.

<http://doi.org/0.1177/1073191112441242>

\*Lockwood, I., Peterson-Badali, M., & Schmidt, F. (2018). The relationship between risk, criminogenic need, and recidivism for Indigenous justice-involved youth. *Criminal Justice and Behavior*, 45(11), 1688-1708. <https://doi.org/10.1177/0093854818785404>

\*Lovatt, K. M. (2020). *An examination of risk, need, and protective factors among Saskatchewan young offenders: Implications for risk management and community reintegration*. [Unpublished doctoral dissertation]. University of Saskatchewan.

\*Lovatt, K. M., Stockdale, K. C., & Olver, M. E. (2022). Dynamic violence risk, protective factors, and therapeutic change in a gender and ethnoculturally diverse sample of court-adjudicated youth. *Psychological Assessment*, 34(2), 105–124. <https://doi.org/10.1037/pas0001071>

\*Lowenkamp, C. T., & Latessa, E. J. (2002). *Level of Service Inventory Revised (LSI-R) validation study: North Dakota*. Unpublished technical report. University of Cincinnati, Cincinnati, Ohio.

\*Luong, D., & Wormith, J. S. (2011). Applying risk/need assessment to probation practice and its impact on the recidivism of young offenders. *Criminal Justice and Behavior*, 38(12), 1177-1199. doi: 10.1177/0093854811421596

Markham, S. (2020). Collaborative risk assessment in secure and forensic mental health settings in the UK. *General Psychiatry*, 33(5), 1-6. <https://doi.org/10.1136/gpsych-2020-100291>

Males, M. (2014, August). *Who are police killing?* Center on Juvenile and Criminal Justice. Retrieved from <http://www.cjcj.org/news/8113>

Matz, A. K., Martinez, A. R., & Kujava, E. (2021). Assessing risk in North Dakota Juvenile Probation: A preliminary examination of the predictive validity of the Youth Assessment and Screening Instrument. *Crime & Delinquency*, 67(4), 551-573. <https://doi.org/10.1177/0011128720950023>

Maxim, R., Akee, R., & Sanchez, G. R. (2022, February 9). For the first time, the government published monthly unemployment data on Native Americans, and the picture is stark. *The Avenue*, Retrieved from <https://www.brookings.edu/blog/the-avenue/2022/02/09/despite-an-optimistic-jobs-report-new-data-shows-native-american-unemployment-remains-staggeringly-high/#:~:text=Prior%20to%20the%20pandemic%2C%20Native,unemployment%20during%20the%20Great%20Depression.>

McCuish, E. C., & Corrado, R. R. (2018). Do risk and protective factors for chronic offending vary across Indigenous and White youth followed prospectively through full adulthood? *Crime & Delinquency*, 64(10), 1247-1270. <https://doi.org/10.1177/0011128717741616>

\*McCuish, E. C., Mathesius, J. R., Lussier, P., & Corrado, R. R. (2017). The cross-cultural generalizability of the Psychopathy Checklist: Youth Version for adjudicated Indigenous Youth. *Psychological Assessment*, 30(2), 192-203. <https://dx.doi.org/10.1037/pas0000468>

\*McKinnon, L. J. (2004). *Predicting risk of violence in a young offender population: The predictive validity of the PCL:YV and the YLS/CMI*. Unpublished master's thesis, Lakehead University, Thunder Bay Ontario, Canada.

MedCalc Statistical Software version 19.0.6 (MedCalc Software bvba, Ostend, Belgium; <https://www.medcalc.org>; 2019)

Meehl, P. E. (1954). *Clinical vs. statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis: University of Minnesota Press.

\*Meyers, J. R., & Schmidt, F. (2008). Predictive validity of the structured assessment for violence risk in youth (SAVRY) with juvenile offenders. *Criminal Justice and Behavior*, 35(3), 344-355.

[https://doi.org/ 10.1177/0093854807311972](https://doi.org/10.1177/0093854807311972)

Miller, H. A. (2006). A dynamic assessment of offender risk, needs, and strengths in a sample of general offenders. *Behavioral Sciences and the Law*, 24(6), 767-782. <https://doi.org/10.1002/bsl.728>

Mills, J. F. (2017). Violence risk assessment: A brief review, current issues, and future directions.

*Canadian Psychology/Psychologie canadienne*, 58(1), 40-49. [https://doi.org/ 10.1037/cap0000100](https://doi.org/10.1037/cap0000100)

Ministry of Justice (2013). *Statistics on race and the criminal justice system 2012*. A Ministry of Justice publication of the Criminal Justice Act 1991. Retrieved from: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/269399/Race-and-cjs-2012.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/269399/Race-and-cjs-2012.pdf)

government/uploads/system/uploads/attachment\_data/file/269399/Race-and-cjs-2012.pdf

Minton, T. (2021, October). *Jails in Indian Country, 2019-2020 and the Impact of COVID-19 on the Tribal jail population*. U.S. Department of Justice, Bureau of Justice Statistics, NCJ 300801.

Monchalin, L. (2016). *The colonial problem: An Indigenous perspective on crime and injustice in Canada*. University of Toronto Press.

Mossman, D. (1994). Assessing predictions of violence: Being accurate about accuracy. *Journal of Consulting and Clinical Psychology*, 62(4), 783-792. <https://doi.org/10.1037/0022-006X.62.4.783>

Motiuk, M. S., Motiuk, L. L., & Bonta, J. (1992). A comparison between self-report and interview-based inventories in offender classification. *Criminal Justice and Behavior*, 19(2), 143-159.

[doi:10.1177/0093854892019002003](https://doi.org/10.1177/0093854892019002003)

\*Muir, N.M., Viljoen, J.L., Jonnson, M.R., Cochrane, D.M., Roger, B.J. (2020). Predictive validity of the Structural Assessment of Violence risk in Youth (SAVRY) with Indigenous and Caucasian female and male adolescents on probation. *Psychological Assessment*, 32(6), 594-607.

<https://dx.doi.org/10.1037/pas000816>

Muir, N. M., Viljoen, J. L., & Shepherd, S. M. (2023). Violence risk assessment tools and Indigenous Peoples: Colonialism as an underlying cause of risk ratings on the SAVRY. *International Journal of Forensic Mental Health*. Advance online publication.

<https://doi.org/10.1080/14999013.2023.2178554>

\*Myer, A. J. (2019). Examining the predictive validity of the Static-99R on Native American sex offenders.

*Justice Evaluation Journal*. <https://doi.org/10.1080/24751979.2019.1636614>

\*Nafekh, M., & Motiuk, L.L. (2002). *The Statistical Information on Recidivism-Revised (SIR-R1) Scale: A*

*psychometric examination*. (Corrections Research Report No. R-126). Ottawa: Correctional Service of Canada.

National Congress of American Indians (2020). *Demographics*. Retrieved from

<https://www.ncai.org/about-tribes/demographics>

National Institute of Justice (2016). *Five things about violence against American Indian and Alaskan*

*Native women and men*. U.S. Department of Justice. Retrieved from

<https://www.ojp.gov/pdffiles1/nij/249815.pdf>

Neal, T. M. S., & Grisso, T. (2014). Assessment practices and expert judgment methods in forensic psychology and psychiatry. *Criminal Justice and Behavior*, 41(12), 1406–1421.

<https://doi.org/10.1177/0093854814548449>

Nellis, A. (2016). *The color of justice: Racial and ethnic disparity in state prisons*. Washington: The Sentencing Project.

Nettlebeck, A., & Smandych, R. (2010). Policing Indigenous peoples on two colonial frontiers: Australia's

mounted police and Canada's North-West mounted police. *The Australian and New Zealand Journal of Criminology*, 43(2), 356-375. <https://journals.sagepub.com/doi/pdf/10.1375/acri.43.2.356>

New Zealand Department of Corrections (2021). *Prison facts and statistics - March 2021*.

[https://www.corrections.govt.nz/resources/statistics/quarterly\\_prison\\_statistics/prison\\_stats\\_march\\_2021](https://www.corrections.govt.nz/resources/statistics/quarterly_prison_statistics/prison_stats_march_2021)

Nuffield, J. (1982). *Parole decision-making in Canada*. Ottawa: Solicitor General Canada.

\*O'Byrne, P. L. (2004). *Exploring the predictive ability of risk needs assessment in Saskatchewan*

*corrections*. [Unpublished master's thesis]. University of Regina.

Olver, M. E. (2016). Some considerations on the use of actuarial and related forensic measures with diverse correctional populations. *Journal of Threat Assessment and Management, 3*(2), 107–121. <https://doi.org/10.1037/tam0000065>

\*Olver, M. E., Cortoni, F., Coupland, R.B.A., Higgs, T., Hogan, N. R., Lewis, K., Gordon, A., & Wong, S.C.P. (2021). [Predictive validity of Violence Risk Scale risk and change scores in a multi-site treated sample of Indigenous and non-Indigenous men.] Unpublished raw data.

\*Olver, M. E., Kingston, D. A., & Sowden, J. N. (2021). An examination of latent constructs of dynamic sexual violence risk and need as a function of indigenous and nonindigenous ancestry. *Psychological Services, 18*(4), 484–496. <https://doi.org/10.1037/ser0000414>

\*Olver, M. E., Lewis, K., & Wong, S. C. P. (2013). [Psychopathy Checklist-Revised and recidivism information for 155 treated men from the Aggressive Behaviour Control Program.] Unpublished raw data. From: Risk reduction treatment of high-risk psychopathic offenders: The relationship of psychopathy and treatment change to violent recidivism. *Personality Disorders, Theory, Research, and Treatment, 4*(2), 160-67. <https://doi.org/10.1037/a0029769>

Olver, M. E., Mundt, J. C., Thornton, D., Beggs Christofferson, S. M., Kingston, D. A., Sowden, J. N., Nicholaichuk, T. P., Gordon, A., & Wong, S. C. P. (2018). Using the Violence Risk Scale-Sexual Offense version in sexual violence risk assessments: Updated risk categories and recidivism estimates from a multisite sample of treated sexual offenders. *Psychological Assessment, 30*(7), 941-955. <https://doi.org/10.1037/pas0000538>

\*Olver, M. E., Neumann, C. S., Sewall, L. A., Lewis, K., Hare, R. D., & Wong, S. C. P. (2018). A comprehensive examination of the psychometric properties of the Hare Psychopathy Checklist-Revised in a Canadian multisite sample of Indigenous and non-Indigenous offenders. *Psychological Assessment, 30*(6), 779-792. <https://doi.org/10.1037/pas0000533>

- \*Olver, M. E., Neumann, C. S., Wong, S. C. P., & Hare, R. D. (2013). The structural and predictive properties of the PCL-R in Canadian Aboriginal and Non-Aboriginal offenders. *Psychological Assessment, 25*(1), 167-179. <https://doi.org/10.1037/a0029840>
- \*Olver, M. E., Nicholaichuk, T. P., Kingston, D. K., & Wong, S. C. P. (2020). [Violence Risk Scale-Sexual Offense version, Static-99R, and recidivism information for 562 treated men from the National Sex Offender Program.] Unpublished raw data. From: A prospective multisite examination of dynamic sexual violence risk: An extension and update to Olver, Nicholaichuk, Kingston, and Wong (2014). *Journal of Consulting and Clinical Psychology, 88*(4), 362-371. <http://dx.doi.org/10.1037/ccp0000478>
- \*Olver, M. E., & Riemer, E. K. (2021). [Psychopathy Checklist-Revised, Structured Assessment of Protective Factors, and recidivism information for 421 treated men from Correctional Service Canada's National Sex Offender Program.] Unpublished raw data. From: High psychopathy men with a history of sexual offending have protective factors too: But are these risk relevant and can they change in treatment? *Journal of Consulting and Clinical Psychology, 89*(5), 406-420. <https://doi.org/10.1037/ccp0000638>
- Olver, M. E., Sewall, L. A., Sarty, G. E., Lewis, K., & Wong, S. C. P. (2015). A cluster analytic examination and external validation of psychopathic offender subtypes in a multisite sample of Canadian federal offenders. *Journal of Abnormal Psychology, 124*(2), 355–371. <https://doi.org/10.1037/abn0000038>
- \*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 VRS-SO risk and change scores in treated Canadian Aboriginal and non-Aboriginal sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 30*(3), 254-275. <https://doi.org/10.1177/1079063216649594>
- Olver, M. E., & Stockdale, K. C. (2021). Sexual violence risk assessment with Indigenous men: Context, controversies, and current findings. *Current Psychiatry Reports, 48*. <https://doi.org/10.1007/s11920-021-01261-9>

Olver, M. E., Stockdale, K. C., Neumann, C. S., Hare, R. D., Mokros, A., Baskin-Sommers, A., Brand, E., Folino, J., Gacono, C., Gray, N. S., Kiehl, K., Knight, R., Leon-Mayer, E., Logan, M., Meloy, J. R., Roy, S., Salekin, R. T., Snowden, R., Thomson, N., Tillem, S., Vitacco, M., & Yoon, D. (2020). Reliability and validity of the Psychopathy Checklist-Revised in the assessment of risk for institutional violence: A cautionary note on DeMatteo et al. *Psychology, Public Policy, and Law*, *26*(4), 490-510.

<https://doi.org/10.1037/law0000256>

Olver, M. E., Stockdale, K. C., Helmus, L. M., Woods, P., Termeer, J., & Prince, J. (2023, September 29).

Forensic Risk Assessment with Indigenous Persons: Meta-Analysis and Systematic Review. Retrieved from [osf.io/48v9d](https://osf.io/48v9d)

\*Olver, M. E., Stockdale, K. C., & Wong, S. C. P. (2012). Short and long-term prediction of recidivism using the youth level of service/case management inventory in a sample of serious young offenders. *Law and Human Behavior*, *36*(4), 331-344. <https://doi.org/10.1037/h0093927>

\*Olver, M. E., Stockdale, K. C., & Wormith, S. J. (2009). Risk assessment with young offenders: A meta-analysis of three assessment measures. *Criminal Justice and Behavior*, *36*(4), 329-353. <https://doi.org/10.1177/0093854809331457>

\*Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2014). Thirty years of research on the Level of Service scales: A meta-analytic examination of predictive accuracy and sources of variability. *Psychological Assessment*, *26*(1), 156-176. <https://doi.org/10.1037/a0035080>

\*Olver, M. E., & Wong, S. C. P. (2015). [Psychopathy Checklist-Revised and recidivism information for 273 men from CSC Prairie Region.] Unpublished raw data. From: Short and long-term recidivism prediction of the PCL-R and the effects of age: A 24-year follow-up. *Personality Disorders: Theory, Research, and Treatment*, *6*(1), 97-105. doi: 10.1037/per0000095

Olver, M. E., & Wong, S. C. P. (2019). Offender risk and need assessment: Theory, research, and applications. In D. Polaschek, A. Day, and C. Hollin (Eds.) *The handbook of correctional psychology*

(pp. 461-475). Wiley.

\*Olver, M. E., Wong, S. C. P., Nicholaichuk, T., & Gordon, A. (2007). [Violence Risk Scale-Sexual Offense version, Static-99R, and recidivism information for 317 treated men from the Clearwater Sex Offender Program.] Unpublished raw data. From: The validity and reliability of the Violence Risk Scale-Sexual Offender version: Assessing sex offender risk and evaluating therapeutic change. *Psychological Assessment, 19*(3), 318-329. <https://doi.org/10.1037/1040-3590.19.3.318>

Partnership with Native Americans (2017, April). Combating food insecurity on Native American reservations. Retrieved from <http://www.nativepartnership.org/site/DocServer/2017-PWNA-NPRA-Food-Insecurity-Project-Grow.pdf?docID=7106>

Peltier, C. (2018) An application of Two-Eye Seeing: Indigenous research methods with participatory action research. *International Journal of Qualitative Methods, 17*(1), 1-12. <https://doi.org/10.1177/1609406918812346>

\*Perley-Robertson, B., Helmus, L. M., & Forth, A. (2019). Predictive accuracy of static risk factors for Canadian Indigenous offenders compared to non-Indigenous offenders: implications for risk assessment scales. *Psychology, Crime, & Law, 25*(3), 248-278. <https://doi.org/10.1080/1068316X.2018.1519827>

Perlin, M. L., & McClain, V. (2009). "Where souls are forgotten": Cultural competencies, forensic evaluations, and international human rights. *Psychology, Public Policy, and Law, 15*(4), 257-277. <https://doi.org/10.1037/a0017233>

Perreault, S. (2011). *Violent victimization of Aboriginal people in the Canadian provinces, 2009* (Statistics Canada Catalogue No. 85-002-X). Retrieved from <http://www.statcan.gc.ca/pub/85-002-x/2011001/article/11415-eng.pdf>

Perrault, R.T., Vincent, G.M., & Guy, L.S. Are risk assessments racially biased? Field study of the SAVRY and YLS/CMI in probation. *Psychological Assessment, 29*(6), 664-678.

<https://dx.doi.org/10.1037/pas0000445>

Perry, B. (2006). Nobody trusts them! Under- and over-policing Native American communities. *Critical Criminology*, 14, 411–444. <https://doi.org/10.1007/s10612-006-9007-z>

Postsecondary National Policy Institute (2021, November). *Factsheets Native American students*.

Retrieved from <https://pnpi.org/native-american-students/>

Polaschek, D. L. L. (2012). An appraisal of the risk-need-responsivity (RNR) model of offender rehabilitation and its application in correctional treatment. *Legal and Criminological Psychology*, 17(1), 1-17. <https://dx.doi.org/10.1111/j.2044-8333.2011.02038.x>

Poverty USA (n.d.). *Poverty facts. The population of poverty USA*. Retrieved from

<https://www.povertyusa.org/facts>

\*Prince, J., Lovatt, K., Stockdale, K. C., & Olver, M. E. (2021, in press). Predictive properties of a general risk-need measure in diverse justice involved youth: A prospective field validity study. *Criminal Justice and Behavior*, 48(11), 1511-1535. <https://doi.org/10.1177/00938548211004669>

Prochaska, J. O., DiClemente, C. C., & Norcross, J. C. (1992). In search of how people change:

Applications to the addictive behaviors. *American Psychologist*, 47(9), 1102-1114.

<https://doi.org/10.1037/0003-066X.47.9.1102>

Public Safety Canada Portfolio Corrections Statistics Committee (2019). *Corrections and conditional release statistical overview: 2018 annual report*. Public Safety Canada. Ottawa.

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting; The state of the art and future directions for psychological research. *Developmental Review*, 41, 71–90.

<https://doi.org/10.1016/j.dr.2016.06.004>

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (1998). *Violent offenders: Appraising and managing risk (first edition)*. Washington, DC: American Psychological Association.

*R. v. Gladue* (1999). 1 SCR 688

*R. v. Ipeelee (2012)*. 1 SCR 433.

\*Rector, B., Wormith, J. S., & Banka, D. (2007). Predictive validity of the LSI-SK Youth Edition. In D. A.

Andrews (Chair), *Level of Service Inventory: Risk/need assessment of female, young and Aboriginal offenders*. Symposium presented at the First North American Correctional and Criminal Justice Psychology Conference, Ottawa, ON.

\*Rettinger, L. J., & Andrews, D. A. (2010). General risk and need, gender specificity, and the recidivism of female offenders. *Criminal Justice and Behavior*, 37(1), 29-46. doi: 10.1177/0093854809349438

Reyhner, J., & Singh, N. K. (2010). Cultural genocide in Australia, Canada, New Zealand, and the United States: The destruction and transformation of Indigenous Cultures. *Indigenous Policy Journal*, XXI(4), 1-26.

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(5), 615-620. <https://doi.org/10.1007/s10979-005-6832-7>

Rice, M. E., Harris, G. T., & Lang, C. (2013). Validation of and revision of the VRAG and SORAG: The Violence Risk Appraisal Guide – Revised (VRAG-R). *Psychological Assessment*, 25(3), 951-965.

Rioja V. B., & Rosenfeld B. (2018). Addressing Linguistic and Cultural Differences in the Forensic Interview. *International Journal of Forensic Mental Health*, 17(4), 377-386.

<https://doi.org/10.1080/14999013.2018.1495280>

Rogers, R. (2000). The uncritical acceptance of risk assessment in forensic practice. *Law and Human Behaviour*, 24, 595-605. doi: 10.1023/A:1005575113507

Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Newbury Park, CA: Sage

\*Rugge, T. (2006). Risk assessment of male aboriginal offenders (Corrections Research Report No. 2006-01). Ottawa: Public Safety Canada.

Santiago Schmitt, S., Mendez-Timmons, M., & Castillo, Y. (2021). *A.D.D.R.E.S.S.I.N.G cultural competence in forensic assessments*. American Psychology-Law Society CE Webinar, March 26, 2021.

- \*Schmidt, F., McKinnon, L., Chattha, H.K., & Brownlee, K. (2006). Concurrent and predictive validity of the Psychopathy Checklist: Youth Version across gender and ethnicity. *Psychological Assessment, 18*(4), 393-401. <https://doi.org/10.1037/1040-3590.18.4.393>
- Schulze, R. (2007). Current methods for meta-analysis: Approaches, issues, and developments. *Zeitschrift für Psychologie/Journal of Psychology, 215*(2), 90-103. <https://doi.org/10.1027/0044-3409.215.2.90>
- \*Schwalbe, C. S. (2009). Risk assessment stability: A revalidation study of the Arizona Risk/Needs Assessment instrument. *Research on Social Work Practice, 19*(2), <https://doi.org/10.1177/1049731508317297>
- Scrim, K. (2010). Aboriginal victimization in Canada: A summary of the literature. *Victims of Crime Research Digest, No. 3*, 15–20.
- \*Sewall, L. A. (2015). [Psychopathy Checklist-Revised and recidivism information for 305 treated men from the Clearwater Sex Offender Program.] Unpublished raw data. From: *The treatment of psychopathic sexual offenders: Exploring the influence of risk, change, subtype, and adaptation on recidivism*. Unpublished dissertation, University of Saskatchewan, Saskatoon, Saskatchewan, Canada.
- Shepherd, S. M. (2015). Finding color in conformity: A commentary on culturally specific risk factors for violence in Australia. *International Journal of Offender Therapy and Comparative Criminology, 59*(12), 1297-1307. <https://dx.doi.org/10.1177/0306624X14540492>
- Shepherd, S.M. (2016). Violence risk instruments may be culturally unsafe for use with Indigenous patients. *Australasian Psychiatry, 24*(6), 565-567. <https://dx.doi.org/10.1177/1039856216665287>
- Shepherd, S. M. (2019). Cultural awareness workshops: Limitations and practical consequences. *BMC Medical Education, 19*(14), Article 14. <https://doi.org/10.1186/s12909-018-1450-5>
- 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(3), 281–293. <https://doi.org/10.1037/law0000017>

Shepherd, S. M., & Anthony, A. (2018). Popping the cultural bubble of violence risk assessment tools.

*The Journal of Forensic Psychiatry & Psychology*, 29(2), 211-220.

<https://doi.org/10.1080/14789949.2017.1354055>

\*Shepherd, S.M., Delgado, R.H., Sherwood, J., & Paradies, Y. (2018). The impact of Indigenous cultural identity and cultural engagement on violent offending. *BMC Public Health*, 18(50), 1-7.

<https://dx.doi.org/10.1186/s12889-017-4603-2>

Shepherd, S.M., Luebbers, S., & Dolan, M. (2012). Gender and ethnicity in juvenile risk assessment.

*Criminal Justice and Behavior*, 40(4), 388-408. <https://doi.org/10.1177/0093854812456776>

Shepherd, S. M., & Lewis-Fernandez, R. (2016). Forensic risk assessment and cultural diversity:

Contemporary challenges and future directions. *Psychology, Public Policy, and Law*, 22(4), 427–438.

<https://doi.org/10.1037/law0000102>

\*Shepherd, S.M., Luebbers, S., Ferguson, M., Ogloff, J.R.P., & Dolan, M. (2014). The utility of the SAVRY across ethnicity in Australian young offenders. *Psychology, Public Policy, and Law*, 20(1), 31-45.

<https://dx.doi.org/10.1037/a0033972>

\*Shepherd, S.M., Singh, J.P., & Fullam, R. (2015). Does the Youth Level of Service/Case Management inventory generalize across ethnicity? *International Journal of Forensic Mental Health*, 14(3), 193-

204. <https://doi.org/10.1080/14999013.2015.1086450>

Shepherd, S. M., & Spivak, B. L. (2020). Finding colour in conformity part II—Reflections on structured

professional judgement and cross-cultural risk assessment. *International Journal of Offender*

*Therapy and Comparative Criminology*, 65(1), 92-99. <https://doi.org/10.1177/0306624X20928025>

\*Shepherd, S. M., & Strand, S. (2016). The PCL: YV and reoffending across ethnic groups. *Journal of*

*Criminal Psychology*, 6 (2), 51-62. <https://doi.org/10.1108/JCP-02-2016-0006>

- Shepherd, S. M., & Willis-Esqueda, C. (2017). Indigenous perspectives on violence risk assessment: a thematic analysis. *Punishment & Society, 0*(0), 1-29. [https://doi.org/ 10.1177/1462474517721485](https://doi.org/10.1177/1462474517721485)
- Shields, I. W., & Simourd, D. J. (1991). Predicting predatory behavior in a population of young offenders. *Criminal Justice and Behavior, 18*(2), 180-194. <https://doi.org/10.1177/0093854891018002006>
- Simmons, J.P., Nelson, L. D., & Simonsohn, U. (October 14, 2012). *A 21 word solution*. Retrieved from: SSRN: <http://dx.doi.org/10.2139/ssrn.2160588>
- Singh, J.P., Grann, M., & Fazel, S. (2011). A comparative study of violence risk assessment tools: A systemic review and metaregression analysis of 68 studies involving 25, 980 participants. *Clinical Psychology Review, 31*(3), 499-513. <https://doi.org/10.1016/j.cpr.2010.11.009>
- Sittner, K. J., & Hautala, D. (2016). Aggressive delinquency among Native American Indigenous adolescents: Trajectories and predictors. *Aggressive Behavior, 42*(3), 274-286. <https://doi.org/1.1002/ab.21622>
- Skeem, J. L., Polaschek, D. L. L., Patrick, C. J., & Lilienfeld, S. O. (2011). Psychopathic personality: Bridging the gap between scientific evidence and public policy. *Psychological Science in the Public Interest, 12*(3), 95-162. <https://doi.org/10.1177/1529100611426706>
- Skewes, M. C., & Bloom, A. W. (2019). Understanding the link between racial trauma and substance use among American Indians. *American Psychologist, 74*(1), 88-100. <https://doi.org/10.1037/amp0000331>
- \*Smallbone, S.W., & Rallings, M. (2013). Short-term predictive validity of the Static-99 and Static-99 R for Indigenous and Nonindigenous Australian sexual offenders. *Sexual Abuse, 25*(3), 302-316. <https://doi.org/10.1177/1079063212472937>
- Smith, A. (n. d.). *Indigenous peoples and boarding schools: A comparative study*. Prepared for the Secretarian of the United Nations Permanent Forum on Indigenous Issues. Retrieved from [https://www.un.org/esa/socdev/unpfii/documents/IPS\\_Boarding\\_Schools.pdf](https://www.un.org/esa/socdev/unpfii/documents/IPS_Boarding_Schools.pdf)

Smith, L. T. (2012). *Decolonizing methodologies: Research and Indigenous peoples* (2nd Edition). Zed Books

Society for the Advancement of Actuarial Risk Needs Assessment (SAARNA) (2023). *Using Static-99R and STABLE-2007 with Indigenous men in Canada, the United States, Australia, and New Zealand*.

Author. Available from [www.saarna.org](http://www.saarna.org)

\*Sowden, J. N., & Olver, M. E. (2017). [Violence Risk Scale-Sexual Offense version, Static-99R, STABLE 2007, and recidivism information for 180 treated men from the Clearwater Sex Offender Program.] Unpublished raw data. From: Use of the Violence Risk Scale—Sexual Offender Version and the Stable 2007 to assess dynamic sexual violence risk in a sample of treated sexual offenders. *Psychological Assessment*, 29(3), 293–303. <https://doi.org/10.1037/pas0000345>

\*Spiranovic, C. (2012). *The Static-99 and Static-99-R Norms Project: Developing norms based on Western Australian sex offenders*. Crawley, Western Australia University of Western Australia. Retrieved from [https://www.law.uwa.edu.au/\\_\\_data/assets/pdf\\_file/0007/2253328/FinalReportForRelease141212\\_Static99Norms\\_141212.pdf](https://www.law.uwa.edu.au/__data/assets/pdf_file/0007/2253328/FinalReportForRelease141212_Static99Norms_141212.pdf)

Spring, J. (1994). *Deculturalization and the struggle for equality: A brief history of the education of dominated cultures in the United States*. McGraw-Hill.

Sterritt, A., & Dickson, C. (2021, June 15). “This is heavy truth”: Tk'emlúps te Secwépemc chief says more to be done to identify unmarked graves. Ground-penetrating radar specialist identifies 200 potential burial sites, says more land still to be surveyed. *CBC News*. Retrieved from <https://www.cbc.ca/news/canada/british-columbia/kamloops-residential-school-findings-1.6084185>

\*Stewart, L. A., Wardrop, K., Wilton, G., Thompson, J., Derkzen, D., & Motiuk, L. (2017). *Reliability and validity of the Dynamic Factors Identification and Analysis – Revised* (Research Report R-395). Ottawa, Ontario: Correctional Service of Canada.

Stewart L., & Wilton, G. (2019). *A comprehensive study of recidivism rates among Canadian federal*

*offenders*. Emerging Research Results Report No. ERR-19-02. Ottawa: Correctional Service Canada.

\*Stockdale, K. C., Olver, M. E., & Wong, S. C. P. (2010). The Psychopathy Checklist: Youth Version and adolescent and adult recidivism: Considerations with respect to gender, ethnicity, and age. *Psychological Assessment*, 22(4), 768–781. <https://doi.org/10.1037/a0020044>

\*Stockdale, K. C., Olver, M. E., & Wong, S. C. P. (2014). The validity and reliability of the Violence Risk Scale – Youth Version in a diverse sample of young violent offenders. *Criminal Justice and Behavior*, 14(1), 114-138. <https://doi.org/10.1177/0093854813496999>

Tamatea, A. J. (2017). Culture is our business: Issues and challenges for forensic and correctional psychologists. *Australian Journal of Forensic Sciences*, 49, 564-478. <https://doi.org/10.1080/00450618.2016.1237549>

\*Thompson, A.P., & McGrath, A. (2012). Subgroup differences and implications for contemporary risk-need assessment with juvenile offenders. *Law and Human Behavior*, 36(4), 345-355. <https://doi.org/10.1037/h0093930>

\*Towers, C. M. (2020). *A revalidation of the Level of Service Inventory-Revised (LSI-R)*. Unpublished master's thesis, University of North Dakota, Fargo, ND, USA.

Travis, J., Western, B., & Redburn, F. S. (2014). *The growth of incarceration in the United States: Exploring causes and consequences*. Committee on Law and Justice, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academies Press.

Truth and Reconciliation Commission of Canada. (2015). *Honouring the truth, reconciling for the future*. Available at: <http://nctr.ca/reports.php>

United Nations (2009). State of the world's Indigenous peoples (ST/ESA/328). Retrieved from [https://www.un.org/esa/socdev/unpfii/documents/SOWIP/en/SOWIP\\_web.pdf](https://www.un.org/esa/socdev/unpfii/documents/SOWIP/en/SOWIP_web.pdf)

United States Sentencing Commission (2017, November). *Demographic differences in sentencing: An update to the 2012 Booker Report*. Washington, DC.

United States Sentencing Commission (2021, July). Quick facts: Native American offenders. Fiscal year 2020. Washington, DC. <https://www.ussc.gov/research/quick-facts>.

van der Put (2019). Predicting domestic violence: A meta-analysis on the predictive validity of risk assessment tools. *Aggression and Violent Behavior, 47*, 100-116.  
<https://doi.org/10.1016/j.avb.2019.03.008>

Venner, S., Sivasubramaniam, D., Leubbers, S., & Shepherd, S. M. (2020). Cross-cultural reliability and rater bias in forensic risk assessment: a review of the literature. *Psychology, Crime, & Law, 27*(2), 105-121. <https://doi.org/10.1080/1068316X.2020.1775829>

Victor, W. (2007). *Indigenous Justice: Clearing Space and Place for Indigenous Epistemologies*. Research paper for the National Centre for First Nations Governance. Retrieved from [http://fngovernance.org/ncfng\\_research/wenona\\_victor.pdf](http://fngovernance.org/ncfng_research/wenona_victor.pdf)

Viljoen, J. L., Vargen, L. M., Cochrane, D. M., Jonnson, M. R., Goossens, I., & Monjazebe, S. (2021). Do structured risk assessments predict violent, any, and sexual offending better than unstructured judgment? An umbrella review. *Psychology, Public Policy, and Law, 27*(1), 79-97. <https://doi.org/10.1037/law0000299>

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

Vose, B., Lowenkamp, C. T., Smith, P., & Cullen, F. T. (2009). Gender and the predictive validity of the LSI-R: A study of parolees and probationers. *Journal of Contemporary Criminal Justice, 25*(4), 459-471.  
<https://doi.org/10.1177/1043986209344797>

\*Watkins, I. (2011). *The utility of Level of Service Inventory-Revised (LSI-R) assessments within NSW correctional environments*. Research Bulletin No. 29, Corrective Services NSW Corporate Research, Evaluation and Statistics.

- \*Wilson, H. A., & Gutierrez, L. (2014). Does one size fit all? A meta-analysis examining the predictive ability of the Level of Service Inventory (LSI) with Aboriginal offenders. *Criminal Justice and Behavior*, 41(2), 196-219. <https://doi.org/10.1177/0093854813500958>
- \*Wilton, G., Stewart, L. A., & Motiuk, L. L. (R-400). *Can the predictive validity of the Dynamic Factors Identification and Analysis – Revised be improved by calculated ratings?* Ottawa, Ontario: Correctional Service of Canada.
- Wong, S. C. P., & Gordon, A. (1999-2003). *Violence Risk Scale*. Regional Psychiatric Centre and the University of Saskatchewan Saskatoon, Saskatchewan, Canada: Author.
- \*Wong, S. C. P., & Gordon, A. (2006). [Violence Risk Scale, Statistical Information on Recidivism scale, and recidivism information for 455 men from the Violent Offender Survey.] Unpublished raw data. From: The validity and reliability of the Violence Risk Scale: A treatment-friendly violence risk assessment tool. *Psychology, Public Policy, and Law*, 12(3), 279-309. <https://doi.org/10.1037/1076-8971.12.3.279>
- Wong, S., Lewis, K., Stockdale, K., & Gordon, A. (2004-2011). *The Violence Risk Scale-Youth Version*. Unpublished Manuscript, Saskatoon, Saskatchewan, Canada
- Wong, S., Olver, M., Nicholaichuk, T., & Gordon, A. (2003, 2017). *Violence Risk Scale: Sexual Offense version*. University of Saskatchewan and Regional Psychiatric Centre, Saskatoon, Saskatchewan, Canada.
- Wormith, J. S. (2011). The legacy of D. A. Andrews in the field of criminal justice: How theory and research can change policy and practice. *International Journal of Forensic Mental Health*, 10, 78–82. <https://doi.org/10.1080/14999013.2011.577138>
- \*Wormith, J. S., Hogg, S. M., & Guzzo, L. (2015). The predictive validity of the LS/CMI with Aboriginal offenders in Canada. *Criminal Justice and Behavior*, 42(5), 481-508. <https://doi.org/10.1177/0093854814552843>

\*Wormith, J. S., & Olver, M. E. (2002). Offender treatment attrition and its relationship with risk, responsivity, and recidivism. *Criminal Justice and Behavior, 29*(4), 447–471.

<https://doi.org/10.1177/0093854802029004006>

Wormith, J. S., & Hogg, S. M. (2012). *The predictive validity of Aboriginal offender recidivism with a general risk/needs assessment inventory*. Saskatoon, SK: Program Effectiveness, Statistics, and Applied Research Unit, Ministry of Community Safety and Correctional.

\*Wormith, J. S., Hogg, S. M., & Guzzo, L. (2015). The predictive validity of the LS/CMI with Aboriginal offenders in Canada. *Criminal Justice and Behavior, 42*(5), 481-508.

<https://doi.org/10.1177/0093854814552843>

Yang, M., Wong, S. C. P., & Coid, J. (2010). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin, 136*(5), 740-767.

<https://doi.org/10.1037/a0020473>

Table 1

*Brief Overview of Assessment Measures Used to Assess Recidivism Risk Included in the Meta-Analysis*

Measure	Brief description	Class	Target population	Primary risk assessed	Item scaling	Primary item type	Reliability [ICC/ $\alpha$ ]
Level of Service (LS) Inventory adult variants	Family of general risk-need tools with item content broadly organized according the Central Eight risk-need domains: criminal history, antisocial attitudes, peers/associates, antisocial pattern, employment/education, family/marital, alcohol/drug, leisure/recreation (e.g., Level of Service Inventory-Revised [LSI-R]; Level of Service/Case Management Inventory [LS/CMI]).	Actuarial	Adult	General	43-54 binary items	Dynamic	ICC = .92 (Rettinger & Andrews, 2010)
Level of Service (LS) Inventory youth variants	Developmentally adapted variants of the LS measures for youth, similar in form and content to adult counterparts (e.g., emphases placed on school, family/parenting relationships, etc.; e.g., Youth Level of Service Case Management Inventory [YLS/CMI]).	Actuarial	Youth	General	42-43 binary items	Dynamic	ICC = .90 (Olver, Stockdale et al., 2012)
Statistical Information on Recidivism (SIR) scale	Static actuarial general recidivism risk tool with differentially weighted item content in criminal history/demographic domains. SIR-Proxy is a parallel form developed for research purposes with Indigenous men and women.	Actuarial	Adult	General	15 items statistically weighted	Static	$\alpha$ = .77-.78 (Nafekh & Motiuk, 2002)
Static Factors Assessment (SFA)	Static recidivism general risk tool with item content reflecting density and severity of criminal history and index offense(s)	SPJ/mechanical	Adult	General	137 binary items	Static	nil
Dynamic Factors Identification and Analysis-Revised (DFIA-R)	General risk-need tool with item coverage of most Central Eight domains to drive correctional planning.	SPJ/mechanical	Adult	General	117 ordinal items	Dynamic	$\alpha$ = .93-.95 (Stewart et al., 2017)
Service Planning Instrument (SPIn)	General risk-need tool with item coverage of most Central Eight domains to drive service planning; assessment of needs also includes corresponding assessment of strengths in a given domain. Has pre-screen and full assessment versions.	Actuarial	Adult	General	35-90 Likert items	Dynamic	nil
Youth Assessment and Screening Instrument (YASI)	Developmental adaptation of SPIn with similar content coverage and scoring/administration platform of parent measure for the assessment of risk, treatment needs, and strengths in youth.	Actuarial	Youth	General	34-90 Likert items	Dynamic	ICC = .63 (Kennealy et al., 2017)

Offender Risk Assessment Management System (ORAMS)	Jurisdiction specific general risk-need tool with differentially weighted ordinal items covering several Central Eight domains and criminal history/demographic indicators.	Actuarial	Adult	General	15 items ordinally weighted	Dynamic	nil
Arizona Risk/Needs Assessment (ARNA)	Jurisdiction specific general risk-need tool with equal representation of criminal history indicators and some Central Eight domains.	Actuarial	Youth	General	10 binary items	Static and dynamic	$\alpha = .54-.60$ (Schwalbe, 2009)
Violence Risk Scale (VRS)	Violence risk assessment and treatment planning tool designed to measure changes in risk through modified application of stages of change model; items have general and violence risk relevance.	Actuarial	Adult	Violent	26 ordinal items	Dynamic	ICC = .84 (Lewis et al., 2013)
Violence Risk Scale-Sexual Offense version (VRS-SO)	Sexual recidivism risk assessment and treatment planning tool with same risk/change platform as the VRS. Dynamic items factor analytically arranged into domains of sexual deviance, general criminality, and treatment responsiveness.	Actuarial	Adult	Sexual	24 ordinal items	Dynamic	ICC = .85 (Olver et al., 2020)
Violence Risk Scale-Youth Version (VRS-YV)	Developmental adaptation of the VRS to assess violence risk/change in youth; same item scaling and assessment platform as parent measure.	Actuarial	Youth	Violent	23 ordinal items	Dynamic	ICC = .90 (Stockdale et al., 2014)
Static-99R	Static actuarial sexual recidivism risk tool with differentially weighted item content reflecting sexual and general criminal history and perpetrator/victim demographics.	Actuarial	Adult	Sexual	10 items statistically weighted	Static	ICC = .89 (McGrath et al., 2012)
Static-2002R	Similar targeted items to Static-99R with expanded coverage on some criminal history/demographic domains and minor variations in scoring.	Actuarial	Adult	Sexual	14 items statistically weighted	Static	ICC = .89 (Jung et al., 2017)
STABLE 2007	Sexual recidivism risk tool comprised of stable dynamic items to link assessment and intervention. Item coverage reflects self-regulation problems in sexual and general risk-domains, emotional functioning, and relationships.	Actuarial	Adult	Sexual	13 ordinal items	Dynamic	ICC = .86 (Fernandez & Helmus, 2017)
Ontario Domestic Assault Risk Assessment (ODARA)	Static actuarial IPV recidivism risk tool with item content reflecting criminal history and perpetrator/victim risks/demographics.	Actuarial	Adult	IPV	13 binary items	Static	ICC = .84 (Hilton et al., 2021)
Spousal Assault Risk Assessment (SARA)	IPV risk assessment and management tool with item content reflecting general/IPV criminal history and concerns in general risk-need and clinical domains.	SPJ/ mechanical	Adult	IPV	20 ordinal items	Dynamic	ICC = .84 (Kropp & Hart, 2000)
Structured Assessment of Violence Risk in Youth (SAVRY)	Youth violence risk assessment measure with item content arranged into static/historical, social/contextual, and individual/clinical domains; has a separate smaller section of binary protective items.	SPJ/ mechanical	Youth	Violent	24 ordinal risk; 6 binary protect	Dynamic	ICC = .91 (Viljoen et al., 2017)

Structured Assessment of Protective Factors (SAPROF)	A measure of protective factors/strengths to mitigate recidivism risk; item content arranged into domains of internal, motivational, and external.	SPJ/ mechanical	Youth	General/ Violent	17 ordinal items	Dynamic	ICC = .76 (Olver & Riemer, 2021)
Structured Assessment of Protective Factors-Youth Version (SAPROF-YV)	Developmental adaptation of the SAPROF to measure protective factors/strengths in justice involved youth; item content arranged into domains of resilience, motivational, relational, and external.	SPJ/ mechanical	Youth	General/ Violent	16 ordinal items	Dynamic	ICC = .82 (Kleeven et al., 2020)
Psychopathy Checklist-Revised (PCL-R)	Numerically scored symptom-construct rating scale designed to clinically assess psychopathy. Item content is factor analytically arranged into interpersonal, affective, lifestyle, and antisocial factors/facets. Not a risk tool per se, but frequently employed in risk evaluations with other measures.	Not a risk tool; ratings summed mechanically	Adult	General/ violent	20 ordinal items	Static and dynamic	ICC = .86 (Hare, 2003)
Psychopathy Checklist: Youth Version (PCL: YV)	Developmentally adapted variant of PCL-R to assess psychopathic features in youth. Same numeric scoring, rating platform, content, and latent structure as parent measure. Also, not a risk tool but frequently used in risk evaluations with other measures.	Not a risk tool; ratings summed mechanically	Youth	General/ violent	20 ordinal items	Static and dynamic	ICC = .90 (Olver & Stockdale, 2010)

*Note:* Class: refers to the type of risk instrument or assessment approach that the measure represents—actuarial (i.e., numeric tools with scores linked to recidivism estimates), structured professional judgment (SPJ; non-numeric tools where pattern of item ratings informs final risk judgment), or mechanical (i.e., items summed to generate a scale total, but scores are not linked to recidivism estimates)—see “Risk Tool Classification” section for further information. Reliability estimates are interrater reliability coefficients (intraclass correlation coefficient, ICC) from a representative study (i.e., based on sample size and midrange coefficient magnitudes across multiple studies) when available; Cronbach alpha ( $\alpha$ ) provided from a representative study when IRR is not available; nil = when no reliability data are available; SPJ: Structured professional judgment; the SPJ/mechanical class designation represents an SPJ tool designed for practice, but for which the item ratings can be summed in a mechanical manner (usually for psychometric research purposes) to yield a total score that is not linked to recidivism estimates; IPV: intimate partner violence. NB: Most “dynamic” tools also contain a smaller subset of static items (< 30%) that are scored separately and combined with dynamic item ratings (> 70%) to yield an overall estimate of recidivism risk; the exceptions are the DFIA-R and STABLE 2007 which have only dynamic items, but are intended to be used in tandem with static tools (SFA and Static-99R/2002R, respectively). PCL-R/YV items are not explicitly arranged into static or dynamic, but rather, items vary in their degree of dynamism from mostly or exclusively static (e.g., early behavior problems, serious criminal behavior) to arguably dynamic (e.g., lacks goals, failure to accept responsibility).

Table 2

*Summary Table of Demographic and Descriptive Variables for Meta-Analysis*

Variable	<i>k</i> (I/NIW)	Overall <i>n</i> / <i>M</i> ( <i>SD</i> )	Indigenous <i>n</i> / <i>M</i> ( <i>SD</i> )	NIW <i>n</i> / <i>M</i> ( <i>SD</i> )
Total studies	91	297,422	59,693	237,729
Age (years)				
Age, overall	60 (34/33)	29.4 (9.2)	30.4 (7.8)	33.2 (8.8)
Age, youth	16 (6/6)	15.6 (1.5)	16.0 (1.6)	16.2 (1.4)
Age, adult	44 (28/27)	34.2 (10.1)	33.4 (9.2)	36.7 (10.5)
Sex				
Male	40/38	161,492	32,413	129,079
Female	11/10	19,754	4,659	15,095
Male/female	26/24	116,176	22,621	93,555
Ethnoracial group				
Non-Indigenous/White (NIW)	75	59,693	59,693	
Indigenous	70	237,729		237,729
Setting				
Custody	42			
Community	9			
Custody/community	36			
Developmental group				
Adult	54/50	270,282	54,464	215,818
Youth	21/20	27,140	5,229	21,911
Design				
Field prospective	48			
Archival retrospective	23			
Face-to-face prospective	5			
Mixed archival/field	7			
Meta-analysis	7			
Year				
2020s	19	<i>M</i> = 2012, <i>Mdn</i> = 2013		
2010s	44			
2000s	22			
1990s	4			
1980s	2			
Source				
Journal article†	57			
Government report	17			
Unpublished thesis/dissertation	11			
Unpublished data/manuscript under review	4			
Conference presentation	2			
Country				
Canada	69			
USA	10			
Australia	8			

Multi-country	4		
Assessment measure			
Level of Service (LS) Inventory adult variants	11	26.9 (8.2)	21.7 (9.1)
Level of Service (LS) Inventory youth variants	11	19.0 (7.0)	15.1 (7.2)
Statistical Information on Recidivism (SIR) scale	8/6	-5.7 (2.8)	-2.5 (2.5)
Static Factors Assessment (SFA)	1	36.9 (13.5)	27.9 (13.7)
Dynamic Factors Identification and Analysis-Revised (DFIA-R)	2	-	-
Service Planning Instrument (SPIn)	1	49.4 (29.8)	32.5 (26.0)
Youth Assessment and Screening Instrument (YASI)	2	34.9 (18.8)	26.2 (17.8)
Offender Risk Assessment Management System (ORAMS)	1	-	-
Arizona Risk/Needs Assessment (ARNA)	4	-	-
Violence Risk Scale (VRS)	4	51.1 (9.1)	46.1 (11.6)
Violence Risk Scale-Sexual Offense version (VRS-SO)	3	34.6 (9.6)	32.5 (9.6)
Violence Risk Scale-Youth Version (VRS-YV)	2	43.8 (10.1)	32.3 (12.3)
Static-99R	13	4.2 (2.2)	3.4 (2.5)
Static-2002R	4	6.0 (2.2)	5.4 (2.8)
STABLE 2007	3	9.6 (4.4)	8.8 (4.8)
Ontario Domestic Assault Risk Assessment (ODARA)	2	6.3 (2.1)	4.95(2.3)
Spousal Assault Risk Assessment (SARA)	2	24.5 (6.5)	22.1 (6.5)
Structured Assessment of Violence Risk in Youth (SAVRY)	5	26.6 (8.6)	21.2 (9.2)
Structured Assessment of Protective Factors (SAPROF)	1	10.6 (4.6)	13.3 (6.0)
Structured Assessment of Protective Factors-Youth Version (SAPROF-YV)	1	6.4 (3.8)	10.3 (6.0)
Psychopathy Checklist-Revised (PCL-R)	5	21.0 (6.0)	19.9 (6.8)
Psychopathy Checklist: Youth Version (PCL: YV)	5	21.3 (6.4)	19.6 (7.2)
Recidivism			
General	70/65	51.5%	38.9%
Violent	38/37	36.6%	23.6%
Sexual	15/15	13.6%	9.6%
Follow-up (years)	63	4.0 (3.7)	

Note: Previously published meta-analyses were excluded from this table (with the exception of the Design category) due to substantial overlapping samples with the present study; I = Indigenous samples, NIW = non-Indigenous (White majority) samples; † includes unpublished secondary analysis of previously published data; Non-overlapping samples were utilized to compute *k* frequencies and *n* for demographic variables (when such information was reported in a given study) and as such, the cell totals do not necessarily sum to *k* = 91 studies. *M* (SD) are unweighted means and concordant SD (for continuous study variables).

Table 3

*Meta-analysis of Indigenous and non-Indigenous Mean Differences on Assessment Measure Scores*

Measure	Random d [95%CI]	Fixed d [95%CI]	k	n	Q	I <sup>2</sup>
LSI (adult)	0.63 [0.56, 0.70]	0.67 [0.66, 0.69]	8	103,795	70.59*	90.08
LSI (youth)	0.58 [0.37, 0.80]	0.39 [0.32, 0.45]	9	4,182	51.47*	84.46
SPIIn/YASI	0.59 [0.45, 0.72]	0.61 [0.53, 0.69]	2	4,120	1.70	41.28
SAVRY total	0.66 [0.28, 1.03]	0.59 [0.46, 0.72]	4	1,086	19.95*	84.97
SAVRY Protect	-0.36 [-0.63, -0.08]	-0.28 [-0.40, -0.15]	4	1,086	10.90*	72.47
SAPROF/YV	-0.56 <sup>NS</sup> [-1.16, 0.04]	-0.43 [-0.60, -0.27]	2	647	10.71*	90.66
VRS-YV total	1.06 [0.52, 1.60]	1.00 [0.75, 1.25]	2	354	4.29*	76.66
VRS-YV dynamic	0.96 [0.48, 1.43]	0.91 [0.66, 1.15]	2	354	3.42	70.75
VRS total	0.46 [0.17, 0.75]	0.53 [0.39, 0.67]	4	879	11.34*	73.53
VRS dynamic	0.37 [0.05, 0.70]	0.46 [0.32, 0.60]	4	879	14.53*	79.36
VRS change	0.06 <sup>NS</sup> [-0.13, 0.26]	0.06 <sup>NS</sup> [-0.13, 0.26]	3	434	0.57	0.00
ODARA	0.65 [0.35, .96]	0.65 [0.35, 0.96]	2	394	0.30	0.00
SARA-V2/3 (mech.)	0.42 [0.29, 0.51]	0.47 [0.40, 0.49]	2	9,244	1.46	30.33
SIR	0.43 [0.26, 0.59]	0.50 [0.46, 0.55]	3	15,561	4.91	59.27
SFA (mech.)	0.63 [0.60, 0.65]	0.63 [0.60, 0.65]	1	64,034	0.00	0.00
VRS-SO total	0.22 [0.02, 0.43]	0.26 [0.13, 0.38]	3	1,044	4.56	56.17
VRS-SO dynamic	0.15 [0.01, 0.30]	0.16 [0.03, 0.29]	3	1,044	2.45	18.42
VRS-SO change	-0.12 <sup>NS</sup> [-0.25, 0.00]	-0.12 <sup>NS</sup> [-0.25, 0.00]	3	1,044	0.70	0.00
STABLE 2007	0.18 [0.01, 0.35]	0.16 [0.09, 0.23]	3	4,312	4.72	57.64
Static-99R	0.36 [0.23, 0.49]	0.40 [0.35, 0.46]	12	7,758	45.48*	75.81
Static-2002R	0.23	0.23	4	989	1.11	0.00

PCL-R	[0.08, 0.38] 0.16 <sup>NS</sup>	[0.08, 0.38] 0.20	5	1,586	19.07*	79.02
PCL: YV	[-0.07, 0.39] 0.23 <sup>NS</sup>	[0.10, 0.31] 0.15	5	1,061	20.30*	80.29
Overall	[-0.09, 0.55] 0.57	[0.02, 0.29] 0.57	90	212,724	1,391.16*	93.60
	[0.55, 0.58]	[0.56, 0.58]				

*Note:* Overall *N* computed by summing a given study/sample once used to generate an effect, hence *ns* do not add to overall *N*. LSI: Level of Service Inventory measures; SFA: Static Factors Assessment; DFIAR: Dynamic Factors Identification and Analysis; SIR: Statistical Information on Recidivism Scale; SFA: Static Factors Assessment; ODARA: Ontario Domestic Assault Risk Assessment; SARA-V2/3 (mech.): Spousal Assault Risk Assessment Versions 2 and 3 mechanical summation of item scores; VRS: Violence Risk Scale (total: combined static and dynamic risk score, change: dynamic risk change score); VRS-SO: Violence Risk Scale-Sexual Offense Version (total: combined static and dynamic risk score, change: dynamic risk change score); VRS-YV Violence Risk Scale-Youth Version; SAVRY: Structure Assessment of Violence Risk for Youth (total: summed score of all risk items, Protect: summed score of all protective items); SPIn: Service Planning Instrument; YASI: Youth Assessment and Screening Instrument; PCL-R: Psychopathy Checklist-Revised; PCL: YV Psychopathy Checklist: Youth Version. NS: standardized mean difference between Indigenous and non-Indigenous group assessment measure scores not significant. \*Q statistic significant at < .05.

Table 4

*Study Details for Meta-Analysis of Level of Service (LS) Measures in the Prediction of Recidivism for Indigenous and Non-Indigenous Samples*

LS measure and study sample	Sex	DG	RC	Indigenous					Non-Indigenous					
				n	BR	r	d	AUC	n	BR	r	d	AUC	
LSI														
Bonta (1989)	M	A	G	48	.438	.35	0.75	-		71	.577	.500	1.17	-
MRNF														
Bonta et al. (1997)	M	A	G	390	.659	.32	0.71	-		513	.478	.360	0.77	-
LSI-OR/LSCMI														
Brews (2009) <sup>a</sup>	F	A	G	253	.283	-	0.91	.74		1502	.283	-	1.02	.76
	F	A	V	253	.148	.15	0.43	-		1502	.145	.123	0.35	-
Hogg (2011)/Wormith et al. (2015) <sup>b</sup>	F	A	G	418	.459	-	0.68	.69		2695	.313	-	1.00	.76
	M	A	G	1274	.607	-	0.83	.72		12951	.400	-	0.91	.79
	M/F	A	G	1692	.570	-	0.82	.72		15646	.385	-	0.92	.74
	F	A	V	418	.153	.20	0.57	-		2695	.145	.262	0.77	-
	M	A	V	1274	.148	.17	0.49	-		12951	.145	.313	0.94	-
	M/F	A	V	1692	.150	.17	0.50	-		15646	.145	.302	0.89	-
Rettinger & Andrews (2010)	F	A	G	48	.500	-	1.19	.80		307	.459	-	1.53	.86
LSI-R														
Hansen (2018)	M/F	A	G	166	.355	-	0.58	.66		-	-	-	-	-
Holsinger et al. (2006)	M/F	A	G	140	.300	.11	0.24	-		263	.255	.230	0.54	-
	M	A	G	100	.310	.19	0.42	-		162	.309	.220	0.50	-
	F	A	G	40	.275	-.13	-0.29	-		101	.287	.260	0.60	-
Hsu et al. (2011) <sup>c</sup>	M	A	G	6023	.279	.18	0.41	-		30791	.201	.210	0.54	-
	F	A	G	2125	.234	.19	0.46	-		6806	.138	.270	0.81	-
	M	A	G	4071	.305	.16	0.35	-		17410	.173	.260	0.71	-
	F	A	G	552	.292	.16	0.36	-		1413	.168	.350	1.00	-
	M	A	G	864	.394	.18	0.37	-		2528	.305	.130	0.29	-
	F	A	G	186	.414	.18	0.37	-		263	.335	.200	0.43	-
Lowenkamp & Latessa (2002)	M/F	A	G	189	.339	.11	0.23	-		311	.309	.230	0.51	-
Towers (2020) <sup>d</sup>	M/F	A	G	798	.472	.24	0.50	-		2900	.382	.27	0.58	-
Watkins (2011) <sup>c</sup>	M	A	G	2465	.311	-	0.57	.66		7555	.337	-	0.71	.69
	F	A	G	417	.311	-	0.35	.60		614	.337	-	0.69	.69

YLS/CMI													
Gossner & Wormith (2007)	M/F	Y	G	62	.613	.345	0.76	-	32	.344	.240	0.52	-
Huang et al. (2021)	M/F	Y	G	205	.824	-	0.21	.56	193	.575	-	0.55	.65
Jung & Rawana (1999) <sup>e</sup>	M/F	Y	G	134	.325	-	0.43	-	129	.282	.746	0.75	-
Lockwood et al. (2014) <sup>f</sup>	M/F	Y	G	37	.703	-	0.70	.69	33	.667	.283	0.63	-
McKinnon (2004)	M/F	Y	G	28	.571	-	0.47	.63	74	.432	-	0.25	.57
	M/F	Y	V	28	.500	-	0.55	.65	74	.176	-	0.25	.57
Olver et al. (2012)	M/F	Y	G	104	.788	-	1.05	.77	40	.475	-	0.07	.52
	M/F	Y	V	104	.548	-	1.00	.76	40	.225	-	0.80	.71
Shepherd et al. (2015)	M/F	Y	G	40	.775	-	0.62	.67	85	.776	-	1.14	.79
	M/F	Y	V	40	.600	-	0.62	.67	85	.612	-	0.80	.71
Thompson & McGrath (2012)	M/F	Y	G	1053	.337	-	0.38	.60	1582	.494	-	0.52	.64
LSI-SK													
Luong & Wormith (2011)	M/F	Y	G	123	.707	-	0.52	.64	69	.478	-	1.16	.79
Prince et al (2022)	M/F	Y	G	203	.683	-	0.50	.64	58	.569	-	0.82	.72
	M	Y	G	148	.762	-	0.40	.61	50	.600	-	1.00	.76
	F	Y	G	55	.473	-	0.62	.67	-	-	-	-	-
	M/F	Y	V	203	.505	-	0.43	.62	58	.328	-	0.87	.73
	M	Y	V	148	.574	-	0.35	.60	50	.340	-	0.87	.73
Rector et al. (2007)	F	Y	V	55	.309	-	0.62	.67	-	-	-	-	-
	M/F	Y	G	628	.537	.370	0.80	-	244	.371	.330	0.72	-

*Note:* Sex: male (M), female (F); DG: developmental group adult (A), youth (Y); RC: recidivism criterion general (G), violent (V); BR: base rate of the recidivism criterion; LSI: Level of Service Inventory; LSI-R: Level of Service Inventory-Revised; LS/CMI: Level of Service/Case Management Inventory; LSI-OR Level of Service Inventory-Ontario Revision; MNRF: Manitoba Risk-Need Form; YLS/CMI: Youth Level of Service/Case Management Inventory (AA: Australian Adaptation); LSI-SK: Level of Service Inventory-Saskatchewan Youth Edition. Sex =biological sex (both men and women or women only) across study samples; a = Brews (2009) did not report base rate of violent recidivism, therefore, rates reported from Hogg (2011)/Wormith et al. (2015) provincial female sample used to generate SE for meta-analysis; b = Hogg (2011)/Wormith et al. (2015) did not report rates of violent recidivism for White or non-Indigenous male and female subgroups; given that base rates of violence were virtually identical for Indigenous men and women, the overall rate of violent recidivism for the White aggregate group was applied to White men and women; c = Hsu et al. (2011) and Watkins (2011) examined LSI-R on large overlapping samples of Australian men and women; the meta-analytic average was taken to generate single effect sizes d (SE) for Indigenous men .520 (.016), Indigenous women .418 (.040), non-Indigenous men .573 (.011), non-Indigenous women .810 (.029); d = Towers (2020) reported four different types of general recidivism outcomes with effect sizes, the reported base rate and effect sizes are the average of these; e = as Jung and Rawana (1999) did not report exact M (SD) of YLS/CMI for Indigenous and non-Indigenous recidivists and nonrecidivists, but presented this graphically, mean scores for Indigenous and non-Indigenous

recidivists and non-recidivists could be extracted from figure 1, and the pooled SD for Indigenous and non-Indigenous groups used to generate Cohen's  $d$ ;  $f$  = Lockwood et al. (2018) did not report AUC for non-Indigenous group, only to state it was not as strong as that for Indigenous subgroup (AUC = .68,  $p$  = .08), as such,  $r_{pb}$  was estimated from  $p > .10$  at  $n = 33$  and then converted to  $d$  using reported recidivism base rate for non-Indigenous subgroup.

Table 5

Updated and Prior Meta-Analyses of Level of Service (LS) Measures in the Prediction of Recidivism for Indigenous and Non-Indigenous Samples

LS group	Updated Meta-Analysis															
	Indigenous								Non-Indigenous							
	k	Random		Fixed		n	Q	I <sup>2</sup>	k	Random		Fixed		n	Q	I <sup>2</sup>
	d [95%CI]	AUC	d [95%CI]	AUC					d [95%CI]	AUC	d [95%CI]	AUC				
General recidivism																
All	22	0.57 [0.49, 0.66]	.66	0.53 [0.51, 0.56]	.65	23,044	85.21*	75.35	21	0.75† [0.64, 0.86]	.70	0.69† [0.67, 0.70]	.69	92,274	422.66*	95.27
Male	6	0.62 [0.44, 0.80]	.67	0.54 [0.51, 0.57]	.65	15,383	29.32*	82.95	6	0.77 [0.56, 0.99]	.71	0.66† [0.64, 0.68]	.68	72,031	236.03*	97.88
Female	6	0.62 [0.36, 0.88]	.67	0.49 [0.42, 0.55]	.63	4,094	24.34*	79.46	5	1.00† [0.82, 1.18]	.76	0.90† [0.86, 0.95]	.74	13,600	42.91*	90.68
Youth	11	0.56 [0.40, 0.72]	.65	0.53 [0.45, 0.61]	.65	2,617	24.28*	58.81	11	0.63 [0.48, 0.79]	.67	0.57 [0.49, 0.65]	.66	2,539	16.97	41.06
Adult	11	0.58 [0.47, 0.70]	.66	0.53 [0.51, 0.56]	.65	20,427	60.93*	83.59	10	0.82† [0.69, 1.01]	.72	0.69† [0.68, 0.71]	.69	89,735	397.69*	97.74
Canada	15	0.66 [0.54, 0.79]	.68	0.73 [0.67, 0.80]	.70	3,955	32.98*	57.56	15	0.81 [0.68, 0.95]	.72	0.91† [0.88, 0.94]	.74	18,881	64.44*	78.28
Australia	4	0.45 [0.37, 0.54]	.63	0.50 [0.47, 0.53]	.64	17,796	9.56*	68.61	4	0.67† [0.51, 0.84]	.68	0.60† [0.58, 0.62]	.66	67,380	64.44*	95.35
USA	4	0.42 [0.27, 0.58]	.62	0.45 [0.33, 0.56]	.62	1,293	4.19	28.36	3	0.57 [0.50, 0.64]	.66	0.57 [0.50, 0.64]	.66	3,474	0.29	0.00
Violent recidivism																
All	6	0.53 [0.39, 0.67]	.65	0.52 [0.41, 0.63]	.64	2,320	6.03	17.05	6	0.66 [0.33, 0.98]	.68	0.84† [0.80, 0.89]	.72	17,405	52.49*	90.47
Canada	5	0.54 [0.37, 0.70]	.65	0.52 [0.41, 0.63]	.64	2,280	5.93	32.58	5	0.63 [0.26, 1.00]	.67	0.84† [0.80, 0.89]	.72	17,320	52.45*	92.37
Female	3	0.53 [0.33, 0.73]	.65	0.53 [0.53, 0.71]	.65	726	0.49	0.00	2	0.56 [0.15, 0.98]	.65	0.62 [0.53, 0.71]	.67	4,197	20.39*	95.10

Prior Meta-Analyses

Olver et al. (2009) <sup>a</sup>	5	-	-	0.75 [0.61, 0.89]	.70	860	2.22	-	5	-	-	0.68 [0.49, 0.86]	.68	462	7.10	-
Olver et al. (2014) <sup>a</sup>	13	0.66 [0.60, 0.71]	.67	0.63 [0.58, 0.69]	.67	5354	61.73*	80.56	24	0.63 [0.61, 0.65]	.67	0.70 [0.68, 0.72]	.68	40989	603.79*	96.19
Wilson & Gutierrez (2014) <sup>g</sup>	13	0.62 [0.46, 0.79]	.67	0.59 [0.55, 0.63]	.66	21058	146.52*	91.81	11	0.82† [NR]	.72	0.74† [NR]	.70	58719	33.06*	69.75
Wilson & Gutierrez (2014) <sup>v</sup>	NR	NR	-	0.49 [0.42, 0.56]	.64	NR	NS	NR	-	-	-	-	-	-	-	-

Note: k = number of studies; d = standardized mean difference with 95% confidence intervals (95%CI) random or fixed effects meta-analytic effect size (0.20 small, 0.50 medium, 0.80 large); AUC = Area under the curve equivalent (.56 small, .64 medium, .71 large) to d (from Rice & Harris, 2005); n = number of individuals across the studies included in the effect size; \* Q significant at < .05; I<sup>2</sup> = measure of effect size variability across studies (25% small, 50% medium, 75% large); † ES significantly larger for non-Indigenous samples; NR = not reported. a = d equivalent in prediction of general recidivism, computed from reported r<sub>pb</sub> and aggregate base rate, reported for consistency and comparative purposes (see Table 1 for r<sub>pb</sub> value); g = first row of Wilson & Gutierrez (2014) prediction of general recidivism, v = second row, violent recidivism.

Table 6

*Study Details for Meta-Analysis of Statistical Information on Recidivism (SIR) Scale Variants in the Prediction of Recidivism for Indigenous and Non-Indigenous Samples*

SIR measure and study sample	Sex	RC	Indigenous					Non-Indigenous				
			n	BR	r	d	AUC	n	BR	r	d	AUC
GSIR												
Hann & Harman (1989)	M	G	49	.571	-	0.82	.72	-	-	-	-	-
Hann & Harman (1993)	M	G	269	.651	-	0.86	.73	-	-	-	-	-
SIR-R1												
Wormith & Olver (2002)	M	G	48	.708	.37	0.88	-	43	.605	.61	1.58	-
		V	48	.438	.29	0.63	-	43	.488	.53	1.25	-
Bonta & Ruge (2004)	M	G	940	.567	-	0.91	.74	4,475	.408	-	1.05	.77
		V	940	.205	-	0.55	.65	4,475	.123	-	0.55	.65
Wong & Gordon (2006)	M	G	165	.297	-	0.48	.65	260	.185	-	0.45	.62
		V	165	.121	-	0.48	.64	260	.058	-	0.78	.71
SIR Proxy/SIR-R1												
Nafekh & Motiuk (2002)	M/F	G	1,211	.280	-	0.67	.68	6,881	.215	-	0.95	.75
		V	1,211	.075	-	0.55	.65	6,881	.034	-	0.79	.71
		S	1,211	.014	-	0.35	.60	6,881	.005	-	0.14	.54
Barnum & Gobeil (2012)	M	G	2,560	.250	-	0.47	.63	12,485	.190	-	0.70	.69
		V	2,560	.083	-	0.26	.57	12,485	.057	-	0.58	.66
		S	2,560	.008	-	0.43	.62	12,485	.005	-	0.26	.57
		G	246	.159	-	0.55	.65	666	.101	-	0.91	.74
		V	246	.041	-	0.75	.70	666	.032	-	1.30	.82

Note: GSIR: General Statistical Information on Recidivism scale; SIR-R1: Statistical Information on Recidivism scale (Revision 1); SIR-Proxy: Statistical Information on Recidivism Proxy scale; RC = recidivism criterion sexual (s), violent (v), or general (g); BR = base rate for a given recidivism criterion; d = standardized mean difference (0.20 small, 0.50 medium, 0.80 large); AUC = area under the curve (.56 small, .64 medium, .71 large).

Table 7

*Meta-Analysis of Statistical Information on Recidivism (SIR) Scale Variants in the Prediction of Recidivism for Indigenous and Non-Indigenous Samples*

SIR group	Indigenous									Non-Indigenous							
	Random			Fixed			n	Q	I <sup>2</sup>	Random			Fixed		n	Q	I <sup>2</sup>
	k	d [95%CI]	AUC	d [95%CI]	AUC	k				d [95%CI]	AUC	d [95%CI]	AUC				
General recidivism																	
All	8	0.69 [0.52, .85]	.69	0.64 [0.58, 0.71]	.68	5,448	31.89*	78.05	6	0.88 [.70, 1.06]	.73	0.83† [.80, .86]	.72	24,810	113.70*	95.60	
Violent recidivism																	
All	6	0.47 [0.31, 0.63]	.63	0.43 [0.34, 0.52]	.62	5,170	10.58*	52.76	6	0.73† [.58, .88]	.70	0.62† [.57, .67]	.67	24,810	22.10*	77.38	

Note: d = standardized mean difference with 95% confidence intervals (95%CI) random or fixed effects meta-analytic effect size (0.20 small, 0.50 medium, 0.80 large); AUC = area under the curve (.56 small, .64 medium, .71 large). \* Q statistic significant < .05; I<sup>2</sup> measure of effect size heterogeneity (25% low, 50% medium, 75% high). † non-Indigenous/White group effect size significantly higher than Indigenous group effect size.

Table 8

*Study Details for Meta-Analysis of Violence Risk Scale (VRS) Measures in the Prediction of Recidivism Outcomes for Indigenous and Non-Indigenous Samples*

VRS measure and study sample	Sex	RC	Indigenous				Non-Indigenous			
			n	BR	d	AUC	n	BR	d	AUC
VRS-SO (total)										
Olver, Sowden, et al. (2018) <sup>†</sup>	M	S	384	.143	0.65	.68	660	.101	0.96	.75
		V	384	.359	0.62	.67	660	.167	0.88	.73
		G	384	.591	0.76	.71	660	.350	0.65	.68
Olver et al. (2007) <sup>a</sup>	M	S	108	.185	0.70	.69	209	.177	1.19	.80
		V	108	.482	0.74	.70	209	.273	0.87	.73
		G	108	.685	0.54	.65	209	.478	0.44	.62
Olver et al. (2014, 2020) <sup>a</sup>	M	S	197	.112	0.73	.70	365	.060	0.78	.71
		V	197	.259	0.58	.66	365	.096	0.91	.74
		G	197	.523	0.89	.74	365	.274	0.79	.71
Sowden & Olver (2017) <sup>a</sup>	M	S	79	.165	0.24	.57	86	.093	0.76	.71
		V	79	.456	0.62	.67	86	.209	0.44	.62
		G	79	.633	0.75	.70	86	.360	0.50	.64
VRS-SO (dynamic)										
Olver, Sowden, et al. (2018) <sup>†</sup>	M	S	384	.143	0.57	.65	660	.101	0.85	.73
		V	384	.359	0.58	.66	660	.167	0.75	.70
		G	384	.591	0.67	.68	660	.350	0.50	.64
Olver et al. (2007) <sup>a</sup>	M	S	108	.185	0.48	.62	209	.177	0.99	.76
		V	108	.482	0.56	.67	209	.273	0.70	.69
		G	108	.685	0.38	.61	209	.478	0.33	.59
Olver et al. (2014, 2020) <sup>a</sup>	M	S	197	.112	0.66	.68	365	.060	0.74	.70
		V	197	.259	0.57	.66	365	.096	0.85	.72
		G	197	.523	0.76	.71	365	.274	0.64	.67
Sowden & Olver (2017) <sup>a</sup>	M	S	79	.165	0.39	.61	86	.093	0.85	.73
		V	79	.456	0.78	.71	86	.209	0.51	.64
		G	79	.633	0.82	.72	86	.360	0.40	.61
VRS-SO (change)										

Olver, Sowden, et al. (2018) <sup>†</sup>	M	S	384	.143	0.33	.59	660	.101	0.44	.62
		V	384	.359	0.37	.60	660	.167	0.37	.60
		G	384	.591	0.06	.52	660	.350	0.17	.55
Olver et al. (2007) <sup>a</sup>	M	S	108	.185	-0.13	.54	209	.177	-0.26	.57
		V	108	.482	-0.08	.52	209	.273	-0.18	.55
		G	108	.685	-0.07	.52	209	.478	-0.10	.53
Olver et al. (2014, 2020) <sup>a</sup>	M	S	197	.112	-0.50	.64	365	.060	-0.40	.61
		V	197	.259	-0.27	.57	365	.096	-0.35	.60
		G	197	.523	-0.05	.51	365	.274	-0.23	.56
Sowden & Olver (2017) <sup>a</sup>	M	S	79	.165	-0.17	.55	86	.093	-0.76	.70
		V	79	.456	-0.55	.65	86	.209	-0.42	.62
		G	79	.633	-0.34	.60	86	.360	-0.19	.55
VRS (total)										
Olver, Cortoni, et al. (2021) <sup>†</sup>	M/F	V	406	.369	0.96	.75	550	.215	1.09	.78
		G	406	.559	0.96	.75	550	.362	1.06	.77
Lewis et al. (2013) <sup>b</sup>	M	V	86	.523	0.58	.66	66	.424	0.57	.66
		G	86	.698	0.48	.63	66	.530	0.44	.62
Coupland & Olver (2020) <sup>b</sup>	M	V	91	.703	0.73	.70	64	.469	0.58	.66
		G	91	.857	0.69	.69	64	.688	1.11	.78
Higgs et al. (2020) <sup>b</sup>	M	V	43	.256	0.03	.51	84	.191	-0.04	.49
		G	43	.535	0.67	.68	84	.476	-0.01	.50
Wong & Gordon (2006) <sup>b</sup>	M	V	173	.121	0.57	.66	272	.065	1.00	.76
		G	173	.289	0.44	.62	272	.184	0.84	.72
VRS (dynamic)										
Olver, Cortoni, et al. (2021) <sup>†</sup>	M/F	V	406	.369	0.91	.74	550	.215	1.12	.79
		G	406	.559	0.88	.73	550	.362	0.98	.76
Lewis et al. (2013) <sup>b</sup>	M	V	86	.523	0.47	.63	66	.424	0.60	.66
		G	86	.698	0.54	.65	66	.530	0.43	.62
Coupland & Olver (2020) <sup>b</sup>	M	V	91	.703	0.56	.65	64	.469	0.76	.70
		G	91	.857	0.51	.64	64	.688	1.30	.82
Higgs et al. (2020) <sup>b</sup>	M	V	43	.256	0.02	.51	84	.191	-0.22	.44
		G	43	.535	0.63	.67	84	.476	0.06	.52
Wong & Gordon (2006) <sup>b</sup>	M	V	173	.121	0.53	.65	272	.065	0.93	.74
		G	173	.289	0.38	.61	272	.184	0.72	.69

VRS (change)										
Olver, Cortoni, et al. (2021) <sup>†</sup>	M/F	V	229	.369	-0.47	.62	272	.215	0.11	.46
		G	229	.559	-0.38	.60	272	.362	0.14	.44
Lewis et al. (2013) <sup>b</sup>	M	V	86	.523	-0.66	.68	66	.424	-0.49	.64
		G	86	.698	-0.46	.63	66	.530	-0.24	.57
Coupland & Olver (2020) <sup>b</sup>	M	V	91	.703	-0.42	.62	64	.469	-0.42	.62
		G	91	.857	-0.64	.67	64	.688	-0.60	.67
Higgs et al. (2020) <sup>b</sup>	M	V	43	.256	-0.65	.68	84	.191	0.11	.53
		G	43	.535	-0.54	.65	84	.476	-0.19	.55
VRS-YV (total)										
Stockdale et al. (2014)	M/F	V	91	.549	0.89	.74	37	.216	1.30	.83
		G	91	.791	0.88	.74	37	.432	0.38	.61
Lovatt et al. (2021)	M/F	V	171	.409	0.49	.64	55	.218	0.48	.63
		G	171	.608	0.42	.61	55	.454	1.01	.76
VRS-YV (dynamic)										
Stockdale et al. (2014)	M/F	V	91	.549	0.84	.72	37	.216	1.10	.78
		G	91	.791	0.91	.74	37	.432	0.33	.59
Lovatt et al. (2021)	M/F	V	171	.409	0.52	.61	55	.218	0.48	.63
		G	171	.608	0.46	.65	55	.454	1.07	.78

Note: VRS-SO: Violence Risk Scale Sexual Offense version; VRS: Violence Risk Scale; VRS-YV: Violence Risk Scale-Youth Version; total = combined static + dynamic score; change = dynamic pre-post change score on a given VRS measure; † Not included in meta-analysis; aggregate sample of individual subsamples included in meta-analysis. a = individual samples from Olver, Sowden, et al. (2018) included in VRS-SO meta-analysis. b = individual samples from Olver, Cortoni, et al. (2021) included in VRS meta-analysis; cell n do not add to aggregate total since Olver, Cortoni et al. (2021) contains an additional sample too small for individual examination (Hogan & Olver, 2019) and AUC based uniformly from unfixed follow-ups for all five samples. RC = recidivism criterion sexual (s), violent (v), or general (g); BR = base rate for a given recidivism criterion; d = standardized mean difference (0.20 small, 0.50 medium, 0.80 large); AUC = area under the curve (.56 small, .64 medium, .71 large). For change score effect sizes, AUC values are reversed for continuity with risk score effect sizes.

Table 9

*Meta-Analysis of Violence Risk Scale (VRS) Measures in the Prediction of Recidivism Outcomes for Indigenous and Non-Indigenous Samples*

VRS measure	Indigenous									Non-Indigenous						
	Random			Fixed			n	Q	I <sup>2</sup>	Random			Fixed		n	Q
k	d [95%CI]	AUC	d [95%CI]	AUC	k	d [95%CI]				AUC	d [95%CI]	AUC				
Sexual recidivism																
VRS-SO total	3	0.61 [0.32, 0.90]	.67	0.61 [.032, 0.90]	.67	384	1.86	0.00	3	0.97 [0.67, 1.27]	.75	0.98 [0.72, 1.25]	.76	660	2.39	16.45
VRS-SO dynamic	3	0.51 [0.22, .080]	.64	0.51 [.022, 0.80]	.64	384	0.75	0.00	3	0.88 [0.62, 1.15]	.74	0.88 [0.62, 1.15]	.74	660	0.75	0.00
VRS-SO change	3	-0.30 [-0.58, -0.01]	.58	-0.30 [-0.58, -0.10]	.58	384	1.44	0.00	3	-0.37 [-0.63, -0.11]	.60	-0.37 [-0.63, -0.11]	.60	660	1.43	0.00
Violent recidivism																
VRS-SO total	3	0.64 [.042, 0.86]	.68	0.64 [0.42, 0.86]	.68	384	0.49	0.00	3	0.80 [0.57, 1.04]	.72	0.81 [0.60, 1.03]	.72	660	2.35	14.78
VRS-SO dynamic	3	0.61 [0.40, 0.83]	.67	0.61 [0.40, 0.83]	.67	384	0.73	0.00	3	0.72 [0.51, 0.93]	.70	0.72 [0.51, 0.93]	.70	660	0.57	0.00
VRS-SO change	3	-0.27 [-0.49, -0.06]	.58	-0.27 [-0.49, -0.06]	.58	384	2.43	17.60	3	-0.28 [-0.49, -0.07]	.58	-0.28 [-0.49, -0.07]	.58	660	0.84	0.00
VRS total	4	0.55 [0.30, 0.79]	.65	0.55 [0.30, 0.79]	.65	393	2.86	0.00	4	0.54 [0.13, 0.94]	.65	0.55 [0.30, 0.81]	.65	486	7.53	60.14
VRS dynamic	4	0.46 [0.22, 0.70]	.63	0.46 [0.22, 0.70]	.63	393	1.87	0.00	4	0.53 [0.04, 1.01]	.64	0.53 [0.04, 1.01]	.64	486	10.61*	71.74
VRS change	3	-0.56 [-0.85, -0.28]	.65	-0.56 [-0.85, -0.28]	.65	220	0.64	0.00	3	-0.28 [-0.64, 0.08]	.58	-0.29 [-0.59, 0.01]	.58	214	2.96	32.36
VRS-YV total	2	0.66 [0.27, 1.05]	.68	0.63 [0.38, 0.88]	.67	262	2.20	54.53	2	0.87 [0.03, 1.70]	.73	0.80 [0.29, 1.31]	.71	92	2.53	60.50
VRS-YV dynamic	2	0.64 [0.34, 0.95]	.68	0.63 [0.38, 0.88]	.67	262	1.40	28.46	2	0.76 [0.20, 1.31]	.70	0.75 [0.24, 1.26]	.70	92	1.16	13.94
VRS/SO/YV total	9	0.61 [0.47, 0.74]	.67	0.61 [0.47, 0.74]	.67	1,039	5.81	0.00	9	0.68 [0.46, 0.74]	.68	0.71 [0.56, 0.87]	.69	1,238	14.90*	46.29

VRS/SO change	6	-0.382 [-0.57, -0.20]	.61	-0.38 [-0.55, -0.20]	.60	604	5.59	10.55	6	-0.28 [-0.46, -0.11]	.58	-0.28 [-0.46, -0.11]	.58	874	3.80	0.00
General recidivism																
VRS-SO total	3	0.77 [.56, .98]	.71	0.77 [.56, .98]	.71	384	1.82	0.00	3	0.60 [0.35, 0.84]	.66	0.62 [0.45, 0.78]	.67	660	3.85	48.04
VRS-SO dynamic	3	0.67 [.46, .88]	.68	0.67 [.46, .88]	.68	384	2.70	25.84	3	0.49 [0.33, 0.63]	.64	0.49 [0.33, 0.66]	.64	660	3.04	32.25
VRS-SO change	3	-0.11 [-.32, .10]	.53	-0.11 [-.32, .10]	.53	384	1.71	0.00	3	-0.18 [-0.34, -0.01]	.55	-0.18 [-0.39, -0.01]	.55	660	0.49	0.00
VRS total	4	0.52 [.29, .75]	.64	0.52 [.29, .75]	.64	393	0.79	0.00	4	0.59 [0.12, 1.05]	.66	0.60 [0.39, 0.81]	.66	486	13.49*	77.76
VRS dynamic	4	0.47 [.24, .70]	.63	0.47 [.24, .70]	.63	393	0.62	0.00	4	0.61 [0.15, 1.06]	.67	.59 [0.38, 0.79]	.66	486	12.85*	76.67
VRS change	3	-0.53 [-.85, -.22]	.65	-0.53 [-.85, -.22]	.65	220	0.23	0.00	3	-0.31 [-0.59, -0.04]	.59	-0.31 [-0.59, -0.04]	.59	214	1.55	0.00
VRS-YV total	2	0.60 [.16, 1.05]	.66	0.54 [.27, .81]	.65	262	2.26	55.78	2	0.72 [0.11, 1.33]	.69	0.74 [0.31, 1.17]	.70	92	1.98	49.54
VRS-YV dynamic	2	0.63 [.20, 1.06]	.67	0.57 [.31, .84]	.66	262	2.12	52.73	2	0.72 [-0.01, 1.44]	.69	0.75 [0.32, 1.18]	.70	92	2.80	64.30
VRS/SO/YV total	9	0.63 [.49, .76]	.67	0.63 [.49, .76]	.67	1039	7.85	0.00	9	0.61 [0.40, 0.82]	.67	0.62 [0.50, 0.75]	.67	1238	19.84*	59.68
VRS/SO change	6	-0.26 [-.46, -.07]	.57	-0.24 [-.41, -.07]	.57	604	6.23	19.77	6	-0.21 [-0.35, -0.07]	.56	-0.21 [-0.35, -0.07]	.56	874	2.74	0.00

Note: VRS-SO: Violence Risk Scale Sexual Offense version; VRS: Violence Risk Scale; VRS-YV: Violence Risk Scale-Youth Version; total = combined static + dynamic score; change = dynamic pre-post change score on a given VRS measure; d = standardized mean difference with 95% confidence intervals (95%CI) random or fixed effects meta-analytic effect size (0.20 small, 0.50 medium, 0.80 large); AUC = area under the curve (.56 small, .64 medium, .71 large). For change score effect sizes, AUC values are reversed for continuity with risk score effect sizes. \* Q statistic significant < .05; I<sup>2</sup> measure of effect size heterogeneity (25% low, 50% medium, 75% high).

Table 10

*Study Details for Meta-Analysis of SAVRY Scores in the Prediction of Recidivism Outcomes for Indigenous and Non-Indigenous Samples*

SAVRY measure and study sample	Sex	RC	Indigenous				Non-Indigenous			
			n	BR	d	AUC	n	BR	d	AUC
<b>SAVRY (total)</b>										
Meyers & Schmidt (2008)	M/F	V	38	.447	1.40	.84	83	.205	0.75	.70
Shepherd et al. (2014)	M	G	32	.862	1.24	.81	84	.725	1.09	.78
		V	32	.690	0.99	.76	84	.580	0.66	.68
Muir et al. (2020)	M	G	219	.557	0.74	.70	357	.384	1.15	.79
		V	219	.288	0.62	.67	357	.165	0.74	.70
	F	G	80	.525	0.99	.76	88	.352	0.70	.69
Lovatt et al. (2021)	M/F	V	80	.250	0.70	.69	88	.125	0.86	.73
		G	171	.608	0.48	.62	55	.454	1.17	.78
		V	171	.409	0.43	.63	55	.218	0.49	.65
<b>SAVRY (Protect)</b>										
Shepherd et al. (2014)	M	G	32	.862	1.88	.91	84	.725	1.20	.80
		V	32	.690	0.62	.67	84	.580	1.05	.77
Muir et al. (2020)	M	G	219	.557	0.40	.61	357	.384	0.82	.72
		V	219	.288	0.28	.58	357	.165	0.54	.65
	F	G	80	.525	0.33	.59	88	.352	0.21	.56
Lovatt et al. (2021)	M/F	V	80	.250	0.33	.59	88	.125	0.11	.53
		G	171	.608	0.42	.59	55	.454	1.06	.75
		V	171	.409	0.36	.59	55	.218	0.50	.64

Note: SAVRY: Structured Assessment of Violence Risk for Youth (total: sum of risk item scores; Protect: sum of protective item scores). For Protect effect sizes, positive d and AUC values represent increased protective factors being associated with decreased recidivism. RC = recidivism criterion violent (v) or general (g); BR = base rate for a given recidivism criterion; d = standardized mean difference (0.20 small, 0.50 medium, 0.80 large); AUC = area under the curve (.56 small, .64 medium, .71 large).

Table 11

*Meta-Analysis of SAVRY Scores in the Prediction of Recidivism Outcomes for Indigenous and Non-Indigenous Samples*

SAVRY measure	Indigenous								Non-Indigenous							
	Random			Fixed			n	Q	I <sup>2</sup>	Random			Fixed			
	k	d [95%CI]	AUC	d [95%CI]	AUC	k				d [95%CI]	AUC	d [95%CI]	AUC	n	Q	I <sup>2</sup>
General recidivism																
Total score	4	0.72 [0.44, 1.00]	.70	0.68 [0.50, 0.87]	.68	502	5.49	45.37	4	1.07 <sup>†</sup> [0.88, 1.26]	.78	1.07 <sup>†</sup> [0.89, 1.25]	.78	584	3.13	4.14
Protect score	4	0.49 [0.18, 0.80]	.64	0.43 [0.25, 0.62]	.62	502	6.94	56.74	4	0.80 [0.42, 1.18]	.71	0.79 <sup>†</sup> [0.62, 0.97]	.71	584	9.88*	69.63
Violent recidivism																
Total score	5	0.69 [0.45, 0.94]	.69	0.65 [0.46, 0.84]	.67	540	5.76	30.53	5	0.71 [0.51, 0.92]	.69	0.71 [0.51, 0.92]	.69	667	0.73	0.00
Protect score	4	0.34 [0.15, 0.53]	.60	0.34 [0.15, 0.53]	.60	502	0.66	0.00	4	0.59 [0.24, 0.93]	.66	0.59 [0.38, 0.81]	.66	584	6.10	50.87

Note: SAVRY: Structured Assessment of Violence Risk for Youth (total: sum of risk item scores; Protect: sum of protective item scores). For Protect effect sizes, positive d and AUC values represent increased protective factors being associated with decreased recidivism; d = standardized mean difference with 95% confidence intervals (95%CI) random or fixed effects meta-analytic effect size (0.20 small, 0.50 medium, 0.80 large); AUC = area under the curve (.56 small, .64 medium, .71 large). \* Q statistic significant < .05; I<sup>2</sup> measure of effect size heterogeneity (25% low, 50% medium, 75% high).

Table 12

*Study Details for Updated Meta-Analysis of Static-99R and Static-2002R Measures in the Prediction of Sexual Recidivism for Indigenous and Non-Indigenous/White Majority Samples*

Static measure and study sample	RC	Country	Indigenous				Non-Indigenous/White					
			n	BR	AUC [SE]	d	n	BR	AUC [SE]	d		
Static-99R												
Boer (2003)	S 10yr	C	53	.075	.82	[.081]	1.03	227	.079	.79	[.056]	1.01
Bonta & Yessine (2005)	S 5yr	C	26	.192	.60	[.184]	0.36	90	.144	.71	[.066]	0.78
Haag (2005)	S 5yr	C	50	.280	.60	[.092]	0.36	141	.177	.70	[.056]	0.74
Spiranovic (2012)	S 5yr	A	162	.247	.58	[.051]	0.29	660	.138	.71	[.031]	0.78
	S 10yr	A	85	.282	.51	[.071]	0.04	472	.216	.67	[.031]	0.62
Smallbone & Rawlings (2013)	S 5yr	A	67	.090	.61	[.082]	0.40	320	.041	.79	[.056]	1.14
Hanson et al. (2015)	S 5yr	C	102	.176	.60	[.084]	0.34	523	.111	.74	[.036]	0.96
Brankley et al. (2017)	S 5yr	C	35	.114	.62	[.178]	0.43	308	.114	.70	[.046]	0.74
	S 10yr	C	35	.200	.60	[.135]	0.36	299	.217	.74	[.033]	0.91
Lee et al. (2018)	S 5yr	C	88	.193	.56	[.077]	0.21	171	.157	.69	[.051]	0.69
	S 10yr	C	66	.288	.66	[.077]	0.59	117	.231	.70	[.059]	0.74
Olver, Sowden et al. (2018) <sup>†</sup>	S 5yr	C	400	.143	.62	[.035]	0.43	665	.096	.71	[.031]	0.78
	S 10yr	C	274	.234	.67	[.038]	0.62	432	.167	.71	[.031]	0.78
Olver et al. (2007) <sup>a</sup>	S 5yr	C	108	.185	.69	[.063]	0.69	200	.170	.72	[.047]	0.83
	S 10yr	C	104	.308	.68	[.056]	0.67	198	.222	.70	[.043]	0.74
Olver et al. (2014, 2020) <sup>a</sup>	S 5yr	C	213	.113	.65	[.014]	0.55	391	.064	.68	[.047]	0.65
	S 10yr	C	118	.153	.68	[.065]	0.65	190	.095	.73	[.047]	0.86
Sowden & Olver (2017) <sup>a</sup>	S 5yr	C	79	.165	.43	[.088]	-0.23	74	.067	.61	[.141]	0.38
	S 10yr	C	52	.269	.57	[.087]	0.23	44	.227	.55	[.091]	0.16
Meyer (2019) <sup>a</sup>	S 5yr	U	123	.090	.64	[.082]	0.50	761	.108	.61	[.032]	0.40
Helmus et al. (2021)	S 5yr	C	943	.064	.68	[.038]	0.65	2662	.044	.71	[.027]	0.78
Static-2002R												
Haag (2005)	S 5yr	C	50	.280	.56	[.097]	0.21	141	.177	.69	[.056]	0.69
Brankley et al. (2017)	S 5yr	C	35	.114	.57	[.153]	0.25	308	.114	.70	[.046]	0.74
	S 10yr	C	35	.200	.64	[.122]	0.50	299	.217	.74	[.031]	0.91
Lee et al. (2018)	S 5yr	C	88	.193	.56	[.071]	0.21	171	.157	.69	[.031]	0.69

Olver et al. (2014, 2020) <sup>a</sup>	S 10yr	C	66	.288	.60	[.066]	0.36	117	.231	.69	[.031]	0.69
	S 5yr	C	108	.148	.71	[.061]	0.78	88	.091	.70	[.058]	0.74
	S 10yr	C	61	.180	.67	[.078]	0.62	44	.136	.89	[.050]	1.75

Note: Study samples are all adult males. † Not included in meta-analysis; aggregate sample of individual subsamples included in meta-analysis. <sup>a</sup> = individual samples from Olver, Sowden, et al. (2018) included in Static-99R meta-analysis; <sup>b</sup> = as Meyer reported both charge and conviction, the averaged base rate and effect size are reported and used for meta-analysis; d = standardized mean difference (0.20 small, 0.50 medium, 0.80 large); AUC = area under the curve (.56 small, .64 medium, .71 large). Country: Canada (C), Australia (A), USA (U). Comparisons in all Canadian and USA samples are between Indigenous and White groups; Australian samples employ a non-Indigenous aggregate.

Table 13

*Prior and Updated Meta-Analyses of Static-99R and Static-2002R Measures in the Prediction of Sexual Recidivism for Indigenous and Non-Indigenous Samples*

Measure and study	Updated Meta-Analysis															
	Indigenous							Non-Indigenous/White								
	k	Random AUC [95%CI]	d	Fixed AUC [95%CI]	d	n	Q	I <sup>2</sup>	k	Random AUC [95%CI]	d	Fixed AUC [95%CI]	d	n	Q	I <sup>2</sup>
Updated Meta-Analysis																
Static-99R (5-yr)	12	.62 [.58, .66]	0.45	.63 [.59, .66]	0.45	1,996	10.01	0.00	12	.70 <sup>+</sup> [.67, .73]	0.74	.70 <sup>+</sup> [.67, .72]	0.74	6,191	13.67	19.53
Static-99R (10-yr)	7	.65 [.58, .72]	0.57	.65 [.60, .71]	0.58	513	9.78	38.68	7	.71 [.68, .74]	0.80	.708 [.68, .74]	0.79	1,547	7.87	23.80
Static-2002R (5-yr)	4	.62 [.54, .71]	0.44	.63 [.55, .70]	0.45	281	3.27	8.19	4	.69 [.65, .74]	0.72	.69 [.65, .74]	0.72	708	0.07	0.00
Static-2002R (10-yr)	3	.63 [.54, .72]	0.47	.63 [.54, .72]	0.47	162	0.43	0.00	3	.77 <sup>+</sup> [.68, .86]	1.02	.74 <sup>+</sup> [.70, .78]	0.93	460	11.93*	83.24
Prior Meta-Analyses																
Babchishin et al. (2012)																
Static-99R	5	.71 [.64, .79]	0.80	.71 [.64, .79]	0.80	319	3.94	-	5	.74 [.70, .78]	0.89	.74 [.70, .78]	0.89	1,269	3.55	-
Static-2002R	3	.61 [.50, .72]	0.39	.61 [.50, .72]	0.39	209	0.39	-	3	.76 <sup>+</sup> [.71, .81]	0.99	.76 <sup>+</sup> [.71, .81]	0.99	955	1.46	-
Lee et al. (2020)																
Static-99R (5-yr)	5	-	-	.61 [.55, .67]	0.40	599	NS	-	5	-	-	.70 [.66, .74]	0.74	1,375	NS	-
Static-99R (10-yr)	3	-	-	.66 [.60, .72]	0.59	375	NS	-	3	-	-	.72 [.68, .76]	0.82	848	NS	-
Static-2002R (5-yr)	3	-	-	.56 [.45, .67]	0.21	173	NS	-	3	-	-	.69 [.64, .75]	0.71	620	NS	-
Static-2002R (10-yr)	2	-	-	.61 [.49, .74]	0.40	101	NS	-	2	-	-	.72 [.67, .78]	0.82	416	NS	-

---

Ahmed et al. (2023)																	
Static-99R	10	.64	0.49	.64	0.49	2043	4.3	0.00	9	.71†	0.76	.68†	0.67	13,756	31.2***	68.0	
		[.60, .67]		[.60, .67]						[.66, .75]		[.66, .70]					

---

Note: Study samples are all adult males. † non-Indigenous/White group effect size significantly higher than Indigenous group effect size; NS = not significant variability but Q statistic not reported; d = standardized mean difference (0.20 small, 0.50 medium, 0.80 large); AUC = area under the curve (.56 small, .64 medium, .71 large). \* Q statistic significant < .05; I<sup>2</sup> measure of effect size heterogeneity (25% low, 50% medium, 75% high). Country: Canada (C), Australia (A), USA (U). Comparisons in all Canadian and USA samples are between Indigenous and White groups; Australian samples employ a non-Indigenous aggregate.

Table 14

*Study Details for Meta-Analysis of Psychopathy Checklist Variants in the Prediction of Recidivism for Indigenous and Non-Indigenous Samples*

PCL measure and study sample	Sex	RC	Indigenous					Non-Indigenous				
			n	BR	r	d	AUC	n	BR	r	d	AUC
PCL-R												
Olver Neumann et al. (2018) <sup>†</sup>	M	G	452	.619	-	0.83	.72	711	.519	-	0.86	.73
		V	452	.460	-	0.80	.71	711	.319	-	0.86	.73
Olver Neumann et al. (2013) <sup>a</sup>	M	G	171	.292	-	0.46	.63	264	.182	-	0.45	.63
		V	171	.122	-	0.53	.65	264	.098	-	0.58	.66
Olver & Wong (2015) <sup>a</sup>	M	G	86	.779	-	0.33	.59	187	.513	-	0.55	.65
		V	86	.453	-	0.56	.65	187	.241	-	0.30	.58
Olver, Lewis, & Wong (2013) <sup>a</sup>	M	G	86	.698	-	0.40	.61	66	.530	-	0.44	.62
		V	86	.523	-	0.28	.58	66	.424	-	0.87	.73
Sewall (2015; Sewall & Olver, 2019) <sup>a</sup>	M	G	109	.835	-	0.62	.67	196	.673	-	0.52	.64
		V	109	.725	-	0.70	.69	196	.474	-	0.71	.69
Olver & Riemer (2021)	M	G	185	.676	-	0.80	.71	236	.419	-	0.52	.64
		V	185	.465	-	0.62	.67	236	.237	-	0.75	.70
PCL: YV												
Schmidt et al. (2006)	M/F	G	39	.539	.44	0.98	-	88	.421	.32	0.68	-
		V	39	.103	.52	1.22	-	88	.239	.21	0.44	-
Stockdale et al. (2010)	M/F	G	101	.802	-	0.82	.72	39	.436	-	0.20	.56
		V	101	.554	-	0.79	.71	39	.205	-	1.09	.78
Shepherd & Strand (2016)	M	G	32	.862	-	0.75	.70	84	.725	-	0.82	.72
		V	32	.690	-	0.59	.66	84	.580	-	1.00	.76
McCuish et al. (2017) <sup>b</sup>	M/F	G	137	.822	-	0.50	.64	312	.706	-	0.35	.60
		V	137	.504	-	0.38	.61	312	.387	-	0.14	.54
Lovatt (2020)	M/F	G	171	.608	-	0.47	.63	55	.454	-	0.88	.73
		V	171	.409	-	0.38	.60	55	.218	-	0.28	.57

Note: PCL-R: Psychopathy Checklist-Revised; PCL: YV: Psychopathy Checklist: Youth Version; † Not included in meta-analysis; aggregate sample of individual subsamples included in PCL-R meta-analysis. <sup>a</sup> = individual samples from Olver, Neumann, et al. (2018) included in PCL-R meta-analysis. <sup>b</sup> = as McCuish et al. (2017) reported both charge and conviction, the averaged base rate and effect size are reported and used for meta-analysis Cell n do not add to PCL-R aggregate total since a small number of cases that were included in multiple samples were removed so that Olver, Neumann et al. (2018) had no duplicate cases. RC = recidivism criterion violent (v) or general (g); BR = base rate for a given recidivism criterion; d = standardized mean difference (0.20 small, 0.50 medium, 0.80 large); AUC = area under the curve (.56 small, .64 medium, .71 large).

Table 15

*Meta-Analysis of Psychopathy Checklist (PCL) Variants in the Prediction of Recidivism for Indigenous and Non-Indigenous Samples*

PCL measure	Indigenous									Non-Indigenous							
	Random			Fixed			n	Q	I <sup>2</sup>	Random			Fixed		n	Q	I <sup>2</sup>
	k	d [95%CI]	AUC	d [95%CI]	AUC	k				d [95%CI]	AUC	d [95%CI]	AUC				
General recidivism																	
PCL-R	5	0.56 [0.38, 0.74]	.65	0.56 [0.38, 0.74]	.65	637	3.88	0.00	4	0.51 [0.37, 0.65]	.64	0.51 [0.37, 0.65]	.64	949	0.31	0.00	
PCL: YV	5	0.61 [0.39, 0.83]	.67	0.61 [0.39, 0.83]	.67	480	2.85	0.00	5	0.56 [0.32, 0.81]	.65	0.54 [0.34, 0.73]	.64	578	5.67	29.82	
PCL-R/YV	10	.58 [0.44, 0.72]	.66	.58 [0.44, 0.72]	.66	1,117	6.84	0.00	10	0.52 [0.40, 0.63]	.64	0.52 [0.40, 0.63]	.64	1,527	6.06	0.00	
Violent recidivism																	
PCL-R	5	0.55 [0.38, 0.73]	.65	0.55 [0.38, 0.73]	.65	637	2.26	0.00	4	0.63 [0.44, 0.83]	.67	0.63 [0.47, 0.79]	.67	949	5.62	28.86	
PCL: YV	5	0.60 [0.32, 0.87]	.66	0.55 [0.36, 0.75]	.65	480	6.53	38.72	5	0.54 [0.13, 0.95]	.65	0.41 [0.21, 0.61]	.61	578	12.60*	68.26	
PCL-R/YV	10	0.55 [0.42, 0.68]	.65	0.55 [0.42, 0.68]	.65	1,117	8.78	0.00	10	0.58 [0.38, 0.79]	.66	0.55 [0.42, 0.67]	.65	1,527	21.15*	57.45	

Note: PCL-R: Psychopathy Checklist-Revised; PCL: YV: Psychopathy Checklist: Youth Version; d = standardized mean difference with 95% confidence intervals (95%CI) random or fixed effects meta-analytic effect size (0.20 small, 0.50 medium, 0.80 large); AUC = area under the curve (.56 small, .64 medium, .71 large). \* Q statistic significant < .05; I<sup>2</sup> measure of effect size heterogeneity (25% low, 50% medium, 75% high).

Table 16

*Study Details for All Other Risk Measures in the Prediction of Recidivism Outcomes for Indigenous and Non-Indigenous Samples*

Measure and study sample	Sex	RC	Indigenous				Non-Indigenous			
			n	BR	d	AUC/C	n	BR	d	AUC/C
SFA										
Helmus & Forrester (2014b)	M	G	1,500	.359	0.10	.53	6,684	.252	0.35	.60
	M	V	1,500	.166	0.16	.54	6,684	.106	0.44	.62
	F	G	149	.222	0.42	.62	377	.103	0.69	.69
	F	V	149	.067	0.68	.68	377	.037	0.94	.75
DFIA-R										
Brown & Motiuk (2005) [SPJ]	M/W	G	2,593	.561	0.25	.57	13,651	.407	0.58 <sup>a</sup>	.66 <sup>a</sup>
Stewart et al. (2017) [SPJ]	M	G	3,710	.529	0.21	.55	12,590	.338	0.43	.61
Wilton et al. (2017) [mech.]	M	G	4,640	.556	0.43	.62	15,487	.359	0.67	.68
ORAMS PRA										
O'Byrne (2004)	M/F	G	1,560	.692	0.75	.70	667	.402	0.91	.74
ARNA										
LeCroy et al. (1998)/Schwalbe (2009)	M/F	G	448	.480	0.96	.75	5,225	.517	1.06	.77
Schwalbe (2009a)	M/F	G	545	.330	0.41	.61	4,464	.330	0.55	.65
Schwalbe (2009b)	M/F	G	494	.321	0.59	.66	4,402	.321	0.58	.66
Schwalbe (2009c)	M/F	G	537	.305	0.47	.63	4,079	.305	0.62	.67
SPIn										
Jones et al. (2014)	M/F	G	635	.332	0.95	.75	3,021	.171	1.00	.76
YASI										
Jones et al. (2016)	M/F	G	120	.292	1.20	.80	344	.180	1.04	.77
	M/F	V	120	.125	1.15	.79	344	.087	1.20	.80
Matz et al. (2021)	M/F	G	20	.700	0.67	.68	93	.484	0.86	.73
SAPROF										
Olver & Riemer (2021)	M	G	185	.676	0.69	.69	236	.419	0.46	.63
	M	V	185	.465	0.68	.67	236	.237	0.86	.73
SAPROF-YV										
Lovatt et al. (2022)	M/F	G	171	.608	0.60	.65	55	.454	0.98	.75
	M/F	V	171	.409	0.44	.63	55	.218	0.50	.66
ODARA										

Buchanan (2009)	M	IPV	53	.292	0.50	.64	41	.073	1.05	.77
Hegel et al. (2021)	M	IPV	278	.464	0.42	.62	22	.294	0.23	.57
	M	V	278	.615	0.47	.63	22	.409	0.68	.68
	M	G	278	.748	0.53	.65	22	.500	0.77	.71
SARA										
Ahmed (2022) [SPJ/mech.]	M	IPV	2,949	.191	0.34/0.37	.59/.60	6,105	.106	0.42/0.50	.62/.64
	M	V	2,949	.311	0.27/0.33	.57/.59	6,105	.177	0.38/0.50	.61/.64
	M	G	2,949	.391	0.27/0.30	.57/.59	6,105	.252	0.36/0.48	.60/.63
Hogan & Corabian (2023) [SPJ/mech.]	M	V	82	.549	0.10/0.19	.53/.55	108	.296	0.14/0.32	.54/.59
STABLE 2007										
Helmus et al. (2012)	M	S	88	.114	0.11	.53	509	.073	0.74	.70
	M	V	88	.318	0.11	.53	509	.126	0.80	.71
	M	G	88	.511	0.29	.58	509	.232	0.82	.72
Sowden & Olver (2017)	M	S	87	.218	0.11	.53	90	.189	0.61	.67
	M	V	87	.540	0.78	.71	90	.367	0.53	.64
	M	G	87	.690	0.62	.67	90	.522	0.27	.58
Helmus et al. (2021)	M	S	921	.064	0.33	.59	2,617	.045	0.72	.69
	M	V	921	.290	0.39	.61	2,617	.151	0.58	.66
	M	G	921	.402	0.52	.64	2,617	.204	0.67	.68

*Note:* SFA: Static Factors Assessment; DFIA-R: Dynamic Factors Identification and Analysis-Revised; ODARA: Ontario Domestic Assault Risk Assessment; SARA: Spousal Assault Risk Assessment Guide; ORAMS PRA: Offender Risk Assessment and Management System Primary Risk Assessment; SPIn: Service Planning Instrument; YASI: Youth Assessment and Screening Instrument. SAPROF: Structured Assessment of Protective Factors; SAPROF-YV: Structured Assessment of Protective Factors-Youth Version; RC = recidivism criterion sexual (s), violent (v), general (g), or intimate partner violence (IPV); BR = base rate for a given recidivism criterion; d = standardized mean difference (0.20 small, 0.50 medium, 0.80 large); AUC = area under the curve (.56 small, .64 medium, .71 large). For SAPROF and SAPROF-YV effect sizes, positive d and AUC values represent increased protective factors being associated with decreased recidivism. <sup>a</sup> Brown and Motiuk (2005) non-Indigenous effect size (d) based on a weighted male-female average for the aggregate sample, compared to effect size for male-female Indigenous persons, given that DFIA-R frequencies for ethnocultural groups were not provided to compute an effect for non-Indigenous men/women; the reported effect size, as a result, is a conservative estimate of the true effect size for the non-Indigenous sample.

Table 17

*Meta-Analysis of Other Risk Measures in the Prediction of Recidivism Outcomes for Indigenous and Non-Indigenous Samples*

Measure and RC	Indigenous									Non-Indigenous						
	Random			Fixed			n	Q	I <sup>2</sup>	Random			Fixed		n	Q
k	d [95%CI]	AUC	d [95%CI]	AUC	k	d [95%CI]				AUC	d [95%CI]	AUC				
<b>STABLE 2007</b>																
sexual	3	0.26 [0.04, 0.49]	.57	0.26 [0.04, 0.49]	.57	1,096	0.80	0.00	3	0.72 <sup>+</sup> [0.56, 0.87]	.69	0.72 <sup>+</sup> [0.56, 0.87]	.69	3,216	0.17	0.00
violent	3	0.42 [0.13, 0.71]	.62	0.40 [0.27, 0.54]	.61	1,096	4.51	55.64	3	0.62 [0.49, 0.75]	.67	0.61 <sup>+</sup> [0.51, 0.70]	.67	3,216	2.38	15.84
general	3	0.51 [0.39, 0.63]	.64	0.51 [0.39, 0.63]	.64	1,096	1.26	0.00	3	0.65 [0.45, 0.86]	.68	0.68 <sup>+</sup> [0.59, 0.76]	.68	3,216	5.53	63.83
<b>DFIA-R</b>																
general	2	0.22 [0.18, 0.26]	.56	0.22 [0.18, 0.26]	.56	6,303	0.76	0.00	2	0.51 <sup>+</sup> [0.36, 0.62]	.64	0.51 <sup>+</sup> [0.49, 0.54]	.64	26,241	33.37*	97.00
<b>SFA</b>																
general	2	0.20 [-0.09, 0.49]	.56	0.12 [0.02, 0.22]	.54	1,649	2.38	58.09	2	0.48 [0.16, 0.80]	.63	0.36 <sup>+</sup> [0.26, 0.36]	.60	7,061	3.85	74.03
violent	2	0.32 [-0.15, 0.79]	.59	.18 [0.05, 0.32]]	.55	1,649	2.35	57.43	2	0.62 [0.06, 0.15]	.67	0.45 <sup>+</sup> [0.37, 0.53]	.63	7,061	3.24	69.11
<b>SPIIn/YASI</b>																
general	3	0.98 [0.82, 1.14]	.76	0.98 [0.82, 1.14]	.76	775	1.49	0.00	3	1.00 [0.91, 1.09]	.76	1.00 [0.91, 1.09]	.76	3,458	0.48	0.00
<b>ARNA</b>																
general	4	0.61 [0.37, 0.85]	.67	0.60 [0.50, 0.69]	.66	2,024	19.18*	84.36	4	0.70 [0.45, 0.95]	.69	0.72 <sup>+</sup> [0.69, 0.76]	.70	18,170	180.97*	98.34
<b>SAPROF/YV</b>																
violent	2	0.56 [0.31, 0.80]	.65	0.56 [0.34, 0.77]	.65	356	1.31	23.60	2	0.77 [0.46, 1.08]	.71	0.79 [0.51, 1.06]	.71	291	1.15	13.23
general	2	0.64 [0.42, 0.87]	.68	0.64 [0.42, 0.87]	.68	356	0.13	0.00	2	0.66 [0.16, 1.16]	.68	0.55 [0.31, 0.79]	.65	291	2.77	63.88
<b>ODARA</b>																
IPV	2	0.43 [0.21, 0.65]	.62	0.43 [.207, .653]	.62	331	0.06	0.00	2	0.58 [-0.22, 1.37]	.66	0.58 [-.22, 1.37]	.66	63	0.99	0.00

SARA																
violent (SPJ)	2	0.27	.58	0.27	.58	3,031	0.55	0.00	2	0.37	.60	0.37	.60	6,213	1.24	19.23
		[0.19, 0.34]		[0.19, 0.34]						[0.31, 0.44]		[0.20, 0.50]				
violent (mech.)	2	0.33	.59	0.33	.59	3,031	0.38	0.00	2	0.50 <sup>†</sup>	.64	0.50 <sup>†</sup>	.64	6,213	0.69	0.00
		[0.25, 0.40]		[0.25, 0.40]						[0.43, 0.56]		[0.43, 0.56]				

*Note:* ODARA: Ontario Domestic Assault Risk Assessment; SPIn: Service Planning Instrument; YASI: Youth Assessment and Screening Instrument; ARNA: Arizona Risk/Needs Assessment; SAPROF/YV: meta-analytic effect for Structured Assessment of Protective Factors (SAPROF) and SAPROF-Youth Version; SFA: Static Factors Assessment; DFIA-R: Dynamic Factors Identification and Analysis-Revised; SARA: Spousal Assault Risk Assessment guide (SPJ/Mech.: Effect sizes reported for SARA SPJ vs. mechanical summed ratings). † non-Indigenous/White group ES significantly higher than Indigenous group ES; RC = recidivism criterion sexual, violent, general, or intimate partner violence (IPV); d = standardized mean difference with 95% confidence intervals (95%CI) random or fixed effects meta-analytic effect size (0.20 small, 0.50 medium, 0.80 large); AUC = area under the curve (.56 small, .64 medium, .71 large). \* Q statistic significant < .05; I<sup>2</sup> measure of effect size heterogeneity (25% low, 50% medium, 75% high).

Table 18

Summary Table of Meta-Analytic Effect Sizes for Risk Measures in the Prediction of General Recidivism for Indigenous and non-Indigenous/White (NIW) Groups

Measure	k	n		<i>d</i> (FE/RE)		AUC (FE/RE)		Overall ES magnitude		Significant ES diff?	FSN <i>d</i> < 0.20 Ind/NIW
		Indigenous	NIW	Indigenous	NIW	Indigenous	NIW	Indigenous	NIW		
General recidivism											
LS (all)	22/21	23,044	92,274	0.53/0.57	0.69/0.75	.65/.66	.69/.70	Medium	Medium	Yes (both)	37/52
LS (male)	6	15,383	72,031	0.54/0.62	0.66/0.77	.65/.67	.68/.71	Medium	Medium	Yes (FE)	11/14
LS (female)	6	4094	13,600	0.49/0.62	0.90/1.00	.63/.67	.74/.76	Small-Medium	Large	Yes (both)	9/18
LS (adult)	11/10	20,427	89,735	0.53/0.58	0.69/0.82	.65/.66	.69/.72	Medium	Medium-Large	Yes (both)	19/28
LS (youth)	11	2,617	2,539	0.53/0.56	0.57/0.63	.65/.65	.66/.67	Medium	Medium	No	19/21
SIR	8/6	5,448	2,4810	0.64/0.69	0.83/0.88	.68/.69	.72/.73	Medium	Large	Yes (FE)	18/19
VRS total	4	393	486	0.52/0.52	0.60/0.59	.64/.64	.66/.66	Medium	Medium	No	7/9
VRS change	3	220	214	-0.53/-0.53	-0.31/-0.31	.65/.65	.59/.59	Medium	Small	No	5/2
VRS-SO total	3	384	660	0.77/0.77	0.62/0.60	.71/.71	.67/.66	Medium	Medium	No	9/7
VRS-SO change	3	384	660	-0.11/-0.11	-0.18/-0.18	.53/.53	.55/.55	Subthreshold	Subthreshold	No	-
VRS-YV	2	262	92	0.54/0.60	0.74/0.72	.65/.66	.70/.69	Medium	Medium	No	-
SAVRY total	4	502	584	0.68/0.72	1.07/1.07	.68/.70	.78/.78	Medium	Large	Yes (both)	10/18
SAVRY protect	4	502	584	0.43/0.49	0.79/0.80	.62/.64	.79/.80	Small	Large	Yes (FE)	5/12
SFA	2	1,649	7,061	0.12/0.20	0.36/0.48	.54/.56	.60/.63	Subthrsh-Small	Small	Yes (FE)	-
DFIA-R	2	6,303	26,241	0.22/0.22	0.51/0.51	.56/.56	.64/.64	Small	Medium	Yes	-
ARNA	4	2,024	18,170	0.60/0.61	0.72/0.70	.66/.67	.70/.69	Medium	Medium	Yes (FE)	8/11
ORAMS	1	1,560	667	0.75	0.91	.70	.74	Medium	Large	No	-
SPIn/YASI	3	775	3,458	0.98/0.98	1.00/1.00	.76/.76	.76/.76	Large	Large	No	12/12
SAPROF/YV	2	356	291	0.64/0.64	0.55/0.66	.68/.68	.65/.68	Medium	Medium	No	-
SARA (SPJ)	1	2,949	6,105	0.27	0.36	.57	.60	Small	Small	Yes	-
SARA (mech.)	1	2,949	6,105	0.30	0.48	.59	.63	Small	Small	Yes	-
STABLE 2007	3	1,096	3,216	0.51/0.51	0.68/0.65	.64/.64	.68/.68	Medium	Medium	Yes (FE)	5/8
PCL-R	5	637	949	0.56/0.56	0.51/0.51	.65/.65	.64/.64	Medium	Medium	No	10/8
PCL: YV	5	480	578	0.61/0.61	0.56/0.54	.67/.67	.65/.65	Medium	Medium	No	11/9

Note: LS: Level of Service measures; SIR: Statistical Information on Recidivism scale; VRS: Violence Risk Scale; VRS-SO: Violence Risk Scale-Sexual Offense version; VRS-YV: Violence Risk Scale-Youth Version; SAVRY: Structured Assessment of Violence Risk for Youth; SFA: Static Factors Assessment; DFIA-R: Dynamic Factors Identification and Analysis-Revised; ARNA: Arizona Risk/Needs Assessment instrument; ORAMS: Offender

Risk Assessment Management System; SPIn/YASI: combined effect for the Service Planning Instrument and the Youth Assessment Screening Inventory; SAPROF/YV: combined effect for Structured Assessment of Protective Factors and its Youth Version; ODARA: Ontario Domestic Assault Risk Assessment; SARA: Spousal Assault Risk Assessment guide (SPJ/Mech.: Effect sizes reported for SARA SPJ vs. mechanical summed ratings); PCL-R: Psychopathy Checklist-Revised; PCL: YV: Psychopathy Checklist: Youth Version;  $k$  = number of samples generating an effect [for unequal  $k$ , Indigenous  $k$ /non-Indigenous White  $k$ ];  $d_{(FE/RE)}$ : respective meta-analytic fixed and random effects for standardized mean difference ( $d$ ) where 0.20 small, 0.50 medium, 0.80 large;  $AUC_{(FE/RE)}$ : respective AUC equivalent to  $d$  for fixed and random effects obtained from Rice and Harris (2005) conversion tables where .56 small, .64 medium, .71 large; single effect reported only for  $k = 1$ ; FSN  $d < 0.20$ : Orwin's fail-safe N, number of studies with a mean  $d = 0.00$  required to reduce the observed fixed effect to below the threshold of a small effect (i.e.,  $d < 0.20$ ) for Indigenous and non-Indigenous/White groups, respectively (required minimum  $k = 3$ ).

Table 19

Summary Table of Meta-Analytic Effect Sizes for Risk Measures in the Prediction of Violent and Sexual Recidivism for Indigenous and non-Indigenous/White (NIW) Groups

Measure	k	n		d (FE/RE)		AUC (FE/RE)		Overall ES magnitude		Significant ES diff?	FSN d < 0.20 Ind/NIW
		Indigenous	NIW	Indigenous	NIW	Indigenous	NIW	Indigenous	NIW		
Violent recidivism											
LS (all)	6	2,320	17,405	0.52/0.53	0.84/0.66	.64/.65	.72/.68	Medium	Medium-Large	Yes (FE)	10/20
LS (female)	3/2	726	4,197	0.53/0.53	0.62/0.56	.65/.65	.67/.65	Medium	Medium	No	5/-
SIR	6	5,170	24,810	0.43/0.47	0.62/0.73	.62/.63	.67/.70	Small	Medium	Yes (both)	7/13
VRS total	4	393	486	0.55/0.55	0.55/0.54	.65/.65	.65/.65	Medium	Medium	No	7/7
VRS change	3	220	214	-0.56/-0.56	-0.28/-0.29	.65/.65	.58/.58	Medium	Small	No	6/2
VRS-SO total	3	384	660	0.64/0.64	0.81/0.80	.68/.68	.72/.72	Medium	Large	No	7/10
VRS-SO change	3	384	660	-0.27/-0.27	-0.28/-0.28	.58/.58	.58/.58	Small	Small	No	2/2
VRS-YV	2	262	92	0.63/0.67	0.80/0.87	.67/.68	.71/.73	Medium	Large	No	-
SAVRY total	5	540	667	0.65/0.69	0.71/0.71	.67/.69	.69/.69	Medium	Medium	No	12/13
SAVRY protect	4	502	584	0.34/0.34	0.59/0.59	.60/.60	.66/.66	Small	Medium	No	3/8
SAPROF/YV	2	356	291	0.56/0.56	0.79/0.77	.65/.65	.71/.71	Medium	Medium-Large	No	-
SFA	2	1,649	7,061	0.18/0.32	0.45/0.62	.55/.59	.63/.67	Subthrsh-Small	Small-Medium	Yes (FE)	-
ODARA (IPV)	2	331	63	0.43/0.43	0.58/0.58	.62/.62	.66/.66	Small	Medium	No	-
SARA (IPV; SPJ)	1	2,949	6,105	0.34	0.42	.59	.62	Small	Small	Yes	-
SARA (IPV; mech.)	1	2,949	6,105	0.37	0.50	.60	.64	Small	Medium	Yes	-
SARA (v; SPJ)	2	3,031	6,213	0.27/0.27	0.37/0.37	.58/.58	.60/.60	Small	Small	No	-
SARA (v; mech.)	2	3,031	6,213	0.33/0.33	0.50/0.50	.59/.59	.64/.64	Small	Medium	Yes (both)	-
STABLE 2007	3	1,096	3,216	0.40/0.42	0.61/0.62	.61/.62	.67/.67	Small	Medium	Yes (FE)	4/7
PCL-R	5	637	949	0.55/0.55	0.63/0.63	.65/.65	.67/.67	Medium	Medium	No	9/11
PCL: YV	5	480	578	0.55/0.60	0.41/0.54	.65/.66	.61/.65	Medium	Small-Medium	No	9/6
Sexual recidivism											
Static-99R (5yr)	12	1,996	6,191	0.45/0.45	0.74/0.74	.63/.62	.70/.70	Small	Medium	Yes	17/32
(10yr)	7	513	1,547	0.58/0.57	0.79/0.80	.65/.65	.71/.71	Medium	Large	No	10/20
Static-2002R (5yr)	4	281	708	0.45/0.44	0.72/0.72	.63/.62	.69/.69	Small	Medium	No	5/11
(10yr)	3	162	460	0.47/0.47	0.93/1.02	.63/.63	.74/.77	Small	Large	Yes	5/11
VRS-SO (total)	3	384	660	0.61/0.61	0.98/0.97	.67/.67	.76/.75	Medium	Large	No	7/12
VRS-SO (change)	3	384	660	-0.30/-0.30	-0.37/-0.37	.58/.58	.60/.60	Small	Small	No	2/3
STABLE 2007	3	1,096	3,216	0.26/0.26	0.72/0.72	.57/.57	.69/.69	Small	Medium	Yes (both)	1/8

*Note:* LS: Level of Service measures; SIR: Statistical Information on Recidivism scale; VRS: Violence Risk Scale; VRS-SO: Violence Risk Scale-Sexual Offense version; VRS-YV: Violence Risk Scale-Youth Version; SAVRY: Structured Assessment of Violence Risk for Youth; SAPROF/YV: combined effect for Structured Assessment of Protective Factors and its Youth Version; SFA: Static Factors Assessment; ODARA: Ontario Domestic Assault Risk Assessment; SARA: Spousal Assault Risk Assessment guide; Effect sizes reported for SARA as SPJ vs. mech. (i.e., mechanical summed ratings); For SARA and ODARA, (IPV) means the recidivism criterion variable is intimate partner violence, otherwise, (v) represents SARA effect sizes for general violence consistent with all other measures; PCL-R: Psychopathy Checklist-Revised; PCL: YV: Psychopathy Checklist: Youth Version;  $d_{(FE/RE)}$ : respective meta-analytic fixed and random effects for standardized mean difference ( $d$ ) where 0.20 small, 0.50 medium, 0.80 large;  $AUC_{(FE/RE)}$ : respective AUC equivalent to  $d$  for fixed and random effects obtained from Rice and Harris (2005) conversion tables where .56 small, .64 medium, .71 large; FSN  $d < 0.20$ : Orwin's fail-safe  $N$ , number of studies with a mean  $d = 0.00$  required to reduce the observed fixed effect below the threshold of a small effect (i.e.,  $d < 0.20$ ) for Indigenous and non-Indigenous/White groups, respectively (required minimum  $k = 3$ ).

Table 20

*Empirically Supported Risk-Need factors and their Composites for Recidivism for Indigenous and Non-Indigenous Groups*

Risk/need domain	Study	RC	ESM	Indigenous				Non-Indigenous			
				<i>k</i>	<i>n</i>	<i>ES</i>	<i>d</i>	<i>k</i>	<i>n</i>	<i>ES</i>	<i>d</i>
Criminal history (general)	Babchishin et al. (2012)	S	AUC	3	209	.54	0.14	3	955	.65	0.54
	Gutierrez et al. (2013) <sup>a</sup>	G	<i>d</i>	24	30,615	0.56	0.56	17	NR	0.80	0.80
		V	<i>d</i>	7	4,514	0.45	0.45	5	NR	0.56	0.56
	Wilson & Gutierrez (2014) <sup>a</sup>	G	<i>d</i>	11	20,294	0.57	0.57	9	NR	0.88	0.88
		V	<i>d</i>	NR	NR	0.52	0.52	3	NR	NR	NR
	Helmus & Forrester (2014b) <sup>d</sup>	G	AUC	2	1,649	.66	0.58	2	7,061	.72	0.82
		V	AUC	2	1,649	.64	0.50	2	7,061	.71	0.80
	Lee et al. (2020)	S	AUC	3	182	.59	0.32	3	641	.61	0.40
	<i>d</i> <sub>med</sub> / <i>d</i> <sub>uwm</sub>	G					0.57/0.57			0.82/0.83	
	<i>d</i> <sub>med</sub> / <i>d</i> <sub>uwm</sub>	V					0.50/0.49			0.68/0.68	
<i>d</i> <sub>med</sub> / <i>d</i> <sub>uwm</sub>	S					0.23/0.23			0.47/0.47		
Antisocial personality pattern/Psychopathy	Gutierrez et al. (2013) <sup>a</sup>	G	<i>d</i>	10	4,840	0.57	0.57	4	NR	0.81	0.81
		V	<i>d</i>	6	4,442	0.36	0.36	3	NR	0.39	0.39
	Wilson & Gutierrez (2014) <sup>a</sup>	G	<i>d</i>	8	6,335	0.60	0.60	5	NR	0.76	0.76
		V	<i>d</i>	NR	NR	**	**	3	NR	NR	NR
	Olver, Neumann et al. (2018)	G	AUC	1	452	.72	0.83	1	711	.73	0.86
		V	AUC	1	452	.71	0.80	1	711	.73	0.86
	Olver, Kingston, & Sowden (2021) <sup>b</sup>	G	AUC	1	385	.73	0.86	1	665	.71	0.80
		V	AUC	1	385	.73	0.88	1	665	.75	0.93
		S	AUC	1	385	.64	0.50	1	665	.70	0.74
	Helmus et al. (2012, 2021) <sup>c</sup>	G	AUC	2	1,003	.61	0.38	2	3,118	.64	0.52
		V	AUC	2	1,003	.58	0.28	2	3,118	.64	0.50
		S	AUC	2	1,003	.55	0.17	2	3,118	.61	0.40
	<i>d</i> <sub>med</sub> / <i>d</i> <sub>uwm</sub>	G					0.60/0.65			0.80/0.75	
<i>d</i> <sub>med</sub> / <i>d</i> <sub>uwm</sub>	V					0.58/0.58			0.68/0.67		
<i>d</i> <sub>med</sub> / <i>d</i> <sub>uwm</sub>	S					0.34/0.34			0.57/0.57		
Education/ employment	Brown & Motiuk (2005)	G	<i>r</i>	1	2,635	.17	0.35	-	-	-	-
	Gobeil (2011)	G	<i>r</i>	1	733	.11	0.37	1	3,229	.11	0.37
	Barnum & Gobeil (2012)	G	<i>r</i>	1	90	-.17	-0.75	1	263	.35	1.96
	Gutierrez et al. (2013) <sup>a</sup>	G	<i>d</i>	19	25,243	0.38	0.38	14	NR	0.39	0.39
		V	<i>d</i>	5	4,352	0.33	0.33	3	NR	-0.13	-0.13
	Wilson & Gutierrez (2014) <sup>a</sup>	G	<i>d</i>	9	6,383	0.52	0.52	9	NR	0.60	0.60
		V	<i>d</i>	NR	NR	**	**	3	NR	NR	NR
	Stewart et al. (2017)	G	<i>e</i> <sup>b</sup>	1	3,785	3.06	0.87	1	12,909	3.81	1.04
	<i>d</i> <sub>med</sub> / <i>d</i> <sub>uwm</sub>	G					0.38/0.29			0.82/0.87	
<i>d</i> <sub>med</sub> / <i>d</i> <sub>uwm</sub>	V					0.33/0.33			-0.13/-0.13		
Substance abuse	Brown & Motiuk (2005)	G	<i>r</i>	1	2,635	.12	0.26	-	-	-	-
	Gobeil (2011)	G	<i>r</i>	1	733	.09	0.30	1	3,229	.17	0.58
	Barnum & Gobeil (2012)	G	<i>r</i>	1	90	.12	0.53	1	263	.38	2.15
	Gutierrez et al. (2013) <sup>a</sup>	G	<i>d</i>	20	25,153	0.33	0.33	15	NR	0.49	0.49
		V	<i>d</i>	5	4,353	0.23	0.23	3	NR	0.36	0.36

	Wilson & Gutierrez (2014) <sup>a</sup>	G	d	11	20,294	0.42	0.42	9	NR	0.54	0.54
		V	d	NR	NR	**	**	3	NR	NR	NR
	Stewart et al. (2017)	G	e <sup>B</sup>	1	3,785	2.45	0.70	1	12,909	3.34	0.94
	Olver, Kingston, & Sowden (2021) <sup>b</sup>	G	AUC	1	389	.63	0.45	1	669	.64	0.50
		V	AUC	1	389	.61	0.38	1	669	.58	0.29
		S	AUC	1	389	.62	0.44	1	669	.56	0.20
		$d_{med}/d_{uwm}$	G				0.42/0.43			0.56/0.87	
	$d_{med}/d_{uwm}$	V				0.34/0.34			0.33/0.33		
	$d_{med}/d_{uwm}$	S				0.44/0.44			0.20/0.20		
Companions	Brown & Motiuk (2005)	G	r	1	2,635	.21	0.43	-	-	-	-
	Gobeil (2011)	G	r	1	733	.15	0.51	1	3,229	.04	0.13
	Barnum & Gobeil (2012)	G	r	1	90	.24	1.05	1	263	.19	1.01
	Gutierrez et al. (2013) <sup>a</sup>	G	d	16	24,397	0.39	0.39	12	NR	0.47	0.47
		V	d	4	4,307	0.23	0.23	-	NR	-	-
	Wilson & Gutierrez (2014) <sup>a</sup>	G	d	11	20,294	0.39	0.39	9	NR	0.52	0.52
		V	d	NR	NR	**	**	3	NR	NR	NR
	Stewart et al. (2017)	G	e <sup>B</sup>	1	3,785	1.62	0.38	1	12,909	2.28	0.64
	Helmus et al. (2012, 2021) <sup>c</sup>	G	AUC	2	1,005	.60	0.34	2	3,114	.64	0.50
		V	AUC	2	1,005	.57	0.24	2	3,114	.64	0.50
S		AUC	2	1,005	.50	0.00	2	3,114	.57	0.24	
	$d_{med}/d_{uwm}$	G				0.39/0.50			0.51/0.55		
	$d_{med}/d_{uwm}$	V				0.24/0.24			0.50/0.50		
	$d_{med}/d_{uwm}$	S				0.00/0.00			0.24/0.24		
Procriminal attitudes	Brown & Motiuk (2005)	G	r	1	2,635	.10	0.21	-	-	-	-
	Gobeil (2011)	G	r	1	733	.19	0.65	1	3,229	.09	0.30
	Barnum & Gobeil (2012)	G	r	1	90	.00	0.00	1	263	-.04	0.17
	Gutierrez et al. (2013) <sup>a</sup>	G	d	15	23,803	0.32	0.32	12	NR	0.40	0.40
		V	d	5	4,395	0.33	0.33	3	NR	0.07	0.07
	Wilson & Gutierrez (2014) <sup>a</sup>	G	d	11	20,293	0.39	0.39	9	NR	0.48	0.48
		V	d	NR	NR	**	**	3	NR	NR	NR
	Stewart et al. (2017)	G	e <sup>B</sup>	1	3,785	1.60	0.37	1	12,909	2.53	0.72
	Olver, Kingston, & Sowden (2021) <sup>b</sup>	G	AUC	1	385	.50	0.00	1	661	.50	0.00
		V	AUC	1	385	.56	0.21	1	661	.57	0.23
		S	AUC	1	385	.53	0.02	1	661	.59	0.30
	Helmus et al. (2021)	G	AUC	1	425	.55	0.19	1	1,051	.56	0.20
		V	AUC	1	425	.55	0.18	1	1,051	.55	0.16
S		AUC	1	425	.55	0.18	1	1,051	.59	0.32	
	$d_{med}/d_{uwm}$	G				0.22/0.27			0.48/0.32		
	$d_{med}/d_{uwm}$	V				0.25/0.24			0.16/0.15		
	$d_{med}/d_{uwm}$	S				0.10/0.10			0.31/0.31		
Community functioning	Brown & Motiuk (2005)	G	r	1	2,635	.13	0.26	-	-	-	-
	Gobeil (2011)	G	r	1	733	.19	0.65	1	3,229	.12	0.41
	Barnum & Gobeil (2012)	G	r	1	90	.50	2.30	1	263	-.21	1.12
	Stewart et al. (2017)	G	e <sup>B</sup>	1	3,785	1.80	0.46	1	12,909	2.98	0.85
	Olver, Sowden et al. (2018) <sup>b</sup>	G	AUC	1	377	.71	0.80	1	651	.68	0.66
		V	AUC	1	377	.68	0.66	1	651	.72	0.81
		S	AUC	1	377	.62	0.42	1	651	.75	0.94

	$d_{med}/d_{uwm}$	G								0.65/0.89		0.76/0.76	
	$d_{med}/d_{uwm}$	V								0.66/0.66		0.81/0.81	
	$d_{med}/d_{uwm}$	S								0.42/0.42		0.94/0.94	
Leisure/recreation	Gutierrez et al. (2013)	G	d	11	18,657	0.33	0.33	8	-	0.34	0.34		
	Wilson & Gutierrez (2014)	G	d	11	20,287	0.34	0.34	8	-	0.37	0.37		
		V	d	-	-	**	**	3	-	-	-		
	$d_{med}/d_{uwm}$	G								0.34/0.34		0.36/0.36	
Family/marital	Brown & Motiuk (2005)	G	r	1	2,635	.06	0.10	-	-	-	-		
	Gobeil (2011)	G	r	1	733	-.09	-0.30	1	3,229	.04	0.13		
	Barnum & Gobeil (2012)	G	r	1	90	.04	0.17	1	263	.36	2.02		
	Gutierrez et al. (2013) <sup>a</sup>	G	d	26	30,330	.19	0.19	19	NR	.30	0.30		
		V	d	8	8,331	.11	0.11	8	NR	.06	0.06		
	Wilson & Gutierrez (2014) <sup>a</sup>	G	d	11	20,294	.24	0.24	9	NR	.28	0.28		
		V	d	NR	NR	.13	0.13	3	NR	NR	NR		
	Stewart et al. (2017)	G	e <sup>B</sup>	1	3,785	1.11	0.08	1	12,909	1.24	0.17		
	Olver, Kingston, & Sowden (2021) <sup>b</sup>	G	AUC	1	302	.56	0.20	1	395	.58	0.27		
		V	AUC	1	302	.56	0.20	1	395	.63	0.45		
		S	AUC	1	302	.52	0.07	1	395	.60	0.34		
	Helmus et al. (2012, 2021) <sup>c</sup>	G	AUC	2	1,008	.58	0.27	2	3,122	.60	0.34		
		V	AUC	2	1,008	.55	0.18	2	3,122	.57	0.27		
		S	AUC	2	1,008	.58	0.27	2	3,122	.60	0.34		
	$d_{med}/d_{uwm}$	G								0.22/.12		0.32/0.50	
	$d_{med}/d_{uwm}$	V								0.18/.16		0.27/0.26	
	$d_{med}/d_{uwm}$	S								0.17/.17		0.34/0.34	
Sexual deviance	Olver, Kingston, & Sowden (2021)	G	AUC	1	385	.54	0.14	1	660	.49	-0.03		
		V	AUC	1	385	.51	0.04	1	660	.53	0.11		
		S	AUC	1	385	.57	0.26	1	660	.61	0.40		
	Babchshin et al. (2012)	S	AUC	3	209	.56	0.20	3	955	.59	0.32		
	Lee et al. (2020)	S	AUC	3	182	.58	0.29	3	641	.65	0.54		
	Helmus et al. (2012, 2021) <sup>c</sup>	G	AUC	2	900	.56	0.20	2	2877	.55	0.17		
		V	AUC	2	900	.54	0.16	2	2877	.55	0.18		
		S	AUC	2	900	.55	0.18	2	2877	.62	0.42		
		$d_{med}/d_{uwm}$	G								0.10/0.10		0.07/.007
		$d_{med}/d_{uwm}$	V								0.10/0.10		0.14/0.14
	$d_{med}/d_{uwm}$	S								0.26/0.23		0.41/0.42	
General risk-need	Olver, Kingston, & Sowden (2021) <sup>b</sup>	G	AUC	1	384	.75	0.95	1	660	.72	0.83		
		V	AUC	1	384	.74	0.91	1	660	.74	0.92		
		S	AUC	1	384	.66	0.58	1	660	.71	0.80		
	Helmus et al. (2021)	G	AUC	1	900	.69	0.68	1	2544	.76	0.99		
		V	AUC	1	900	.66	0.57	1	2544	.74	0.92		
		S	AUC	1	900	.61	0.38	1	2544	.67	0.63		
		$d_{med}/d_{uwm}$	G								0.82/0.82		0.91/0.91
	$d_{med}/d_{uwm}$	V								0.74/0.74		0.92/0.92	
	$d_{med}/d_{uwm}$	S								0.48/0.48		0.72/0.72	
Personal/emotional	Brown & Motiuk (2005)	G	r	1	2,635	.07	0.14	-	-	-	-		
	Gobeil (2011)	G	r	1	733	.11	0.37	1	3,229	.12	0.40		
	Barnum & Gobeil (2012)	G	r	1	90	.10	0.40	1	263	.02	0.09		

	Stewart et al. (2017)	G	e <sup>B</sup>	1	3,785	1.58	0.36	1	12,909	1.92	0.51
	$d_{med}/d_{uwm}$	G					0.37/0.32			0.40/0.33	
Emotional distress	Gutierrez et al. (2013) <sup>a</sup>	G	d	10	9,574	.18	0.18	-	-	-	-
		V	d	4	3,793	.06	0.06	-	-	-	-
Prior victimization	Gutierrez et al. (2013) <sup>a</sup>	G	d	5	4,152	.23	0.23	-	-	-	-
		V	d	4	3,727	.03	0.03	-	-	-	-
Cultural engagement	Shepherd et al. (2018)	V	e <sup>B</sup>	1	84	0.80	-0.17	-	-	-	-
Cultural identity	Shepherd et al. (2018)	V	e <sup>B</sup>	1	84	0.97	-0.02	-	-	-	-
Cultural connectedness	Hansen (2018)	G	e <sup>B</sup>	1	166	0.89	-0.16	-	-	-	-

Note:  $d_{med}$ : Median  $d$  value computed by taking the midpoint of ranked effect sizes, in which single studies ( $k = 1$ ) are given a single unweighted mean and ranked relative to meta-analytic effects ( $k \geq 2$ ) for a given risk-need domain and recidivism criterion;  $d_{uwm}$ : Unweighted mean  $d$  value; RC = recidivism criterion, general (G), violent (V), and sexual (S); ESM: original effect size metric reported in the corresponding study; ES: the original effect size of corresponding to a given metric from that study;  $d$ : the corresponding Cohen's  $d$  for the original reported effect size where 0.20 is small, 0.50 is medium, and 0.80 is large (Rice & Harris, 2005). AUCs converted to  $d$  using tables from Rice and Harris (2005); Hazard ratios converted to a  $d$  equivalent based on Azuero (2016)  $d = \ln(HR) \times \sqrt{6}/\pi$

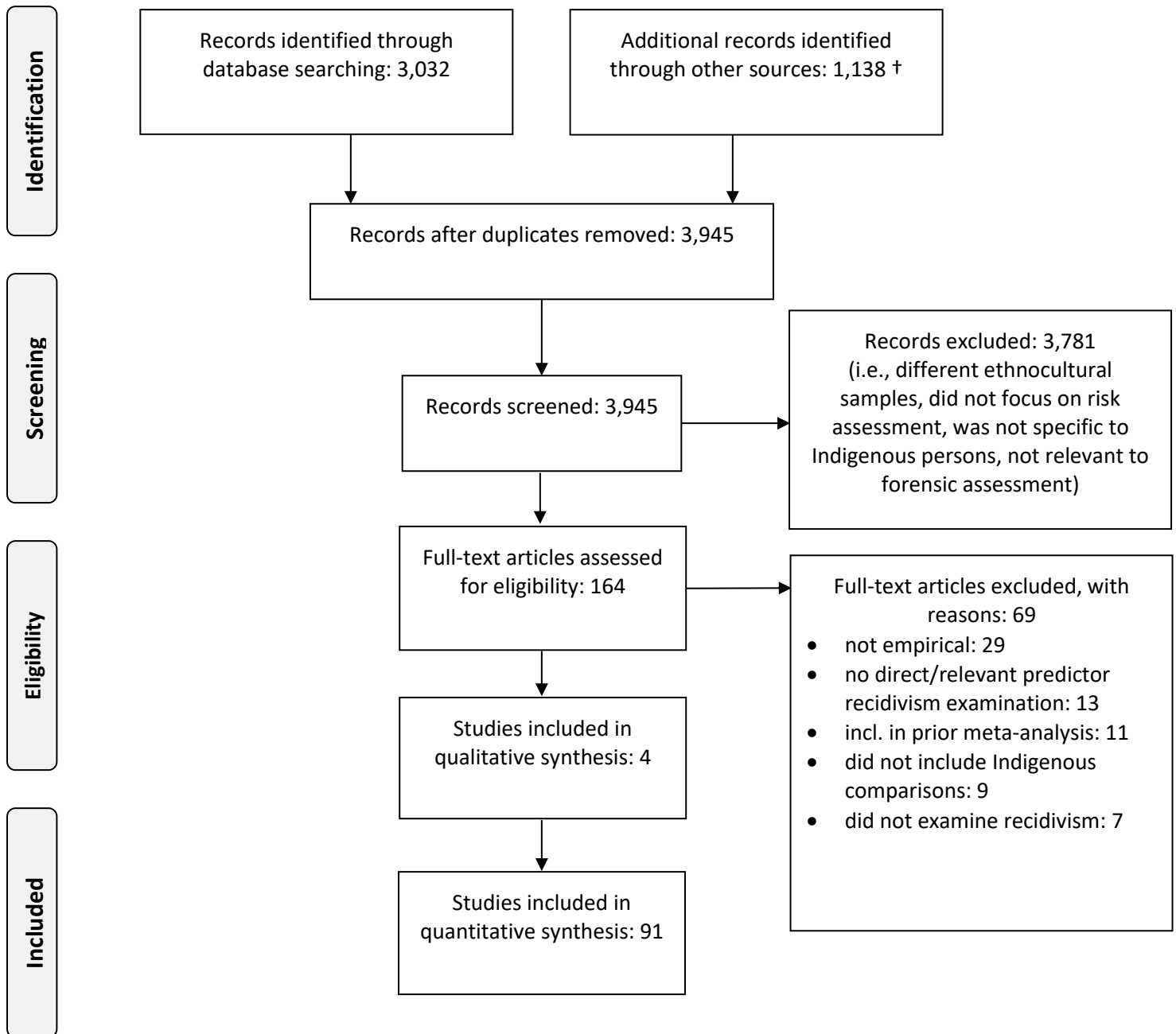
<sup>a</sup> \*\*significant effect reported but value not reported in original report; NR = not reported; Wilson and Gutierrez (2014) did not report most LS subscale effect sizes for violent recidivism for Indigenous samples and did not report any for non-Indigenous samples. For both Gutierrez et al. (2013) and Wilson and Gutierrez (2014), non-Indigenous effects were not reported for any LS subscales, but rather the mean diff was reported, which was added to the Indigenous sample effect size to generate an estimate for non-Indigenous samples (random effects reported).

<sup>b</sup> Violence Risk Scale Sexual Offense version (VRS-SO) Psychopathy composite (D4 criminal personality, D6 interpersonal aggression, D13 impulsivity, D14 compliance community supervision); D9 substance abuse; Cognition composite (D5 cognitive distortions and D8 insight); Community composite (D10 community support, D11 released to high risk situations, D14 compliance community supervision); D17 intimacy deficits.

<sup>c</sup> mean AUC taken from STABLE 2007 domain and converted to  $d$  and weighted by sample size; STABLE 2007 psychopathy composite includes lack of concern for others, impulsive acts, poor problem solving, negative emotionality, and cooperation with community supervision; family/marital composite includes rejection/loneliness and relationship stability; attitudes composite includes sexual entitlement, child molester attitudes, and rape attitudes from STABLE 2000.

<sup>d</sup> separate effects for male and female Indigenous and non-Indigenous samples combined through meta-analysis of AUC values to generate single effect.

Figure 1: Indigenous Risk Assessment Systematic Review and Meta-Analysis PRISMA Flow Diagram



† Includes reviewing reference sections of records culled from the database searches

Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine*, 6 (7): e1000097. <https://doi.org/10.1371/journal.pmed.1000097>

Supplemental Table S1

Detailed Summary Table of Studies Included in Systematic Review and Meta-Analyses

Study	Design	Source	Sample	Country	Measure	DG	Sex	ECG	Measure					
									M [SD]	N	RC	BR	ESM	ES
Ahmed (2022)	Field prospective	UTD	Correctional sample with IPV convictions on community supervision, administered the SARA	Canada	SARA-V2 (mechanical)	A	M	I	15.7 [6.8]	2949	G	.391	C	.59
							M	W	12.7 [6.6]	6105	G	.252	C	.63
							M	I	-	2949	V	.311	C	.59
							M	W	-	6105	V	.177	C	.64
							M	I	-	2949	IPV	.191	C	.60
					M	W	-	6105	IPV	.106	C	.64		
					SARA-V2 (SPJ)	A	M	I	-	2949	G	.391	C	.57
							M	W	-	6105	G	.252	C	.60
							M	I	-	2949	V	.311	C	.57
							M	W	-	6105	V	.177	C	.61
M	I	-	2949	IPV			.191	C	.59					
M	W	-	6105	IPV	.106	C	.62							
Ahmed et al. (2023)	Meta-analysis	JA	Convicted sexual offending samples from correctional/ forensic psychiatric settings (Indigenous $k = 10$ , $N = 2043$ ; White $k = 10$ $N = 13756$ )	Canada, Australia, USA	Static-99R	A	M	I	3.8 [2.2]	2043	S	.117	AUC	.64
							M	W	2.4 [4.0]	13756	S	.062	AUC	.68
Allan & Dawson (2002)	Archival retrospective	GR	Correctional sample referred for violence or sexual offense program	Australia	3-Predictor model	A	M	I	-	109	V	.358	AUC	.97
							M	NI	-	92	V	.413	AUC	1.0
Babchishin et al. (2012)	Meta-analysis	JA	Convicted sexual offending samples from correctional/ forensic psychiatric settings (Indigenous $k = 3-5$ , $N =$	Canada	Static-99/R	A	M	I	4.0 [2.3]	319	S	.160	AUC	.71
							M	NI	3.1 [2.7]	1269	S	.121	AUC	.74
					Static-2002/R	A	M	I	4.7 [2.3]	209	S	.160	AUC	.61
							M	NI	3.8 [2.7]	955	S	.121	AUC	.76
Sexual deviance	A	M	I	-	209	S	.160	AUC	.56					

			209-319; non-Indigenous k = 3-5 N = 955-1269)		Gen criminality	A	M	NI	-	955	S	.121	AUC	.59
										209	S	.160	AUC	.56
										955	S	.121	AUC	.65
Barnum & Gobeil (2012a)	Mixed automated and field ratings	GR	Federally sentenced adult male and female prison inmates	Canada	SIR Proxy	A	M	I	3.8 [1.4]	2560	G	.250	AUC	.63
					SIR-R1	A	M	NI	3.0 [1.6]	12485	G	.190	AUC	.69
					SIR Proxy	A	F	I		246	G	.159	AUC	.65
					SIR-R1	A	F	NI		666	G	.101	AUC	.74
					SIR Proxy	A	M	I		2560	V	.083	AUC	.57
					SIR-R1	A	M	NI		12485	V	.057	AUC	.66
					SIR Proxy	A	F	I		246	V	.041	AUC	.70
					SIR-R1	A	F	NI		666	V	.032	AUC	.82
					SIR Proxy	A	M	I		2560	S	.008	AUC	.62
					SIR-R1	A	M	NI		12485	S	.005	AUC	.57
Barnum & Gobeil (2012b)	Field prospective	GR	Federally sentenced adult female prison inmates (DFIAR ratings)	Canada	Empl/educ	A	F	I	-	90	G	.056	r	-.17
					Empl/educ	A	F	NI	-	263	G	.038	r	.35
					Companions	A	F	I	-	90	G	.056	r	.24
					Companions	A	F	NI	-	263	G	.038	r	.19
					Subst abuse	A	F	I	-	90	G	.056	r	.12
					Subst abuse	A	F	NI	-	263	G	.038	r	.38
					Attitudes	A	F	I	-	90	G	.056	r	.00
					Attitudes	A	F	NI	-	263	G	.038	r	-.04
					Comm func	A	F	I	-	90	G	.056	r	.50
					Comm func	A	F	NI	-	263	G	.038	r	-.21
					Fam/marital	A	F	I	-	90	G	.056	r	.04
					Fam/marital	A	F	NI	-	263	G	.038	r	.36
					Pers/emotional	A	F	I	-	90	G	.056	r	.10
					Pers/emotional	A	F	NI	-	263	G	.038	r	.02
Boer (2003; from Babchishin et al., 2012)	Archival retrospective	UTD	Federally sentenced men convicted for sex offenses	Canada	Static-99R	A	M	I	3.0 [2.5]	53	S	.075	AUC	.82
								W	2.7 [2.8]	227	S	.079	AUC	.79
Bonta (1989)	Field prospective	JA	Provincially incarcerated men, Ontario Corrections	Canada	LSI	A	M	I	24.8 [7.8]	48	G	.438	r	.35
							M	NI	23.8 [7.8]	71	G	.577	r	.50

Bonta et al. (1997)	Field prospective	JA	Adult male and female probationers (Manitoba)	Canada	MRNF	A	M	I	-	390	G	.659	r	.32
							M	NI	-	513	G	.478	r	.36
Bonta & Rugge (2004; from Rugge, 2006)	Field prospective	GR	Federally sentenced adult male prison inmates	Canada	SIR-R1	A	M	I	-	940	G	.567	AUC	.74
							M	NI	-	4475	G	.408	AUC	.77
							M	I	-	940	V	.205	AUC	.65
							M	NI	-	4475	V	.123	AUC	.65
Bonta & Yessine (2005; from Lee et al., 2020)	Field prospective	GR	Federally sentenced men convicted for sex offenses flagged as potential dangerous offender (DO)	Canada	Static-99R	A	M	I	5.1 [2.1]	26	S	.192	AUC	.60
								W	5.0 [2.2]	90	S	.144	AUC	.71
Brankley et al. (2017; from Lee et al., 2020)	Field prospective	UD	Men convicted for sex offenses under community supervision	Canada	Static-99R	A	M	I	4.7 [2.0]	35	S 5yr	.114	AUC	.62
							M	W	4.2 [2.4]	308	S 5yr	.114	AUC	.70
							M	I		35	S 10yr	.200	AUC	.60
							M	W		299	S 10yr	.217	AUC	.74
					Static-2002R	A	M	I	6.3 [2.1]	35	S 5yr	.114	AUC	.57
							M	W	5.9 [2.5]	308	S 5yr	.114	AUC	.70
							M	I		35	S 10yr	.200	AUC	.64
							M	W		299	S 10yr	.217	AUC	.74
Brews (2009)	Field prospective	UTD	Adult women, Ontario Provincial corrections.	Canada	LSI-OR	A	F	I	20.0 [8.5]	253	G	.283	AUC	.74
							F	W	14.8 [9.1]	1502	G	.283	AUC	.76
							F	I		253	V	.148	r	.15
							F	W		1502	V	.145	r	.12
Brown & Motiuk (2005)	Field prospective	GR	Federally sentenced adult male and female prison inmates	Canada	DFIAR SRR	A	M/F	I	-	2635	G	.561	φ	.13
					Empl/educ	A	M/F	I	-	2635	G	.561	φ	.17
					Companions	A	M/F	I	-	2635	G	.561	φ	.21
					Subst abuse	A	M/F	I	-	2635	G	.561	φ	.12
					Attitudes	A	M/F	I	-	2635	G	.561	φ	.10
					Comm func	A	M/F	I	-	2635	G	.561	φ	.13
					Fam/marital	A	M/F	I	-	2635	G	.561	φ	.06
					Pers/emotional	A	M/F	I	-	2635	G	.561	φ	.07
Buchanan (2009)		UTD		Canada	ODARA	A	M	I	6.6 [2.0]	53	IPV	.292	AUC	.64

	Archival retrospective		Provincially sentenced men convicted for IPV offenses		ODARA	A	M	NI	5.1 [2.1]	41	IPV	.073	AUC	.77
Coupland & Olver (2020)	Archival retrospective	JA	Treated high risk-need adult male violent offending correctional sample	Canada	VRS total	A	M	I	56.4 [6.1]	91	G	.857	d	.69
								NI	49.4 [13.2]	64	G	.688	d	1.11
								I		91	V	.703	d	.73
								NI		64	V	.469	d	.58
								NI		64	V	.469	d	.76
					VRS dynamic	A	M	I	42.5 [5.0]	91	G	.857	d	.51
								NI	38.4 [8.7]	64	G	.688	d	1.30
								I		91	V	.703	d	.56
								NI		64	V	.469	d	.76
								NI		64	V	.469	d	.76
VRS change	A	M	I	4.8 [3.2]	91	G	.857	d	-.64					
			NI	4.4 [2.6]	64	G	.688	d	-.60					
			I		91	V	.703	d	-.42					
			NI		64	V	.469	d	-.42					
			NI		64	V	.469	d	-.42					
Gobeil (2011)	Field prospective	GR	Federally sentenced adult male prison inmates (DFIAR ratings)	Canada	Empl/educ	A	M	I	-	733	G	.132	r	.11
					Empl/educ	A	M	NI	-	3,229	G	.094	r	.11
					Companions	A	M	I	-	733	G	.132	r	.15
					Companions	A	M	NI	-	3,229	G	.094	r	.04
					Subst abuse	A	M	I	-	733	G	.132	r	.09
					Subst abuse	A	M	NI	-	3,229	G	.094	r	.17
					Attitudes	A	M	I	-	733	G	.132	r	.19
					Attitudes	A	M	NI	-	3,229	G	.094	r	.09
					Comm func	A	M	I	-	733	G	.132	r	.19
					Comm func	A	M	NI	-	3,229	G	.094	r	.12
					Fam/marital	A	M	I	-	733	G	.132	r	-.09
					Fam/marital	A	M	NI	-	3,229	G	.094	r	.04
					Pers/emotional	A	M	I	-	733	G	.132	r	.11
					Pers/emotional	A	M	NI	-	3,229	G	.094	r	.12
Gossner & Wormith (2007)	Archival retrospective	JA	Court adjudicated youth	Canada	YLS/CMI	Y	M/F	I	18.2 [6.2]	62	G	.613	r	.35
								NI	11.6 [6.6]	32	G	.344	r	.24
Gutierrez et al. (2013)	Meta-analysis	JA	Adult and youth male and female samples from		Criminal history	A/Y	M/F	I	-	30615	G	.559	d	.56

community and correctional setting (Indigenous $k = 43$ , $N = 57,315$ ; non-Indigenous $k = 32$ , $N = 204,997$ )	Australia, Canada, USA	Criminal history	A/Y	M/F	NI	-	-	G	.384	d	.80
		Criminal history	A/Y	M/F	I	-	4514	V	.477	d	.52
		Criminal history	A/Y	M/F	NI	-	-	V	.316	d	.63
		Antisocial PP	A/Y	M/F	I	-	4840	G	.559	d	.57
		Antisocial PP	A/Y	M/F	NI	-	-	G	.384	d	.81
		Antisocial PP	A/Y	M/F	I	-	4442	V	.477	d	.36
		Antisocial PP	A/Y	M/F	NI	-	-	V	.316	d	.39
		Empl/educ	A/Y	M/F	I	-	25243	G	.559	d	.38
		Empl/educ	A/Y	M/F	NI	-	-	G	.384	d	.39
		Empl/educ	A/Y	M/F	I	-	4352	V	.477	d	.33
		Empl/educ	A/Y	M/F	NI	-	-	V	.316	d	-.13
		Companions	A/Y	M/F	I	-	24397	G	.559	d	.39
		Companions	A/Y	M/F	NI	-	-	G	.384	d	.47
		Companions	A/Y	M/F	I	-	4307	V	.477	d	.23
		Subst abuse	A/Y	M/F	I	-	25153	G	.559	d	.33
		Subst abuse	A/Y	M/F	NI	-	-	G	.384	d	.49
		Subst abuse	A/Y	M/F	I	-	4353	V	.477	d	.23
		Subst abuse	A/Y	M/F	NI	-	-	V	.316	d	.36
		Attitudes	A/Y	M/F	I	-	23803	G	.559	d	.32
		Attitudes	A/Y	M/F	NI	-	-	G	.384	d	.40
		Attitudes	A/Y	M/F	I	-	4395	V	.477	d	.33
		Attitudes	A/Y	M/F	NI	-	-	V	.316	d	.07
		Leisure/rec	A/Y	M/F	I	-	18657	G	.559	d	.33
		Leisure/rec	A/Y	M/F	NI	-	-	G	.384	d	.34
		Fam/marital	A/Y	M/F	I	-	30330	G	.559	d	.19
		Fam/marital	A/Y	M/F	NI	-	-	G	.384	d	.30
		Fam/marital	A/Y	M/F	I	-	8331	V	.477	d	.11
		Fam/marital	A/Y	M/F	NI	-	-	V	.316	d	.06
		Emot distress	A/Y	M/F	I	-	9574	G	.559	d	.18
		Emot distress	A/Y	M/F	I	-	3793	V	.477	d	.06
		Prior victimiz	A/Y	M/F	I	-	4152	G	.559	d	.23

					Prior victimiz	A/Y	M/F	I	-	3727	V	.477	d	.03		
Haag (2005; from Lee et al., 2020)	Archival retrospective	UTD	Federally sentenced men convicted for sexual offenses	Canada	Static-99R	A	M	I	4.2 [1.9]	50	S	.280	AUC	.60		
					Static-2002R	A	M	I	5.8 [1.9]	50	S	.280	AUC	.56		
Hann & Harman (1989)	Field prospective	GR	Federally sentenced men in custody	Canada	GSIR	A	M	I	-	49	G	.571	AUC	.72		
Hann & Harman (1993)	Field prospective	GR	Federally sentenced men in custody	Canada	GSIR	A	M	I	-	269	G	.651	AUC	.73		
Hansen (2018)	Field prospective	UTD	Released male/female correctional sample to tribal reservation	USA	LSI-R	A	M/F	I	28.7 [6.6]	166	G	.355	AUC	.66		
					Cultural Conn.	A	M/F	I	15.1 [3.5]	166	G	.355	e <sup>B</sup>	.89		
Hanson et al. (2015)	Field prospective	JA	Multisite nationwide sample of men convicted for sex offenses under community supervision	Canada	Static-99R	A	M	I	3.3 [2.2]	102	S	.176	AUC	.60		
							M	W	2.3 [2.4]	523	S	.111	AUC	.74		
Hegel et al. (2021)	Archival retrospective	JA	Community sample of men charged/convicted for IPV	Canada	ODARA	A	M	I	6.0 [2.1]	278	G	.748	AUC	.65		
							M	I	-	278	V	.615	AUC	.63		
							M	I	-	278	IPV	.464	AUC	.62		
							M	NI	4.8 [2.5]	22	IPV	.227	AUC	.57		
Helmus et al. (2012)	Field prospective	JA	Multisite nationwide sample of men convicted for sex offenses under community supervision	Canada	STABLE 2007	A	M	I	8.8 [4.9]	88	S	.114	AUC	.53		
							M	NI	6.9 [5.1]	509	S	.073	AUC	.70		
							M	I	-	88	V	.318	AUC	.53		
							M	NI	-	509	V	.126	AUC	.71		
							M	I	-	88	G	.511	AUC	.58		
							M	NI	-	509	G	.232	AUC	.72		
							M	I	-	1500	G	.359	AUC	.53		
Helmus & Forrester (2014a,b)/Perley Robertson et al. (2019)	Field prospective	GR/JA	Nationwide sample of federally sentenced men	Canada	SFA SRR	A	M	NI	-	6684	G	.252	AUC	.60		
							M	I	-	1500	V	.166	AUC	.54		
							M	NI	-	6684	V	.106	AUC	.62		
							F	I	-	149	G	.222	AUC	.62		
							F	NI	-	377	G	.103	AUC	.69		
							F	I	-	149	V	.067	AUC	.68		
							F	NI	-	377	V	.037	AUC	.75		

					SFA total	A	M/F	I	36.9 [13.5]	12,265	-	-	-	-
							M/F	NI	27.9 [13.7]	51,769	-	-	-	-
					Criminal history	A	M	I	-	1500	G	.359	AUC	.65
							M	NI	-	6684	G	.252	AUC	.72
							M	I	-	1500	V	.166	AUC	.62
							M	NI	-	6684	V	.106	AUC	.71
					Offense severity	A	M	I	-	1500	G	.359	AUC	.47
							M	NI	-	6684	G	.252	AUC	.54
							M	I	-	1500	V	.166	AUC	.54
							M	NI	-	6684	V	.106	AUC	.64
Helmus et al. (2021)	Field prospective	UD	Provincially sentenced men convicted for sex offenses	Canada	Static-99R	A	M	I	3.1 [2.3]	943	S	.064	C	.68
							M	W	2.2 [2.6]	2662	S	.044	C	.71
					STABLE 2007	A	M	I	8.2 [4.7]	921	S	.064	C	.59
							M	W	7.5 [4.9]	2617	S	.045	C	.69
							M	I		921	V	.290	C	.61
							M	W		2617	V	.151	C	.66
							M	I		921	G	.402	C	.64
							M	W		2617	G	.204	C	.68
					Sexual deviance	A	M	I	1.6 [2.1]	900	S	.074	C	.55
							M	W	2.2 [2.1]	2877	S	.051	C	.62
							M	I		900	V	.336	C	.54
							M	W		2877	V	.138	C	.55
							M	I		900	G	.429	C	.56
							M	W		2877	G	.208	C	.55
					Gen. criminality	A	M	I	3.5 [2.1]	900	S	.064	C	.61
							M	W	2.4 [2.1]	2544	S	.045	C	.67
							M	I		900	V	.326	C	.66
							M	W		2544	V	.140	C	.74
							M	I		900	G	.404	C	.69
							M	W		2544	G	.201	C	.76

Higgs, Olver, et al. (2020)	Field prospective	JA	Treated high risk-need adult male violent offending correctional sample	Canada	VRS total	A	M	I	51.0 [8.5]	43	G	.535	d	.67		
							M	NI	49.7 [8.1]	84	G	.476	d	-.01		
							M	I		43	V	.256	d	.03		
							M	NI		84	V	.191	d	-.04		
							VRS dynamic	A	M	I	37.2 [7.5]	43	G	.535	d	.63
									M	NI	36.4 [7.2]	84	G	.476	d	.06
					VRS change	A	M	I		43	V	.256	d	.02		
							M	NI		84	V	.191	d	-.22		
							VRS change	A	M	I	7.4 [4.6]	43	G	.535	d	-.54
									M	NI	7.8 [3.9]	84	G	.476	d	-.19
							M	I		43	V	.256	d	-.65		
							M	NI		84	V	.191	d	.11		
Hogan & Corabian (2023)	Field prospective	CP	Law enforcement referred cases for violence risk assessment		SARA-V3 (mechanical)	A	M	I	33.2 [6.2]	82	V	.549	AUC	.55		
							M	NI	31.5 [6.3]	108	V	.296	AUC	.59		
							M	I		82	V	.549	AUC	.53		
Hogg (2011)/ Wormith et al. (2015)	Field prospective	UTD/JA	Adult men and women, Ontario Provincial corrections.	Canada	LSI-OR	A	F	I		418	G	.459	AUC	.68		
							F	W		2695	G	.313	AUC	.76		
							M	I		1274	G	.607	AUC	.72		
							M	W		12951	G	.400	AUC	.74		
							M/F	I	20.7 [9.6]	1692	G	.570	AUC	.72		
							M/F	W	13.5 [8.7]	15646	G	.385	AUC	.74		
							F	I		418	V	.153	r	.20		
							F	W		2695	V	.145	r	.26		
							M	I		1274	V	.148	r	.17		
							M	W		12951	V	.145	r	.31		
							M/F	I		1692	V	.150	r	.17		
							M/F	W		15646	V	.145	r	.30		
Holsinger et al. (2006)	Field prospective	JA	Custodial and community correctional sample	USA	LSI-R	A	F	I	-	40	G	.275	r	-.29		
							F	W	-	101	G	.287	r	.26		
							M	I	-	100	G	.310	r	.19		
							M	W	-	162	G	.309	r	.22		
							M/F	I	-	140	G	.300	r	.11		
							M/F	W	-	263	G	.255	r	.23		

Hsu et al. (2010, 2011)	Field prospective	JA	Custodial, community, and custodial + community adult male and female correctional samples	Australia	LSI-R	A	M	I		6023	G	.279	r	.18
							M	NI		30791	G	.201	r	.21
							M	I		4071	G	.305	r	.16
							M	NI		17410	G	.173	r	.26
							M	I		864	G	.394	r	.18
							M	NI		2528	G	.305	r	.13
							M	I	27.8 [9.1]	10958	-	-	-	-
							M	NI	21.0 [10.0]	50729	-	-	-	-
							F	I		2125	G	.234	r	.19
							F	NI		6806	G	.138	r	.27
							F	I		552	G	.292	r	.16
							F	NI		1413	G	.168	r	.35
							F	I		186	G	.414	r	.18
							F	NI		263	G	.335	r	.20
Huang (2021)	Field prospective	JA	Court adjudicated youth	Canada	YLS/CMI	Y	M/F	I	16.7 [7.4]	205	G	.824	AUC	.56
							M/F	NI	12.8 [6.5]	193	G	.575	AUC	.65
Jones et al. (2014)	Field prospective	JA	Community sentenced men and women, Alberta Provincial Corrections	Canada	SPIn	A	M/F	I	49.4 [29.8]	635	G	.332	AUC	.75
							M/F	NI	32.5 [26.0]	3021	G	.171	AUC	.76
Jones et al. (2016)	Field prospective	JA	Court adjudicated youth	Canada	YASI	Y	M/F	I	34.9 [18.8]	120	G	.292	AUC	.80
							M/F	NI	26.2 [17.8]	344	G	.180	AUC	.77
							M/F	I		120	V	.125	AUC	.79
							M/F	NI		344	V	.087	AUC	.80
Jung & Rawana (1999)	Field prospective	JA	Court adjudicated youth	Canada	YLS/CMI	Y	M/F	I	12.7 [8.4]	134	G	.325	d	.43
							M/F	NI	10.1 [8.1]	129	G	.282	d	.75
Lee et al. (2018; from Lee et al., 2020)	Field prospective	UD	Men convicted of sex offenses with actual or potential dangerous offender/long term offender designation	Canada	Static-99R	A	M	I	5.4 [2.1]	88	S 5yr	.193	AUC	.56
							M	W	4.8 [2.7]	171	S 5yr	.157	AUC	.69
							M	I		66	S 10yr	.288	AUC	.66
					M		W		117	S 10yr	.231	AUC	.70	
					Static-2002R		M	I	6.4 [2.1]	88	S 5yr	.193	AUC	.56
							M	W	5.8 [2.8]	171	S 5yr	.157	AUC	.69

Lee e al. (2020)	Meta-analysis	JA	Convicted sexual offending samples from correctional/ forensic psychiatric settings (Indigenous Static-99R $k = 3-5$ , $N = 375-599$ , Static-2002R $k = 2-3$ $N = 101-173$ ; White Static-99R $k = 3-5$ , $N = 955-1375$ , Static-2002R $k = 2-3$ $N = 416-620$ )	Canada	Static-99R	A	M	I		66	S 10yr	.288	AUC	.60
							M	W		117	S 10yr	.231	AUC	.59
							M	I	4.8 [0.9]	599	S 5yr	.162	AUC	.61
							M	W	4.3 [1.1]	1375	S 5yr	.119	AUC	.74
					Static-2002R	A	M	I		375	S 10yr	.240	AUC	.66
							M	W		848	S 10yr	.193	AUC	.72
							M	I	6.1 [1.2]	173	S 5yr	.202	AUC	.56
							M	W	5.7 [1.5]	620	S 5yr	.140	AUC	.69
							M	I		101	S 10yr	.257	AUC	.61
							M	W		416	S 10yr	.221	AUC	.72
Sexual deviance	A	M	I	-	182	S	-	AUC	.58					
		M	W	-	641	S	-	AUC	.65					
		M	I	-	182	S	-	AUC	.59					
Gen criminality	A	M	W	-	641	S	-	AUC	.61					
		M	I	-	182	S	-	AUC	.59					
Lewis, Olver, et al. (2013)	Archival retrospective	JA	Treated high risk-need adult male violent offending correctional sample	Canada	VRS total	A	M	I	57.6 [8.2]	86	G	.698	d	.48
							M	NI	55.7 [10.2]	66	G	.530	d	.44
							M	I		86	V	.523	d	.58
							M	NI		66	V	.424	d	.57
					VRS dynamic	A	M	I	42.3 [7.4]	86	G	.698	d	.54
							M	NI	41.8 [7.1]	66	G	.530	d	.43
							M	I		86	V	.523	d	.47
							M	NI		66	V	.424	d	.60
					VRS change	A	M	I	4.7 [3.4]	86	G	.698	d	-.46
							M	NI	4.8 [2.4]	66	G	.530	d	-.24
M	I		86	V			.523	d	-.66					
M	NI		66	V			.424	d	-.49					
Lockwood et al. (2018)	Field prospective	JA	Court adjudicated youth	Canada	YLS/CMI	Y	M/F	I	19.6 [7.8]	37	G	.703	AUC	.69
							M/F	NI	18.2 [7.3]	33	G	.667	r	.28
Lowenkamp & Latessa (2002)	Field prospective	GR	Correctional sample on community supervision	USA	LSI-R	A	M/F	I	28.0	189	G	.339	r	.11
								W	22.2	311	G	.309	r	.23
Lovatt (2020)	Archival retrospective	UTD	Court adjudicated youth	Canada	PCL: YV	Y	M/F	I	19.6 [6.3]	171	G	.608	d	.47
							M/F	NI	17.3 [7.9]	55	G	.454	d	.88
							M/F	I		171	V	.409	d	.38

							M/F	NI		55	V	.218	d	.28
Lovatt et al. (2022)	Archival retrospective	JA	Court adjudicated youth	Canada	VRS-YV total	Y	M/F	I	39.5 (10.4)	171	G	.608	AUC	.61
							M/F	NI	30.6 (13.0)	55	G	.454	AUC	.76
							M/F	I		171	V	.409	AUC	.64
							M/F	NI		55	V	.218	AUC	.62
							M/F	I	32.9 [8.8]	171	G	.608	AUC	.65
							M/F	NI	26.0 [11.1]	55	G	.454	AUC	.78
							M/F	I		171	V	.409	AUC	.61
							M/F	NI		55	V	.218	AUC	.63
						M/F	I	28.3 (7.1)	171	G	.608	AUC	.62	
						M/F	NI	19.5 (8.7)	55	G	.454	AUC	.78	
						M/F	I		171	V	.409	AUC	.63	
						M/F	NI		55	V	.218	AUC	.65	
						M/F	I	1.1 (1.2)	171	G	.608	AUC	.59	
						M/F	NI	1.9 (1.5)	55	G	.454	AUC	.75	
						M/F	I		171	V	.409	AUC	.59	
						M/F	NI		55	V	.218	AUC	.64	
	M/F	I	6.4 (3.8)	171	G	.608	AUC	.65						
	M/F	NI	10.3 (6.0)	55	G	.454	AUC	.75						
	M/F	I		171	V	.409	AUC	.63						
	M/F	NI		55	V	.218	AUC	.66						
Luong & Wormith (2011)	Field prospective	JA	Court adjudicated youth	Canada	LSI-SK	Y	M/F	I	22.8 [8.9]	123	G	.707	AUC	.64
Matz et al. (2021)	Field prospective	JA	Court adjudicated youth	USA	YASI	Y	M/F	I	15.9 [8.4]	69	G	.478	AUC	.79
							M/F	W	-	93	G	.484	AUC	.73
McCuish et al. (2017)	Face to face prospective	JA	Court adjudicated youth	Canada	PCL: YV	Y	M/F	I	21.9 (6.4)	137	G	.88/.76	AUC	.64
							M/F	W	22.4 (6.5)	312	G	.81/.61	AUC	.60
							M/F	I		137	V	.63/.38	AUC	.61
							M/F	W		312	V	.44/.33	AUC	.54
McKinnon (2004)	Field prospective	UTD	Court adjudicated youth	Canada	YLS/CMI	Y	M/F	I	-	28	G	.571	AUC	.63
							M/F	NI	-	74	G	.432	AUC	.57
							M/F	I	-	28	V	.500	AUC	.65
							M/F	NI	-	74	V	.176	AUC	.57

Muir et al. (2020)	Field prospective	JA	Court adjudicated youth	Canada	SAVRY total	Y	M	I	22.4 (9.3)	219	G	.557	AUC	.70		
							M	W	18.3 (9.4)	357	G	.384	AUC	.79		
							F	I	26.6 (8.2)	80	G	.525	AUC	.76		
							F	W	19.7 (9.5)	88	G	.352	AUC	.69		
							M	I		219	V	.288	AUC	.67		
							M	W		357	V	.165	AUC	.70		
							F	I		80	V	.250	AUC	.69		
							F	W		88	V	.125	AUC	.73		
							SAVRY Protect	Y	M	I	3.3 (1.8)	219	G	.557	AUC	.61
									M	W	3.5 (1.8)	357	G	.384	AUC	.72
									F	I	2.5 (1.6)	80	G	.525	AUC	.59
									F	W	3.4 (1.9)	88	G	.352	AUC	.56
									M	I		219	V	.288	AUC	.58
									M	W		357	V	.165	AUC	.65
			F	I		80	V	.250	AUC	.59						
			F	W		88	V	.125	AUC	.53						
Meyers & Schmidt (2008)	Archival retrospective	JA	Court adjudicated youth	Canada	SAVRY	Y	M/F	I	-	38	V	.447	AUC	.84		
							M/F	NI	-	83	V	.205	AUC	.70		
Myer, A. J. (2019)	Field prospective	JA		USA	Static-99R	A	M	I	-	123	S	.08/.10	AUC	.64		
							M	W	-	761	S	.10/.12	AUC	.61		
Nafekh & Motiuk (2002)	Mixed automated and field ratings	GR	Federally sentenced adult male and female prison inmates	Canada	SIR Proxy	A	M/F	I	-	1211	G	.280	AUC	.68		
					SIR-R1	A	M/F	NI	-	6881	G	.215	AUC	.75		
					SIR Proxy	A	M/F	I	-	1211	V	.075	AUC	.65		
					SIR-R1	A	M/F	NI	-	6881	V	.034	AUC	.71		
					SIR Proxy	A	M/F	I	-	1211	S	.014	AUC	.60		
					SIR-R1	A	M/F	NI	-	6881	S	.005	AUC	.54		
O'Byrne, P. L. (2004).	Field prospective	UTD	Adult men and women, Saskatchewan Provincial corrections.	Canada	ORAMS PRA	A	M/F	I	-	1560	G	.692	AUC	.70		
							M/F	NI	-	667	G	.402	AUC	.74		
Olver, Cortoni et al. (2021)	Mixed archival retrospective and field prospective	UD	Mixed treated and untreated sample of federally sentenced adult male prison	Canada	VRS total	A	M	I	48.5 [13.5]	406	G	.559	AUC	.75		
							M	NI	38.9 [16.7]	550	G	.362	AUC	.77		
							M	I		406	V	.369	AUC	.75		
							M	NI		550	V	.215	AUC	.78		
							VRS dynamic	A	M	I	36.1 [10.7]	406	G	.559	AUC	.73

			inmates with a history of violent offending			M	NI	29.3 [12.9]	550	G	.362	AUC	.76
						M	I		406	V	.369	AUC	.74
					VRS change	M	NI		550	V	.215	AUC	.79
						A	M	I	5.2 [3.7]	G	.755	AUC	.60
							M	NI	4.9 [3.7]	G	.544	AUC	.44
							M	I		V	.554	AUC	.62
							M	NI		V	.367	AUC	.46
Olver, Lewis, et al. (2013)	Archival retrospective	JA	Treated high risk-need adult male violent offending correctional sample	Canada	PCL-R	A	M	I	25.4 [4.9]	G	.688	d	.40
							M	NI	26.7 [5.9]	G	.530	d	.44
							M	I		V	.523	d	.28
							M	NI		V	.424	d	.87
Olver, Kingston, & Sowden (2021)	Mixed archival retrospective and field prospective	JA	Treated federally sentenced men convicted for sexual offenses	Canada	Sexual deviance	A	M	I	5.6 [3.6]	S	.143	AUC	.57
							M	NI	6.7 [3.9]	S	.101	AUC	.61
							M	I		V	.359	AUC	.51
							M	NI		V	.167	AUC	.53
							M	I		G	.591	AUC	.54
							M	NI		G	.350	AUC	.49
					Gen risk-need	A	M	I	8.9 [3.4]	S	.143	AUC	.66
							M	NI	6.8 [3.7]	S	.101	AUC	.71
							M	I		V	.359	AUC	.74
							M	NI		V	.167	AUC	.74
							M	I		G	.591	AUC	.75
							M	NI		G	.350	AUC	.72
Olver, Nicholaichuk, et al. (2014, 2020)	Field prospective	JA	Multisite treated sample of federally sentenced men convicted for sexual offenses	Canada	VRS-SO total	A	M	I	32.1 [11.7]	S	.112	d	.73
							M	NI	27.7 [11.2]	S	.060	d	.78
							M	I		V	.259	d	.58
							M	NI		V	.096	d	.91
							M	I		G	.523	d	.89
							M	NI		G	.274	d	.79
					VRS-SO dynamic	A	M	I	22.7 [8.3]	S	.112	d	.66
							M	NI	20.6 [8.2]	S	.060	d	.74
							M	I		V	.259	d	.57

						M	NI		365	V	.096	d	.85
						M	I		197	G	.523	d	.76
					VRS-SO change	M	NI		365	G	.274	d	.64
						M	I	3.8 [2.6]	197	S	.112	d	-.50
						M	NI	4.0 [2.8]	365	S	.060	d	-.40
						M	I		197	V	.259	d	-.27
						M	NI		365	V	.096	d	-.35
						M	I		197	G	.523	d	-.05
					Static-99R	M	NI		365	G	.274	d	-.23
						M	I	4.3 [2.6]	213	S 5yr	.113	AUC	.65
						M	NI	2.6 [2.8]	391	S 5yr	.064	AUC	.68
						M	I		118	S 10yr	.153	AUC	.68
						M	NI		190	S 10yr	.095	AUC	.73
Olver, Neumann et al. (2013)	Archival retrospective	JA	Federally sentenced adult male prison inmates	Canada	PCL-R	M	I	15.6 [6.9]	171	G	.292	AUC	.63
						M	NI	12.6 [7.9]	264	G	.182	AUC	.63
						M	I		171	V	.122	AUC	.65
						M	NI		264	V	.098	AUC	.66
Olver, Neumann et al. (2018)	Archival retrospective	JA	Federally sentenced adult male prison inmates	Canada	PCL-R	M	I	20.5 [7.4]	452	G	.619	AUC	.72
						M	NI	18.5 [8.8]	711	G	.519	AUC	.73
						M	I		452	V	.460	AUC	.71
						M	NI		711	V	.319	AUC	.73
Olver & Riemer (2021)	Archival retrospective	JA	Treated federally sentenced men convicted for sexual offenses	Canada	PCL-R	M	I	18.6 [6.4]	185	G	.676	d	.80
						M	NI	17.5 [6.8]	236	G	.419	d	.52
						M	I		185	V	.465	d	.62
						M	NI		236	V	.237	d	.75
					SAPROF	M	I	14.8 [5.4]	185	G	.676	d	.69
						M	NI	16.3 [5.9]	236	G	.419	d	.46
						M	I		185	V	.465	d	.68
						M	NI		236	V	.237	d	.86
Olver, Sowden et al. (2018) <sup>a,b</sup>	Mixed archival retrospective and field prospective	JA	Treated federally sentenced men convicted for sexual offenses	Canada	VRS-SO total	M	I	34.0 [10.6]	384	S	.143	AUC	.68
						M	NI	30.5 [11.0]	660	S	.101	AUC	.75
						M	I		384	V	.359	AUC	.67
						M	NI		660	V	.167	AUC	.73
						M	I		384	G	.591	AUC	.71

							M	NI		660	G	.350	AUC	.68
					VRS-SO dynamic	A	M	I	23.7 [7.7]	384	S	.143	AUC	.65
							M	NI	22.0 [8.0]	660	S	.101	AUC	.73
							M	I		384	V	.359	AUC	.66
							M	NI		660	V	.167	AUC	.70
							M	I		384	G	.591	AUC	.68
					VRS-SO change	A	M	NI		660	G	.350	AUC	.64
							M	I	3.4 [2.7]	384	S	.143	AUC	.59
							M	NI	3.7 [2.8]	660	S	.101	AUC	.62
							M	I		384	V	.359	AUC	.60
							M	NI		660	V	.167	AUC	.60
							M	I		384	G	.591	AUC	.52
							M	NI		660	G	.350	AUC	.55
					Static-99R	A	M	I	4.6 [2.4]	400	S 5yr	.143	AUC	.62
							M	W	3.2 [2.9]	665	S 5yr	.096	AUC	.71
							M	I		274	S 10yr	.234	AUC	.67
							M	W		432	S 10yr	.167	AUC	.71
	Archival retrospective	UD			Static-2002R	A	M	I	5.5 [2.7]	108	S 5yr	.148	AUC	.71
							M	W	4.5 [3.2]	88	S 5yr	.091	AUC	.70
							M	I		61	S 10yr	.180	AUC	.67
							M	W		44	S 10yr	.136	AUC	.89
Olver, Stockdale, et al. (2012)	Mixed archival retrospective and field prospective	JA	Court adjudicated youth	Canada	YLS/CMI	Y	M/F	I	27.6 [6.6]	104	G	.788	AUC	.77
							M/F	W	19.8 [8.2]	40	G	.475	AUC	.52
							M/F	I	-	104	V	.548	AUC	.76
							M/F	W	-	40	V	.225	AUC	.71
Olver, Stockdale, et al. (2009)	Meta-analysis	JA	Court adjudicated youth samples ( <i>k</i> = 5)	Canada	LSI variants	A/Y	M/F	I	-	860	G		r	.35
							M/F	NI	-	462	G		r	.35
Olver, Stockdale, et al. (2014)	Meta-analysis	JA	Adult and youth correctional and forensic samples (Indigenous <i>k</i> = 13, White <i>k</i> = 24)	Australia, Canada, USA	LSI variants	A/Y	M/F	I	-	5354	G	[.360	r	.30
							M/F	NI	-	40989	G	overall]	r	.29
		JA		Canada	PCL-R	A	M	I	23.3 [5.1]	86	G	.779	d	.33

Olver & Wong (2015)	Archival retrospective		Federally sentenced adult male prison inmates from institutions throughout Canadian Prairies			M	NI	23.6 [5.7]	187	G	.513	d	.55	
						M	I		86	V	.453	d	.56	
						M	NI		187	V	.241	d	.30	
Olver, Wong, et al. (2007)	Archival retrospective	JA	Treated high risk-need adult male sexual offending correctional sample	Canada	VRS-SO total	A	M	I	33.3 [9.1]	108	S	.185	d	.70
							M	NI		32.1 [10.2]	209	S	.177	d
							M	I	108		V	.482	d	.74
							M	NI	209	V	.273	d	.87	
							M	I	108	G	.685	d	.54	
							M	NI	209	G	.478	d	.44	
					VRS-SO dynamic	A	M	I	22.8 [6.9]	108	S	.185	d	.42
							M	NI		22.3 [7.6]	209	S	.177	d
							M	I	108		V	.482	d	.56
							M	NI	209	V	.273	d	.70	
							M	I	108	G	.685	d	.38	
							M	NI	209	G	.478	d	.33	
					VRS-SO change	A	M	I	2.3 [1.8]	108	S	.185	d	-.13
							M	NI		2.7 [2.2]	209	S	.177	d
							M	I	108		V	.482	d	-.08
							M	NI	209	V	.273	d	-.18	
							M	I	108	G	.685	d	-.07	
							M	NI	209	G	.478	d	-.10	
Static-99R	A	M	I	4.8 [1.9]	108	S 5yr	.185	AUC	.69					
		M	NI		4.4 [2.5]	200	S 5yr	.170	AUC	.72				
		M	I	104		S 10yr	.308	AUC	.68					
		M	NI	198	S 10yr	.222	AUC	.70						
Prince et al. (2021)	Field prospective	JA	Court adjudicated youth	Canada	LSI-SK	Y	M/F	I	5.7 [2.1]	203	G	.683	AUC	.64
							M/F	NI		3.7 [2.8]	58	G	.569	AUC
							M	I	148		G	.762	AUC	.61
							M	NI	50	G	.600	AUC	.76	
							F	I	55	G	.473	AUC	.67	
							M/F	I	203	V	.505	AUC	.62	
							M/F	NI	58	V	.328	AUC	.73	

							M	I		148	V	.574	AUC	.60
							M	NI		50	V	.340	AUC	.73
							F	I		55	V	.309	AUC	.67
Rector et al. (2007)	Field prospective	CP	Court adjudicated youth	Canada	LSI-SK	Y	M/F	I	-	628	G	.537	r	.37
							M/F	NI	-	244	G	.371	r	.33
Rettinger & Andrews (2010)	Field prospective	JA	Adult women, Ontario Provincial corrections.	Canada	LS/CMI	A	F	W	-	48	G	.500	AUC	.80
Schmidt et al. (2006)	Archival retrospective	JA	Court adjudicated youth	Canada	PCL: YV	Y	F	I	-	307	G	.459	AUC	.86
							M/F	I	23.0 (6.0)	40	G	.539	r	.44
							M/F	W	21.6 (6.6)	90	G	.421	r	.32
							M/F	I		40	V	.103	r	.52
							M/F	W		90	V	.239	r	.21
Schwalbe (2009)	Field prospective	JA	Community samples of court adjudicated youth over three fiscal cohorts	USA	ARNA	Y	M/F	I	-	545	G	.330	AUC	.61
							M/F	W	-	4464	G	.330	AUC	.65
							M/F	I	-	494	G	.321	AUC	.66
							M/F	W	-	4402	G	.321	AUC	.66
							M/F	I	-	537	G	.305	AUC	.63
							M/F	W	-	4079	G	.305	AUC	.67
Schwalbe (2009)/ LeCroy et al. (1998)	Field prospective	JA/GR	Community sample of court adjudicate youth	USA	ARNA	Y	M/F	I	-	448	G	.480	AUC	.75
Sewall (2015)	Archival retrospective	UTD	Treated high risk-need adult male sexual offending correctional sample	Canada	PCL-R	A	M/F	W	-	5225	G	.517	AUC	.77
							M	I	22.2 [6.7]	109	G	.835	d	.62
							M	NI	18.9 [7.9]	196	G	.673	d	.52
							M	I		109	V	.725	d	.70
							M	NI		196	V	.474	d	.71
Shepherd et al. (2018)	Fact to face prospective	JA	Men or women sentenced or remanded in custody	Australia	Cultural engmt Cultural identity	A	M/F	I	9.5 [2.6]	84	V	.417	e <sup>B</sup>	0.80
							M/F	I	21.1 [3.4]	84	V	.417	e <sup>B</sup>	0.97
Shepherd et al. (2014)	Fact to face prospective		Court adjudicated youth remanded or sentenced to custody	Australia	SAVRY total	Y	M	I	29.2 [9.6]	32	G	.862	AUC	.81
							M	W	27.1 [9.3]	84	G	.725	AUC	.78
							M	I		32	V	.690	AUC	.76
							M	W		84	V	.580	AUC	.68
					SAVRY protect	Y	M	I	1.3 [1.6]	32	G	.862	AUC	.91
							M	W	1.7 [1.9]	84	G	.725	AUC	.80
							M	I		32	V	.690	AUC	.67
							M	W		84	V	.580	AUC	.77

Shepherd et al. (2015)	Fact to face prospective	JA	Court adjudicated youth remanded or sentenced to custody	Australia	YLS/CMI	Y	M/F	I	27.4 [6.4]	40	G	.775	AUC	.67		
							M/F	W	25.9 [7.3]	85	G	.776	AUC	.79		
							M/F	I		40	V	.600	AUC	.67		
							M/F	W		85	V	.612	AUC	.71		
Shepherd & Strand (2016)	Fact to face prospective	JA	Court adjudicated youth remanded or sentenced to custody	Australia	PCL: YV	Y	M	I	18.2 [6.7]	32	G	.862	AUC	.70		
							M	W	18.9 [6.6]	84	G	.725	AUC	.72		
							M	I		32	V	.690	AUC	.66		
							M	W		84	V	.580	AUC	.76		
Smallbone & Ralling (2013)	Field prospective	JA	Adult males convicted for sexual offending released from custody	Australia	Static-99R	A	M	I	3.7 [2.3]	67	S	.090	AUC	.61		
							M	NI	2.2 [2.7]	320	S	.041	AUC	.79		
							M	I		67	V	.373	AUC	.65		
							M	NI		320	V	.059	AUC	.75		
Sowden & Olver (2017)	Archival retrospective	JA	Treated high risk-need adult male sexual offending correctional sample	Canada	VRS-SO total	A	M	I	38.5 [8.1]	79	S	.165	d	.24		
							M	NI	37.8 [7.5]	86	S	.093	d	.76		
							M	I		79	V	.456	d	.62		
							M	NI		86	V	.209	d	.44		
							M	I		79	G	.633	d	.75		
							M	NI		86	G	.360	d	.50		
							VRS-SO dynamic	A	M	I	27.1 [6.2]	79	S	.165	d	.39
									M	NI	26.9 [6.1]	86	S	.093	d	.85
					M	I				79	V	.456	d	.78		
					M	NI				86	V	.209	d	.51		
					VRS-SO change	A	M	I	3.9 [3.2]	79	S	.165	d	-.17		
							M	NI	4.4 [3.3]	86	S	.093	d	-.76		
							M	I		79	V	.456	d	-.55		
							M	NI		86	V	.209	d	-.42		
							M	I		79	G	.633	d	-.34		
							M	NI		86	G	.360	d	-.19		
STABLE 2007	A	M	I	11.9 [3.7]			87	S	.218	d	.11					
		M	NI	11.9 [4.5]			90	S	.189	d	.61					
		M	I		87	V	.540	d	.78							
		M	NI		90	V	.367	d	.53							

						M	I		87	G	.690	d	.62	
						M	NI		90	G	.522	d	.27	
					Static-99R	A	M	I	5.1 [2.2]	S 5yr	.165	AUC	.43	
							M	NI	4.6 [2.4]	S 5yr	.067	AUC	.61	
							M	I		S 10yr	.269	AUC	.57	
							M	NI		S 10yr	.227	AUC	.55	
Spiranovic (2012)	Field prospective	GR	Adult males convicted for sexual offending on parole or community supervision	Australia	Static-99R	A	M	I	4.1 [1.9]	S 5yr	.247	AUC	.58	
							M	NI	2.1 [2.6]	S 5yr	.138	AUC	.71	
							M	I		S 10yr	.282	AUC	.51	
							M	NI		S 10yr	.216	AUC	.67	
							M	I		V 5yr	.494	AUC	.66	
							M	NI		V 5yr	.200	AUC	.73	
							M	I		V 10yr	.612	AUC	.66	
							M	NI		V 10yr	.292	AUC	.70	
Stewart et al. (2017)	Field prospective	GR	Federally sentenced adult male prison inmates	Canada	DFIAR total	A	M	I	74.9%H	3710	G	-	C	.56
					DFIAR total	A	M	NI	61.2%H	12590	G	-	C	.62
					Empl/educ	A	M	I	13.3%H	3229	G	-	e <sup>B</sup>	3.06
					Empl/educ	A	M	NI	6.6%H	12909	G	-	e <sup>B</sup>	3.81
					Companions	A	M	I	28.8%H	3229	G	-	e <sup>B</sup>	1.62
					Companions	A	M	NI	23.7%H	12909	G	-	e <sup>B</sup>	2.28
					Subst abuse	A	M	I	60.5%H	3229	G	-	e <sup>B</sup>	2.45
					Subst abuse	A	M	NI	34.4%H	12909	G	-	e <sup>B</sup>	3.34
					Attitudes	A	M	I	37.8%H	3229	G	-	e <sup>B</sup>	1.60
					Attitudes	A	M	NI	38.3%H	12909	G	-	e <sup>B</sup>	2.53
					Comm func	A	M	I	8.5%H	3229	G	-	e <sup>B</sup>	1.80
					Comm func	A	M	NI	5.1%H	12909	G	-	e <sup>B</sup>	2.98
					Fam/marital	A	M	I	17.9%H	3229	G	-	e <sup>B</sup>	1.11
					Fam/marital	A	M	NI	10.8%H	12909	G	-	e <sup>B</sup>	1.24
					Pers/emotional	A	M	I	59.5%H	3229	G	-	e <sup>B</sup>	1.58
					Pers/emotional	A	M	NI	42.9%H	12909	G	-	e <sup>B</sup>	1.92
Stockdale et al. (2010)	Mixed archival retrospective	JA	Court adjudicated youth	Canada	PCL: YV	Y	M/F	I	23.8 [6.4]	101	G	.802	AUC	.72
							M/F	W	18.0 [8.3]	39	G	.436	AUC	.56
							M/F	I		101	V	.554	AUC	.71
							M/F	W		39	V	.205	AUC	.78

Stockdale et al. (2014)	and field prospective	JA	Court adjudicated youth with current or prior violent offenses	Canada	VRS-YV	Y	M/F	I	48.0 [9.8]	91	G	.791	AUC	.74	
	Archival retrospective						M/F	NI	34.0 [11.6]	37	G	.432	AUC	.61	
							M/F	I		91	V	.549	AUC	.74	
							M/F	NI		37	V	.216	AUC	.83	
					VRS-YV total		Y	M/F	I	38.9 [8.4]	91	G	.791	AUC	.74
							M/F	NI	27.9 [10.5]	37	G	.432	AUC	.59	
							M/F	I		91	V	.549	AUC	.72	
							M/F	NI		37	V	.216	AUC	.78	
Thompson & McGrath (2012)	Field prospective	JA	Court adjudicated youth	Australia	YLS/CMI-AA	Y	M/F	I	20.6 [9.5]	1053	G	.337	AUC	.60	
							M/F	NI	18.0 [9.7]	1582	G	.494	AUC	.64	
	Towers (2020)	Field prospective	UTD	Released custody and probation adult male and female correctional sample	USA	LSI-R	A	M/F	I	-	798	G	.472	r	.24
								M/F	W	-	2900	G	.382	r	.27
Watkins (2011)	Field prospective	GR	Custody, community, and custody + community adult male and female correctional samples	Australia	LSI-R	A	M	I	33.0 [7.6]	2465	G	.311	AUC	.66	
							M	NI	27.7 [9.2]	7555	G	.337	AUC	.69	
							F	I	34.2 [6.3]	417	G	.537	AUC	.60	
							F	NI	30.2 [8.7]	614	G	.311	AUC	.69	
Wilson & Gutierrez (2014)	Meta-analysis	JA	Adult and youth correctional and forensic samples (Indigenous $k = 13$ , $N = 21,807$ ; non-Indigenous $k = 11$ , $N = 42,515$ )	Australia, Canada, USA	LSI variants	A/Y	M/F	I	-	21508	G	.555	d	.62	
							M/F	NI	-	58719	G	.378	d	.82	
							M/F	I	-	-	V	.324	d	.49	
							M/F	I	-	20294	G	.555	d	.57	
							M/F	NI	-	-	G	.378	d	.88	
							M/F	I	-	-	V	.324	d	.52	
							M/F	I	-	6335	G	.555	d	.60	
							M/F	NI	-	-	G	.378	d	.76	
							M/F	I	-	6383	G	.555	d	.52	
							M/F	NI	-	-	G	.378	d	.60	

					Companions	A/Y	M/F	I	-	20294	G	.555	d	.39
					Companions	A/Y	M/F	NI	-	-	G	.378	d	.52
					Subst abuse	A/Y	M/F	I	-	20294	G	.555	d	.42
					Subst abuse	A/Y	M/F	NI	-	-	G	.378	d	.54
					Attitudes	A/Y	M/F	I	-	20293	G	.555	d	.39
					Attitudes	A/Y	M/F	NI	-	-	G	.378	d	.48
					Leisure/rec	A/Y	M/F	I	-	20287	G	.555	d	.34
					Leisure/rec	A/Y	M/F	NI	-	-	G	.378	d	.37
					Fam/marital	A/Y	M/F	I	-	20294	G	.555	d	.24
					Fam/marital	A/Y	M/F	NI	-	-	G	.378	d	.28
					Fam/marital	A/Y	M/F	I	-	-	V	.324	d	.13
Wilton et al. (2017)	Field prospective	GR	Federally sentenced adult male and female prison inmates	Canada	DFIA-R mechanical	A	M	I	51%H 9%L	4640	G	.559	C	.62
						A	M	NI	44%H	15487	G	.359	C	.68
									21%L					
Wong and Gordon (2006)	Archival retrospective	JA	Federally sentenced adult male prison inmates convicted for violent offenses from multiple institutions throughout Canadian Prairies	Canada	GSIR	A	M	I	-3.7 [10.4]	165	G	.297	d	.48
								NI	-0.7 [10.5]	260	G	.185	d	.45
								I		165	V	.121	d	.48
								NI		260	V	.058	d	.78
					VRS total	A	M	I	39.3 [13.4]	173	G	.289	d	.44
								NI	29.6 [14.8]	272	G	.184	d	.84
								I		173	V	.121	d	.57
								NI		272	V	.065	d	1.00
					VRS dynamic	A	M	I	29.4 [11.0]	173	G	.289	d	.38
								NI	21.8 [11.6]	272	G	.184	d	.72
								I		173	V	.121	d	.53
								NI		272	V	.065	d	.93
Wormith & Olver (2002)	Archival retrospective	JA	Treated high risk-need adult male violent offending correctional sample	Canada	GSIR	A	M	I	-7.6 [6.6]	48	G	.708	r	.37
								NI	-4.3 [9.1]	43	G	.605	r	.61
								I		48	V	.438	r	.29
								NI		43	V	.488	r	.53

*Note:* Design: field prospective (ratings completed in field as routine service delivery and individuals followed up in real time), archival retrospective (ratings completed from past institutional file blind to outcome on individuals currently released), face to face prospective (ratings completed on individuals face to face in a research context and individuals followed up in real time), meta-analysis (quantitative review); Source: GR (government report), JA (refereed journal article), UTD (unpublished thesis/dissertation), UD (unpublished data from exiting study),

CP (conference presentation); DG = developmental group, adult (A), youth (Y), adult and youth combined (Y/A); Sex: male (M), female (F), male and female combined (M/F); ECG: ethnocultural group, Indigenous (I), non-Indigenous majority White (NI), White (W); Measure M (SD): mean and standard deviation of an assessment measure or construct for a given ethnocultural group; N: cell N of ethnocultural group for a given outcome analysis; RC = recidivism criterion, general (G), violent (V), sexual (S), intimate partner violence (IPV); BR: recidivism base rate for associated RC within sex, developmental group, and recidivism category; ESM: effect size metric employed in study or ad hoc analysis and ES: effect size where: standardized mean difference (d, .20 small, .50 medium, .80 large), point biserial (r) or phi ( $\phi$ ) correlation (.10 small, .24 medium, .37 large), area under the curve (AUC, .56 small, .64 medium, .71 large). LSI: Level of Service Inventory; LSI-R: Level of Service Inventory-Revised; LS/CMI: Level of Service/Case Management Inventory; LSI-OR Level of Service Inventory-Ontario Revision; MNRF: Manitoba Risk-Need Form; YLS/CMI: Youth Level of Service/Case Management Inventory (AA: Australian Adaptation); LSI-SK: Level of Service Inventory-Saskatchewan Youth Edition; SFA: Static Factors Assessment; DFIA-R: Dynamic Factors Identification and Analysis; SIR: Statistical Information on Recidivism Scale; ORAMS PRA: Offender Risk Assessment and Management System Primary Risk Assessment; ARNA: Arizona Risk/Needs Assessment Instrument; SPIn: Service Planning Instrument; YASI: Youth Assessment and Screening Instrument; ODARA: Ontario Domestic Assault Risk Assessment; SARA; Spousal Assault Risk Assessment Guide; VRS: Violence Risk Scale; VRS-SO: Violence Risk Scale-Sexual Offense Version; VRS-YV Violence Risk Scale-Youth Version; SAVRY: Structured Assessment of Violence Risk for Youth; SAPROF: Structured Assessment of Protective Factors; SAPROF-YV: Structured Assessment of Protective Factors-Youth Version; PCL-R: Psychopathy Checklist-Revised; PCL: YV Psychopathy Checklist: Youth Version.

## Supplemental Table S2

*Summary of Studies Included in Systematic Review and Meta-Analyses*

Study	Design	Source	Sample	Country	Measures	DG	Sex	I n	NIW n	RC
Ahmed (2022)	FP	UTD	IPV provincial correctional sample on community supervision (BC)	C	SARA-V2	A	M	2,949	6,105	IPV, V, G
Ahmed et al. (2023)	MA	JA	Convicted sexual offending samples from correctional/forensic settings	C, A, U	Static-99R	A	M	2,043	13,756	S
Allan & Dawson (2002)	AR	GR	Correctional sample referred for violence or sexual offense program	A	3-Predictor model	A	M	109	92	V
Babchishin et al. (2012)	MA	JA	Convicted sexual offending samples from correctional/ forensic psychiatric settings	C	Static-99/R, Static-2002/R, Sexual deviance, Gen crim	A	M	319	1,269	S
Barnum & Gobeil (2012a)	ATM/FP	GR	Federally sentenced men/women in custody	C	SIR Proxy, SIR-R1	A	M	2,560	12,485	G, V, S
							F	246	666	G, V, S
Barnum & Gobeil (2012b)	FP	GR	Federally sentenced female inmates	C	E/E, Comp., SA, Att., CF, F/M, P/E	A	F	90	263	G
Boer (2003; from Babchishin et al., 2012)	AR	UTD	Men federally sentenced for sex offenses	C	Static-99R	A	M	53	227	S
Bonta (1989)	FP	JA	Provincially incarcerated men (ON)	C	LSI	A	M	48	71	G
Bonta et al. (1997)	FP	JA	Male and female probationers (MB)	C	MRNF	A	M	390	513	G
Bonta & Rugge (2004; from Rugge, 2006)	FP	GR	Federally sentenced men in custody	C	SIR-R1	A	M	940	4,475	G, V
Bonta & Yessine (2005; from Lee et al., 2020)	FP	GR	Men federally sentenced for sex offenses flagged as possible dangerous offender	C	Static-99R	A	M	26	90	S
Brankley et al. (2021; from Lee et al., 2020)	FP	UD	Men convicted for sex offenses under community supervision	C	Static-99R, Static-2002R	A	M	35	308	S
Brews (2009)	FP	UTD	Provincially sentenced women (ON)	C	LSI-OR	A	F	253	1,502	G, V
Brown & Motiuk (2005)	FP	GR	Federally sentenced men/women in custody	C	DFIA-R, E/E, Comp., SA, Att., CF, F/M, P/E	A	M/F	2,635	-	G
Buchanan (2009)	AR	UTD	Provincially sentenced IPV sample (SK)	C	ODARA	A	M	53	41	IPV
Coupland & Olver (2020)	AR	JA	Treated high risk-need adult male violent offending correctional sample	C	VRS total, dynamic, change	A	M	91	64	G, V
Gobeil (2011)	FP	GR	Federally sentenced women in custody	C	E/E, Comp., SA, Att., CF, F/M, P/E	A	M	733	3,229	G
Gossner & Wormith (2007)	AR	JA	Court adjudicated youth	C	YLS/CMI	Y	M/F	62	32	G
Gutierrez et al. (2013)	MA	JA	Adult/youth male/female samples from community and correctional settings	C, A, U	CH, E/E, Comp., SA, Att., AP, CF, F/M, L/R, P/E, Em dis, Prior vic	A/Y	M/F	57,315	204,997	G

Haag (2005; from Lee et al., 2020)	AR	UTD	Men federally sentenced for sex offenses	C	Static-99R, Static-2002R	A	M	50	141	S
Hann & Harman (1989)	FP	GR	Federally sentenced men in custody	C	GSIR	A	M	49	-	G
Hann & Harman (1993)	FP	GR	Federally sentenced men in custody	C	GSIR	A	M	269	-	G
Hansen (2018)	FP	UTD	Custodial releasees to tribal reservation	U	LSI-R, Cultural Connectedness	A	M/F	166	-	G
Hanson et al. (2015)	FP	JA	Multisite nationwide sample of men on community supervision for sex offending	C	Static-99R	A	M	102	523	S
Hegel et al. (2021)	AR	JA	Community sample w/ current/prior IPV	C	ODARA	A	M	278	22	IPV
Helmus et al. (2012)	FP	JA	Multisite nationwide sample of men on community supervision for sex offending	C	STABLE 2007	A	M	88	509	S, V, G
Helmus & Forrester (2014a,b)/ Perley Robertson et al. (2019)	FP	GR/JA	Nationwide sample of federally sentenced men and women	C	SFA total, SRR, Crim hist, Off severity	A	M	1,500	6,684	G, V
						F		149	377	G, V
						M/F		12,265	51,769	-
Helmus et al. (2021)	FP	UD	Provincially sentenced men convicted for sexual offenses	C	Static-99R, STABLE 2007, Sexual deviance, Gen crim, and risk-need	A	M	943	2,877	S, V, G
Higgs, Olver, et al. (2020)	FP	JA	Treated high risk-need adult male violent offending corrections sample	C	VRS total, dynamic, change	A	M	43	84	G, V
Hogan & Corabian (2023)	FP	CP	Law enforcement referred cases for violence risk assessment	C	SARA-V3	A	M	82	108	V
Hogg (2011)/ Wormith et al. (2015)	FP	UTD/JA	Provincially sentenced men/women (ON)	C	LSI-OR	A	M	1,274	12,951	G, V
						F		418	2,695	G, V
Holsinger et al. (2006)	FP	JA	Custody/community corrections sample	U	LSI-R	A	M	100	162	G
						F		40	101	G
Hsu et al. (2010, 2011)	FP	JA	Custody, community, and hybrid adult male and female correctional samples	A	LSI-R	A	M	10,958	50,729	G
						F		2,863	8,482	G
Huang (2021)	FP	JA	Court adjudicated youth	C	YLS/CMI	Y	M/F	205	193	G
Jones et al. (2014)	FP	JA	Community sentenced men/women (AB)	C	SPIIn	A	M/F	635	3,021	G
Jones et al. (2016)	FP	JA	Court adjudicated youth	C	YASI	Y	M/F	120	344	G, V
Jung & Rawana (1999)	FP	JA	Court adjudicated youth	C	YLS/CMI	Y	M/F	134	129	G
Lee et al. (2018; from Lee et al., 2020)	FP	UD	Sex offense sample with actual/potential dangerous/long term offender status	C	Static-99R, Static-2002R	A	M	88	171	S
Lee e al. (2020)	MA	JA	Convicted sexual offending samples from correctional/ forensic psychiatric settings	C	Static-99R, Static-2002R, Sexual deviance, Gen criminality	A	M	599	1,375	S
Lewis et al. (2013)	AR	JA	Treated high risk-need adult male violent offending correctional sample	C	VRS total, dynamic, change	A	M	86	66	G, V
Lockwood et al. (2018)	FP	JA	Court adjudicated youth	C	YLS/CMI	Y	M/F	37	33	G
Lowenkamp & Latessa (2002)	FP	GR	Correctional community supervision	U	LSI-R	A	M/F	189	311	G
Lovatt (2020)	AR	UTD	Court adjudicated youth	C	PCL: YV	Y	M/F	171	55	G, V

Lovatt et al. (2022)	AR	JA	Court adjudicated youth	C	VRS-YV total, dynamic, SAVRY total, protect, SAPROF-YV	Y	M/F	171	55	G, V
Luong & Wormith (2011)	FP	JA	Court adjudicated youth	C	LSI-SK	Y	M/F	123	69	G
Matz et al. (2021)	FP	JA	Court adjudicated youth	U	YASI	Y	M/F	20	93	G
McCuish et al. (2017)	FTF-P	JA	Court adjudicated youth	C	PCL: YV	Y	M/F	137	312	G, V
McKinnon (2004)	FP	UTD	Court adjudicated youth	C	YLS/CMI	Y	M/F	28	74	G, V
Muir et al. (2020)	FP	JA	Court adjudicated youth	C	SAVRY total, protect	Y	M	219	357	G, V
							F	80	88	G, V
Meyers & Schmidt (2008)	AR	JA	Court adjudicated youth	C	SAVRY	Y	M/F	38	83	V
Myer, A. J. (2019)	FP	JA	Releasees convicted for sex offenses	U	Static-99R	A	M	123	761	S
Nafekeh & Motiuk (2002)	ATM/FP	GR	Federally sentenced men and women	C	SIR Proxy, SIR-R1	A	M/F	1,211	6,881	G, V, S
O'Byrne, P. L. (2004).	FP	UTD	Provincially sentenced men/women (SK)	C	ORAMS PRA	A	M/F	1,560	667	G
Olver, Cortoni et al. (2021)	AR/FP	UD	Mixed treated/untreated sample of men federally sentenced for violent offenses	C	VRS total, dynamic, change	A	M	406	550	G, V
Olver, Lewis, et al. (2013)	AR	JA	Treated high risk-need adult male violent offending corrections sample	C	PCL-R	A	M	86	66	G, V
Olver, Kingston, & Sowden (2021)	AR/FP	JA	Multisite treated sample of men federally sentenced for sexual offenses	C	Sexual deviance, Gen criminality	A	M	384	660	S, V, G
Olver, Nicholaichuk, et al. (2014, 2020)	FP	JA/UD	Multisite treated sample of men federally sentenced for sexual offenses	C	VRS-SO total, dynamic, change, Static-99R, Static 2002-R	A	M	197	365	S
Olver, Neumann et al. (2013)	AR	JA	Federally sentenced male prison inmates	C	PCL-R	A	M	171	264	G, V
Olver, Neumann et al. (2018)	AR	JA	Federally sentenced male prison inmates	C	PCL-R	A	M	452	711	G, V
Olver & Riemer (2021)	AR	JA	Treated federally sentenced men convicted for sexual offenses	C	PCL-R, SAPROF	A	M	185	236	G, V
Olver, Sowden et al. (2018) <sup>a,b</sup>	AR/FP	JA/UD	Treated federally sentenced men convicted for sexual offenses	C	VRS-SO total, dynamic, change, Static-99R, Static-2002R	A	M	384	660	S
Olver, Stockdale, et al. (2012)	AR/FP	JA	Court adjudicated youth	C	YLS/CMI	Y	M/F	104	40	G, V
Olver et al. (2009)	MA	JA	Court adjudicated youth	C	LSI youth variants	A/Y	M/F	860	462	G
Olver, Stockdale, et al. (2014)	MA	JA	Adult/youth corrections/forensic samples	C, A, U	LSI variants	A/Y	M/F	5,354	40,989	G
Olver & Wong (2015)	AR	JA	Multisite federal male custody sample	C	PCL-R	A	M	86	187	G, V
Olver et al. (2007)	AR	JA	Treated high risk-need adult male sexual offending correctional sample	C	VRS-SO total, dynamic, change, Static-99R	A	M	108	209	S, V, G
Prince et al. (2021)	FP	JA	Court adjudicated youth	C	LSI-SK	Y	M	148	50	G, V
							F	55	-	G, V
Rector et al. (2007)	FP	CP	Court adjudicated youth	C	LSI-SK	Y	M/F	628	244	G
Rettinger & Andrews (2010)	FP	JA	Adult women, Provincial corrections (ON)	C	LS/CMI	A	F	48	307	G
Schmidt et al. (2006)	AR	JA	Court adjudicated youth	C	PCL: YV	Y	M/F	40	90	G, V

Schwalbe (2009)	FP	JA	Court adjudicated youth community samples (3 fiscal cohorts)	U	ARNA	Y	M/F	545	4,464	G
								494	4,402	G
								537	4,079	G
Schwalbe (2009)/LeCroy et al. (1998)	FP	JA/GR	Court adjudicated youth (community)	U	ARNA	Y	M/F	448	5,225	G
Sewall (2015)	AR	UTD	Treated high risk-need adult male sexual offending correctional sample	C	PCL-R	A	M	109	196	G, V
Shepherd et al. (2018)	FTF-P	JA	Adults sentenced/remanded to custody	A	Cultural engagement, identity	A	M/F	84	-	V
Shepherd et al. (2014)	FTF-P	JA	Youth sentenced/remanded to custody	A	SAVRY total, protect	Y	M	32	84	G, V
Shepherd et al. (2015)	FTF-P	JA	Youth sentenced/remanded to custody	A	YLS/CMI	Y	M/F	40	85	G, V
Shepherd & Strand (2016)	FTF-P	JA	Youth sentenced/remanded to custody	A	PCL: YV	Y	M	32	84	G, V
Smallbone & Ralling (2013)	FP	JA	Men convicted for sexual offenses released from custody	A	Static-99R	A	M	67	320	S, V
Sowden & Olver (2017)	AR	JA	Treated high risk-need adult male sexual offending correctional sample	C	VRS-SO total, dynamic, change, STABLE 2007, Static-99R	A	M	79	86	S, V, G
Spiranovic (2012)	FP	GR	Men convicted for sexual offenses on parole/community supervision	A	Static-99R	A	M	162	660	S
Stewart et al. (2017)	FP	GR	Federally sentenced men in custody	C	DFIA-R, E/E, Comp., SA, Att., CF, F/M, P/E	A	M	3710	12,590	G
Stockdale et al. (2010)	AR/FP	JA	Court adjudicated youth	C	PCL: YV	Y	M/F	101	39	G, V
Stockdale et al. (2014)	AR	JA	Court adjudicated youth (violence)	C	VRS-YV	Y	M/F	91	37	G, V
Thompson & McGrath (2012)	FP	JA	Court adjudicated youth	A	YLS/CMI-AA	Y	M/F	1,053	1,582	G
Towers (2020)	FP	UTD	Custody and probation correction sample	U	LSI-R	A	M/F	798	2,900	G
Watkins (2011)	FP	GR	Custody, community, and hybrid adult male and female correctional samples	A	LSI-R	A	M	2,465	7,555	G
							F	417	614	G
Wilson & Gutierrez (2014)	MA	JA	Adult/youth correctional/forensic samples	C, A, U	LSI variants, CH, AP, E/E, Comp, F/M, L/R, SA, Att.	A/Y	M/F	21,508	58,719	G, V
Wilton et al. (2017)	FP	GR	Federally sentenced men in custody	C	DFIA-R (mechanical)	A	M	4640	15,487	G
Wong and Gordon (2006)	AR	JA	Federal custody multisite sample of men in sentenced for violent offenses	C	GSIR, VRS total, dynamic	A	M	173	272	G, V
Wormith & Olver (2002)	AR	JA	Treated high risk-need adult male violent offending corrections sample	C	GSIR	A	M	48	43	G, V

*Note:* Design: FP = field prospective (ratings completed in field as routine service delivery and individuals followed up in real time), AR = archival retrospective (ratings completed from past institutional file blind to outcome on individuals currently released), AR/FP = hybrid sample of archivally coded and clinically rated cases, FTF-P = face-to-face prospective (ratings completed on individuals face-to-face in a research context and followed up in real time), MA = meta-analysis (quantitative review), ATM-FP = hybrid sample of automated (SIR-Proxy) and field scored (SIR-R1) measures; Source: GR (government report), JA (refereed journal article), UTD (unpublished thesis/dissertation), UD (unpublished data from

exiting study), CP (conference presentation); DG = developmental group, adult (A), youth (Y), adult and youth combined (Y/A); Sex: male (M), female (F), male and female combined (M/F); I n = Indigenous n, NIW n = non-Indigenous White majority n; RC = recidivism criterion, general (G), violent (V), sexual (S), intimate partner violence (IPV); Measures: LSI: Level of Service Inventory; LSI-R: Level of Service Inventory-Revised; LS/CMI: Level of Service/Case Management Inventory; LSI-OR Level of Service Inventory-Ontario Revision; MNRF: Manitoba Risk-Need Form; YLS/CMI: Youth Level of Service/Case Management Inventory (AA: Australian Adaptation); LSI-SK: Level of Service Inventory-Saskatchewan Youth Edition; SFA: Static Factors Assessment; DFIA-R: Dynamic Factors Identification and Analysis; SIR: Statistical Information on Recidivism Scale; ORAMS PRA: Offender Risk Assessment and Management System Primary Risk Assessment; ARNA: Arizona Risk/Needs Assessment Instrument; SPIn: Service Planning Instrument; YASI: Youth Assessment and Screening Instrument; ODARA: Ontario Domestic Assault Risk Assessment; SARA: Spousal Assault Risk Assessment Guide; VRS: Violence Risk Scale; VRS-SO: Violence Risk Scale-Sexual Offense Version; VRS-YV Violence Risk Scale-Youth Version; SAVRY: Structured Assessment of Violence Risk for Youth; SAPROF: Structured Assessment of Protective Factors; SAPROF-YV: Structured Assessment of Protective Factors-Youth Version; PCL-R: Psychopathy Checklist-Revised; PCL: YV Psychopathy Checklist: Youth Version. CH: criminal history; E/E: education/employment; Comp.: companions; SA: substance abuse; Att.: procriminal attitudes; AP: antisocial pattern; CF: community functioning; F/M: family/marital; L/R: leisure/recreation; P/E: personal/emotional; Em dis: emotional distress; Prior vic: prior victimization

Supplemental Table S3

*Unweighted Recidivism Rates for Indigenous and non-Indigenous as a Function of Assessment Measure*

Measure	Indigenous			Non-Indigenous		
	General (k)	Violent (k)	Sexual (k)	General (k)	Violent (k)	Sexual (k)
LSI (adult)	.400 (22)	.150 (3)	-	.333 (21)	.145 (3)	-
LSI (youth)	.618 (12)	.506 (6)	-	.499 (11)	.336 (6)	-
DFIA-R	.559 (2)	-	-	.359 (1)	-	-
SFA	.359 (1)	.166 (1)	-	.252 (1)	.106 (1)	-
SIR	.453 (8)	.161 (6)	.011 (2)	.284 (6)	.132 (6)	.005 (2)
SPIIn/YASI	.441 (3)	.125 (1)	-	.278 (3)	.087 (1)	-
SAPROF/YV	.642 (2)	.437 (2)	-	.437 (2)	.228 (2)	-
SAVRY	.638 (4)	.417 (5)	-	.479 (4)	.259(5)	-
VRS-YV	.700 (2)	.479 (2)	-	.443 (2)	.217 (2)	-
VRS	.594 (4)	.401 (4)	-	.470 (4)	.287 (4)	-
ODARA	.748 (1)	.615 (V, 1); .378 (IPV, 2)	-	.500 (1)	.409 (V,1); .184 (IPV, 2)	-
SARA-V2/3	.391 (1)	.430 (V, 2); .191 (IPV, 1)	-	.252 (1)	.237 (V,2); .106 (IPV, 1)	-
ORAMS PRA	.692 (1)	-	-	.402 (1)	-	-
ARNA	.480 (1)	-	-	.517 (1)	-	-
VRS-SO	.614 (3)	.399 (3)	.154 (3)	.371 (3)	.193 (3)	.110 (3)
STABLE 2007	.467(2)	.304 (2)	.089 (2)	.218 (2)	.139 (2)	.059 (2)
Static-99R	-	-	.159 (5yr, 10); .216 (10yr, 5)	-	-	.113 (5yr, 10); .182 (10yr, 5)
Static-2002R	-	-	.183 (5yr, 4); .233 (10yr, 3)	-	-	.135 (5yr, 4); .195 (10yr, 3)
PCL-R	.656 (5)	.457 (5)	-	.463 (5)	.295 (5)	-
PCL: YV	.726 (5)	.452 (5)	-	.548 (5)	.326 (5)	-
Overall	.515 (70)	.366 (38)	.136 (15)	.389 (65)	.236 (37)	.096 (15)

Note: Non-overlapping samples are included in unweighted mean computation, hence individual instrument k total does not sum to overall k total; Recidivism criteria are general, violent, and sexual; k = number of studies; LSI: Level of Service Inventory measures; SFA: Static Factors Assessment; DFIA-R: Dynamic Factors Identification and Analysis; SIR: Statistical Information on Recidivism Scale; ORAMS PRA: Offender Risk and Management System Primary Risk Assessment; ARNA: Arizona Risk/Needs Assessment instrument; ODARA: Ontario Domestic Assault Risk Assessment; SARA-V2/3: Spousal Assault Risk Assessment guide versions 2 and 3; VRS: Violence Risk Scale; VRS-SO: Violence Risk Scale-Sexual Offense Version; VRS-YV Violence Risk Scale-Youth Version; SAVRY: Structured Assessment of Violence Risk for Youth; SAPROF/YV: Structured Assessment of Protective Factors and the SAPROF-Youth Version; SPIIn: Service Planning Instrument; YASI: Youth Assessment and Screening Instrument; PCL-R: Psychopathy Checklist-Revised; PCL: YV Psychopathy Checklist: Youth Version.