USE OF DYNAMIC RISK INSTRUMENTS TO ASSESS SEXUAL VIOLENCE RISK IN A
COMMUNITY-SUPERVISED SAMPLE OF MEN WITH SEXUAL OFFENSE CONVICTIONS
A Dissertation Submitted to the College of
Graduate and Postdoctoral Studies
In Partial Fulfillment of the Requirements
For the Degree of Doctor of Philosophy
In the Department of Psychology
University of Saskatchewan
Saskatoon

By

Farron E. Wielinga

Permission to Use

In presenting this dissertation in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this dissertation in any manner, in whole or in part, for the purpose of scholarly purposes may be granted by the professor or professors who supervised my dissertation work it, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this dissertation thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my dissertation.

Requests for permission to copy or to make other uses of materials in this dissertation in whole or in part should be addressed to:

Head of the Department of Psychology and Health Studies
9 Campus Drive
University of Saskatchewan
Saskatoon, Saskatchewan S7N 5A5
Canada

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan S7N 5C9
Canada

Abstract

The present study examined the predictive validity and psychometric properties of several actuarial risk measures developed to estimate likelihood of sexual recidivism, and one protective factor measure developed to assess protective factors related to desistance from sexual offending. Each of Static-99R (Helmus et al., 2012), Violence Risk Scale - Sexual Offense Version (VRS-SO; Wong et al., 2003-2017), STABLE-2007 (Hanson et al., 2007), Sex Offender Treatment Intervention Progress Scale (SOTIPS; McGrath et al., 2012), and Structured Assessment of Protective Factors for Violence Risk-Sexual Offence version (SAPROF-SO; Willis et al., 2017-2020) was rated based on file information of 200 community-supervised men with sexual offense convictions who were court mandated to receive assessments (and often treatment) at an outpatient forensic clinic in Edmonton, Alberta. Recidivism information was available for 172 men; mean follow-up time was 8.6 years. Nine percent of the sample was charged or convicted of a new sexual offense, 18.5% for any new violent (including sexual) charge or conviction, and 33% for any new charge or conviction. Predictive validity for all tools was obtained with respect to sexual, violent (including sexual), and general recidivism. All measures significantly predicted sexual (AUC = .65- .72) recidivism and dynamic measures were sometimes incremental to static measures in the prediction of sexual recidivism, depending on the pairing of predictors. An exploration of the structural properties of the VRS-SO dynamic items revealed a three-factor solution isomorphic to previous research (Olver et al., 2007; Olver & Eher, 2019). Further discrimination and calibration findings will be discussed, including implications for assessment and treatment of men convicted of sexual offenses.

Acknowledgements

Mark, your enthusiasm for this research from our first brainstorming session formed the impetus to help me see this dissertation through. Thank you for your direction in this venture, and for all you have done to support your students and Steve Wormith's students over the past years. Thank you to my committee: Drs. John Hansen, Paulette Hunter, Kiera Stockdale, and external examiner Dr. Alberto Choy. Your time, thoughtful questions, and contributions greatly improved this research.

I acknowledge the following financial support for this research: Joseph Armand Bombardier Canada Graduate Scholarship Masters, Dean's scholarship, Wanda Young scholarship, and University of Saskatchewan Centre for Forensic Behavioural Sciences and Justice Studies.

To Sandy – forensic psychology was not an obvious choice when I was looking for a research area in my undergraduate degree – I was drawn to you as a supervisor, and the research interest came later! Your encouragement and guidance were foundational in getting me to where I am today. I think of you with the utmost fondness.

To my parents and siblings – you have seen me through some peaks and some very dark and long valleys before I started postsecondary. Thank you for sticking by me and waiting for brighter days to emerge; I am immensely privileged to have you all. Mum and Dad – your support, especially in this past year – has made life much richer and less stressful. To Matthew, Aleah, and Annelise, thank you for the many conversations and diversionary trips along the way.

August, Robi, Carissa, Ashley, and Kendall: In the thick of the work, you all made this journey sparkle in a way that is only attained by having meaningful personal connections. Thank you for seeing me through and helping me maintain perspective. A specific thank you to Carissa – you have provided so much guidance for me over the years, and I am so grateful for your friendship. It is so special to now share our journeys in motherhood. Thank you also to Daina, Roberta, Erica, Deepthi, Alana, and Roxanne, whom I love dearly. You have all greatly enriched my life over this past decade.

Martina, you were my research assistant during this undertaking, but more importantly, you are my friend. I cannot think of many other people that I could have spent months of my life reading assessment files with (in a poorly ventilated, cold, and harshly lit room, listening to your endless chewing noises). Thank you also to Christine Sribney for your excitement and commitment to this project and your guidance during data collection at FACS.

My greatest thanks go out to my partner and children, who help me maintain perspective. Shane, your unwavering care, and unreserved confidence in my ability has been one of your greatest gifts to me.

"We grow. It hurts at first." – Sylvia Plath, from the Collected Poems. "Witch Burning," c. Oct 1961.

Dedication

For Shane: You show me every day what it means to be both held and free.

F + I: You are – and will always remain – my greatest teachers.

Table of Contents

List of Figures

List of Appendices

Table of Abbreviations

Abbreviation	Explanation
2SLGBTQ	Two-Spirit, Lesbian, Gay, Bisexual, Transgender, and Queer
AUC	Area Under the Curve
CPIC	Canadian Police Information Centre
FNIM	First Nations, Inuit, and Métis
GLM	Good Lives Model
GPCSL	General Personality and Cognitive Social Learning perspective
ICCs	Intraclass Correlation Coefficients
MFM	Motivation Facilitation Model
RNR	Risk-Need-Responsivity
ROC	Receiver Operating Characteristic Curve
SAPROF-SO	Structured Assessment of Protective Factors for Violence Risk – Sexual Offence Version
SOTIPS	Sex Offender Treatment Intervention and Progress Scale
SPJ	Structured Professional Judgment
VRAG-R	Violence Risk Appraisal Guide – Revised
VRS-SO	Violence Risk Scale – Sexual Offending version

Note. All abbreviations listed also have explanations included in text as they appear.

Use of Dynamic Risk Instruments to Assess Sexual Violence Risk in a Community-Supervised Sample of Men with Sexual Offense Convictions

CHAPTER ONE: INTRODUCTION

1.1 General Introduction

Despite increased efforts at prevention, sexual violence continues to be a prevalent social problem, with recent statistics indicating that sexual violence rates are consistent with those reported a decade ago (Conroy & Cotter, 2017). Women and children are at increased risk; one in three women are sexually assaulted at some point in their lifetimes, and one in ten children face sexual abuse before the age of 18 (Luce et al., 2010; Townsend & Rheingold, 2013). Risk of victimization is also elevated for people from ethnocultural minority groups, those who identify as 2SLGBTQ, and individuals with disabilities (Conroy & Cotter, 2017; Logie, et al., 2014). Though sexual violence varies in its level of severity (Sinha, 2013), victimization often contributes to deleterious consequences including physical complaints (e.g., physical injury, chronic pain, irritable bowel syndrome, sexually transmitted infections; Luce et al., 2010), mental health concerns (e.g., suicide attempts, Post Traumatic Stress Disorder, substance abuse; Luce et al., 2010), and relationship difficulties (e.g., intimate relationships with abusive partners; Dube et al., 2005). Society is also impacted through the potential for intergenerational trauma, as well as lost productivity in employment, and costs to the health care system resulting from long-term physical and mental health issues faced by those who have been victimized. The detrimental impact of sexual offending on its victims and society highlights the importance of working to prevent sexual violence from occurring in the first place. However, when it does occur, providing effective assessment and treatment for individuals who have perpetrated sexual violence is an important part of preventing future harm. Considering the above findings, the present research was designed to investigate risk and protective tools used with men who have sexually offended.

1.2 Development of Risk Assessment

1.2.1 Models of Persistence/Desistence from Offending and Effective Rehabilitation

Several models focus on effectively rehabilitating individuals with offending histories; two of the most utilized and debated models are the Risk-Need-Responsivity model (RNR; Andrews et al., 1990) and Good Lives Model (Willis et al., 2013). As the RNR model focuses

more heavily on risk and GLM focuses more on positive psychology, much debate exists between service providers as to the appropriate approach to management of individuals with sexual offending histories. These two models will be outlines and the debate described.

The RNR model states that reductions in sexual offending and its consequences are improved through appropriate assessment, treatment, and management of individuals at risk for perpetrating sexual violence. RNR bases its interventions on three principles: 1) Risk – the level of intervention should match the individual's level of risk, with the highest intensity of intervention reserved for the highest risk individuals; 2) Need – the focus of treatment should be targeting dynamic (i.e., changeable) factors that are empirically associated with risk (i.e., criminogenic needs); and 3) Responsivity – treatment should be delivered using modalities that maximize ability to benefit from intervention, tailoring interventions to an individual's learning style, motivations, abilities, and strengths.

The theory underpinning RNR is the General Personality and Cognitive Social Learning theory of criminal behaviour (GPCSL; Andrews & Bonta, 2006). GPCSL indicates that criminal behaviour reflects a personality predisposition and the learning of criminal behaviour directed by the expectations an individual holds, as well as the actual consequences of their behaviour. Behaviour that is rewarded or comes with the expectation of reward is likely to occur, while behaviour that is punished or comes with the expectation of punishment is unlikely to occur. The likelihood of criminal behaviour increased when rewards and costs for crime outweigh the rewards and costs for prosocial behaviour. Rewards and costs can be internal (e.g., emotions of self-efficacy or shame), others (e.g., family, friends, employers), or at times arise from the behaviour itself (e.g., obtaining income from selling stolen property, feelings of euphoria after ingesting a drug). Assessment of the "central eight" risk factors relevant to GPCSL lays the foundation for effective intervention by prioritizing service delivery towards risk factors that are linked to an individual's criminal behaviour. These "central eight" risk factors include criminal history, antisocial attitudes and peers, as well as substance use, leisure/recreation activities, family/marital relationships, and employment/education status.

RNR's effectiveness in reducing general recidivism is well-established (Andrews & Bonta, 2010), and evidence for its efficacy in reducing sexual recidivism is mounting. Two meta-analyses have found that adherence to RNR principles reduces sexual recidivism (Gannon, et al., 2019; Hanson et al., 2009). Hanson et al. (2009) found that the greatest reductions in sexual and

general recidivism occur when higher-risk individuals meaningfully engage in programs that focus on changing their criminogenic needs, with programs that adhere to all three principles generating the strongest effect in relation to decreased sexual recidivism (odds ratio [OR] = 0.21), followed by adherence to two (OR = 0.63), and one (OR = 0.64), versus none (1.17). Gannon et al.'s (2019) meta-analysis also found that higher quality treatment (i.e., using RNR or evidence-based principles) was associated with the greatest sexual recidivism reductions.

RNR principles are also relevant in community-based interventions and risk management; one study reported that adherence to all three principles translated to an average recidivism difference of 17% between treated and nontreated individuals in custody and 35% between treated and nontreated individuals in the community (Andrews & Bonta, 2006). Implementing RNR principles also assists in devising appropriate post-release risk management plans where evidence-based case management paired with community-based aftercare resulted in incremental reductions in post-release offending (Wilson et al., 2009), and clients whose probation officers utilize more RNR-based skills in community-based supervision sessions have lower recidivism rates (Bonta et al., 2010).

The Good Lives Model (GLM; Ward & Stewart, 2003; Ward, 2010) posits that criminal behaviour arises when individuals lack the internal and external resources to satisfy their needs prosocially, ultimately representing a maladaptive attempt to attain life goods (Ward & Steward, 2003). GLM posits that building strengths and capabilities in individuals with offending histories reduces recidivism risk. According to GLM, individuals with offense histories, just as all humans, value certain states of mind, personal characteristics, and experiences, defined by GLM as primary human goods. These human goods are proposed to fall into eleven categories, including relatedness (intimate, family, and friend relationships), excellence in agency and work, and creativity. It is assumed humans all seek to attain goods in each category but may prioritize them differently. Primary goods represent more conceptual things for which people strive, while secondary goods represent activities the individual engages in to achieve these ends. Within GLM, criminogenic or noncriminogenic needs exert control on secondary goods. Specifically, individuals may rely on antisocial means to achieve primary goods, which presents barriers towards satisfying primary goods.

Overall, as an approach to rehabilitation that draws from desistance research and positive psychology, GLM is a strength-based approach for individuals with sexual offending histories

that aims to equip individuals with skills, capabilities, and capacities to acquire primary goods and basic human needs in a personally fulfilling and socially acceptable way (Ward & Stewart, 2003). GLM focuses on providing opportunities to explore common human values, goals, and characteristics associated with psychological well-being rather than focusing on addressing deficits in functioning. This approach strives to support individuals in accessing positive psychological experiences such as well-being, safety, and social connectedness, opportunities which authors note are theoretically and empirically related to positive outcomes for individuals with offending histories (Ward & Stewart, 2003; Willis et al., 2013).

There has been a spirited debate between advocates of the RNR and GLM models. GLM was created, in part, as a response to the deficits-based focus of other rehabilitation models, such as RNR and the Relapse Prevention model. Ward and colleagues opposed the exclusive focus on risk management (i.e., a deficits-based approach) and asserted that rehabilitation instead needed to assist individuals with offending histories in achieving valuable life goods through construction of a meaningful life plan. Ward and colleagues criticized the RNR model, specifically focusing on how the model conceptualizes the needs of individuals it seeks to treat, stating that criminogenic needs are value laden, that criminogenic needs do not inform what clinicians do in therapy, that the RNR model does specify the relationship between the principles contained herein, that insufficient explanation exists in the model pertaining to why criminogenic needs arise, and RNR ignores noncriminogenic needs (Ward & Stewart, 2003). The major criticism that has been levelled against RNR is that it overemphasizes risk factors at the expense of helping individuals with offending histories live more fulfilling lives.

Proponents of RNR had many criticisms of GLM in response: namely, that the model is based on theory and the personal views of the authors rather than evidence, and that the popular appeal of a strength-based approach did not overcome the lack of evidence for the model's efficacy in rehabilitation of individuals with offending histories (Bonta & Andrews, 2003). These criticisms have continued over time, with advocates of each model continuing to voice their concerns with each model. Andrews et al. (2011) argued that RNR encompasses many features of the GLM model and that GLM's only contribution is weak assessment approaches and confusion in service planning and delivery. In response, Ward and colleagues (2012) attempted to explain the incremental value of GLM and conclude that the "model contains all the valuable elements of the RNR model, though the reverse is not true" (p.107), ultimately asserting that

GLM has wider scope and applicability than the RNR model. A second response from Wormith and colleagues (2012) highlights the weak empirical record of the GLM, in comparison to the RNR model, which has garnered robust research support, as indicated above.

1.2.2 Risk Assessment Generations

Adherence to the RNR model is predicated on accurate assessment of risk for future violence. Assessment tools used to inform risk for future violence have evolved significantly over the past half century and the approaches are often described as having four distinct generations.

The first generation of risk assessment involved unstructured clinical judgement and relied on the opinions and experience of professionals and correctional staff. This form of clinical decision making was fraught with bias and is commonly referred to as unstructured clinical judgment (UCJ), where decisions about an individual's risk level were based on an appraisal of case history information and the assessor's own professional experience. Criticism of the sole use of UCJ as a clinical decision-making strategy first emerged in 1954 when Paul Meehl compared clinical and statistical methods of prediction and concluded that clinicians were incapable of predicting behaviour more reliably than actuarial methods (Meehl, 1954). Nearly thirty years later, Monahan (1981) reasserted these criticisms, stating that forensic evaluators using UCJ were incorrect twice as often as they were correct. UCJ has a limited ability to distinguish between high and low risk individuals with offending histories; clinicians make these discernments with no more accuracy than intelligent but untrained laypeople (Hanson & Morton-Bourgon, 2009). These critiques and various U.S. court cases (e.g., *Baxstrom v. Herold*, 1966) provided an impetus to develop more empirically rigorous methods for risk assessment – namely, actuarial tools. This impetus was strengthened by recommendations to employ more stringent statistical analyses to predict violence that controlled for base rates and clinician bias (Mossman, 1994).

The second generation of risk assessment, utilizing static empirical-actuarial tools, addressed the limitations of previous assessment methods by focusing on more objective and reliable indicators that are statistically predictive of recidivism. Tools such as Static-99 (Hanson & Thornton, 2000), and Static-2002 (Hanson & Thornton, 2003) evaluate static (i.e., historical and unchangeable) risk factors, but are subject to their own criticisms, as such measures cannot be used to identify appropriate treatment targets or capture changes in risk as a result of

treatment (Andrews et al., 1990), and some tools were atheoretical (Bonta, 2002). Actuarial tools are still widely used to estimate the likelihood of recidivism by assigning a numeric value to factors empirically associated with offending, and the resulting scores provide probabilistic estimates of the individual's likelihood of reoffending.

Third generation assessment tools demonstrated several advancements beyond the second generation; most significantly, these tools incorporate dynamic (i.e., changeable) risk factors. Besides greatly increasing the comprehensiveness of a risk tool, inclusion of dynamic risk items addressed criticism of the inability of second generation tools to capture change that may occur within an individual as a product of time or treatment (Andrews & Bonta, 2006). By including dynamic risk factors, these tools are often referred to as "risk-need" tools, some of which are actuarial in nature given that they generate numeric ratings that are linked to recidivism estimates. However, they also incorporate treatment-relevant information, such as factors that require attention in treatment, and are malleable. Examples of third generation risk assessment instruments include the Structured Assessment of Risk and Treatability (START; Webster et al., 2006) and the Level of Service Inventory-Revised (LSI-R; Andrews & Bonta, 1995), both of which combine static risk factors with criminogenic needs (i.e., dynamic risk factors).

Fourth generation tools mark a shift from risk assessment to risk management, which incorporates not only an evaluation of risk level, but also an identification of interventions that can help lower that risk level; these measures too may be actuarial in nature. These tools are designed to be administered at various points during an individual's supervision in the justice system (Campbell et al., 2009). Two examples of fourth generation instruments are the Level of Service/Case Management Inventory (LS/CMI; Andrews et al., 2004) and the Violence Risk Scale – Sexual Offense Version (Wong et al., 2003-2017). Both the LS/CMI and VRS-SO are based on RNR principles. The LS/CMI has been found to predict general and sexual recidivism for individuals with sexual offending histories (Wormith et al., 2012), while the VRS-SO has moderate to high predictive validity for sexual and violent recidivism (Beggs & Grace, 2010; Olver et al., 2014; Olver, Mundt et al., 2018).

1.2.3 Methods for Determining Final Risk Classification

The advancing field of risk assessment has historically debated the comparative efficacy of actuarial and structured professional judgment (SPJ) approaches (Mills, 2017). Two important differences distinguish an SPJ approach from an actuarial one. First, while actuarial (particularly

static actuarial) approaches tend to use items that have an empirical relationship with risk, SPJ instruments include items that are either theoretically or empirically related to risk. Thus, some SPJ items, rather than being robust predictors of risk, are selected based on clinical experience and intuition (Hilton et al., 2006). Second, while actuarial approaches evaluate risk by summing items and comparing scores to norm groups, SPJ users are directed to assign idiosyncratic weights to individual items and label with a description of risk: either low, medium, or high risk (Hilton et al., 2006). The final appraisal of risk is grounded in the professional judgment of the evaluator rather than the numeric risk score. SPJ approaches have substantial flexibility for the consideration of additional, unmeasured factors in the final risk appraisal. While this flexibility is argued to be a strength of the approach, others have voiced concerns that this subjectivity is a limitation of the approach (Northcott, 2012).

Actuarial approaches have been shown to be more accurate over time, various settings, and decision-making scenarios (Ægisdóttir et al., 2006), and they tend to outperform SPJ guides in the prediction of sexual reoffending (Hanson & Morton-Bourgon, 2009). Mills (2017) notes that in SPJ validation studies, however, SPJ risk factors are summed and treated like actuarial measures, leading to a lack of clarity on the relative accuracy of actuarial and SPJ approaches, though Hanson and Morton-Bourgon's meta-analysis accounts for this, as summed SPJ tools were evaluated as a separate effect from true SPJ. Actuarial measures, however, are frequently typecast as placing emphasis on static factors compared to SPJ approaches, though such a criticism is misplaced given that actuarial tools may also include dynamic factors as with third and fourth generation tools.

Some tools may be used in an actuarial or SPJ fashion, such as the Structured Assessment of Protective Factors – Sexual Offence Version (SAPROF-SO; Willis, et al., 2018). There is considerable variability among risk assessment techniques and practices, with the most commonly used risk tools coming from different generations and using different approaches. However, best practice for sexual recidivism risk assessment considers factors that are specifically risked to sexual offending, using a validated actuarial risk scale that incorporates static and dynamic risk factors (Hanson & Bussière, 1998; Hanson & Morton-Bourgon, 2009).

1.2.4 Five-Level Council of State Governments Risk Scheme

A need for clarity in risk communication to assist in decision making and risk management created an impetus for development of standardized risk categories for criterion

referenced risk assessment tools. For several decades, these risk assessment tools typically categories risk level as low, medium, or high. Yet each measure has varying factors that are weighed differently from other measures, which poses a challenge as to how assessments conducted with different measures should be compared (Hanson et al., 2017). Given that there have previously been no specifications or standards that describe risk and need categories across measures, risk levels may have different definitions from one risk measure to the next, leading to a lack of consensus on what various levels mean regarding probability of reoffending (Hanson et al., 2017). Barbaree et al. (2006) found that, when five actuarial sexual offending risk instruments were compared, 3% of individuals were assessed as low risk across all instruments and 4% were assessed as high risk by all five measures. Jung et al. (2013) compared the number of individuals convicted of sexual offenses who were placed into risk categories on four sexual offense risk assessment instruments. Using percentile-derived categories, they found that the percentage of the sample in the low-risk category varied from 17.7% to 83.3%, the medium-risk varied from 12.2% to 67.3%, and the high risk varied from 2.7% to 15% (Jung et al., 2013).

In the past decade, the Council of State Governments (CSG) has made substantial efforts to adopt a common language for communicating the results of criterion-referenced risk assessment tools (Council of State Governments, 2014; Hanson, Bourgon et al., 2016). Hanson et al. (2017) adapted these guidelines for sex offense-specific factors, such as atypical sexual preferences, emotional identification with children, and sexualized coping (Barbaree et al. 2006). By integrating information about absolute risk (i.e., logistic regression derived risk estimates) and relative risk (i.e., percentile ranks and risk ratios), five risk levels were generated that represent practically meaningful and reliable differences between the categories. Level I (Very Low Risk) represent generally prosocial individuals with scant to nonexistence criminal backgrounds who represent the bottom 5% of the distribution in scores and have an estimated 5-year sexual recidivism rate of approximately 1%. Level II (Below Average Risk) represents a group that is higher risk than individuals who do not offend, but lower risk than the average individual who offenders, with some transient criminogenic needs that may require minimal intervention. Level III (Average Risk) represents a typical offending individual with criminogenic needs in several areas that require meaningful intervention to decrease recidivism risk. Level IVA (Above Average Risk) represents individuals who are meaningfully higher risk

than average and who often have chronic histories of rule violation, adverse child experiences, and criminogenic needs.

The CGS guidelines also have a Level V category, which encompasses individuals generally found in high-security units who require significant interventions for their current antisocial behaviour (Hanson et al., 2017). Their likelihood of reoffending is 85%, a base rate rarely observed for sexual violence. Instead, Hanson et al. (2017) derived a Level IVB (Well Above Average Risk) category, which corresponds to the top 5-10% of the scoring distribution, with 5-year sexual recidivism rates ranging from 20-50% depending on the sample. Each risk category is associated with successively higher criminogenic needs which require greater intervention requirements to manage risk.

The CSG five-level risk category framework has been used to generate updated risk categories for Static-99R, Static-2002R, and the VRS-SO (Hanson et al., 2017; Olver et al. 2018). The five-level system has been found to demonstrate 5% greater agreement for both risk-level placement and recidivism rates (Kroner & Derrick, 2020). Concern has been voiced about application of the five-level system to sexual offending risk tools using sexual recidivism as the outcome of interest, given that the five-level system was created with general reoffending as the outcome of interest, which means the criterion-based meaning of the labels has deviated from definitions presented in the original white paper (Hogan, 2021). Additional research finds that the five-level system leads to substantial increase in the number of individuals who received a higher risk rating, raising important questions about resource deployment and intensity of intervention delivered (Hogan & Sribney, 2019).

1.2.5 Risk Factors for Sexual Offending

In their seminal meta-analysis that investigated predictors of sexual recidivism, Hanson and Bussière (1998) reviewed 61 follow-up studies that had an average follow-up time of four to five years. Using reconviction as the recidivism criterion, they found an average sexual offense recidivism rate of 13.4% (18.9% for individuals with adult victims; 12.7% for those with child victims). For nonsexual violent recidivism, the average recidivism rate was 12.2% (22.1% for individuals who sexually offended against adults, 9.9% for those who sexually offended against children); average general recidivism rates were 36.3% (46.2% for individuals who sexually offended against adults, 36.9% for those who sexually offended against children). Results were quantified using correlations (i.e., *r*) with values between .10 and .20 considered small and

values greater than .20 and .30 considered moderate and large, respectively. Risk factors were broken down by into the following groups of variables: demographic, general criminality, sexual criminal history, sexual deviancy, clinical presentation and treatment history, developmental history, psychological maladjustment, and other psychological problems.

Demographic and general criminality variables such as younger age ($r = -.13$), single marital status ($r = .11$), antisocial personality disorder ($r = .17$), and prior offenses ($r = .12$) were found to be modest predictors of sexual recidivism, as were sexual criminal history variables such as prior sexual offenses ($r = .19$), having stranger victims ($r = .15$), and early onset of sexual offending ($r = .12$). The strongest predictors of sexual recidivism were measures of sexual deviancy such as phallometric assessment that showed arousal to children ($r = .32$) and any deviant sexual preference ($r = .22$). For the clinical presentation and treatment history, and psychological maladjustment groups of variables, failure to complete treatment ($r = .17$) and personality disorders ($r = .16$) were modest predictors, though the personality disorder variable included an unidentifiable number of individuals with antisocial personality disorder.

For nonsexual violent recidivism, significant demographic predictors included younger age ($r = -.24$), minority ethnicity ($r = .23$), and single marital status ($r = .10$), while significant general criminality predictors included juvenile delinquency ($r = .22$), and prior violent offenses, ($r = .21$), antisocial personality disorder ($r = .19$). Individuals who had previously sexually offended against a child (extrafamilial or intrafamilial) were less likely to violently recidivate ($r = -.16$ and $-.12$ respectively), than individuals with previous sexual offenses against adults ($r = .23$). There were no significant predictors for nonsexual violent recidivism from other groups of variables.

General recidivism could also be predicted by age ($r = -.16$), minority ethnicity ($r = .10$), single marital status ($r = .11$), nonsexual criminal history ($r = .20$ to $.28$), and antisocial personality disorder ($r = .16$). Failure to complete treatment ($r = .20$) and alcohol use during the offense ($r = .12$) were also predictors of general recidivism.

Hanson and Bussière (1998) also correlated each category of predictor with categories of recidivism. Criminal lifestyle was related to nonsexual violent ($r = .16$) and general recidivism ($r = .21$). Sexual deviance was related to sexual recidivism ($r = .19$), negative clinical presentation was only related to general recidivism ($r = .15$), and failure to complete treatment was related to

sexual (r = .17) and general recidivism (r = .20). Psychological maladjustment was unrelated to any form of recidivism.

A second meta-analysis was conducted as an update to Hanson and Bussière's 1998 study and was based on 82 studies. The average sexual recidivism rate was 13.7%; violent recidivism rate (including sexual and nonsexual violent) was 14.3% and general recidivism was 36.2%; average follow-up time was five to six years (Hanson & Morton-Bourgon, 2005). The authors considered factors from seven broad categories of risk factors: sexual deviancy, antisocial orientation, sexual attitudes, intimacy deficits, adverse childhood environment, general psychological problems, and clinical presentation. This meta-analysis focused in part on identifying possible dynamic risk factors in light of ongoing research that stressed the importance of targeting changeable risk factors (i.e., criminogenic needs) in order to reduce recidivism. The factors associated with recidivism were quantified according to recidivism type using Cohen's (1988) d values of .20, .50, and .80 to quantify small, moderate, and large effect sizes, respectively. The most promising dynamic risk factors were related to sexual deviancy, such as any deviant sexual interest (d = .31), sexual preoccupation (d = .39), and antisocial orientation, such as antisocial personality disorder (d = .21), psychopathy as measured by the Psychopathy Checklist revised (PCL-R, Hare, 2003) (d = .29), general self-regulation problems (d = .37), employment instability (d = .22), and hostility (d = .17).

Importantly, Hanson and Morton-Bourgon (2005) also outlined factors that seem intuitively connected to sexual recidivism but have trivial predictive utility. These potentially misleading risk factors include negative family background (e.g., childhood abuse or neglect), internalization of psychological problems (e.g., loneliness, low self-esteem), and poor clinical presentation (e.g., lack of victim empathy, denial, and low motivation). Thus, the two broad domains most strongly associated with sexual recidivism are sexual deviancy and antisocial orientation/lifestyle instability.

1.2.6 Dynamic Risk

The justice system is predicated on the notion that risk can be managed, contained, and ultimately changed; in essence, that risk is dynamic. Even risk factors that are considered static (e.g., age, criminal history) change in their relationship to recidivism as a function of time, with an individual's likelihood of recidivism showing an inverse relationship with age. Dynamic risk factors (or what Andrews and Bonta, 2010, term criminogenic needs) can be established if

change occurs following exposure to the change agent, and if the changes in the risk factor are associated with changes in the outcome (i.e., recidivism; Douglas & Skeem, 2005). Measurement of these potential dynamic predictors must occur at a minimum of two time points prior to measurement of the outcome variable.

While both static and dynamic risk tools have comparable predictive accuracy for sexual offense recidivism (Hanson & Morton-Bourgon, 2009) as they are markers for risk relevant propensities (Mann et al., 2010), dynamic risk tools have important added advantages. These tools are a collection of predictor variables that reflect important domains of psychological functioning (e.g., atypical sexual interests, distorted attitudes, relationship pathology) that are incremental to static measures in sexual violence prediction (Hanson et al., 2015). Dynamic risk tools speak specifically to criminogenic needs and are designed to inform dosage or intensity of treatments services (Bonta, 2002). Many clinicians find them useful for this purpose (Briken & Muller, 2014; Walker & O'Rourke, 2013). Dynamic measures can also be used to assess changes in risk through treatment or other change agents (e.g., aging, improvement in social supports; Andrews & Bonta, 2010). However, dynamic tools typically require more information and time to score, as evaluators need to integrate information from interviews, file review, and other collateral sources to assess the individual's overall functioning. They also require greater professional expertise, background knowledge, and training (Brankley et al., 2019).

There is strong empirical evidence for the existence of genuinely dynamic sexual violence risk factors. In their updated prospective multisite examination, Olver et al. (2020) evaluated the predictive accuracy of the VRS-SO dynamic items by computing change scores of 570 men before and after their participation in high-intensity treatment that specifically targeted sexual offending. Cox regression survival analyses revealed that change scores in the dynamic items were consistently associated with decreases in sexual and violent recidivism after controlling for baseline risk (e^B = .878 to .938), supporting the dynamism of sexual violence risk.

Additionally, in a 2018 meta-analysis, van den Berg et al. found that risk tools containing putatively dynamic variables incrementally predicted recidivism after controlling for static, criminal history variables. Based on a small number of studies (k = 6), they found that change scores on the VRS-SO (k = 4), SOTIPS (k = 1), and STABLE-2007 (k = 1) tools incrementally predicted sexual recidivism after controlling for both a) static risk factors, and b) the initial

scores of the dynamic risk tool with fixed-effect weighted d for change scores of 0.26, 95% CI [0.10, 0.42], $N = 2,043$.

In a 2019 meta-analysis, Brankley et al. evaluated the ability of STABLE-2007 to discriminate between individuals who recidivated and those who did not, and the extent to which STABLE-2007 improves prediction over and above Static-99R. Based on 21 studies (12 unique samples, $N = 6955$), the authors found that STABLE-2007 was significantly and incrementally related to sexual recidivism, violent (nonsexual) recidivism, violent (including sexual) recidivism, and any crime. Scores on individual STABLE-2007 items discriminated between individuals who sexually reoffended and those who did not.

Drawing attention to the assessment of acute (i.e., rapidly changing) dynamic risk factors, Babchishin and Hanson (2020) conducted a prospective investigation of 632 individuals who were on parole or probation for a sexually motivated offense in Canada or the United States (Alaska or Iowa). The ACUTE-2007 (Hanson et al., 2007), an interview- or file-review-based instrument designed to assess acute dynamic (i.e., rapidly changing) risk factors for sexual recidivism was used. The ACUTE-2007 assesses items such as access to victims, emotional collapse, changes in social supports, hostility, rejection of supervision, sexual preoccupations, and substance abuse. Using Cox regression survival analyses, the authors found that the most recent assessment was the best fitting model (BIC = 590.66) for sexual recidivism. Their results also demonstrated that risk to reoffend changes over time, that the best predictor of recidivism is the latest score or a rolling average of scores rather than the first, highest, or lowest score, and that risk levels can predict different patterns of change. Overall, this and the previously discussed studies highlight the dynamic nature of risk, and the importance of reassessing risk to obtain an accurate indication of an individual's current risk level. Additionally, there is significant evidence that tools utilizing dynamic risk factors to assess and track changes in risk status achieved greater predictive power than their counterparts that assessed only static risk.

1.2.7 Protective Factors for Desistance from Sexual Offending

Protective factors are receiving increasing empirical attention and are now sometimes included as a part of risk assessment, though it has been argued they have long been included as an under addressed component of RNR (Andrews & Bonta, 2006). This increased focus on protective factors marks a shift from a strictly deficits-based approach to a more holistic approach that includes strengths (Heffernan & Ward, 2019; Kelley et al.; 2020). Protective

factors can be defined as any characteristic of an individual, or their environment/situation, that reduces recidivism risk (de Vogel et al., 2012). They are theorized to protect an individual from stressors, prevent development of antisocial behaviour, or increase the likelihood of prosocial functioning (de Vries Robbé & Willis, 2017). Several potential protective factors have garnered empirical support with adult offending populations, such as prosocial support networks (Lodewijks et al., 2010; Ullrich & Coid, 2011), emotional support networks (Lodewijks et al., 2010), structured leisure activities (Bouman et al., 2010), stable accommodation and housing upon release (Ullrich & Coid, 2011), and strong coping skills (Lodewijks et al., 2010, Rennie & Dolan, 2010).

Just as some risk factors are unique to sexual offending, emerging evidence suggests that some protective factors, such as prosocial interests and healthy sexual interests, uniquely support desistance from sexual offending (de Vries Robbé et al., 2015; Nolan et al., 2022). Protective factors have demonstrated predictive utility for inpatient violence and self-harm, and sexual recidivism after treatment, with the combined use of risk and protective factors demonstrating incremental predictive validity for violent and sexually violent recidivism over the use of risk factors (de Vries Robbè et al., 2015; Judges, 2016). Yet there is significant controversy surrounding the consideration of protective factors in risk assessment, as some researchers maintain that protective factors are subsumed within factors that have been empirically and theoretically linked to recidivism and are unlikely distinct constructs from risk factors (Harris & Rice, 2015; Helmus, 2018). In a qualitative review of risk and protective factors for sexual violence, Tharp et al. (2013) found that all identified protective factors could be rephrased as risk factors (e.g., relational stability versus relational instability).

There is an intuitive appeal to including both risk and protective factors in intervention planning for individuals who are supervised in the community given that they will have more opportunity to access supports and work on increasing protective factors in tandem with reducing risk factors. Additionally, Miller (2006) has argued that broadening the focus of risk assessment to include protective factors may decrease negative bias in the assessing clinicians, potentially leading to more balanced reports and consideration of possible positive outcomes (e.g., getting a job, furthering one's education, or building prosocial connections). Assessing protective factors may also add incrementally to the accuracy of risk tools and recidivism prediction (de Vries Robbé et al., 2011; Coupland & Olver, 2020). Some researchers argue that given the low base

14

rates of recidivism among high-risk individuals, protective factors may partially differentiate between high-risk individuals who do and do not go on to reoffend (de Vries Robbé & Willis, 2017). Inclusion of protective factors may also have important implications for risk management, recidivism prevention, promotion of positive outcomes, and community reintegration. (Coupland & Olver, 2020). A recent survey found more than half of clinician respondents report completing empirically guided assessments of protective factors even though they are not completing a formal instrument (Kelley et al., 2020). Given that empirical work involving protective factors against sexual recidivism is still limited, further investigation is warranted.

1.2.8 Domains of Risk and Need/Pathways to Sexual Violence

Sexual violence typically has no single cause and rather is enacted as a result of multiple causal factors and pathways. In 1999, Knight presented a theory involving two dimensions for sexual coercion (not limited to recidivistic acts) based on various studies completing using U.S. participants. He delineated two independent etiological factors for sexual coercion: (a) sexual deviance/promiscuity, and (b) hypermasculinity; the latter factor includes characteristics such as hostile orientation, achieving gratification from domination, and viewing violence towards women as appropriate. This etiological model is largely consistent with the meta-analytic findings of Hanson and Bussiére (1998) and Hanson and Morton-Bourgon (2005) that support two pathways to sexual offending: sexual deviance, and antisocial orientation.

Doren's (2004) systematic review also found support for a multidimensional model for sexual recidivism that outlined several risk dimensions and the risk factors within them. The first dimension is sexual deviance, which Doren notes can be assessed with a pedophilia diagnosis, the Rapid Risk Assessment for Sexual Offense Recidivism (RRASOR; Hanson, 1997), and penile plethysmography. The psychopathy/general criminality dimension can be assessed using an antisocial personality disorder diagnosis, the Psychopathy Checklist Revised (Hare, 1991; 2003), and Static-99 (Hanson & Thornton, 1999). Some evidence was also found for a third pathway, which Doren describes as "Detachment" or "Immature" and includes variables such as: young age, single relationship status, stranger victim, nonsexual violence during index offense, and treatment resistance/failure. Doren concludes that sexual recidivism risk is multidimensional and is best assessed by using separate instruments to examine each dimension.

Based on a literature review, Thornton (2002) identified four domains of dynamic sexual violence risk: Sexual Interests (i.e., the direction and strength of sexual interests), Distorted

Attitudes (i.e., sets of beliefs about offenses, sexuality, or victims that serve to justify sexual offending), Socioaffective Functioning (i.e., ways of relating to others and to the motivating emotions felt in these interpersonal interactions, such as emotional congruence with children and lack of emotionally intimate relationships), and Self-Management (i.e., planning, problem solving, and impulse regulation in the service of long-term goal achievement). Thornton (2002) indicates that Structured Risk Assessment (SRA) includes Static Assessment, Initial Deviance Assessment (IDA), Evaluation of Progress, and Risk Management. The four identified dynamic risk domains are evaluated by the IDA. When Thornton (2002) validated three of the four identified domains on a sample of 158 individuals with sexual offenses against children, he found that those who sexually recidivated showed significantly more distorted attitudes, socioaffective dysfunction, and poorer self-management than individuals who did not recidivate. Further, he incorporated these psychometric indicators into a Deviance Classification and classified participants as Low, Moderate, or High Deviance. When participants were scored on Static-99, the Deviance Classification significantly predicted sexual recidivism (AUC = .78), as did Static-99 (AUC =.92).

In another study that evaluated all four domains and the SRA with 119 adult males convicted of sexual offenses, the Sexual Interests domain and Psychological Deviance Index predicted sexual recidivism independent of the Static-99, and the Distorted Attitudes and Self-Management domains demonstrated moderate predictive accuracy (Craig et al., 2007). Allan and colleagues (2007) also examined the relationship between dynamic risk factors and sexual recidivism in a New Zealand sample of 495 individuals with child sexual offense convictions. The individuals were all scored on Static-99 and completed a pre- and posttreatment test battery. Test battery data was factor analyzed and four dimensions emerged: Social Inadequacy, Sexual Interests, Anger/Hostility, and Pro-Offending Attitudes, all of which are highly consistent with previous literature (Thornton, 2002). When these dimensions were combined into an Overall Deviance Score, each dimension, including the overall score, significantly predicted sexual recidivism (AUC values from .60 to .76; when controlling for static risk, Sexual Interests, Pro-Offending Attitudes, and Overall Deviance remained significant predictors. The authors state that the results provide support for the following: the identified dimensions represent dynamic sexual risk factors, that psychometric measures can validly assess such factors, and dynamic factors add incrementally to static factors in the prediction of sexual recidivism. Overall, these

studies strengthen the case for combining dynamic and static risk factors to improve predictive accuracy for sexual offending.

More recently, Seto (2019) has outlined another theory outlining factors associated with onset of sexual offending – the Motivation-Facilitation Model (MFM). He highlights the notion that onset and persistence factors are not necessarily the same, offering the example of childhood sexual abuse, which is a factor associated with onset of sexual offending, but is not a significant predictor of sexual recidivism outcomes (Hanson & Bussière, 1998). This model is influenced by Finkelhor's (1984) preconditions model and Gottfredson and Hirschi's (1990) general theory of crime. Finkelhor (1984) identified three factors reflecting motivations to sexually offend against children: pedophilia, emotional congruence with children, and a perception of inability to have one's needs (emotional and sexual) met in adult relationships. These three factors are influenced by a fourth – overriding inhibitions to commit a sexual offense. In theory general theory of crime, Gottfredson and Hirshi (1990) posit that individuals who are low in self-control commit crime when opportunities exist – in the case of sexual offending against children, access to victims and situational factors.

MFM expands on the previously mentioned model and theory by specifying both primary sexual motivations and delineating trait and state facilitation factors. The model ultimately identified traits of paraphilia, high sex drive, and intense mating effort as primary motivations for sexual offending, and outlines trait (e.g., self regulation, hostile masculinity) and state (e.g., negative affect, alcohol use) factors that facilitate acting on sexual offending opportunities when opportunities are presented. These situational factors include access to a victim, presence of a guardian, location, and time.

Though MFM was developed to explain contact offending against children, it has been extended to explain child sexual exploitative material offending and online solicitation of adolescents. Seto argues the model is also relevant to explain sexual offending against adults and noncontact offenses such as exhibitionism and voyeurism, noting that different levels of motivation and facilitation factors may contribute to different types of sexual offending.

1.3 Review of Risk/Protective Measures for Sexual Offending

Building on the findings of Hanson and Bussière's (1998) meta-analysis, Hanson and Morton-Bourgon (2009) conducted a meta-analysis that examined the predictive accuracy of four different approaches to structuring risk instruments for sexual offending: empirical actuarial,

clinically adjusted actuarial, mechanical, and structured professional judgment (SPJ). They also investigated the predictive accuracy of measures that include factors intended to guide case management. Using 118 distinct samples from 110 studies, they found that the most accurate approaches for prediction of sexual recidivism were empirical actuarial measure specifically designed to predict sexual recidivism and any recidivism, and mechanical measures designed to predict sexual recidivism; SPJ approaches were found to be intermediate between actuarial approaches and unstructured professional judgment.

The authors emphasize that advancing the field of sexual violence risk assessment requires development of actuarial measures that capture causal, clinically relevant risk factors is necessary. They define these causal risk factors as dynamic variables that predict an outcome and, when manipulated, are associated with changes in that outcome. Hanson and Morton-Bourgon (2009) note that the development of fully actuarial measures containing causal, clinically relevant (i.e., dynamic) risk factors is essential for the advancement of sexual offending risk assessment. They indicate that the tools that most closely approximate this aim are the Vermont Sex Offender Treatment Needs and Progress Scale (McGrath et al., 2005) and the Violence Risk Scale – Sexual Offense Version (VRS-SO; Wong et al., 2003-2017).

Tully and colleagues echo the importance of dynamic risk assessment in their 2013 systematic review of 43 predictive accuracy studies of 15 actuarial and SPJ sexual violence risk assessment tools including the Static-99 (Hanson & Thornton, 1999), Static-2002 (Hanson & Thornton, 2003), VRS-SO (Wong et al., 2003, 2017), and SRA (Thornton, 2002). All reviewed tools produced at least a moderate effect size, with the SRA and VRS-SO evidencing AUCs of .74 and .76, respectively, which translate to large effect sizes. The authors note that these results tentatively support Hanson and Morton-Bourgon's (2009) statement that development of actuarial measures containing clinically relevant causal factors are the future of sexual offending risk assessment.

van den Berg et al. (2018) conducted a meta-analysis that examined the predictive properties, incremental predictive validity, and predictive validity of change scores for dynamic sexual offending risk instruments. These risk instruments and coding protocols included the Structured Risk Assessment – Forensic Version (SRA-FV; Thornton & Knight, 2015), STABLE-2007 (Hanson et al., 2007), Risk for Sexual Violence Protocol (RSVP; Hart et al., 2003), the Sex Offender Treatment Intervention and Progress Scale (McGrath et al., 2012), and the VRS-SO

(Wong et al., 2007). This meta-analysis found that dynamic tools predicted recidivism among men convicted of sexual offenses, though they did not consider the predictive accuracy of individual tools. Some of the dynamic risk measures were also demonstrated to be incremental to static factors (k = 52, N = 13 446). The meta-analysis considered sexual, violent, and nonviolent recidivism. However, there was significant variability between studies for violent recidivism and the incremental effect for any crime. For sexual recidivism, van den Berg et al. (2018) found that risk tools containing putatively dynamic variables incrementally predicted recidivism after controlling for static, criminal history variables.

1.3.1 Static-99R

Static-99R (Helmus et al., 2012) is an empirically derived actuarial scale comprised of historical factors that are linked to sexual recidivism risk. It serves as a revised version of Static-99, with adjusted weights for the "age at release" item that account for the relationship between age and recidivism. Static-99/R is the most commonly used actuarial scale for the prediction of sexual recidivism risk, owing to its ease of use by diverse professionals across a number of settings (e.g., correctional, community, civil commitment), using widely available information (e.g., criminal history, age of individual being assessed). Two related tools include Static-2002R (a 14-item static sexual offense risk measure, see Helmus et al., 2012) and, to assess risk for nonsexual recidivism among men with sexual offense histories, the Brief Assessment of Recidivism Risk 2002R (BARR-2002R; see Babchishin et al., 2016); the present review will focus on Static-99R, however, given its centrality in the present program of research. Static-99/R has been extensively researched, with both measures demonstrating strong interrater reliability (de Vogel et al., 2014; Hanson et al., 2014). Static-99R also demonstrates moderate predictive accuracy for sexual recidivism across a variety of samples, with area under the curve (AUC) values ranging from .73 in a Canadian sample (Brouillette-Alarie & Prouxl, 2013), .74 for Dutch sample (Smid et al., 2014), and .81 (Lee & Hanson, 2017) in an American sample.

Though First Nations, Inuit, Métis (FNIM), Black, and Asian individuals convicted of sexual offenses score higher and Hispanic individuals score lower on Static-99R than white individuals convicted of sexual offenses, these differences are not necessarily indicative of Static-99R predicting sexual recidivism with less accuracy between ethnocultural groups. Static-99R demonstrates good discrimination between individuals who recidivate and those who do not in ethnoculturally diverse populations, with AUCs ranging from .61 to .74 for individuals who

are Black, East Asian, Hispanic, Native American, and FNIM (Babchishin et al., 2012; Lee & Hanson, 2017; Lee et al., 2018; Myer, 2019; Smallbone & Rallings, 2013). However, Static-99R authors note that the variability of results should be a consideration when using risk tools with individuals from diverse ethnocultural groups who have sexually offended. Additionally, Static-99R is acceptable for use for individuals with developmental delays (AUC = .79; Hanson et al., 2013).

Factor analysis of Static-99R and 2002R items suggests three factors: Sexual Criminality, Detachment, and General Criminality; this result has been similarly replicated with two samples (Roberts et al., 2002; Brouillette-Alarie & Proulx, 2013).Though the two dimensions of sexual deviance and general criminality are well-supported elsewhere in the literature, authors from both studies note that the Detachment factor may simply reflect the effect of age, given that the three items loading on this factor (young age, absence of long-term cohabitation with an intimate partner, and sexual abuse of an unrelated/stranger victim). These items may speak to the lower likelihood that younger individuals are unlikely to be in long-term intimate relationships or have access to intrafamilial victims (Brouillette-Alarie & Proulx, 2013).

1.3.2 Violence Risk Scale – Sexual Offense Version (VRS-SO)

The VRS-SO (Wong et al., 2003–2017) is an empirically and conceptually derived actuarial scale containing static and dynamic items linked to sexual recidivism. As a fourth-generation risk assessment, the VRS-SO is designed to assess risk for sexual recidivism, identify treatment targets or risk management interventions, and to evaluate possible changes in risk from treatment or other change agents. The VRS-SO is the only sexual offending risk tool that provides empirically derived adjustments to initial risk ratings based on how the individual progresses (or not) in treatment (Olver, Beggs Christofferson et al., 2014). The VRS-SO is also unique in its defined method for assessing change across multiple administrations of the tool. Using a modified application of the transtheoretical model of change (TTM; Prochaska, et al., 1992), the VRS-SO rates dynamic items that have been deemed a treatment target according to the five stages of change (Precontemplation, Contemplation, Preparation, Action, and Maintenance). Positive pre-posttreatment changes (quantified using a rating system that incorporates TTM) have been associated with lower rates of sexual, violent, and general recidivism even after controlling for pretreatment risk level and length of follow-up (Beggs & Grace, 2011; Olver, Sowden, et al., 2018; Olver et al., 2020).

VRS-SO has demonstrated very good interrater reliability for total pre-and posttreatment dynamic scores (ICC = .90 and 0.92, respectively; Beggs & Grace, 2010). Predictive accuracy of the VRS-SO is also shown to be in the moderate range for samples with individuals who have committed sexual offenses against both adults and children for 5- and 10-year sexual (AUC = .73-.75) and violent (AUC = 70-.72) recidivism (Olver, Mundt et al., 2018) and with individuals who have sexually offended against children (AUC = .76; Eher et al., 2015). VRS-SO also has good discrimination between recidivists and non-recidivists in Indigenous populations, with AUCs of .67 and .68 for 5- and 10-year rates of sexual recidivism respectively (Olver, Sowden, et al., 2018). The dimensionality of VRS-SO dynamic items has been replicated across Canadian, New Zealand, and Austrian samples, and finds three factors: Sexual Deviance, Criminality, and Treatment Responsivity (Beggs & Grace, 2010; Olver & Eher, 2019; Olver et al., 2007).

1.3.3 STABLE-2007

STABLE-2007 (Hanson et al., 2007) is an empirically derived measure of dynamic risk for sexual recidivism; items were selected by identifying supervision problems that discriminated between adult males who had been convicted of a sexual offense who recidivated while on community supervision from those with similar static risk scores who did not recidivate (Hanson & Harris, 2000). STABLE-2007 is the most widely used dynamic sexual recidivism risk measure in North America (Bourgon et al., 2018) and has widespread support as a tool that assists in guiding case management and treatment decisions for individuals convicted of sexual offenses (Briken & Muller, 2014; Ryan et al., 2014; Walker & O'Rourke, 2013).

However, there are some inconsistent findings regarding the predictive validity of STABLE-2007. While STABLE-2007 has evidenced moderate predictive accuracy for sexual recidivism in some studies (AUC = .63 and .67, respectively; Etzler et al., 2020; Hanson, et al., 2015), Sowden and Olver (2017) found that STABLE-2007 predicted violent and general recidivism but was not significant for sexual recidivism for a group of high-risk incarcerated men who were receiving sex-offense specific treatment (AUC = .57). When combined with the Static-99R, the STABLE-2007 predicts incrementally over the Static-99R to varying degrees for sexual, violent, and any recidivism in community samples (AUC = .67, .66, and .66, respectively; Hanson et al., 2007; and AUC = .55, .68, and .66, respectively; Veith, 2018).

A recent meta-analysis (Brankley et al., 2019) sought to evaluate the predictive and incremental validity of STABLE-2007 and found that the measure predicted all recidivism types

21

and was incremental to Static-99R. As well, each STABLE-2007 item predicted sexual recidivism, bolstering its support as a measure that can inform treatment and supervision targets. Investigation of the dimensionality of STABLE-2007 reveals a three-factor solution: Antisociality, Sexual Deviance, and Hypersexuality; one item, Cooperation with Supervision, did not load on any factor and was excluded from further analysis (Etzler et al., 2020).

1.3.4 Sex Offender Treatment and Intervention Progress Scale (SOTIPS)

The SOTIPS (McGrath et al., 2012) is a dynamic sexual recidivism risk scale designed specifically to aid clinicians and probation or parole officers in identifying the supervision and treatment needs of men who are being supervised in the community following a sexual offense conviction. The SOTIPS has also been used as a collaborative treatment planning tool for men convicted of sexual offenses and their therapists (Lasher et al., 2015), and can be scored at multiple timepoints (i.e., at intake and thereafter in 6-month intervals).

In the development study, the SOTIPS showed moderate predictive accuracy for sexual and other violent recidivism (AUC = .70 and .66, respectively; McGrath et al., 2012). When combined with the Static-99R, predictive accuracy increases (AUC = .74 and .70, respectively; McGrath et al., 2012). Change scores on the SOTIPS (i.e., reductions in total scores over time) have also been associated with decreased likelihood of committing new sex offenses (McGrath et al., 2012), and improvements in aggression, sexual offending, and indicators of treatment compliance and change (Stinson et al., 2017). Hanson et al. (2020) prospectively evaluated the predictive validity of the SOTIPS over multiple time points, as scored by probation officers, and found that static SOTIPS scores (i.e., the first SOTIPS score) predicted all types of recidivism (sexual, violent, and general), with AUC values ranging from 0.59 to 0.70. Dynamic SOTIPS scores (i.e., at various time points), however, only improved the prediction of general recidivism, and only when high powered statistical analyses were used (Cox regression with time dependent covariates; Hanson et al., 2021). Initial exploration of the SOTIPS structural properties revealed three distinct factors: Sexual deviance, criminality, and social stability and support (McGrath et al., 2012). However, a recent replication study supported a two factor solution: Sexual risk and Antisocial opposition (Miner et al., 2022).

1.3.5 Structured Assessment of Protective Factors– Sexual Offence version (SAPROF-SO)

The SAPROF-SO (Willis et al., 2017-2020) is a measure designed to provide clinicians with a structured way of assessing hypothesized protective factors against sexual recidivism risk

tools; this tool is designed to be used in combination with commonly used risk assessment tools. The authors hypothesize that protective factors may differentiate between individuals assessed as high risk who go on to reoffend and those who do not. The developers intend for the SAPROF-SO to eventually be used as either an actuarial or an SPJ tool, but note further validation efforts are required before the SAPROF-SO is ready for use in actuarial form (Kelley et al., 2022).

There are two known studies that speak to the predictive accuracy of the SAPROF-SO. Nolan and colleagues (2022) found that the SAPROF-SO has very good predictive validity for sexual recidivism (AUC = .81) in a sample of 210 men convicted of sexual offenses against children in New Zealand and was incrementally predictive over Static-99R. de Vries Robbé and Olver(2021) also found that protective factors, as measured in the SAPROF family of tools, predict decreased recidivism, predict increased positive outcomes, and are often incrementally predictive of decreased recidivism beyond formalized risk measures. Further research has examined construct validity, interrater reliability (Willis et al., 2020), and factor structure of the SAPROF-SO (Willis & Thornton, 2021).

Validation studies have also been conducted using the Structured Assessment of Protective Factors Against Violence Risk (SAPROF; de Vogel et al., 2012), an SPJ tool from which the SAPROF-SO is derived. These studies demonstrate predictive validity for sexual and violent recidivism in a forensic psychiatric sample of men with histories of violence or sexual violence after controlling for risk assessment scores (de Vries Robbé et al., 2011; de Vries Robbé et al., 2015). For men with sexual offense histories, AUC values for total SAPROF scores were significant for follow-up periods of three years (AUC = .76) and fifteen years (AUC = .71), and uniquely predicted risk above the Historical Clinical Risk Management-20 (HCR-20; Webster et al., 1997). Coupland and Olver (2020) also found SAPROF to have incremental predictive validity to the Violence Risk Scale (VRS; Wong & Gordon, 2006).

1.4 Sexual Offense Recidivism Rates

The public often assumes that most individuals with general offense histories can stop offending and be reintegrated into the community, but there is a well-entrenched belief that individuals with sexual offending histories present a lifelong risk to public safety (Harris & Socia, 2016). This notion extends to policymakers (Sample & Kadleck, 2008), and those working in the criminal justice system (Zevitz & Farkas, 2000). However, this notion is not empirically supported. Though observed rates will always underestimate the actual rates of

sexual recidivism, research shows that sexual recidivism risk decreases over time (Harris & Hanson, 2004; Hanson et al., 2018) with sexual recidivism risk roughly halving for every five years an individual is in the community without being charged for a new sexual offense (Hanson et al., 2014).

The definition of recidivism is another important consideration in studies that aim to capture sexual recidivism rates. Some studies define recidivism as a reconviction for a sexual offense, while other studies include charges for sexual offenses, regardless of whether the individual went on to be convicted (Falshaw et al., 2003). Other studies elect to include informal indicators, such as arrests, self-report, reports to child protection agencies, conditional release violations (Marshall & Barbaree, 1988). As would be expected, broader definitions garner higher reported rates of recidivism, making it essential to specify recidivism criteria.

Harris and Hanson (2004) considered the question of sexual recidivism using ten individual samples from Canada, the U.S., England, and Wales ($N = 4724$) and included new charges or convictions for sexual offenses as their criteria, comparing sexual recidivism rates by victim type. Survival analyses indicated that overall recidivism rates were 14% after five years, 20% after ten years, and 24% after fifteen years. These sexual recidivism rates were found to be similar for those with adult victims (14%, 21%, and 24% for 5, 10, and 15 years, respectively) and those with child victims (13%, 18%, and 23% for 5, 10, and 15 years, respectively). However, there were significant differences in sexual recidivism rates between those with child victims, with the highest rates among those with extrafamilial male child victims (35% after 15 years), and the lowest for those with intrafamilial victims (13% after 15 years). The authors also found that individuals with prior sexual offenses had the highest reconviction rate, more than double that of individuals who had committed their first sexual offense (37% versus 19% after 15 years). Additionally, older individuals reoffended at roughly half the rate of younger individuals (12% versus 26% after 15 years).

In 2018, Hanson and colleagues used discrete time survival analysis to estimate hazard rates for a large, aggregated sample of individuals convicted of sexual offenses from diverse settings and risk levels ($N > 7000$). The samples came from Canada, the U.S., Austria, Denmark, Germany, New Zealand, and Sweden and were followed for up to 25 years. Life-table survival analyses indicated overall recidivism rates of 9.1% at five years, 13.3% at ten years, 16.2% at fifteen years, 18.2% at twenty years, and 18.5% at twenty-five years. The authors note that

though the cumulative recidivism rate increased, the 5-year hazard rate decreased at 9.1% up to five years, 4.1% between 5 and 10 years, 2.9% between 10 and 15 years, 2% between 15 and 20 years, and 0.3% between 20 and 25 years. They also report that each year an individual was offense-free was associated with a 12% decrease in the odds of recidivism (e [-.131] = .877), and that declines were observed for individuals at all risk levels, with the majority of individuals with sexual offending histories being no more likely to commit a new sexual crime than individuals with previous convictions for nonsexual crimes who have never been convicted of a sexual crime (1 to 2% after 5 years; Kahn et al., 2017).

1.5 Treatment for Individuals with Sexual Offense Convictions

Early studies examining psychological treatment's effectiveness at recidivism reduction for individuals convicted of sexual offenses indicate some degree of treatment effectiveness, but studies were often plagued with methodological concerns that prevented clear conclusions from being drawn (Alexander, 1999; Gallagher et al., 1999; Hall, 1995). Meta-analyses present one solution to the methodological shortfalls of single studies and overcome the limitation of low sexual recidivism base rates, given that statistical power can be increased by aggregating samples. To date, there are at least four comprehensive meta-analyses that best address the question of sex offense-specific treatment effectiveness.

Hanson et al. (2002) conducted a meta-analysis that examined effectiveness of specialized and non-specialized treatment in reducing recidivism in 43 studies (N = 9454), mostly from North America, with five studies from the U.K., and one from New Zealand. In 23 studies, treatment took place in an institution, 17 studies investigated community-based treatment, and three examined treatment effectiveness in both settings. Treatment effectiveness was comparable across settings. Hanson and colleagues found a significant unweighted average reduction in sexual recidivism (12.3% for treated group versus the comparison group 16.8%, OR = 0.81), and a similar pattern for general recidivism (27.9% for the treated versus 39.2% for comparison group, OR = 0.38). Cognitive behavioural (k =13) and systemic (k =2) treatment was associated with the greatest reductions in sexual (17.4% to 9.9%) and general recidivism (51% to 32%), while older treatments (prior to 1980) showed little effect. However, only four included studies provided sex offense specific treatment (OR = 0.61, 95% CI [0.45-0.82], Q = 32.79, df = 3, p < .001, n = 866). Nevertheless, this meta-analysis still supports Cognitive Behavioural

Treatment as significantly related to lower rates of sexual and general recidivism for individuals convicted of sexual offenses.

Lösel and Schmucker (2005) examined 69 studies evaluating treatment for individuals who had sexually offended ($N = 22\ 181$), most of which were conducted in North America. Over half of the treatment programs took place in institutions. Biological treatments such as surgical castration ($k = 8$, OR = 15.34, 95% CI 7.34-32.05) and hormonal medication ($k = 6$, OR = 3.08, 95% CI 1.40-6.79) had higher efficacy rates than psychological interventions Cognitive Behavioural ($k = 35$, OR = 1.45, 95% CI 1.12-1.86) and classic behavioural ($k = 7$, OR = 2.19, 95% CI 1.22-3.92), and only sexual offense specific treatment (k = 56, OR = 1.56, 95% CI 1.27-1.93) and outpatient treatment ($k = 27$, OR = 1.93, 95% CI 1.35-2.77) had significant effects. The effect size for sexual recidivism was highly significant (OR = 1.70), with an absolute difference is sexual recidivism between the treatment and comparison group of 6.4 percentage points, which translates to a 37% reduction from the base rate of the comparison group. The effect size for violent recidivism was also highly significant (OR = 1.90), and the average recidivism rate for the treatment group was 5.2 percentage points lower than for the comparison groups: a 44% reduction. For general recidivism, the mean effect size was OR = 1.67, with an average recidivism rate that was 11.1 percentage points lower than the comparison group, which translated into a 31% reduction.

In 2015, Schmucker and Lösel updated their meta-analysis, this time only including the highest quality research design (i.e., studies that were quasi-experimental with between group equality), which excluded any biological treatments. This meta-analysis included 27 studies (N = 10 387). Much smaller mean treatment effects for sexual recidivism were found (OR = 1.41), with recidivism reductions of 10.1% treated versus 13.7% untreated, translating to a relative reduction of 26.3%. For general recidivism, treatment effects were also small (OR = 1.45), with recidivism reductions of 32.6% for treated versus 41.2% untreated groups, and a relative recidivism reduction of 26.4%. Cognitive behavioural treatment had modest effectiveness (k = 20, OR = 1.38), though it was dwarfed by the large effects of Multisystemic treatment (k = 2, OR = 21.76). However, Multisystemic treatment was conducted for adolescents with convictions for sexual offenses, and targeted higher risk, younger individuals. Community programs, but not prison programs, significantly reduced sexual recidivism.

A meta-analysis conducted by Gannon et al. (2019) examined specialized psychological treatment and its association with reductions in offense-specific and non-offense specific recidivism, as well as the staff and program variables in the treatment programs. Sexual offending-specific treatment programs (k =44) evidenced stable and significant treatment effects, across both random (OR = 0.64, 95% CI = 0.53, 0.76) and fixed effects (OR = 0,65, 95% = 0.59, 0.72), though there was significant heterogeneity (Qs = 118.75, p < .001). The average follow-up time was 76.2 months (*SD* = 34.2), and sexual recidivism rates were 9.5% for the treated groups and 14.1% for untreated groups, with a relative decrease in recidivism of 32.6%. The authors also found larger positive treatment effects when a qualified psychologist was consistently present, with treatments across institutions and the community producing comparable recidivism reductions. Group based treatment was found to produce the greatest reductions in sexual recidivism, with treatment effects generally observed across programs of various length, though 100-200-hour programs (i.e., lower intensity) generated smaller effects. Finally, Gannon et al. (2019) found that programs provided in New Zealand, Australia, and Canada produced the most substantial reductions in sexual recidivism.

1.6 Treatment Change for Individuals Convicted of Sexual Offenses

Given the dynamic nature of sexual violence risk, it is both practical and ethical to document an individual's risk-related changes (Olver & Stockdale, 2020). Changes (i.e., decreases) in established dynamic risk factors tend to be associated with decreased recidivism for individuals with sexual offending histories (Olver et al., 2020). Though sexual offense specific treatment is the most prominent change agent, its effect depends on the quality of intervention (addresses salient risk factors; Hanson et al.,2009) the individual's response to treatment (responsivity principle; Olver et al., 2016), and having interventions that are facilitated by qualified psychological staff who receive regular supervision while facilitating treatment (Gannon et al., 2019).

Olver, Mundt et al. (2018) demonstrated how risk change information can be integrated into sexual violence risk assessments to non-arbitrarily adjust risk appraisals. Using logistic regression, modeling of 5- and 10- year estimates of sexual and violent recidivism employed specific risk and/or change information across six regression models, with each predictor incrementally predicting the target outcome. Using four combined samples, a robust change

effect was demonstrated: 10% to 14% reduction in odds of sexual violence for ever 1-point increase in change score, controlling for baseline risk.

Latent structure of change items in risk instruments may also differ from the latent structure of pretreatment dynamic risk. Olver et al. (2022) conducted factor analysis and item-level prediction analyses on VRS-SO pretreatment, posttreatment, and change ratings obtained from a sample of men convicted and treated for sexual offenses. They found that a two factor structure (Sexual Self-Management and Regulation of Antisociality) was more appropriate than the three factor structure of pretreatment items. Additionally, change scores from two of the three VRS-SO factors (Sexual Deviance and Criminality) uniquely predicted sexual recidivism. These results speak to the importance of measuring treatment change and the usefulness of the VRS-SO as a treatment planning tool.

Emerging evidence suggests that protective factors can also be used to evaluate treatment change. de Vries Robbé et al. (2015) provide an example of how protective factors can also be associated with decreases in violent (including sexual) recidivism. Pre- and posttreatment assessments of risk (using the HCR-20) and protective factors (using the Structured Assessment of Protective Factors; SAPROF; de Vogel et al., 2012) were compared for a sample of 108 individuals who had been discharged from forensic psychiatric hospital. The authors found that improvements on risk and protective factors during treatment showed moderate predictive validity for violent recidivism in short- (1 year) and long-term (11 year) follow-up.

Coupland and Olver (2020) also investigated the impact of protective factors and change scores on recidivism. 178 individuals who were federally incarcerated for a violent offense and had received high-intensity violence reduction programming were scored on a variety of measures, including the SAPROF, Protective Factors List (PF; a list of seven protective factors developed by the investigators), VRS, and the HCR-20. Changes on protective factors were measured at three time points and linked to possible changes in recidivism. They found significant associations between positive changes in protective factors and reductions in violent and general recidivism after controlling for baseline scores. Additionally, positive protective factor change scores (particularly internal and motivational factors) were associated with positive community outcomes. The authors concluded that protective factors are more than the inverse or risk factors and may be beneficial to include in violence risk assessment and treatment planning.

1.7 Evaluating Individuals who are Community Supervised

Evaluating and monitoring risk is especially relevant in community settings, yet the majority of sexual violence risk assessment research has been conducted with individuals who are incarcerated, illuminating the importance of extending this research into community-based samples. Community supervision offers individuals unique opportunities that are not available when incarcerated. Though there is certainly increased opportunity for reoffending, there is also increased opportunity for skill building and attainment of valuable life goods (i.e., stable housing, meaningful employment, healthy relationships) that may decrease risk for reoffending.

From a strictly risk-based perspective, community supervision can use recent information concerning an individual's community adjustment to predict recidivism. Babchishin and Hanson (2020) note that it is best practice to update assessments and adjust community supervision practices based on the client's most recent assessment, or the average of previous assessments. Given that risk may fluctuate more quickly for individuals who are community supervised, continued dynamic sexual violence risk assessment validation and replication efforts are necessary.

1.8 Present Program of Research

The purpose of the present program of research was to conduct a community-based evaluation of risk and its association with recidivism outcomes on Static-99R, VRS-SO, STABLE-2007, SOTIPS, and SAPROF-SO in a sample of community-supervised men participating in a sexual offense (SO) treatment program. This study is an extension of previous research (Sowden & Olver, 2017) and is the first to examine the association of the VRS-SO and SAPROF-SO to recidivism in a sample of non-incarcerated men with sexual offense convictions. The aim of this study is to contribute to the literature on dynamic sexual violence risk and protective factors, and examine risk-related, protective-related, and broader psychometric properties of four such dynamic instruments on a sample of men convicted of sexual offenses, the majority of whom subsequently attended SO treatment in the community.

1.9 Hypotheses

Based on our current empirical understanding of the relationship between risk to reoffend and recidivism, the following hypotheses were proposed for this program of research.

1. Construct (convergent) validity:

a. Static-99R, VRS-SO, STABLE-2007, and SOTIPS total scores will be moderately (r = .30-.49) to highly (r = .50+) positively correlated with each other.

b. SAPROF-SO total scores will be moderately to highly negatively correlated with Static-99R, VRS-SO, STABLE-2007, and SOTIPS total scores.

c. VRS-SO dynamic scores will be moderately to highly positively correlated with the STABLE-2007 and SOTIPS

d. VRS-SO dynamic scores will be moderately to highly negatively correlated with the SAPROF-SO

e. VRS-SO static scores will be moderately to highly positively correlated with Static-99R

2. Predictive validity:

a. Static-99R, VRS-SO, STABLE-2007, SOTIPS, and SAPROF-SO total scores will show good properties of discrimination for sexual recidivism (e.g., AUCs $\geq$ .64).

b. VRS-SO static, dynamic, factor, and total scores will each significantly discriminate sexual recidivists from non-recidivists.

c. The above measures are also anticipated to significantly discriminate violent (including sexual) and general (i.e., any) recidivists from non-recidivists

d. Static-99R, VRS-SO, STABLE-2007, and SOTIPS will also show good properties of calibration. That is, increasing scores and risk bands representing risk categories will be associated with successively higher rates of sexual recidivism.

3. Incremental predictive validity:

a. The VRS-SO dynamic total will demonstrate incremental predictive validity beyond Static-99R for sexual recidivism.

b. The VRS-SO dynamic total will demonstrate incremental predictive validity beyond the VRS-SO static total for sexual recidivism.

c. Each dynamic measure (VRS-SO dynamic total, STABLE-2007, SOTIPS, and SAPROF-SO) will demonstrate incremental predictive validity beyond Static-99R for sexual recidivism.

d. The SAPROF-SO will demonstrate incremental predictive validity beyond Static-99R and the VRS-SO dynamic total for sexual recidivism.

e. The above measures are anticipated to add incrementally to the prediction of violent (including sexual) and general (i.e., any) recidivism.

4. Construct (factorial) validity

a. A confirmatory factor analysis of VRS-SO will reveal a multidimensional factor structure resembling the three factor domains identified by previous research.

CHAPTER TWO: METHOD

2.1 Ethics

The present research was archival in nature, eliminating the necessity of the men's active participation. Ethical and operational approval to conduct the present research involved a four-step process. First, ethics approval was obtained from the University of Saskatchewan Behavioural Research Ethics Board (Beh-REB #907) and, second, granted from the University of Alberta Human Research Ethics Board (H-REB Pro00088518). Third, prior to implementing the research programme, operational approval to conduct research at Alberta Health Services was obtained from the Northern Alberta Clinical Trials and Research Centre (NACTRC) RA 90921. Lastly, a request was made to the RCMP National Headquarters to obtain federal criminal records following verification of Beh-REB, University of Alberta H-REB, and NACTRC approval.

The assessment and treatment files were property of Alberta Health Services – Forensic Assessment and Community Services (FACS) and were maintained onsite. All data collection and storage took place at FACS. A list of men who received outpatient services related to their sexual offending history (i.e., assessment and treatment) from 2009 to 2015 was created and their records were accessed to code the necessary information. A non-identifying study number was assigned to each man's file to ensure confidentiality. The master list containing the men's names, FACS file number, and study number was created to link raw data to the study data. The raw coded data of study measures are currently stored at the University of Saskatchewan – Forensic Lab in a locked filing cabinet. A database stripped of identifying information, but containing the study numbers, was maintained by the student investigator. Criminal records were obtained from the Canadian Police Information Centre (CPIC) through the Royal Canadian Mounted Police (RCMP). Raw CPIC data were securely stored under lock and key at the Forensic Lab under the supervision of the student investigator's research supervisor.

2.2 Current Sample

This sample included 200 community-supervised adult males who had been convicted of a sexual offense and received assessments at FACS, in Edmonton, Alberta between 2009 and 2015. For inclusion in this sample, case files needed to contain an assessment report with adequate information to code risk measures. Most of these men ($n = 195$) participated in sex offense-specific treatment at FACS, and available information indicated that just over half completed treatment ($n = 102$).

Participant characteristics are reported in Table 2.1. The participant mean age at the time of their index offense was 37.1 years (SD = 14.4), the mean age upon assessment at FACS was 39.6 years ($SD = 14.4$), and the mean age upon completion of treatment was 39.53 ($SD = 13.4$). Approximately 57% of the men were of White/European descent, 14.5% were Indigenous, and 28.5% encompassed a diverse number of ethnocultural groups including South Asian (5.0%), Central American (2.5%), Southeast Asian (2.5%), West Asian/Middle Eastern (2.5%), African (1.5%), and unidentified or mixed ancestry (14.5%).

In all, 36.5% of the sample were unemployed at the index offense, 57% were employed at the time of index, and this information was missing for 6.5% of the men. 37% of the sample had minimal or no problems with employment (i.e., full-time employment or school with stability and satisfaction), 15% had some problems (i.e., full-time employment with moderate dissatisfaction, two job changes in the last six months, or part-time work), 5.5% had considerable problems (i.e., three or more job changes in the last six months or unemployed more than half the time), and 41% had serious problems with employment (i.e., unemployed more than 80% of the time or if unable to work, uses free time unproductively). The average educational level obtained was 10.8 ($SD = 1.8$). Approximately 70% of the sample were assessed to have average cognitive functioning (i.e., full scale intelligence quotient scores greater than or equal to 90 and no evidence of cognitive impairment), 13.5% had low-average intelligence or average intelligence with cognitive impairments, and 14% had borderline or extremely low intelligence or marked cognitive impairment.

Overall, 40.5% of the sample were diagnosed as having a major mental disorder (not including substance use disorders), according to the Diagnostic and Statistical Manual of Mental Disorders. Of this subsample, 24% were assessed as having a non-sexual mental disorder (e.g., bipolar, depression, schizophrenia, PTSD) and 16.5% were assessed as having a paraphilia (e.g.,

pedophilia, pedohebephilia, exhibitionism). Fourteen percent of the sample were assessed as having a neurodevelopmental disorder (e.g., intellectual disability, Fetal Alcohol Spectrum Disorder, Attention Deficit Hyperactivity Disorder). Approximately 16% of the sample was assessed as having a substance use disorder (e.g., alcohol dependence or abuse); 9.5% as having a personality disorder or clinically significant traits of a personality disorder (e.g., most commonly antisocial personality disorder or borderline personality disorder); and 4% were specifically diagnosed with antisocial personality disorder or antisocial traits.

Regarding criminal history, 21.5% of the sample had at least one previous sexual offense conviction and 63.5% had no previous conviction, 15% of the sample had no information on file. The mean number of previous sexual convictions was 0.64 ($SD = 1.41$), with a range of 0-8. For individuals with a previous conviction, 23% had at least one previous violent conviction, 62% had no previous violent conviction, and 15% of the sample had no information on file. The mean number of previous violent convictions was 0.64 ($SD = 1.36$), with a range of 0-7. Of those with previous convictions, 55.5% had at least one previous non-sexual, non-violent conviction, 44.5% had no previous non-sexual, non-violent convictions. The mean number of previous non-sexual, non-violent convictions was 3 ($SD = 6.1$), with a range of 0-39. Regarding prior sentencing dates. 43% had at least one prior sentencing date for any offense, 41.5% had no prior sentencing date, and 15.5% did not have information on file. The mean number of prior sentencing dates was 2.3 ($SD = 3.9$), with a range of 0-20.

The index sexual offense (offense that brought individual into contact with FACS during 2009-2015 period) was coded according to the Criminal Code of Canada conviction. 41.5% of the sample had an index offense of sexual assault, 25.5% sexual interference, 8% was indecent act or exhibitionism, 3% was voyeurism, 8.5% child sexual abuse material-related offenses, 2.5% sexual exploitation, 3.5% luring a child, 2% sexual interference and sexual assault, 2% invitation to sexual touching and sexual interference, 1% possession of child sexual abuse material and luring a child, 1% sexual assault causing bodily harm, and for 1% of the sample the exact Criminal Code charge was not stated.

When considering risk level as conceptualized by Static-99R total scores, 4% of the sample was very low risk, 22% were below average risk, 44.5% were average risk, 23.5% were above average risk, and 6% were classified as well above average risk. 62% of the sample attended sex-offense treatment at FACS, with 51% successfully completing. Of the 11% who did

not complete, 6.5% of clients were discharged for program infractions (e.g., poor attendance, low motivation or effort, disrespectful behaviour, security concerns), 3% ceased attendance after probation expiry or because they requested to stop attending, and 1.5 % did not complete because they incurred new charges (e.g., breached, or a new offense and were returned to custody). The mean length of time in treatment was 6.3 months ($SD = 4.2$).

2.2.1 Sexual Offense Classification

Participants were classified into one of four groups in accordance with Violence Risk Scale – Sexual Offense Version descriptions of sexual offense victim profiles, that is, based on victim characteristics. Individuals with extrafamilial adult/teen victims; according to the VRS-SO, "adult/teen" is defined as 14 years of age or older (i.e., corresponding to emergence of secondary sex characteristics), individuals with extrafamilial child victims (i.e., approximately 13 years old or younger, corresponding to a relative absence of secondary sex characteristics), individuals with intrafamilial victims, and individuals with mixed adult/teen and child victims. According to these classifications, 46% of individuals had an extrafamilial adult/teen victim profile, 23.5% extrafamilial child victims, 25.5% had intrafamilial victims, and 4% had mixed adult/teen and child victim profiles.

Table 2.1

FACS Program Participant Characteristics

Measure	*M (SD)*	Frequency (*n*)
Demographics		
Age at index	37.0 (14.4)	-
Age at assessment	40.0 (14.4)	-
Age at treatment completion	40.0 (13.4)	-
Predominantly unemployed at index	-	73
Single/ never married	-	72
Highest grade completed	10.8 (1.8)	
Impaired cognitive abilities	-	28
Mental Health		
Major mood disorder	-	48
Paraphilia	-	33
Substance use disorder	-	32
Personality disorder	-	17
Antisocial personality disorder/traits	-	8
Neurological/developmental disorder	-	23
Criminal history		
Age at 1st sexual offense		-
Prior sexual offenses		-
Prior non-sexual violent		-
Prior non-sexual non-violent		-
Prior sentencing dates		-
Offense-related		
Above average/well above average risk	-	16 (VRS-SO)
		59 (Static-99R)
Adult victim	-	92
Extrafamilial child victim	-	47
Mixed victim	-	8
Intrafamilial child victim	-	51
Treatment-related		
Treatment length (months)	6.3 (4.2)	-
Successful completion	-	102
Unsuccessful completion	-	22

2.2.1.1 Comparison of Current Sample to VRS-SO Normative Sample. Compared to the VRS-SO normative sample, the current sample had mean Static-99R scores and mean VRS-SO dynamic and total pretreatment scores that fall one half standard deviation and over one standard deviation below the normative data, respectively. The normative and current sample are similar regarding victim age composition, with both samples evidencing approximately half of the men having child and half having adult victims. Additionally, individuals in the present sample may have participated in treatment while completing a custodial sentence prior to being released for community supervision, however VRS-SO ratings were made prior to receipt of any

community-based sexual offense-specific treatment. Comparatively, the VRS-SO normative data describes a treated sample. Sample composition also varies, with VRS-SO normative sample containing more Indigenous men (34.1% versus 14.5% in the present sample) but less ethnocultural diversity overall. Overall, the present sample is lower risk, untreated, with less criminogenic needs, similar proportions regarding victim profiles, a smaller proportion of Indigenous men, but a greater proportion of men from ethnoculturally diverse backgrounds than the normative VRS-SO sample.

2.2.2 FACS Sexual Offense Services

FACS conducts assessments and provides treatment services for individuals who are involved in the criminal justice system. There are several potential referral sources: Correctional Services division (i.e., probation officers); Correctional Service of Canada (i.e., parole officers); the Edmonton Police Service (specifically, the Behavioural Assessment Unit [BAU]), and self-referrals or community referrals from individuals who may not have legal involvement but have been deemed at risk for sexual offending. Services are provided by a multidisciplinary team, including nursing staff, medical doctors, psychologists, social workers, and occupational therapists and are designed to reduce and manage risk posed by individuals who have sexual offending histories or are at risk to sexually offend.

Treatment utilized within the Sexual Violence program is designed to reduce criminogenic needs and risk factors identified during the assessment process. Aligning with literature on effective treatment for sexually abusive behaviour (Gannon et al., 2019; Hanson et al., 2009; Schmucker & Lösel, 2015), treatment programming at FACS for sexually abusive behaviour follows behavioural and cognitive behavioural strategies and RNR principles. Treatment providers seek to help clients identify and make observable changes to the environmental, cognitive, affective, and interpersonal factors that increase their risk to engage in sexually abusive behaviours. With this, dynamic risk factors are addressed through interventions that may include education, prosocial monitoring (both from facilitators and other group members), graduated skill practice, behavioural rehearsal (in group, but also between sessions), redirection, role playing, positive reinforcement, interpersonal feedback, and detailed verbal guidance (e.g., suggestions regarding high-risk thoughts and restructuring such thoughts, resource provision, etc.). Most importantly, clients are assisted in developing and maintaining an explicit plan to avoid recidivism.

Most treatment interventions at FACS follow a group format, barring any responsivity issues (e.g., language barriers, acute psychiatric symptoms, active substance use concerns, personality/behavioural issues that are disruptive to group cohesion, etc.). In cases where any of the former concerns are present, alternative treatment plans are arranged, such as individual therapy. Individuals who attend Sexual Violence group at FACS are placed in one of four groups. Quest I and Quest II are cognitive-behavioural, skill-based groups designed for individuals with lower levels of intellectual functioning, learning disabilities, and/or problems in adaptive functioning. Quest I is reserved for individuals with low cognitive abilities and low adaptive functioning (e.g., low IQ and may not be able to live independently), while Quest II is designed for clients with low cognitive abilities and low to moderate adaptive functioning (e.g., low IQ or the presence of a significant learning disability yet may be able to live independently and participate in employment). Programming covers three components: feelings, thoughts, and behaviours. Both Quest I and II offer a minimum of 80 hours of treatment based on ten-month programs that meet for two hours weekly.

Sexual Offense Relapse Prevention Program (SORPP) divides into two groups depending on risk level. SORPP I is available to individuals assessed to be the in moderate to high-risk range based on a combination of dynamic and static factors. The group adopts a continuous intake format and runs year-round, one evening per week for two hours. Discharge dates for this group are based on participants' successful completion of modules corresponding to their specific criminogenic needs. SORPP II is available to individuals assessed to be in the low to low-moderate risk range when both static and dynamic risk factors are considered. This group incorporates approximately 30 hours of treatment, based on a nine-week, three hour per week format, with an individual pre- and post-group session.

2.3 Measures

Client information, in the form of hardcopy files, was accessed manually from the file room at FACS. Each file varied in content, potentially containing demographic information, general documentation about the perpetrator and victims, police reports, criminal records, presentence reports, and other criminal justice and/or mental health reports. This information was used to code the data collection protocol, forensic measures (Static-99R, VRS-SO, STABLE-2007, SOTIPS, and SAPROF-SO). Recidivism data was accessed through the national criminal record database, the Canadian Police Information Centre (CPIC).

2.3.1 Static-99R

Static-99R (Helmus et al., 2012), is an empirically derived actuarial tool designed to predict sexual recidivism risk in men convicted of sexual offenses. This scale has ten items assessing criminal history, victim characteristics, age, and relationship history. Individuals are placed into one of five standardized risk categories concerning absolute risk based on their total score (ranging from -3 to 12) with the categories numbered and labelled as follows: I – Very low risk (-3, -2), II – Below average risk (-1, 0), III – Average risk (1 to 3), IVA – Above average risk (4, 5), and IVB – Well above average risk (scores of 6+; Hanson et al., 2017). Risk categories correspond to the following extrapolated sexual recidivism rates at 10-years follow-up: Very low (1.2-1.8%), below average (2.5-2.6%), average (5.1-10.0%), above average (13.8-18.8%), and well above average (25.0-49.9%; Lee and Hanson, 2021).

2.3.2 Violence Risk Scale – Sexual Offense Version (VRS-SO)

The VRS-SO (Wong et al., 2003, 2017) is a clinician-rated scale designed to assess pre- and posttreatment risk for sexual recidivism. It is comprised of 7 static and 17 dynamic factors that are combined to form a total score. The Dynamic scale breaks down into three factors: Sexual Deviance, Criminality, and Treatment Responsivity. Each dynamic item is scored on a 4-point ordinal scale (ranging from 0 to 3), with higher scores indicating a stronger relationship with sexual offending. Dynamic items that are rated as a 2 or 3 are considered appropriate treatment targets. Recidivism risk is most frequently estimated using the total VRS-SO score (sum of static and dynamic variable ratings), although risk bands also exist for the static and dynamic scores. Total scores are translated into five standardized risk categories, in accordance with the common language risk categories: I – Very low risk <15), II – Below average risk (15-23.5), III – Average risk (24-39.5), IVA – Above average risk (40-49.5), and IVB – Well above average risk (50-67+; Olver et al., 2018). The risk categories correspond to the following base rates of sexual recidivism at 10-years follow up: I – Very low (1.0-3.2%), II – Below average (3.3-6.6%), III – Average (6.9-21.7%), IVA – Above average (22.4-39.3%), and IVB – Well above average (40.3-74.1%; Olver et al., 2018).

Static items are assessed at a single time point (pre-treatment), and dynamic items are assessed pre- and posttreatment. Changes in risk level due to treatment are assessed using a modified version of Prochaska et al.'s (1992) stages of change model. Each dynamic variable identified as a treatment target is rated both pre- and posttreatment to assess which stage of

38

change (i.e., precontemplation, contemplation, preparation, action, or maintenance) an individual is at. Change is quantified by comparing the stages of change rating at pretreatment with that at posttreatment. Progression from one stage to the next indicates a positive change and is scored as a 0.5 reduction in the pretreatment risk rating of the item. A reduction is not provided for progression from Precontemplation to Contemplation as no behavioural change is evident with this shift. The total change score is obtained by summing the change scores for all the dynamic items. The total VRS-SO scores obtained pre- and posttreatment represent an individual's level of risk to reoffend at two different points of time.

2.3.3 STABLE-2007

STABLE-2007 (Hanson et al., 2007) is an empirical actuarial risk tool of sex offense-specific criminogenic needs. It contains 13 items related to sexual self-regulation (e.g., sexualized coping, deviant sexual interests), general self-regulation (e.g., poor cognitive problem solving), social relationships (e.g., lack of concern for others), and cooperation with supervision. These items are each scored using a three-point rating system with 0 indicating "no problem", 1 indicating "some concern/slight problem", and 2 indicating "present/definite concern." Total scores are calculated by summing all item scores. Given that the emotional identification with children item is scored only for individuals with child victims, total scores can range from 0 to 26 individuals with at least one victim under the age of 14 years old, and 0 to 24 for individuals with other sexual offense types. Individuals will be assigned to one of three risk categories based on density of their criminogenic needs: low density (0-3), moderate density (4-11), or high density (12+). The STABLE-2007 risk categories can be combined with a static measure (i.e., Static-99R) to determine a risk rating that aligns with the five standardized risk categories (Brankley et al., 2017).

2.3.4 Sex Offender Treatment and Intervention Progress Scale (SOTIPS)

The SOTIPS (McGrath et al., 2012) is a measure specifically designed to assess sexual recidivism risk and treatment-related changes in community-supervised men convicted of sexual offenses. It contains 16 items grouped into three categories: (a) sexual deviance (e.g., offense responsibility, sexual behaviour, sexual attitudes, sexual interests, risk management, and stage of change), (b) criminality (e.g., criminal and rule breaking behaviour, cooperation with treatment, cooperation with supervision, and impulsivity), and (c) social stability and support (e.g., emotion management, problem solving, employment, residence, and social influences). This scale was

designed to be scored at intake and thereafter as often as every six months on a four-point scale ranging from *minimal need for improvement* to *very considerable need for improvement.* Total scores are calculated by summing all item totals and range from 0 to 48. Individuals will be assigned to one of three risk/need categories based on the density of their criminogenic needs: low (0-10), moderate (11-20), and high (21-48).

2.3.5 Structured Assessment of Protective Factors for Violence Risk– Sexual Offence version (SAPROF-SO)

The SAPROF-SO (Willis et al., 2017-2020) is a measure currently in development designed to assess protective factors in individuals convicted of sexual offenses. It contains 24 items that assess protective factors in five domains: Internal capacity (e.g., intact cognitive functioning, coping), prosocial identity (e.g., prosocial sexual interests, motivation for managing risk), prosocial connection (e.g., leisure activities, emotional connection to adults), stability (e.g., housing, finances), and professionally provided support (e.g., sex offense-specific treatment, therapeutic alliance). Items are rated on a five-point scale, with higher ratings indicating a greater degree of protection. Ratings are made based on information from the previous 6 months (unless otherwise indicated) to make predictions about the upcoming 6-12 months. The SAPROF-SO is derived from the Structured Assessment of Protective Factors for violence risk (SAPROF; de Vogel et al., 2012), a tool that assesses protective factors against violence in adults.

2.3.6 Data Collection Protocol

To collect information relevant to this study, a data collection protocol (Appendix A) was developed with variables that included all forensic measures, as well as client demographic information, victim information, substance use and psychiatric information, criminal history, treatment related information, criminal history, and recidivism information.

2.3.7 Recidivism

All the measures (Static-99R, VRS-SO, STABLE-2007, SOTIPS, and SAPROF-SO) were designed to assess the likelihood of new sexually motivated offenses for individuals who had already been charged or convicted of such an offense against an identifiable victim. General and violent recidivism were also considered given that nonsexual recidivism is more common among individuals with sexual offending histories, and predictors of general and sexual recidivism overlap (Hanson & Bussière, 1998). In this program of research, recidivism is defined as any charge or conviction for a new sexual, non-sexual violent, or non-sexual, non-violent

offense that followed their index sexual offense conviction. Sexual recidivism is defined as any new offense that was clearly sexually motivated or sexual in nature (e.g., sexual assault, possession of child exploitative material); non-sexual violent recidivism was defined as any offense against a person that was not sexually motivated (e.g., assault, forcible confinement, uttering threats); and nonviolent recidivism was defined as any offense that was neither sexual nor violent (e.g., substance-related offenses, theft, breaches). Recidivism variables were coded in both a binary (i.e., yes – no reoffended) and continuous (total number of new offenses) manner. In this study, three categories of recidivism were analysed: 1) sexual recidivism, 2) violent (including sexual) recidivism, and 3) any (sexual, violent, and nonviolent) recidivism. Aligning with previous research (Sowden & Olver, 2017; Wakeling et al., 2013, and given the low base rates of sexual recidivism and corresponding low statistical power to yield significant results, non-sexual violent and sexual recidivism were combined into an aggregate violent recidivism variable. Using charges and convictions as the recidivism criteria strikes a balance between using rearrest and reincarceration, which overestimates and underestimates recidivism, respectively (Bonta et al., 2003).

2.3.8 Procedure

The data collection protocol (which included all risk assessment and protective factor measures) was coded from archival client file information obtained for the entire sample, with pre-treatment ratings obtained for entire sample. Twenty randomly selected cases were coded by two raters to establish interrater reliability. Outcome data were retrieved by April 1, 2021, through CPIC. The research supervisor and student researcher coded and entered CPIC outcome data following the rating of study risk measures and key variables.

2.3.9 Data Preparation

Several pre-analytic statistical procedures were conducted to describe and summarize the data and prepare it for inferential statistics. First, to conduct survival analyses for individuals who recidivated, length of time to recidivism was calculated by subtracting the assessment date from the recidivism date (for a new sexual, violent, or nonviolent offense) or the data collection end date (i.e., for nonrecidivists). For individuals with no new offenses, the treatment completion date was subtracted from the CPIC date. Second, to ensure fidelity of instrument rating and integrity of data collection, reliability analyses were conducted. Specifically, the interrater reliability of the risk and protective factor assessments was assessed using single measure

intraclass correlation coefficients (ICCs, single measure, absolute agreement). The internal consistency of Static-99R, VRS-SO Static, Dynamic, and Total (Static and Dynamic combined) scales, STABLE-2007, SOTIPS, and SAPROF-SO was also examined using Cronbach's alpha. Third, scales with missing items were prorated. Fourth, descriptive statistics were obtained for the total sample, including means, variances, standard deviations, ranges, maximum and minimum scores, and frequencies. Overall group differences were examined using one-way analysis of variance (ANOVA); for instance, level of risk and VRS-SO factor scores according to sexual offense type.

Interrater reliability analyses, descriptive analyses, validity analyses, and survival analyses were conducted using SPSS. Confirmatory Factor Analysis (CFA) was conducted using Mplus 7.11 (Muthén & Muthén, 2013); this software was selected as it offers alternative estimation methods less likely to produce inaccurate parameter estimates and fit values, as well as relatively concise output compared to other software programs (Brown, 2015).

2.4 Data Analytic Plan

2.4.1 Convergent Validity

1. Pearson correlation coefficients were computed between all the risk assessment measures. The greatest interest was on the convergent associations between the VRS-SO (static, dynamic, and total) and other measures in this study: Static-99R, STABLE-2007, SOTIPS, and SAPROF-SO.

Correlation coefficients quantify the magnitude of a relationship between two variables. They range from -1.0 to +1.0, with both values indicating a perfect relationship, and a value of 0 reflecting no relationship. Positive correlations indicate that as one variable increases, the other increases as well (e.g., risk and recidivism). Negative correlations indicate that as one variable increases, the other decreases (e.g., protective measure score and risk measure score). Correlation strength between two continuous measures can be interpreted as follows: .10 = small, .30 = medium, and .50 = large (Cohen, 1992).

2.4.2 Predictive Validity (Discrimination and Calibration)

2.

   a. Correlation coefficients and area under the receiver operating curve (*AUC*) statistics were calculated for the Static-99R, VRS-SO, STABLE-2007, SOTIPS, and SAPROF-SO total scores for general, violent, and sexual recidivism.

42

Relative risk, or discrimination, states individuals with higher risk measure scores are more likely to reoffend than individuals with lower risk measure scores. The extent of these differences is described by area under the curve (AUC) values derived from receiver operating characteristics (ROC) analyses. For the present analyses, AUC values represent the probability that a randomly selected individual who has reoffended will obtain a higher score on a risk measure (e.g., Static-99R, VRS-SO, STABLE-2007, or SOTIPS) or a lower score on a protective measure (e.g., the SAPROF-SO) than a randomly selected individual who did not reoffend. AUC values range from 0 to 1.0, with a value of .50 indicating a predictive accuracy equivalent to chance and a value of 1.0 indicating perfect predictive accuracy. A small effect ($d = .20$) corresponds to an AUC between .56-.63, and a r_{pb} of .10; a medium effect size ($d = .50$) corresponds to an AUC between .64-.70 and a r_{pb} of .24; and a large effect ($d = .80$) corresponds to an AUC of .71-1.0 and a r_{pb} of .37 (Rice & Harris, 2005).

AUC values are less influenced by the outcome variable (e.g., recidivism base rates) and are generally preferred for comparing results across samples. However, AUC values are influenced by variance of the predictor variable (e.g., risk assessment scores) and are smaller in samples with less variance (e.g., all high or low risk; Helmus et al., 2012). Correlation coefficients are influenced by variance in the predictor and base rate of the outcome variable. As the variance in the predictor or outcome variable decreases, the coefficient decreases as well (Hanson, 2008). The impact is strongest when the outcome variable is dichotomous (e.g., recidivated/did not recidivate) and the base rate is less than .50 (Hanson, 2008).

      b. Each measure was subdivided according to risk bins and trajectories of recidivism for the different risk groups were examined using survival analyses.

Kaplan-Meier survival analyses depict the time it takes for events (e.g., recidivism) to occur by providing a graphical representation of the survival rate as a function of time. Survival analysis corrects for unequal follow-up times by estimating the expected recidivism rate of the sample for a specific follow-up period (Helmus et al., 2012). In this study, survival analyses were used to compare recidivism (i.e., the time to first new offense) among groupings of individuals (e.g., risk categories) over the total follow-up period (the maximum follow-up period was 16.5 years). Individuals who did not recidivate are said to "survive." The cumulative proportion of individuals who did recidivate, out of the group being examined, was computed at each time interval (i.e., every 2.5 years). Graphically, this produces a curve that descends as

individuals recidivate, typically more steeply in the initial years of follow-up, and often levelling out over time as recidivism slows.

    c. Chi square analyses were conducted with each risk measure to determine if significant differences existed in observed recidivism frequencies between risk bands.

Chi-square analysis captures the size of association between two variables and is suitable for comparing multiple chi-square test statistics and generalizes across contingency tables of varying sizes. As it is not impacted by sample size, it is useful when it is suspected that statistically significant chi-square relationships are a result of large sample size. Cramer's V is interpreted as a measure of the relative strength of an association between two variables. Coefficients range from 0 to 1.

    d. E/O index was calculated to examine the correspondence between sexual recidivism rates attached to VRS-SO risk categories in the current sample, and in the normative sample.

The E/O index is the ratio of the expected number of individuals who recidivated (E) to the observed number of individuals with that outcome (O). This E/O estimator is considered unbiased when complete follow-up information is available (i.e., 5-year fixed follow-up periods, with predicted recidivism rates calculated using logistic regression). The confidence intervals for the E/O index were calculated using the Poisson variance for the logarithm of the observed number of cases (O):

$$95\% \text{ CI } (E/O) = (E \, / \, O) \exp (\pm 1.96 \, \sqrt{1- O})$$

E/O index can be used as a measure of effect size for fit within a group. An E/O index of 1.0 indicates perfect calibration, meaning the number of observed recidivists in a comparison sample matches the number of recidivists expected from a reference or normative sample. When the ratio is below 1.0, the number of expected recidivists is lower than the number observed (i.e., risk scale underpredicts recidivism). If the ratio is above 1, the expected number of recidivists exceeds the number observed, (i.e., the risk scale overpredicts recidivism). Magnitude is interpreted as a ratio: an E/O index of 0.40 could be interpreted as predicting 40% of the observed recidivism rate, while an index of 3 indicates that the scale predicted three times the number of actual recidivists (Hanson, 2017).

To calculate the E/O index in the present sample, the observed number of individuals who reoffended at 5 years was compared to the predicted number of individuals who reoffended from both raw observed numbers and logistic regression analyses at 5 years. In logistic regression, expected probabilities are calculated in the same way as traditional regression analyses, where the predicted value for a given score on the independent variable (VRS-SO) is obtained by adding the intercept (the predicted value for a score of 0; the B_0) to the product of the slope and the independent variable. The only difference is that the predicted values are in the unit of logits (the natural logarithm of the data that is used in logistic regression). Logits were then transformed back into probabilities, where:

$$\text{Probability} = \frac{e^{LOGIT}}{1 + e^{LOGIT}}$$

2.4.3 Incremental Validity:

3.

   a. Cox regression analyses were used to examine the incremental contributions of the VRS-SO dynamic items in predicting sexual recidivism over and above the VRS-SO static total.

   b. In separate analyses for each Static-99R and dynamic measure pairing, Cox regression analyses were used to examine the incremental contributions of each dynamic measure (VRS-SO dynamic total, STABLE-2007, SOTIPS, and SAPROF-SO) in predicting sexual recidivism over and above Static-99R.

   c. Cox regression analyses were used to examine the incremental contributions of the SAPROF-SO in predicting sexual recidivism over and above Static-99R and the VRS-SO dynamic total.

   d. Cox regression analyses were used to examine the incremental contributions of each VRS-SO factor (Sexual Deviance, Criminality, and Treatment Responsivity) in predicting sexual recidivism

   e. Cox regression analyses were used to examine the incremental contributions of each VRS-SO factor (Sexual Deviance, Criminality, and Treatment Responsivity) in predicting sexual recidivism over and above Static-99R.

   f. The above analyses were also run to determine respective incremental prediction of violent (sexual and nonsexual) and general (i.e., any) recidivism.

Cox regression analyses provide a hazard ratio (HR), called the exponentiated beta coefficient (e^B), which is an indicator of the proportional increase in the hazard of a given outcome (e.g., recidivism) per one-unit change in the predictor (e.g., risk score). Values greater than 1.0 indicate that as the predictor increases, the hazard of that outcome occurring increases (e.g., higher risk, more recidivism). Values less than 1.0 indicate that as the predictor increases, the hazard of the outcome occurring decreases (e.g., more protective factors, less recidivism).

For every e^B there is a corresponding Wald chi square statistic, computed as $(B/SE)^2$, used to indicate whether the e^B is significantly associated with the prediction of outcome, with larger Wald statistic values indicating greater predictive ability (Field, 2009). Cox regression is not influenced by variability in follow-up time and therefore tends to provide more stable estimates of predictive accuracy than AUC values when the follow-up period is not fixed; however, it assumes the rates of recidivism over time (i.e., the shape of the survival curve) are approximately the same across samples, even after controlling for baseline risk (Helmus et al., 2012).

2.4.4   Construct Validity:

4.

  a) A confirmatory factor analysis (CFA) of the VRS-SO dynamic items was used to establish constructs that underpin dynamic sexual violence risk as measured by this tool.

CFA is a hypothesis-driven type of Structural Equation Modeling (SEM) commonly based on the results of an Exploratory Factor Analysis (EFA), and theoretical merit (Brown, 2015; Costello & Osborne, 2005). CFA is similar to EFA in that it tests the relationship between indicators (e.g., items on a risk measure) and hypothesized latent constructs. Like EFA, it is based on the common (linear) factor model. Unlike EFA, CFA is used to establish the ability of predicted models, in which every parameter is specified in advance by the researcher, to fit a particular data sample (Brown, 2015). Model fit is determined based on several goodness of fit indices; it is generally recommended that researchers report a minimum of three indices that represent a cross-section of index types (i.e., sample-based indices, population-based indices, relative, and absolute indices; Brown, 2015; Sun, 2005).

Choosing which fit indices to report depends on the initial estimation method used; an estimation method is a mathematical operation applied to minimize the difference between the

observed data variance-covariance matrix (i.e., the sample data) and the hypothesized variance-covariance matrix (i.e., the hypothesized model). Indices are also affected by sample and model characteristics, as well as the accuracy of the factor loading specifications (Brown, 2015). Nunnally (1978) suggests having at least 10 times as many subjects as variables, whereas Cliff (1987) states that with approximately 40 variables, a group of 150 is the minimum, though 500 is preferable.

Absolute fit indices only evaluate the hypothesized model, while relative fit indices compare the hypothesized model to the baseline model (e.g., a baseline model for comparison could be one in which all the correlations among the indicators are set to zero). Essentially, a relative fit index is comparative. Though there are numerous restrictions and limitations that each index is subject to, it is most important to note that cut-off values for any index are suggestive and represent different implications for various models under various circumstances. The present research utilizes the suggestions of Marsh et al. (2010), stating that CFI values of .90 to .95, respectively, represent acceptable and excellent fits to the data, while RMSEA values below .05 and .08 represent close and reasonable fits to the data.

CHAPTER THREE: RESULTS

3.1 Base Rates of Recidivism

Recidivism data were available for 172 of the 200 participants. The 172 individuals were followed for a mean of 8.6 years ($SD = 2.3$), with a range of 1.3 to 16.2 years. Recidivism was defined as any new charge or conviction for a sexual, non-sexual violent, or non-violent offense following the FACS assessment date. One-third of the sample (33%) had at least one reoffense and the average number of reoffenses was $M = 2.13$ ($SD = 6.78$). The base rates of recidivism were: 9% of the sample had at least one new sexual charge or conviction, 18.5% had at least one new violent (including sexual) charge or conviction, 13.5% had at least one new nonsexual violent charge or conviction, and 33% had at least one new charge or conviction (i.e., any new charge or conviction). Table 3.1 illustrates the recidivism base rate by sex offense type.

47

Table 3.1

Recidivism Base Rate Percentages by Sexual Offense Victim Type

	EF child	Adult	Mixed	IF child
Sexual Recidivism	12.8	9.8	0.0	5.9
Violent (including sexual) Recidivism	23.4	19.6	12.5	13.7
Nonsexual Violent Recidivism	10.6	17.6	12.5	9.8
Any Recidivism	38.3	37.0	37.5	21.6

Note: EF = extrafamilial child victim; IF = intrafamilial child victim.

3.2 Descriptive Statistics

Table 3.2 outlines the means and standard deviations for the following forensic measures: Static-99R, VRS-SO Static scale, VRS-SO Dynamic scale, VRS-SO total, VRS-SO factors (Sexual Deviance, Criminality, and Treatment Responsivity), STABLE-2007, SOTIPS total, SOTIPS factors (Sexual Deviance, Criminality, and Social Stability and Support), SAPROF-SO total, and SAPROF-SO factors (Internal Capacity, Prosocial Identity, Prosocial Connection, Stability, and Professionally Provided Support). These results are briefly summarized below.

The mean Static-99R total score fell within the Average risk category ($M = 2.19$) and was very close to the mean score reported by Hanson et al., 2007 ($M = 2.4$) in a community-supervised sample. The mean VRS-SO total score fell within the Below average risk category ($M =20.94$) and was significantly lower than the mean score reported by Olver et al. (2018); the mean VRS-SO dynamic score and mean VRS-SO static scores both fell within the Below average risk category ($M = 15.12$ and 5.82, respectively); these were again lower than the Olver et al. (2018) study ($M = 24.9$ and 8.9, respectively). Given that Olver et al. (2018) utilized correctional samples, this difference in mean scores is not surprising. The STABLE-2007 total score fell within the Moderate risk/need category ($M=7.36$) which is highly similar to Hanson et al., 2015 ($M = 7.4$); and the mean SOTIPS total score fell within the Moderate risk/need category ($M =14.96$) which is comparable to the mean score of a community-supervised sample ($M = 14.1$; Hanson et al., 2020). The mean SAPROF-SO total score in the current study was $M =45.76$, which is comparable to the mean score in a correctional sample that had undergone treatment ($M = 39.08$; Nolan et al., 2022).

Table 3.2

Forensic Measures: Means and Standard Deviations

Measure	*M (SD)*
Static-99R	2.19 (2.28)
VRS-SO Static	5.82 (3.73)
VRS-SO Dynamic	15.12 (8.69)
VRS-SO Sexual Deviance	4.33 (4.47)
VRS-SO Criminality	4.05 (3.42)
VRS-SO Treatment Responsivity	6.59 (3.27)
VRS-SO Total	20.94 (11.03)
STABLE 2007	7.36 (4.64)
SOTIPS Sexual Deviance	5.81 (3.54)
SOTIPS Criminality	6.16 (4.84)
SOTIPS Social Stability & Support	3.08 (2.41)
SOTIPS Total	14.96 (8.63)
SAPROF-SO Internal Capacity	14.96 (8.40)
SAPROF-SO Prosocial Identity	8.42 (5.84)
SAPROF-SO Prosocial Connection	6.38 (5.31)
SAPROF-SO Stability	5.02 (2.49)
SAPROF-SO Professionally Provided Support	11.59 (3.84)
SAPROF-SO Total	45.76 (18.08)

Note: scale totals are prorated to account for missing items where applicable.

3.3 Group Differences by Sexual Offense Type

A series of one-way ANOVA analyses was conducted using Bonferroni's post-hoc multiple comparisons to compare the four categories of sex offense type on level of risk and VRS-SO, SOTIPS, and SAPROF-SO factor scores. Given that equal variances could not be assumed for the VRS-SO Static total and Sexual Deviance Factor (as per Levene's test), Games-Howell post-hoc tests were run. Table 3.3 illustrates the differences on the measures by sex offense type.

On Static-99R, individuals with mixed adult/teen and child offenses had the highest mean score, followed by extrafamilial adult/teen, extrafamilial child, and intrafamilial offenses. There were statistically significant differences between Static-99R mean scores for each offense group. The VRS-SO Static total, Dynamic total, Total score, STABLE-2007, SOTIPS Social Stability and Support Factor, and SOTIPS Total score results were similar to Static-99R: individuals with mixed adult/teen and child victims had the highest mean score, followed by extrafamilial adult/teen, extrafamilial, and intrafamilial child victims. There were statistically significant differences between groups for all above-listed mean factor and total scores.

The VRS-SO Sexual Deviance, Criminality, and Treatment Responsivity factors followed a slightly different pattern of mean scores. Individuals with a mixed victim profile had the highest mean scores followed by those with extrafamilial child victims for each VRS-SO factor, but for the VRS-SO Sexual Deviance factor, individuals with the adult/teen victim profile had a lower mean score than those with intrafamilial child victims, and differences between groups were statistically significant. For the VRS-SO Criminality factor, individuals with adult/teen victims had higher mean scores than individuals with extrafamilial child victims and those with intrafamilial victims; differences between groups were not statistically significant. For the VRS-SO Treatment Responsivity factor, individuals with the intrafamilial profile had higher mean scores than individuals with the extrafamilial child profile; differences between groups were not statistically significant.

For the SOTIPS Sexual Deviance and Criminality factors, individuals with the mixed victim profile again had the highest mean score for each factor, but for SOTIPS Sexual Deviance, individuals with intrafamilial child victims had higher mean scores than individuals with extrafamilial adult/teen victims. For the SOTIPS Criminality Factor, those with extrafamilial adult/teen victims had higher mean scores than individuals with extrafamilial child victims and those with intrafamilial child victims. Differences between groups were statistically significant for both SOTIPS Sexual Deviance and Criminality Factors.

For the SAPROF-SO factor scores, individuals with the mixed victim profile had the lowest mean score for all factors except the SAPROF-SO Professionally Provided Support factor, where they had the highest mean score, followed by those with extrafamilial child victims, extrafamilial adult/teen victims, and intrafamilial child victims. For the SAPROF-SO Internal Capacity, those with extrafamilial adult/teen victims had the highest mean score, followed by those with intrafamilial child victims, extrafamilial child victims, and a mixed victim profile. For the SAPROF-SO Prosocial Identity, Prosocial Connection, and Stability factors, individuals with intrafamilial child victims had the highest mean scores, followed by extrafamilial adult/teen victims, extrafamilial child victims, and a mixed victim profile. Group differences were only statistically significant for the SAPROF-SO Prosocial Identity, Prosocial Connection, and total scores.

50

Table 3.3

Forensic Measure Group Means, Standard Deviations, and Differences by Sex Offense Type

Measure	Intrafamilial	Extrafamilial adult/teen	Extrafamilial child	Mixed adult/child	F
			M (SD)		
Static-99R	0.41 (1.60)	2.79 (2.19)	2.51 (2.13)	4.38 (1.60)	19.49***
VRS-SO Static	2.75 (2.53)	6.18 (2.95)	7.26 (3.76)	12.63 (2.00)	34.74***
VRS-SO Dynamic	14.01 (7.13)	14.09 (8.76)	16.75 (9.24)	25.05 (6.97)	5.03**
VRS-SO Sexual Deviance	4.28(4.17)	3.36 (3.98)	5.40 (5.00)	9.63 (4.50)	6.55***
VRS-SO Criminality	3.38 (2.66)	4.22 (3.91)	4.00 (2.96)	6.63 (3.25)	2.29
VRS-SO Tx Responsivity	6.60 (3.21)	6.50 (3.26)	6.62 (3.43)	7.96 (2.55)	0.49
VRS-SO Total	16.76 (8.67)	20.28 (10.73)	24.00 (11.11)	37.67 (7.86)	11.49***
STABLE 2007	6.61 (3.43)	6.86 (5.17)	8.13 (3.85)	13.25 (5.18)	5.94**
SOTIPS Sexual Deviance	5.49 (3.49)	5.42 (3.31)	6.51 (3.83)	8.63 (3.32)	2.90*
SOTIPS Criminality	5.03 (3.18)	6.58 (5.61)	5.72 (3.73)	10.85 (6.47)	4.00**
SOTIPS Soc. Stability/Support	2.28 (2.23)	3.01 (2.37	3.78 (2.37)	4.94 (2.18)	5.11**
SOTIPS Total	12.79 (7.16)	14.92 (9.39)	15.87 (7.21)	23.96 (8.98)	4.42**
SAPROF-SO Internal Capacity	15.21 (8.41)	15.64 (8.65)	14.09 (7.64)	10.21 (8.64)	1.25
SAPROF-SO Prosocial Identity	9.82 (6.35)	8.76 (5.57)	7.08 (5.55)	4.64 (4.48)	2.80*
SAPROF-SO Prosocial Connection	7.34 (5.42)	7.06 (5.43)	4.61 (4.39)	1.46 (2.14)	5.04**
SAPROF-SO Stability	5.35 (2.52)	5.20 (2.53)	4.43 (2.14)	4.14 (3.39)	1.57
SAPROF-SO Prof. Provided Support	11.22 (3.67)	11.48 (4.00)	11.88 (3.58)	12.48 (4.19)	0.41
SAPROF-SO Total	47.99 (17.37)	44.79 (19.61)	41.37 (14.23)	32.55 (13.60)	3.13*

Note: *** = $p < .001$; ** = $p < .01$; * = $p < .05$. Tx = Treatment; Soc. = Social; Prof = Professionally

3.4 Interrater Reliability

The interrater reliability (IRR) of the forensic measures was assessed on 20 randomly selected cases and analysed using absolute agreement single measures intraclass correlation coefficients (ICC_{A1}). Differences between raters of a magnitude that warranted removing outliers was not observed in any of the randomly selected cases. Using Cicchetti and Sparrow's (1981) interrater reliability magnitude criteria, ICC_{A1} values for all forensic tools were "excellent" (i.e., $ICC_{A1} \geq .75$) at the $p < 0.001$ level. Almost all factor/domain scores were also significant at the $p < 0.001$ level and demonstrated excellent agreement with reliability ranging from $ICC_{A1} = 0.77$ to 0.96. The exception to this was the SOTIPS Criminality factor, which evidenced fair reliability (i.e., ICC_{A1} of .40 to .59). Additionally, the SAPROF-SO Prosocial Connection factor was significant at the $p < 0.01$ level.

The ICC_{A1} values for the VRS-SO total score and Static-99R were consistent with past research ($ICC = .95$ and .97, respectively; Sowden & Olver, 2017). The ICC_{A1} for the STABLE-

2007 in the current study was much higher than in previous research (*ICC* = .46; Sowden & Olver, 2017). The SOTIPS total and factor (Sexual Deviance, Social Stability and Support) scores were also higher in the present study than in past research (ICC_{A1} = .77, .68, and .69, respectively; McGrath et al., 2013), except for the SOTIPS Criminality score, which was lower in this study than in past research (ICC_{A1} = .76; McGrath et al., 2013). Only the SAPROF-SO total and Prosocial Connection and Stability factor scores were calculable in the current study. These scores were consistent with past research (ICC_{A1} > .75; Willis et al., 2020). Generally, the current research demonstrated higher ICC_{A1} values overall, except for the SOTIPS Criminality factor.

Table 3.4

Interrater Reliability of Forensic Measures Total Scores and Factor Scores: Absolute Agreement Single Measures Intraclass Correlation Coefficients

Measure	Cronbach's Alpha	$ICC_{A,1}$
Static-99R	.99	.98**
VRS-SO		
Static	.98	.96**
Dynamic	.92	.86**
Total	.95	.89**
Sexual Deviance	.91	.84**
Criminality	.91	.83**
Treatment Responsivity	-	-
STABLE 2007	.99	.98**
SOTIPS		
Total	.93	.87**
Sexual Deviance	.98	.96**
Criminality	.56	.39
Social Stability/Support	.94	.88**
SAPROF-SO		
Total	.94	.89**
Internal Capacity	-	-
Prosocial Identity	-	-
Prosocial Connection	.89	.81*
Stability	.87	.77**
Professionally Provided Support	-	-

Note: ICC = intraclass correlation coefficient. ** = $p < .001$; * = $p < .01$

3.4.1 Internal Consistency Reliability

The internal consistency reliability of the Static-99R, VRS-SO, STABLE-2007, SOTIPS, and SAPROF-SO was examined using Cronbach's alpha (Table 3.5). The Static-99R showed

low internal consistency ($\alpha =.36$) indicating heterogeneity of item content. The VRS-SO Static total demonstrated much higher internal consistency ($\alpha = .67$). The dynamic measures evidenced higher internal consistency, ranging from good to excellent ($\alpha = .76$ to $.93$).

Table 3.5

Scale Reliability: Internal Consistency of Forensic Measures

	Cronbach's Alpha
Static-99R	.36
VRS-SO	
Static	.67
Dynamic	.92
Total	.93
STABLE 2007	.76
SOTIPS	.83
SAPROF-SO	-

3.5 Validity of Risk/Protective Measures

3.5.1 Convergent Validity of Forensic Measures

Static-99R, VRS-SO Static, VRS-SO Dynamic, VRS-SO Total, STABLE-2007, SOTIPS, and SAPROF-SO were correlated to each other to assess the convergent validity of the forensic measures (see Table 3.6). According to Cohen's conventions, significant point biserial correlation values of .10, .30, and .50 are considered small, medium, and large in magnitude, respectively (1988). Correlations for most measures except the SAPROF-SO were positive, significant at the $p < .001$ level, and large in magnitude as per Cohen (1988); exceptions include the correlations between the SOTIPS total and Static-99R and VRS-SO static total, and between VRS-SO static total and STABLE-2007, which were all positive, moderate in magnitude, and significant at the $p < .001$ level. The correlations between the SAPROF-SO total score and the risk measures were all negative, moderate to large in magnitude, and significant at the $p < .001$ level, except with the VRS-SO Static total, which was negative, small in magnitude, and significant at the $p < .001$ level.

Factor scores for the VRS-SO dynamic items (Sexual Deviance, Criminality, and Treatment Responsivity), SOTIPS items (Sexual Deviance, Criminality, and Social Support and Stability), and SAPROF-SO items (Internal Capacity, Prosocial Identity, Prosocial Connection, Stability, and Professionally Provided Support) were also correlated to each other to assess convergent validity (see Table 3.7). Correlations between the risk tool factor scores were mostly

positive, small to moderate in magnitude, and significant at the $p < .001$ level. Exceptions included the correlations between SOTIPS Sexual Deviance and VRS-SO Treatment Responsivity, VRS-SO Criminality and SOTIPS Criminality, and SOTIPS Criminality and SOTIPS Social Stability and Support, which were all positive, large in magnitude, and significant at the $p < .001$ level, and correlations between VRS-SO Sexual Deviance and SOTIPS Criminality, and VRS-SO Treatment Responsivity and SOTIPS Social Stability and Support, which were positive, small in magnitude, and significant at the $p < .01$ level. Additionally, there was no significant relationship between the VRS-SO Sexual Deviance factor and the VRS-SO Treatment Responsivity factors.

Most correlations between the SAPROF-SO factors and the risk tool factors were negative, small to large in magnitude, significant at the $p < .001$ level. Exceptions include the correlations between SAPROF-SO Internal Capacity and VRS-SO Sexual Deviance and Treatment Responsivity factors, which were negative, small in magnitude, and significant at the $p < .01$, and the following correlations which were negative, small in magnitude, and significant at the $p < .05$ level: SAPROF-SO Stability factor and the VRS-SO Sexual Deviance and Treatment Responsivity factors, and SAPROF-SO Professionally Provided Support and the VRS-SO Sexual Deviance factors. There was no significant relationship between the SAPROF-SO Prosocial Connection and VRS-SO Sexual Deviance factors, or between SAPROF-SO Professionally Provided Support and any of the other risk tool or protective tool factors.

Table 3.6

Convergent Validity: Correlations among Forensic Measure Total Scores

Measure	2	3	4	5	6	7
1.Static-99R	.77	.51	.66	.52	.43	-.36
2. VRS-SO Static	-	.50	.73	.47	.39	-.29
3. VRS-SO Dynamic		-	.96	.78	.70	-.57
4. VRS-SO Total			-	.77	.68	-.54
5. STABLE 2007				-	.69	-.64
6. SOTIPS Total					-	-.81
7. SAPROF-SO Total						-

Note: all = *p* < .001

Table 3.7

Convergent Validity: Correlations among Forensic Measure Factor/Domain Scores

Measure	2	3	4	5	6	7	8	9	10	11
VRS-SO										
1. Sexual Deviance	.29***	.13	.35***	.24**	.25***	-.22**	-.33***	-.22	-.16*	-.20*
2. Criminality	-	.34***	.37***	.69***	.58***	-.46***	-.48***	-.54***	-.43***	-.02
3. Txt Responsivity		-	.55***	.26***	.23**	-.23**	-.38***	-.28***	-.15*	-.05
SOTIPS										
4. Sexual Deviance			-	.45***	.41***	-.37***	-.59***	-.43***	-.29***	-.08
5. Criminality				-	.50***	-.55***	-.54***	-.49***	-.39***	-.07
6. Soc. Stability and Support					-	-.54***	-.47***	-.76***	-.65***	-.12
SAPROF-SO										
7. Internal Capacity						-	.44***	.57***	.35***	.08
8. Prosocial Identity							-	.53***	.33***	.08
9. Prosocial Connection								-	.52***	.04
10. Stability									-	.00
11. Prof. Provided Support										-

Note: Txt = Treatment; Soc. = Social; Prof. = Professionally. *** = *p* < .001; ** = *p* < .01; * = *p* < .001

3.5.2 Discrimination Predictive Validity of Risk/Protective Measures

   3.5.2.1 Correlations and Area Under the Curve (AUC). The predictive validity of the forensic measures was examined with respect to sexual, violent (including sexual), and any recidivism following completion of the assessment at FACS. Separate analyses were conducted to examine all recidivism outcomes with no fixed follow-up time, and a five-year fixed follow-up time. Predictive validity was examined using two methods: 1) Point biserial correlation coefficients (i.e., r_{pb}), and 2) AUC values. AUC values are interpreted as follows: small (.56 – .63), medium (.64 – .70), and large (.71 – 1.0) effect (Rice & Harris, 2005). Correlations between total and factor/domain scores and the total number of charges/convictions are reported in the table only.

   3.5.2.1.1 No Fixed Follow-up. As reported in Table 3.8, SOTIPS Sexual Deviance and SAPROF-SO Stability had large effects in the prediction of sexual recidivism while VRS-SO Sexual Deviance had medium effects. Further, VRS-SO Dynamic and Total scores, STABLE-2007, SOTIPS Social Stability and Support, and SAPROF-SO Prosocial Identity and Total score all had medium effects. Finally, Static-99R, VRS-SO Static, Criminality, Treatment Responsivity, SOTIPS Criminality, SAPROF-SO Internal Capacity, Prosocial Connection, and Professionally Provided Support all had non-significant associations and small effects (or lower) for this outcome.

   For violent (including sexual) recidivism prediction, Static-99R, VRS-SO Criminality, and SOTIPS total had medium effects. SAPROF-SO Stability and Total score also had medium effects. VRS-SO Static, Dynamic, and Total scores, STABLE-2007, SOTIPS Criminality and Social Stability and Support scores, SAPROF-SO Prosocial Identity and Prosocial Connection scores all evidenced a small effect for this outcome. Finally, VRS-SO Sexual Deviance and Treatment Responsivity, SOTIPS Sexual Deviance, SAPROF-SO Internal Capacity and Professionally Provided Support all had nonsignificant and small effects (or lower) for violent (including sexual) recidivism.

   In the prediction of any recidivism, VRS-SO Criminality and SAPROF-SO Prosocial Identity had large effects, while VRS-SO Static, Dynamic, and Total scores and SOTIPS Criminality and Total score had medium effects. Static-99R, STABLE-2007, SOTIPS Social Stability and Support, SAPROF-SO Stability and Total score demonstrated a medium effect for this outcome, and SAPROF-SO Prosocial Connection evidenced a small effect for this outcome. The following measures had nonsignificant associations and small effects (or lower) in the prediction of any recidivism: VRS-SO Sexual Deviance and Treatment Responsivity, SOTIPS Sexual Deviance, and SAPROF-SO Internal Capacity.

Table 3.8

Predictive Validity: Correlations and AUCs between the Forensic Measures and Outcome, No Fixed Follow-up Time

Measure	Sexual recidivism			Violent recidivism			General recidivism		
	r_{pb}	AUC	95% CI	r_{pb}	AUC	95% CI	r_{pb}	AUC	95% CI
Static-99R	.14*	.63	.48, .77	.20**	.65**	.55, .74	.24**	.65**	.57, .73
VRS-SO									
Static	.09	.58	.44, .72	.14	.61*	.52, .71	.23**	.64**	.56, .73
Dynamic	.19**	.69*	.56, .82	.11	.61*	.51, .70	.24**	.65**	.56, .73
Sexual deviance	.22**	.70**	.57, .83	.03	.51	.41, .62	.05	.51	.42, .60
Criminality	.09	.63	.51, .75	.15*	.65**	.56, .74	.34***	.71***	.64, .79
Txt Responsivity	-.03	.47	.34, .62	.00	.50	.40, .59	.09	.55	.47, .64
Total	.18*	.68*	.54, .82	.14	.63*	.54, .72	.26***	.67***	.59, .75
STABLE-2007	.18**	.66*	.51, .80	.14	.61*	.51, .71	.25***	.65**	.56, .73
SOTIPS									
Sexual deviance	.24**	.71**	.57, .85	.12	.57	.46, .68	.13	.55	.47, .64
Criminality	.08	.56	.42, .70	.16*	.61*	.51, .71	.36***	.67***	.59, .75
Soc stability/support	.18*	.69*	.56, .82	.15	.61*	.51, .72	.21***	.65**	.56, .73
Total	.20**	.68*	.55, .81	.20**	.64**	.54, .74	.33***	.70***	.62, .78
SAPROF-SO									
Internal capacity	-.09	.58	.43, .73	-.10	.59	,49, .70	-.10	.58	.49, .66
Prosocial identity	-.12	.65*	.52, .78	-.15	.63*	.53, .73	-.15	.63**	.54, .71
Prosocial connection	-.12	.63	.52, .75	-.11	.61*	.51, .70	-.11	.62**	.54, .71
Stability	-.20**	.72**	.59, .85	-.21**	.67**	.57, .78	-.21**	.64**	.56, .73
Prof. provided support	.10	.59	.44, .73	.08	.53	.42, .63	.08	.47	.38, .55
Total	-.14	.65*	.51, .78	-.15	.64*	.54, .73	-.15	.65**	.57, .73

Note: Txt = Treatment; Soc. = Social; Prof. = Professionally *** = $p < .001$; ** = $p < .01$; * = $p < .05$

   ***3.5.2.1.2 Five-year Fixed Follow-up*.** When five-year fixed follow-up times were employed (Table 3.9), the following measures and domain/factor scores evidenced large effects in the prediction of sexual recidivism: VRS-SO Dynamic, Sexual Deviance, and Total score, STABLE-2007, and SOTIPS Social Stability and Support. Static-99R, SOTIPS Sexual Deviance and Total score, SAPROF-SO Prosocial Identity, Prosocial Connection, Stability, and Total score all had medium effects. SAPROF-SO Professionally Provided Support had a small effect. VRS-SO Static, Criminality, and Treatment Responsivity score, SOTIPS Criminality, and SAPROF-SO Internal Capacity had nonsignificant associations and small effects (or lower) in the prediction of this outcome.

   For violent (including sexual) recidivism prediction, VRS-SO-SO Dynamic and Criminality score, Static-99R, STABLE-2007, SOTIPS Social Stability and Support and Total score, and SAPROF-SO Prosocial Identity, Prosocial Connection, Total, and Stability score demonstrated medium effects. The following measures had nonsignificant associations and small effects (or lower) in the

prediction of this outcome: VRS-SO Sexual Deviance and Treatment Responsivity scores, SOTIPS Sexual Deviance and Criminality scores, and SAPROF-SO Internal Capacity and PPS scores.

　　　For the prediction of any recidivism, VRS-SO Criminality and SOTIPS Total evidenced large effects. VRS-SO Static, Dynamic, and Total score, Static-99R, STABLE-2007, SOTIPS Criminality and Social Stability and Support, and SAPROF-SO Prosocial Connection, Stability, and Total score had medium effects. SAPROF-SO Internal Capacity and Prosocial Identity evidenced small effects in the prediction of this outcome. VRS-SO Sexual Deviance, Treatment Responsivity, SOTIPS Sexual Deviance, and SAPROF-SO Professionally Provided Support had nonsignificant associations and small effects (or lower) in the prediction of this outcome.

Table 3.9

Predictive Validity: Correlations and AUCs between the Forensic Measures and Outcome, 5-year Fixed Follow-up Time

Measure	Sexual recidivism			Violent recidivism			General recidivism		
	r_{pb}	*AUC*	*95% CI*	r_{pb}	*AUC*	*95% CI*	r_{pb}	*AUC*	*95% CI*
Static-99R	.17*	.66*	.51, .81	.24**	.69**	.59, .79	.25***	.66***	.58, .75
VRS-SO									
Static	.12	.62	.47, .76	.16*	.64*	.55, .74	.23**	.65**	.57, .73
Dynamic	.23**	.72**	.59, .86	.20**	.67**	.57, .77	.24**	.65**	.57, .74
Sexual deviance	.27***	.76**	.64, .88	.11	.59	.47, .70	.05	.50	.41, .60
Criminality	.08	.64	.52, .75	.19**	.68**	.59, .78	.35***	.72***	.64, .80
Txt Responsivity	-.02	.49	.35, .64	.04	.53	.43, .64	.08	.56	.47, .64
Total	.22**	.72*	.57, .86	.21**	.69*	.60, .79	.27***	.68***	.60, .76
STABLE-2007	.23**	.71*	.57, .85	.21**	.67*	.57, .77	.24**	.65**	.57, .73
SOTIPS									
Sexual deviance	.21**	.69*	.54, .85	.15*	.61	.49, .73	.15*	.59	.50, .68
Criminality	.03	.53	.39, .68	.15*	.61	.50, .72	.35***	.70***	.62, .79
Soc stability/support	.19*	.71**	.59, .83	.20**	.67**	.57, .77	.24**	.66*	.57, .74
Total	.17**	.67*	.53, .81	.22**	.68**	.57, .78	.34***	.71***	.63, .79
SAPROF-SO									
Internal capacity	-.10	.60	.44, .76	-.10	.60	.48, .72	-.10	.60*	.51, .69
Prosocial identity	-.12	.67*	.54, .80	-.11	.65*	.54, .75	-.15	.62*	.53, .71
Prosocial connection	-.13	.65*	.54, .76	-.13	.64*	.55, .73	-.16*	.65**	.57, .73
Stability	-.16*	.70*	.56, .84	-.17*	.69**	.58, .80	-.20*	.65**	.56, .73
Prof provided support	.18*	.66*	.52, .79	.18*	.59	.48, .70	.01	.48	.39, .57
Total	-.12	.65*	.52, .78	-.11	.65*	.54, .75	-.19	.66***	.58, .74

Note: Txt = Treatment; Soc. = Social; Prof. = Professionally *** = $p < .001$; ** = $p < .01$; * = $p < .05$

3.5.3 Calibration/Absolute Risk

The next set of analyses examined what rates of recidivism are associated with scores on Static-99R, VRS-SO Total and Dynamic Total, STABLE-2007, and SOTIPS (i.e., absolute risk). This was examined using three approaches: 1) Kaplan-Meier survival analysis of risk bins for each measure examining trajectories of recidivism over time; and 2) chi square analysis (including Cramer's V as a measure of effect size) examining rates of recidivism employing 5-year fixed follow-ups for the categories; 3) E/O index, computing the ratio between expected 5-year sexual recidivism rates based on the VRS-SO normative sample and the observed 5-year sexual recidivism rates in the current sample. Each analysis will be described hereafter.

3.5.3.1 Kaplan-Meier Survival Analyses. Kaplan-Meier survival charts are presented for each set of analyses, with log rank chi square pairwise comparisons reported in tabular form immediately below each chart. Survival curves for all measures and recidivism outcomes were consistent with the anticipated recidivism trajectories; there are higher and faster failure rates (i.e., recidivism rates) with each successive increase in risk level. However, differences between risk groups were not always significant. Significant results for each risk measure band are outlined below.

First, Static-99R total scores were arranged into five risk bands: I – very low risk (-3, -2), II – below average risk (-1, 0), III – average risk (1 to 3), IVA – above average risk (4, 5), and IVB – well above average risk (scores of 6), as described by Hanson et al. (2017). As displayed in Figure 3.1 and its accompanying table (Table 3.10), Level IVB had significantly faster and higher rates of sexual recidivism compared to the II, III, and IVA risk groups, but did not significantly differ from the Level I risk group. The remaining groups did not significantly differ from one another. For violent (including sexual) recidivism, (Figure 3.2 and table 3.11), none of the groups significantly differed from one another. For any recidivism (Figure 3.3 and table 3.12), the Level IVB group significantly differed from the I, II, and III risk groups, but not the IVA risk group. Level III and IVA groups also significantly differed for any recidivism; there were no significant differences between remaining groups.

Second, VRS-SO total scores were arranged into four risk bands: I – Very low risk <15), II – Below average risk (15-23.5), III – Average risk (24-39.5), and IVA & IVB – Above average risk/Well above average risk (40-67+) as described in Olver, Mundt et al., 2018. The two highest risk groups were collapsed into one risk band owing to the low number of group members. As displayed in Figure 3.4 and its accompanying table (Table 3.13), for sexual recidivism, the only significant difference was between Level I and III. For violent (including sexual) recidivism (Figure 3.5 and table 3.14), there were significant differences in VRS-SO total scores between Level I and II, and I and III. For any

59

recidivism (Figure 3.6 and table 3.15), there were significant differences between Level I and all other groups.

Third, VRS-SO dynamic scores were arranged into four risk bands: I – Very low risk < 10.5), II – Below average risk (11-16.5), III – Average risk (17-27.5), and IVA & IVB – Above average risk/Well above average risk (28-51; as described in Olver et al., 2021). Again, the two highest risk groups were collapsed into one risk band owing to the low number of group members. As displayed in Figure 3.7 and its accompanying table (Table 3.16), for sexual recidivism, Level I and II significantly differed in VRS-SO dynamic total score from both Level III and Level IVA & IVB. For violent (including sexual) recidivism (Figure 3.8 and table 3.17), Level I evidenced significant differences in VRS-SO dynamic total scores from Level III and IVA & IVB. For any recidivism (Figure 3.9 and table 3.18), significant differences occurred between Level I and Level III and IVA & IVB, and Level II and IVA & IVB.

Fourth, STABLE-2007 total scores were arranged into three risk bands: Low (0-3), moderate (4-11), and high (12 and above). For sexual recidivism (Figure 3.10 and table 3.19), violent (including sexual) recidivism (Figure 3.11 and table 3.20), and any recidivism (Figure 3.12 and table 3.21), STABLE-2007 total scores were significantly different between the low and high, and medium and high risk groups, but not low and high groups.

Fifth, SOTIPS total scores were arranged into three risk bands: Low (0-10), moderate (11-20), and high (21-48) As displayed in Figure 3.13 and its accompanying table (table 3.22), for sexual recidivism, SOTIPS total scores were significantly different between the low and high risk groups; there were no significant differences between remaining groups. For violent (including sexual) recidivism (Figure 3.14 and table 3.23), and any recidivism (Figure 3.15 and table 3.24), there were significant differences in SOTIPS total scores between the low and high, and medium and high risk groups, but not low and medium risk groups.

Figure 3.1: *Kaplan-Meier Survival Analysis: Rates of New Sexual Charges or Convictions Over Time for Static-99R Risk Groups*

Table 3.10

Log Rank Chi Square Pairwise Comparison Static-99R New Sexual Charge or Conviction

Static-99R risk band	Pairwise comparison (log rank χ^2) reference group			
	Level II	Level III	Level IVa	Level IVb
Level I: Very low risk (-3, -2, $n = 8$)	0.74	0.56	0.72	3.27
Level II: Below average risk (-1, 0, $n = 44$)	-	0.35	0.02	5.76*
Level III: Average risk (1-3, $n = 89$)	-	-	0.14	10.89***
Level IVa: Above average risk (4-5, $n = 47$)	-		-	6.15*
Level IVb: Well above average risk (6+, $n = 12$)	-	-	-	-

Note: *** $p < .001$, ** $p < .01$, * $p < .05$

Figure 3.2: *Kaplan-Meier Survival Analysis: Rates of New Violent (Including Sexual) Charges or Convictions Over Time for Static-99R Risk Groups*

Table 3.11

Log Rank Chi Square Pairwise Comparison Static-99R New Violent (Including Sexual) Charge or Conviction

Static-99R risk band	Pairwise comparison (log rank χ^2) reference group			
	Level II	Level III	Level IVa	Level IVb
Level I: Very low risk (-3, -2, $n = 8$)	1.18	1.37	2.62	3.27
Level II: Below average risk (-1, 0, $n = 44$)	-	0.02	2.76	3.47
Level III: Average risk (1-3, $n = 89$)	-	-	3.24	3.51
Level IVa: Above average risk (4-5, $n = 47$)	-		-	0.46
Level IVb: Well above average risk (6+, $n = 12$)	-	-	-	-

Note: *** $p < .001$, ** $p < .01$, * $p < .05$

Figure 3.3: *Kaplan-Meier Survival Analysis: Rates of Any New Charges or Convictions Over Time for Static-99R Risk Groups*
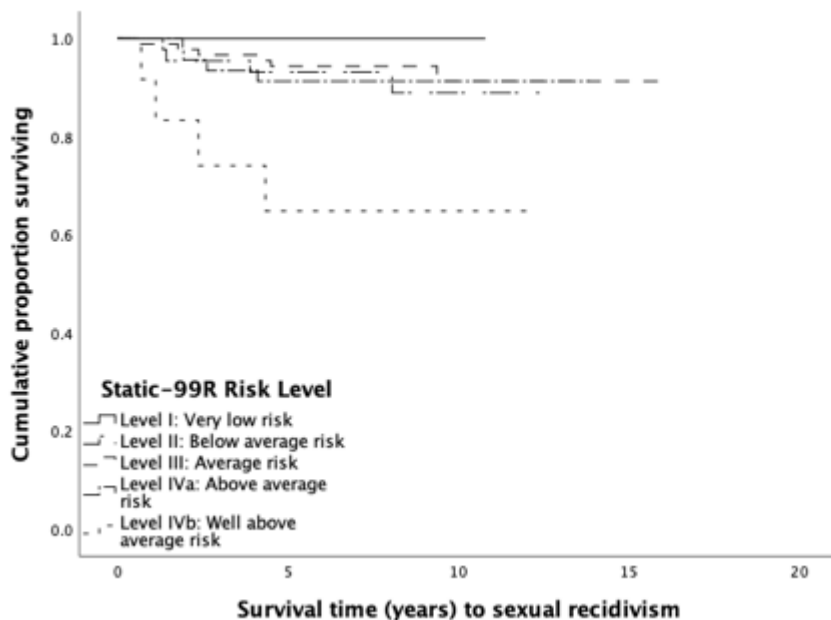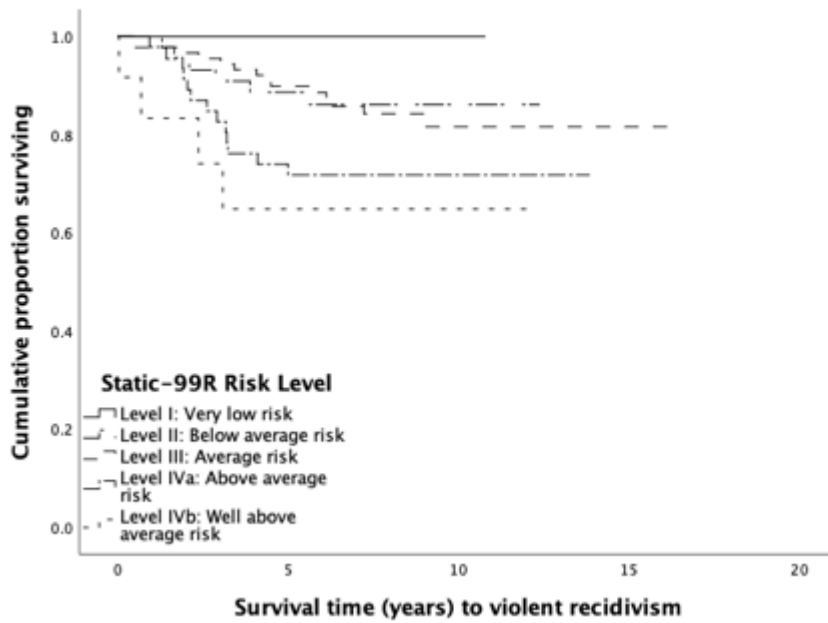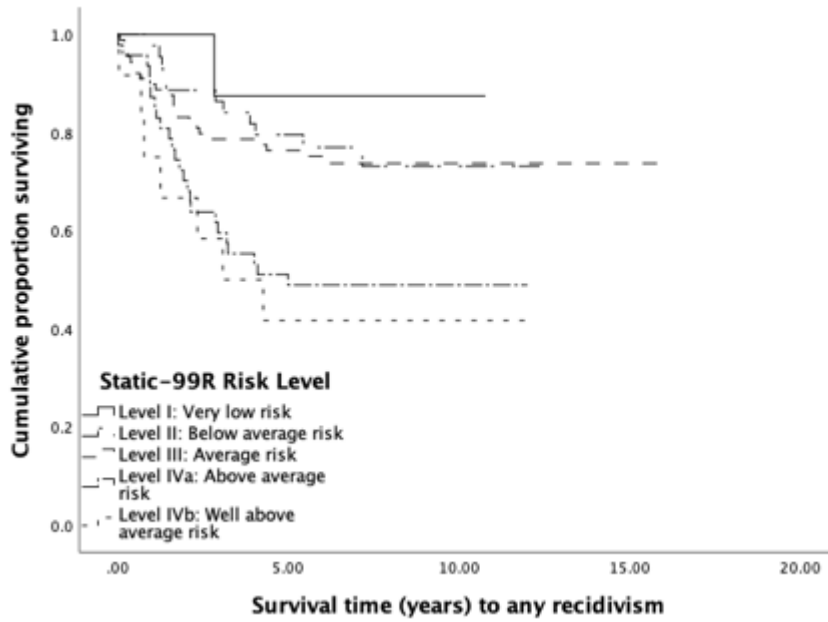

Table 3.12

Log Rank Chi Square Pairwise Comparison Static-99R Any New Charge or Conviction

Static-99R risk band	Pairwise comparison (log rank χ^2) reference group			
	Level II	Level III	Level IVa	Level IVb
Level I: Very low risk (-3, -2, *n* = 8)	0.59	0.74	3.44	3.99*
Level II: Below average risk (-1, 0, *n* = 44)	-	0.04	7.04	6.61*
Level III: Average risk (1-3, *n* = 89)	-	-	8.27**	6.00*
Level IVa: Above average risk (4-5, *n* = 47)	-		-	0.29
Level IVb: Well above average risk (6+, *n* = 12)	-	-	-	-

Note: *** *p* < .001, ** *p* < .01, * *p* < .05

Figure 3.4: *Kaplan-Meier Survival Analysis: Rates of New Sexual Charges or Convictions Over Time for VRS-SO Risk Groups*

Table 3.13

Log Rank Chi Square Pairwise Comparison VRS-SO Risk Group New Sexual Charge or Conviction

VRS-SO risk band	Pairwise comparison (log rank χ^2) reference group		
	Level II	Level III	Level IVa & IVb
Level I: Very low risk (<15, $n = 71$)	0.60	5.38*	1.82
Level II: Below average risk (15-23.5, $n = 65$)	-	-	0.49
Level III: Average risk (24-39.5, $n = 48$)	-	-	0.05
Level IVa & IVb: Above average risk/Well above average risk (40-67+, $n = 16$)	-	-	-

Note: *** $p < .001$, ** $p < .01$, * $p < .05$

Figure 3.5: *Kaplan-Meier Survival Analysis: Rates of New Violent (Including Sexual) Charges or Convictions Over Time for VRS-SO Risk Groups*
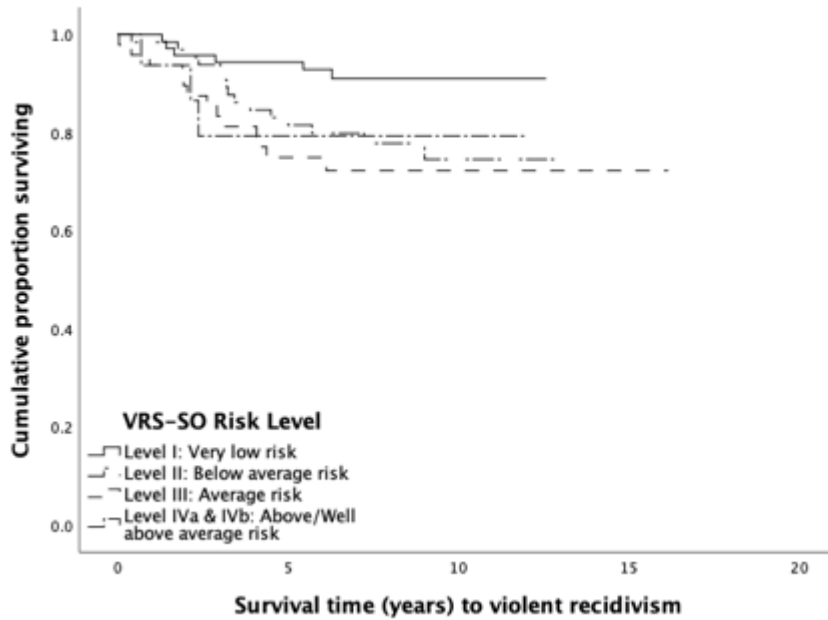
Table 3.14

Log Rank Chi Square Pairwise Comparison VRS-SO Total Score New Violent (Including Sexual) Charge or Conviction

VRS-SO risk band	Pairwise comparison (log rank χ^2) reference group		
	Level II	Level III	Level IVa & IVb
Level I: Very low risk (<15, *n* = 71)	5.04*	7.57**	2.03
Level II: Below average risk (15-23.5, *n* = 65)	-	-	0.01
Level III: Average risk (24-39.5, *n* = 48)	-	-	0.19
Level IVa & IVb: Above average risk/Well above average risk (40-67+, *n* = 16)	-	-	-

Note: *** $p < .001$, ** $p < .01$, * $p < .05$

Figure 3.6: *Kaplan-Meier Survival Analysis: Rates of Any New Charges or Convictions Over Time for VRS-SO Risk Groups*

Table 3.15

Log Rank Chi Square Pairwise Comparison VRS-SO Any New Charge or Conviction

VRS-SO risk band	Pairwise comparison (log rank χ^2) reference group		
	Level II	Level III	Level IVa & IVb
Level I: Very low risk (<15, *n* = 71)	4.93*	9.08**	11.72***
Level II: Below average risk (15-23.5, *n* = 65)	-	0.79	2.98
Level III: Average risk (24-39.5, *n* = 48)	-	-	1.09
Level IVa & IVb: Above average risk/Well above average risk (40-67+, *n* = 16)	-	-	-

Note: *** *p* < .001, ** *p* < .01, * *p* < .05

Figure 3.7: *Kaplan-Meier Survival Analysis: Rates of New Sexual Charges or Convictions Over Time for VRS-SO Dynamic Risk Groups*
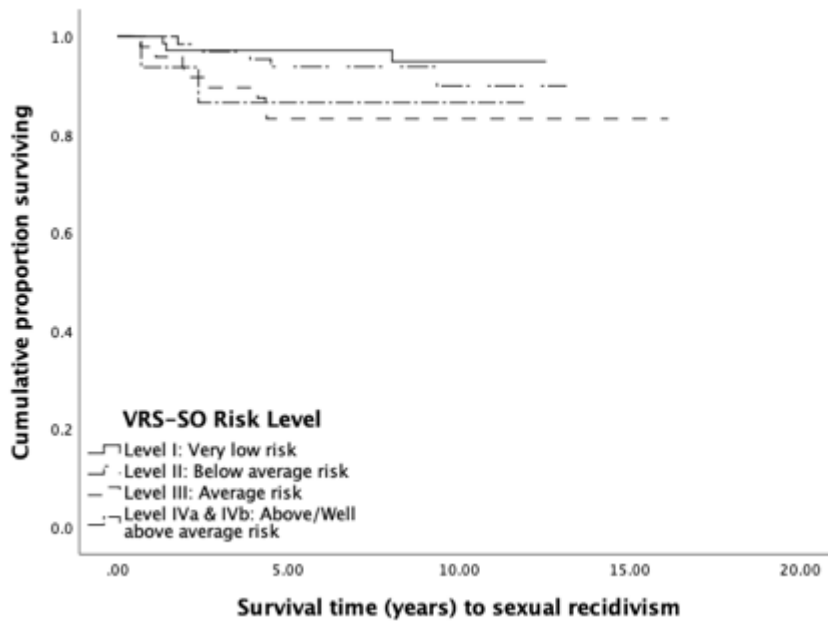

Table 3.16

Log Rank Chi Square Pairwise Comparison VRS-SO Dynamic New Sexual Charge or Conviction

VRS-SO risk band	Pairwise comparison (log rank χ^2) reference group		
	Level II	Level III	Level IVa & IVb
Level I: Very low risk (< 10.5, *n* = 77)	0.859	4.205*	5.409*
Level II: Below average risk (11-16.5, *n* = 49)	-	6.391*	8.354**
Level III: Average risk (17-27.5, *n* = 55)	-	-	0.318
Level IVa & IVb: Above average risk/Well above average risk (28-51, *n* = 19)	-	-	-

Note: *** $p < .001$, ** $p < .01$, * $p < .05$

Figure 3.8: *Kaplan-Meier Survival Analysis: Rates of New Violent (Including Sexual) Charges or Convictions Over Time for VRS-SO Dynamic Risk Groups*
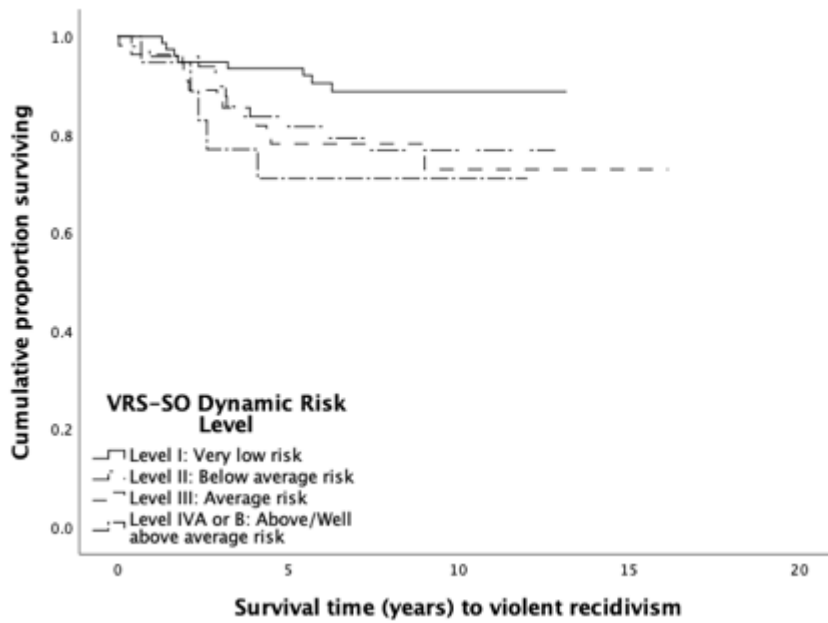
Table 3.17

Log Rank Chi Square Pairwise Comparison VRS-SO Dynamic New Violent (Including Sexual) Charges or Convictions

VRS-SO risk band	Pairwise comparison (log rank χ^2) reference group		
	Level II	Level III	Level IVa & IVb
Level I: Very low risk (0-10.5, $n = 77$)	3.05	4.05*	4.21*
Level II: Below average risk (11-16.5, $n = 49$)	-	0.08	0.41
Level III: Average risk (17-27.5, $n = 55$)	-	-	0.13
Level IVa & IVb: Above average risk/Well above average risk (28-51, $n = 19$)	-	-	-

Note: *** $p < .001$, ** $p < .01$, * $p < .05$

68

Figure 3.9: *Kaplan-Meier Survival Analysis: Rates of Any New Charges or Convictions Over Time for VRS-SO Dynamic Risk Groups*
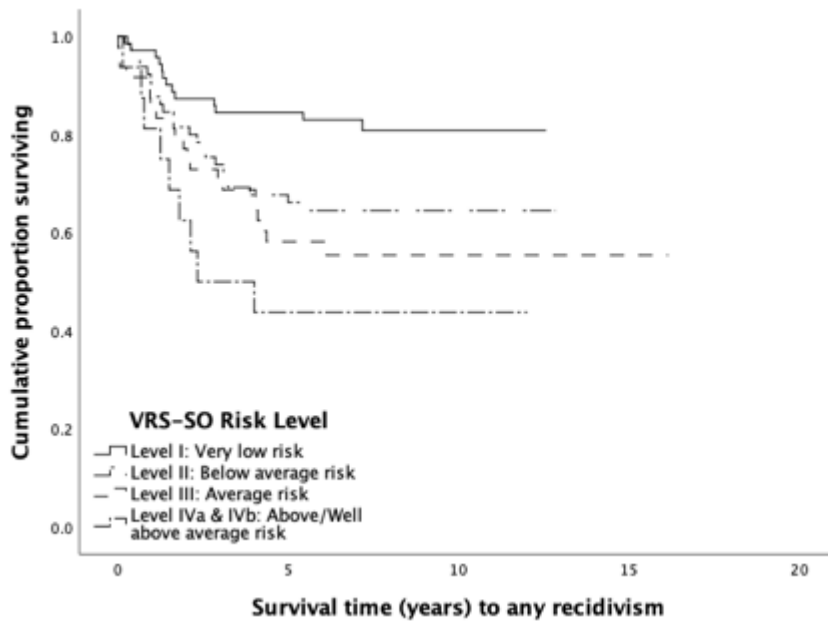

Table 3.18

Log Rank Chi Square Pairwise Comparison VRS-SO Dynamic: Any New Charge or Conviction

VRS-SO risk band	Pairwise comparison (log rank χ^2) reference group		
	Level II	Level III	Level IVa & IVb
Level I: Very low risk (0-10.5, $n = 77$)	2.61	4.61*	4.41*
Level II: Below average risk (11-16.5, $n = 49$)	-	0.21	5.82*
Level III: Average risk (17-27.5, $n = 55$)	-	-	-
Level IVa & IVb: Above average risk/Well above average risk (28-51, $n = 19$)	-	-	-

Note: *** $p < .001$, ** $p < .01$, * $p < .05$

Figure 3.10: *Kaplan-Meier Survival Analysis: Rates of New Sexual Charges or Convictions Over Time for STABLE-2007 Risk/Need Groups*
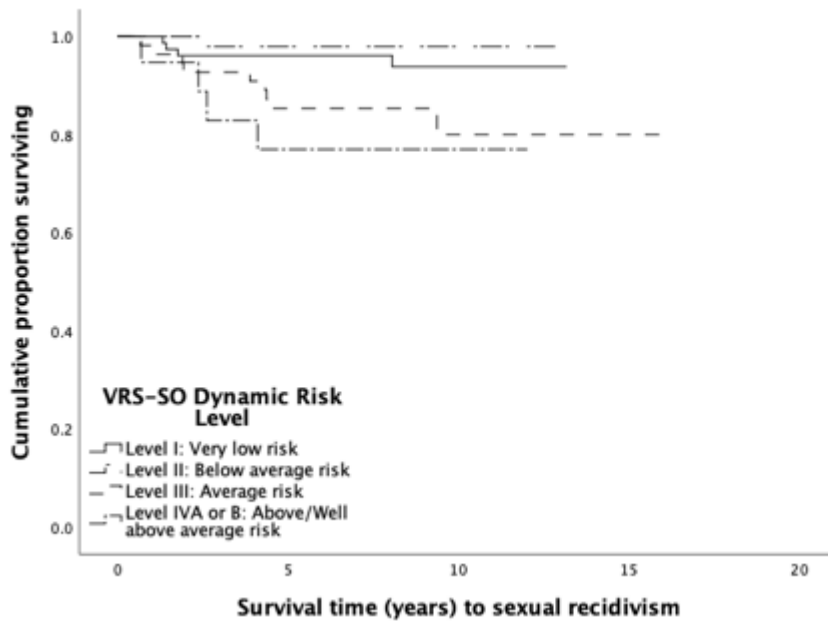

Table 3.19

Log Rank Chi Square Pairwise Comparison STABLE-2007 New Sexual Charge or Conviction

STABLE-2007 risk band	Pairwise comparison (log rank χ^2) reference group	
	Medium (4-11, $n = 119$)	High (12+, $n = 39$)
Low (0-3, $n = 42$)	0.03	6.23*
Medium (4-11, $n = 119$)	-	10.42***

Note: *** $p < .001$, ** $p < .01$, * $p < .05$

Figure 3.11: *Kaplan-Meier Survival Analysis: Rates of New Violent (Including Sexual) Charges or Convictions Over Time for STABLE-2007 Risk/Need Groups*
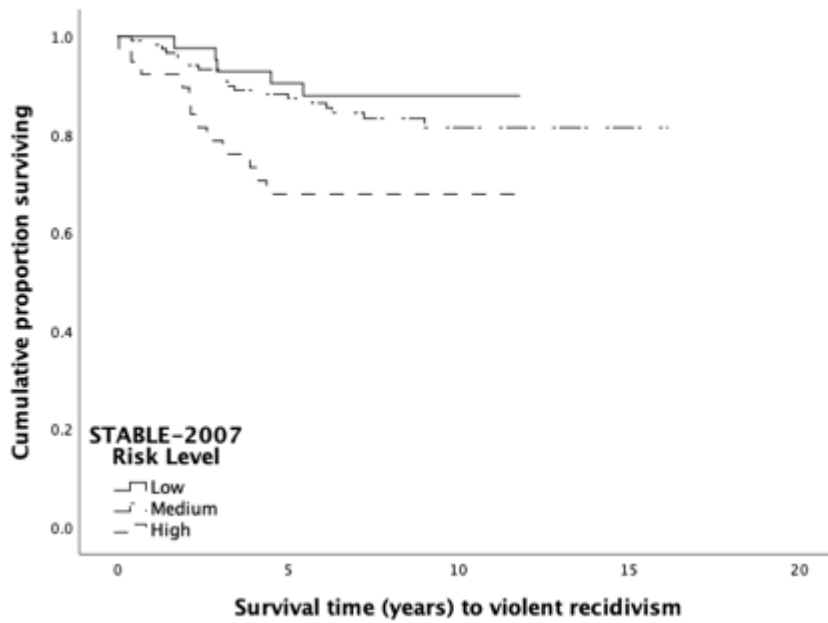
Table 3.20

Log Rank Chi Square Pairwise Comparison STABLE-2007 New Violent (Including Sexual) Charge or Conviction

STABLE-2007 risk band	Pairwise comparison (log rank χ^2) reference group	
	Medium (4-11, $n = 119$)	High (12+, $n = 39$)
Low (0-3, $n = 42$)	0.38	5.18*
Medium (4-11, $n = 119$)	-	4.51*

Note: *** $p < .001$, ** $p < .01$, * $p < .05$

Figure 3.12: *Kaplan-Meier Survival Analysis: Rates of Any New Charges or Convictions Over Time for STABLE-2007 Risk/Need Groups*

Table 3.21

Log Rank Chi Square Pairwise Comparison STABLE-2007 Any New Charge or Conviction

STABLE-2007 risk band	Pairwise comparison (log rank χ^2) reference group	
	Medium (4-11, n = 119)	High (12+, n = 39)
Low (0-3, n = 42)	1.51	7.21*
Medium (4-11, n = 119)	-	4.67*

Note: *** $p < .001$, ** $p < .01$, * $p < .05$

Figure 3.13: *Kaplan-Meier Survival Analysis: Rates of New Sexual Charges or Convictions Over Time for SOTIPS Risk/Need Groups*
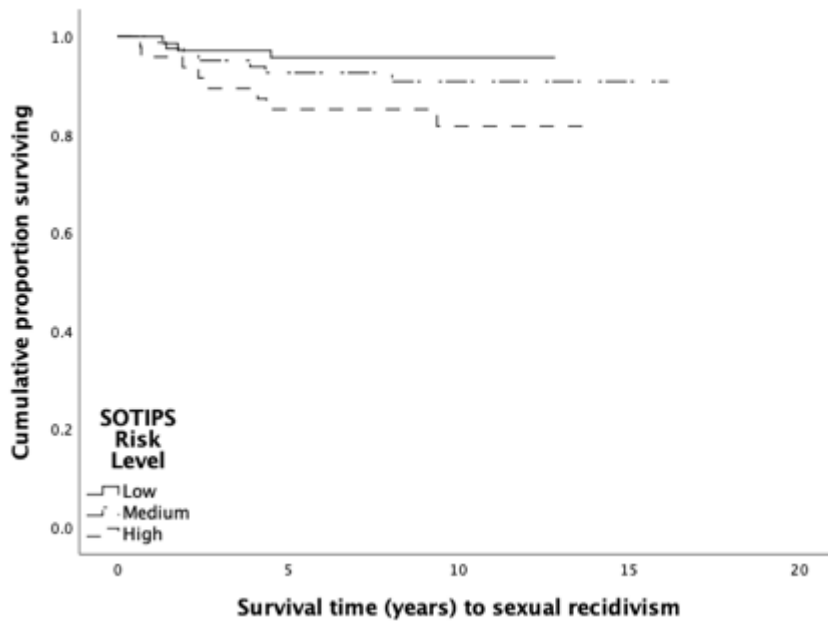
Table 3.22

Log Rank Chi Square Pairwise Comparison SOTIPS New Sexual Charge or Conviction

SOTIPS risk band	Pairwise comparison (log rank χ^2) reference group	
	Medium (11-20, *n* = 82)	High (21-48, *n* = 48)
Low (0-10, *n* = 70)	1.04	4.66*
Medium (11-20, *n* = 82)	-	1.79

Note: *** *p* < .001, ** *p* < .01, * *p* < .05

Figure 3.14: *Kaplan-Meier Survival Analysis: Rates of New Violent (Including Sexual) Charges or Convictions Over Time for SOTIPS Risk/Need Groups*
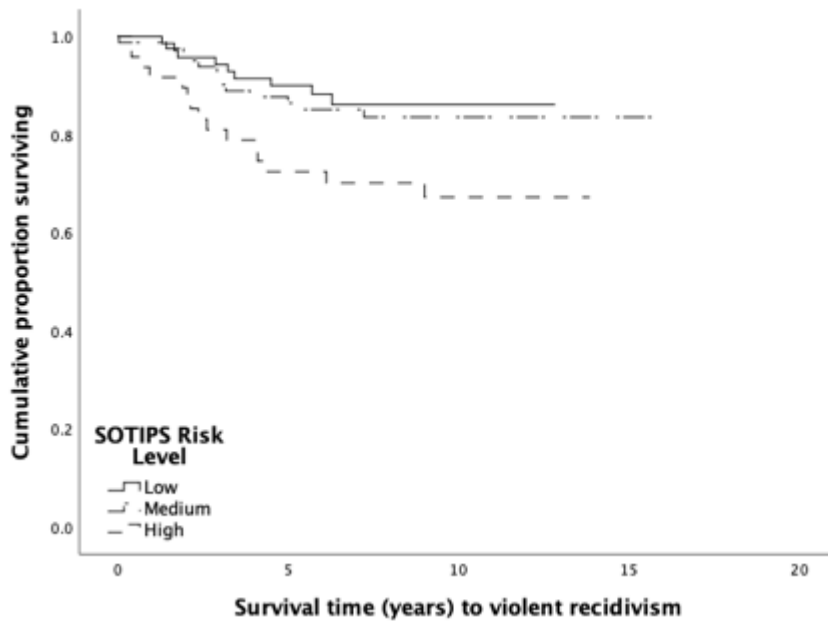
Table 3.23

Log Rank Chi Square Pairwise Comparison SOTIPS New Violent (Including Sexual) Charge or Conviction

SOTIPS risk band	Pairwise comparison (log rank χ^2) reference group	
	Medium (11-20, n = 82)	High (21-48, n = 48)
Low (0-10, n = 70)	0.25	5.64*
Medium (11-20, n = 82)	-	4.27*

Note: *** $p < .001$, ** $p < .01$, * $p < .05$

74

Figure 3.15: *Kaplan-Meier Survival Analysis: Rates of Any New Charges or Convictions Over Time for SOTIPS Risk/Need Groups*
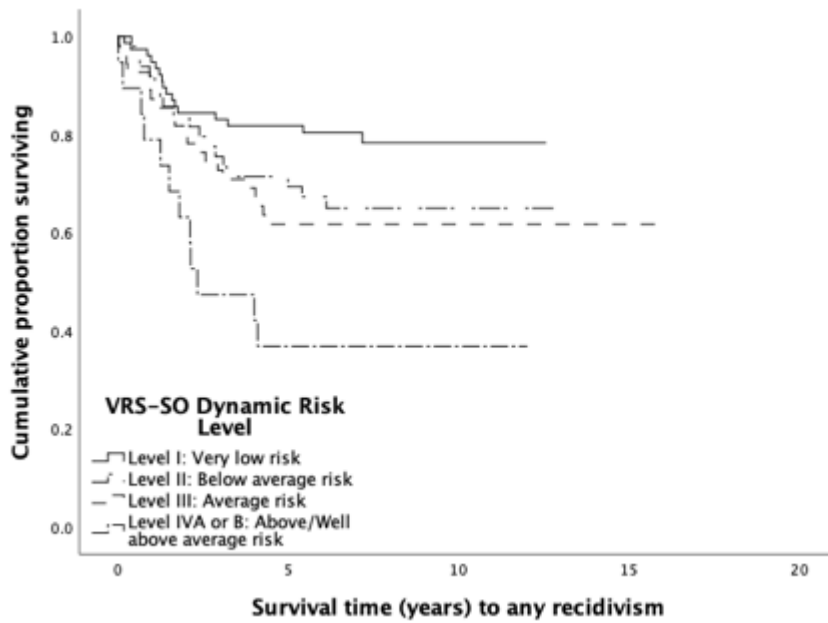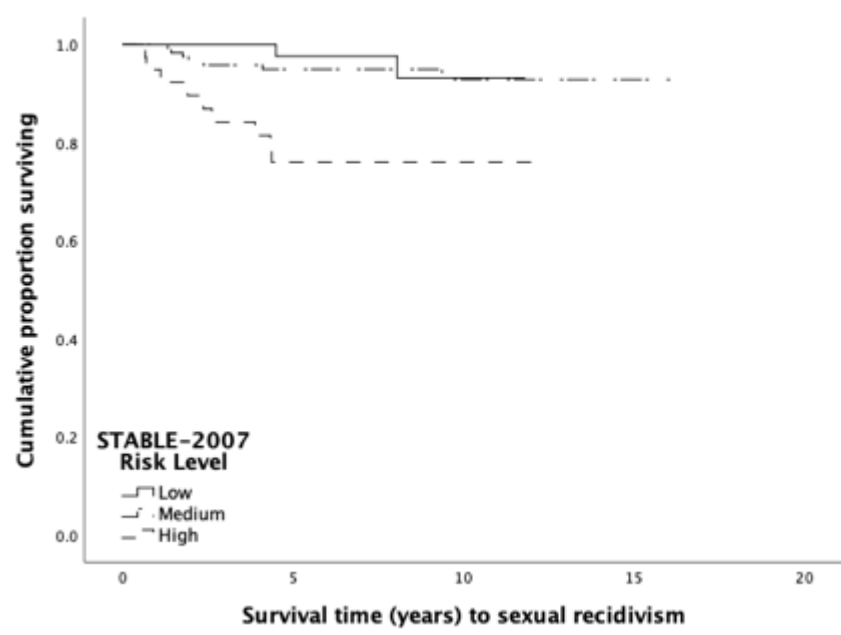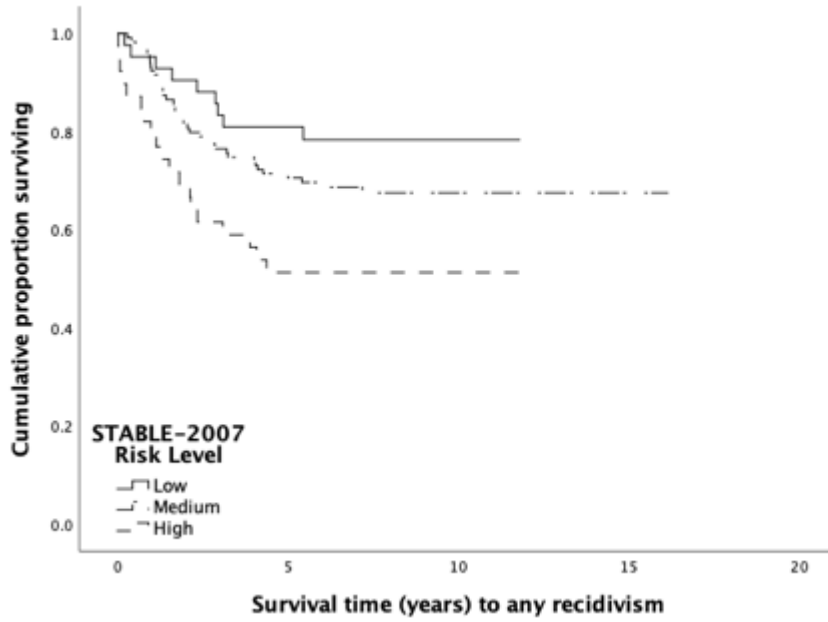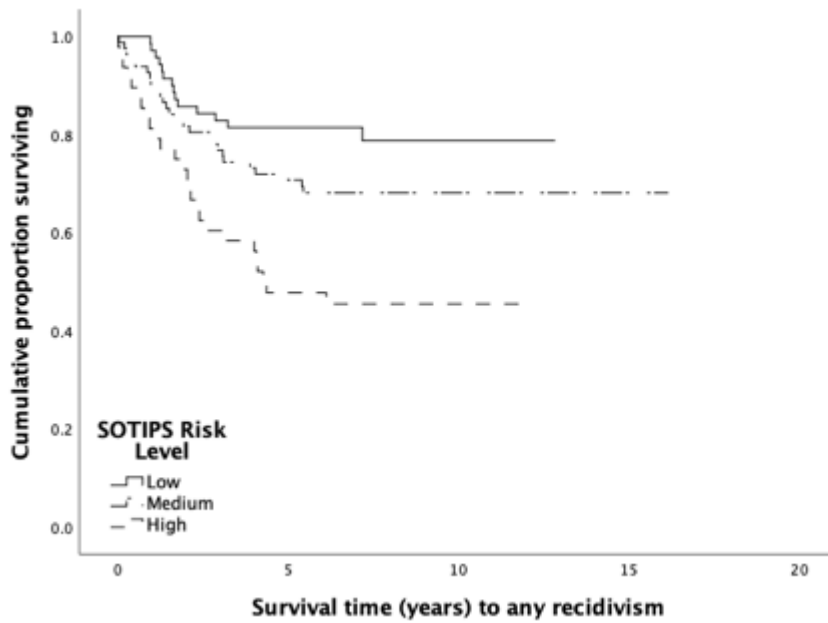

Table 3.24

Log Rank Chi Square Pairwise Comparison SOTIPS Any New Charge or Conviction

SOTIPS risk band	Pairwise comparison (log rank χ^2) reference group	
	Medium (11-20, $n = 82$)	High (21-48, $n = 48$)
Low (0-10, $n = 70$)	2.55	14.85***
Medium (11-20, $n = 82$	-	6.52*

Note: *** $p < .001$, ** $p < .01$, * $p < .05$

3.5.3.2 Chi Square and Cramer's V. The second set of calibration analyses examined rates of sexual, violent (including sexual), and any recidivism for each risk measure employing 5-year fixed follow-up using chi square models and Cramer's V measure of association. This produces actual observed rates of recidivism associated with a given risk band while mechanically controlling for time at risk in the community. As V represents associations between a categorical variable and a binary outcome, the Rice and Harris (2005) conventions for interpreting correlation magnitude (i.e., for a point biserial correlation) were used.

Risk/need bands for nearly all risk instruments were significantly associated with fixed 5-year rates of sexual, violent (including sexual), and any recidivism; effects were broadly medium in magnitude (Table 3.25). The one exception: SOTIPS risk/need bands were significant for violent and any recidivism, but not sexual recidivism; effect sizes for all outcomes ranged from small to medium.

Table 3.25

Associations Between Risk Measure Category and 5-Year Fixed Follow-up

Measure	5-year		
	χ^2	p	V
	Sexual		
Static-99R	13.12	.011	0.26
VRS-SO Total	8.48	.037	0.21
VRS-SO Dynamic	12.65	.005	0.26
STABLE-2007	16.04	.000	0.29
SOTIPS	4.42	.110	0.15
	Violent		
Static-99R	13.94	.007	0.27
VRS-SO Total	11.04	.012	0.24
VRS-SO Dynamic	9.10	.028	0.22
STABLE-2007	10.90	.004	0.24
SOTIPS	8.39	.015	0.21
	General		
Static-99R	16.97	.002	0.30
VRS-SO Total	14.48	.002	0.27
VRS-SO Dynamic	13.55	.004	0.26
STABLE-2007	7.70	.021	0.20
SOTIPS	15.97	.000	0.29

3.5.3.3 E/O Index. The final calibration analysis employed the E/O index to estimate rates of sexual recidivism associated with each VRS-SO risk band using 5-year fixed follow-ups. The current sample was divided into five risk bands. Inconsistent with the VRS-SO normative sample, the highest concentration of men in the current sample were classified as Level I (Very low risk), followed by

Level II (below average risk), and then Level III (Average risk), whereas in the normative sample, the highest concentration of men were classified as Level III (Average risk), with successively smaller proportions in the more extreme bands at each end.

Logistic regression (LR) was conducted to estimate the rates of recidivism associated with specific VRS-SO scores for each risk band over a 5-year follow-up for sexual recidivism. The Hosmer-Lemeshow goodness of fit test was nonsignificant ($\chi^2 = 5.901$, $df = 8$, p = .658), suggesting that the logistic distributions provided a reasonable approximation of 5-year sexual recidivism rates to permit modelling. The risk band ns in the present sample were then multiplied by the recidivism frequencies reported in Olver et al. (2018) to obtain the expected number of recidivists for calibration analyses. The observed recidivism rates were the predicted rates of 5-year sexual recidivism derived from logistic regression for each risk band in the present sample. No E/O index could be generated for the Level I (very low risk) band, as there were no expected recidivists. Results of the E/O index calibration analyses are reported in Table 3.26 with the VRS-SO normative sample.

Examining Table 3.26, when logistic regression estimates were employed, none of the differences between observed and expected recidivism rates were statistically significant, as indicated by 95% confidence intervals that overlap with 1. The recidivism norms for the VRS-SO underpredicted the total number of individuals who recidivated in this sample. Specifically, the scale predicted 38% less recidivists compared to what was observed. Looking at each risk category individually, calibration accuracy increased as the level of risk increased, with Level IVb achieving nearly perfect calibration. The VRS-SO was also well calibrated for individuals in Level IVa, where the scale predicted about 10% less individuals who recidivated than what was observed. The scale predicted about half of the observed number of individuals who recidivated in Level II, and about three-quarters of the observed individuals who recidivated in Level III. Overall, the VRS-SO tends to underestimate recidivism for lower risk individuals, though these differences lessen as the risk level increases, and none of the differences are statistically significant.

Table 3.26

Calibration analyses: Logistic Regression Base E/O Index VRS-SO comparing current sample to normative sample on rates of 5-year sexual recidivism

VRS-SO common language metric		Total score	*n*	Expected VRS-SO norms		Observed Community sample		E/O Index	95% CI
Risk Category				%	*n*	%	*n*		
I	Very low	<15	69	0.0	0	3.0	2.1	-	-
II	Below average	15-23.5	65	3.1	2.0	6.2	4.0	0.50	0.19, 1.32
III	Average	24-39.5	47	9.3	4.4	13.0	6.1	0.73	0.32, 1.58
IVa	Above average	40-49.5	11	22.6	2.5	25.1	2.8	0.90	0.28, 2.93
IVb	Well above average	50-72	3	36.7	1.1	37.2	1.1	0.98	0.15, 6.27
Total		0-72	195		10.0		16.1	0.62	0.38, 1.01

Note. CI = Confidence Interval

Examining Table 3.27, when raw recidivism numbers were employed, none of the differences between observed and expected recidivism rates were statistically significant, as indicated by 95% confidence intervals that overlap with 1. The recidivism norms for the VRS-SO underpredicted the total number of individuals who recidivated in this sample. Specifically, the scale predicted 37% less recidivists compared to what was observed. Examining each risk category in kind, calibration accuracy increased between Level II and III, with each level predicting half and slightly more than half of the observed number of individuals who recidivated., respectively. For Level IVa, the scale predicted 250% more individuals who recidivated than what was observed. The VRS-SO was well calibrated for Level IVb, with the scale predicting about 10% more individuals who recidivated than what was observed. Overall, the VRS-SO tends to underestimate recidivism for lower risk individuals and overpredict recidivism (either slightly or considerably) for higher risk individuals, however, none of these differences are statistically significant.

Table 3.27

Calibration analyses: E/O Index VRS-SO comparing current sample to normative sample on rates of 5-year sexual recidivism

VRS-SO common language metric				Expected VRS-SO norms		Observed Community sample			
Risk Category		Total score	*n*	%	*n*	%	*n*	E/O Index	95% CI
I	Very low	<15	69	0.0	0	2.9	2.0	-	-
II	Below average	15-23.5	65	3.1	2.0	6.2	4.0	0.50	0.19, 1.33
III	Average	24-39.5	47	9.3	4.4	17.0	8.0	0.55	0.28, 1.10
IVa	Above average	40-49.5	11	22.6	2.5	9.0	1.0	2.50	0.35, 17.75
IVb	Well above average	50-72	3	36.7	1.1	33.3	1.0	1.10	0.16, 7.81
Total		0-72	195		10.0		16.0	0.63	0.38, 1.02

Note. CI = Confidence Interval

3.5.4 Incremental Validity

3.5.4.1 Cox Regression Survival Analyses. Cox regression survival analyses were conducted to examine whether the dynamic forensic measures demonstrate unique and incremental validity over static measures in the prediction of sexual, violent (including sexual), and any recidivism over time (see Table 3.28, 3.29, and 3.30, respectively) Similarly, Cox regression survival analysis was also used to examine whether VRS-SO factor scores added incrementally to the prediction of each recidivism outcome over time, controlling for static risk.

For sexual recidivism static-dynamic pairings (Table 3.28), VRS-SO dynamic scores significantly uniquely predicted this outcome over VRS-SO static scores (model 1), while neither the Static-99R and VRS-SO dynamic scores (model 2) or Static-99R and STABLE-2007 pairing (model 3) were significantly predictive. In the Static-99R and SOTIPS pairing, SOTIPS significantly uniquely predicted sexual recidivism (model 4), while neither Static-99R or SAPROF-SO were uniquely incrementally predictive of this outcome (model 5). When SAPROF-SO was entered with Static-99R and VRS-SO dynamic scores, none of the measures emerged as significantly uniquely predictive of sexual recidivism. When the three VRS-SO factors were entered together, Sexual Deviance was incrementally predictive of this outcome (model 7); this remained true when Static-99R was added in (model 8).

For violent (including sexual) recidivism (Table 3.29), VRS-SO static and VRS-SO dynamic scores did not significantly and uniquely predict this outcome (model 9), while only Static-99R significantly uniquely predicted violent recidivism when paired with VRS-SO dynamic scores (model 10) and STABLE-2007 (model 11). In the Static-99R and SOTIPS pairing, neither measure

significantly uniquely predicted this outcome, though Static-99R approached significance (model 12); Static-99R was the unique predictor of violent recidivism when paired with SAPROF-SO (model 13), and with VRS-SO dynamic scores and SAPROF-SO (model 14). Of the three VRS-SO factors, only Criminality incrementally predicted violent recidivism (model 15), and in tandem with Static-99R, only Static-99R was predictive of this outcome (model 16).

Regarding static-dynamic pairings for any recidivism (Table 3.30), neither VRS-SO static scores or VRS-SO dynamic scores were uniquely predictive of this outcome (model 17), while Static-99R was significantly uniquely predictive when entered with VRS-SO dynamic scores (model 18). The dynamic measures were significantly uniquely predictive for any recidivism over Static-99R for both STABLE-2007 (model 19) and SOTIPS (model 20), while both Static-99R and SAPROF-SO were incrementally predictive (model 21). When Static-99R, VRS-SO dynamic, and SAPROF-SO were entered together, only SAPROF-SO was incrementally predictive of any recidivism. Of the three VRS-SO factors, only Criminality incrementally predicted this outcome (model 23), and in tandem with Static-99R, both Static-99R and VRS-SO criminality factor incrementally predicted any recidivism (model 24).

Table 3.28

Incremental Validity: Cox Regression Survival Analyses for the Forensic Measures and Sexual Recidivism

Regression Models	Sexual Recidivism					
	B	SE	Wald	e^B	*p*	95% CI
Model 1						
VRS-SO Static	-.01	.07	.04	.99	.844	.85, 1.14
VRS-SO Dynamic	.07	.03	5.39	1.07	**.020**	1.01, 1.13
Model 2						
Static-99R	.11	.13	.66	1.11	.415	.86, 1.44
VRS-SO Dynamic	.05	.03	2.83	1.05	.092	.99, 1.11
Model 3						
Static-99R	.11	.13	.71	1.12	.401	.86, 1.45
STABLE-2007	.13	.06	2.61	1.10	.106	.98, 1.22
Model 4						
Static-99R	.12	.12	.95	1.13	.330	.89, 1.43
SOTIPS	.06	.03	3.93	1.06	**.047**	1.00, 1.12
Model 5						
Static-99R	.17	.12	2.10	1.19	.148	.94, 1.50
SAPROF-SO	-.02	.02	2.11	.98	.146	.95, 1.01
Model 6						
Static-99R	.09	.13	.49	1.10	.484	.85, 1.43
VRS-SO Dynamic	.04	.03	1.42	1.04	.233	.98, 1.11
SAPROF-SO	-.01	.02	.71	.99	.399	.96, 1.02
Model 7						
VRS-SO Sexual deviance	.142	.05	8.29	1.15	**.004**	1.05, 1.27
VRS-SO Criminality	.063	.07	.80	1.07	.371	.93, 1.22
VRS-SO Treatment responsivity	-.106	.09	1.55	.90	.213	.76, 1.06
Model 8						
Static-99R	.096	.12	.60	1.10	.437	.86, 1.40
VRS-SO Sexual deviance	.127	.05	5.68	1.14	**.017**	1.02, 1.26
VRS-SO Criminality	.043	.07	.33	1.04	.564	.90, 1.21
VRS-SO Treatment responsivity	-.106	.09	1.54	.90	.214	.76, 1.06

Note: Significant p-values are bolded.

Table 3.29

Incremental Validity: Cox Regression Survival Analyses for the Forensic Measures and Violent (Including Sexual) Recidivism

	B	SE	Wald	e^B	p	95% CI
			Violent (Including sexual) Recidivism			
Regression Models						
Model 9						
VRS-SO Static	.07	.05	1.69	1.07	.194	.97, 1.18
VRS-SO Dynamic	.02	.02	.67	1.02	.415	.98, 1.06
Model 10						
Static-99R	.22	.09	5.78	1.25	**.016**	1.04, 1.50
VRS-SO Dynamic	.004	.02	.04	1.00	.852	.97, 1.04
Model 11						
Static-99R	.20	.10	4.74	1.23	**.029**	1.02, 1.47
STABLE-2007	.02	.04	.24	1.02	.622	.94, 1.10
Model 12						
Static-99R	.16	.09	3.55	1.17	.060	.99, 1.38
SOTIPS	.04	.02	3.03	1.04	.082	1.00, 1.08
Model 13						
Static-99R	.18	.08	4.68	1.19	**.031**	1.02, 1.40
SAPROF-SO	-.02	.01	2.85	.98	.091	.96, 1.00
Model 14						
Static-99R	.21	.09	4.91	1.23	**.027**	1.02, 1.48
VRS-SO Dynamic	-.02	.02	.46	.98	.497	.94, 1.03
SAPROF-SO	-.02	.01	3.35	.98	.067	.96, 1.00
Model 15						
VRS-SO Sexual deviance	.004	.04	.01	1.00	.924	.93, 1.08
VRS-SO Criminality	.10	.05	4.42	1.11	**.036**	1.01, 1.22
VRS-SO Treatment responsivity	-.04	.06	.51	.96	.477	.86, 1.08
Model 16						
Static-99R	.28	.10	7.80	1.33	**.005**	1.09, 1.62
VRS-SO Sexual deviance	-.05	.04	1.26	.95	.263	.88, 1.04
VRS-SO Criminality	.05	.05	1.04	1.05	.307	.95, 1.17
VRS-SO Treatment responsivity	-.02	.06	.08	.98	.782	.88, 1.10

Note: Significant p-values are bolded.

Table 3.30

Incremental Validity: Cox Regression Survival Analyses for the Forensic Measures and Any Recidivism

	B	SE	Any Recidivism Wald	e^B	*p*	95% CI
Regression Models						
Model 17						
VRS-SO Static	.07	.04	3.42	1.07	.064	1.00, 1.15
VRS-SO Dynamic	.03	.02	3.34	1.03	.067	1.00, 1.06
Model 18						
Static-99R	.14	.07	4.39	1.15	**.036**	1.01, 1.31
VRS-SO Dynamic	.03	.02	2.71	1.03	.100	1.00, 1.06
Model 19						
Static-99R	.12	.07	3.17	1.13	.075	.99, 1.28
STABLE-2007	.06	.03	4.17	1.07	**.041**	1.00, 1.13
Model 20						
Static-99R	.08	.06	1.70	1.08	.192	.96, 1.22
SOTIPS	.06	.02	12.85	1.06	**.000**	1.03, 1.10
Model 21						
Static-99R	.14	.06	5.49	1.15	**.019**	1.02, 1.29
SAPROF-SO	-.02	.01	6.66	.98	**.010**	.97, 1.00
Model 22						
Static-99R	.12	.07	3.48	1.13	.062	.99, 1.29
VRS-SO Dynamic	.01	.02	.22	1.01	.637	.97, 1.04
SAPROF-SO	-02	.01	4.35	.98	**.037**	.97, 1.00
Model 23						
VRS-SO Sexual deviance	-.02	.03	.40	.98	.530	.93, .1.04
VRS-SO Criminality	.18	.04	26.67	1.20	**<.001**	1.12, 1.29
VRS-SO Treatment responsivity	-.03	.04	.58	.97	.447	.89, 1.05
Model 24						
Static-99R	.16	.07	5.11	1.17	**.024**	1.02, 1.34
VRS-SO Sexual deviance	-.05	.03	2.23	.95	.135	.89, 1.02
VRS-SO Criminality	.15	.04	14.91	1.16	**<.001**	1.08, 1.25
VRS-SO Treatment responsivity	.02	.04	.193	.98	.661	.90, 1.07

Note: Significant p-values are bolded.

3.5.5 Construct Validity

3.5.5.1 Confirmatory Factor Analysis of VRS-SO Dynamic Items. An EFA was initially conducted to assess the model fit for the current data using the three-factor model from previous research (Olver et al., 2007), which indicated an appropriate fit. Given that item D15 (Treatment Compliance) could not be reliably coded (missing n = 179), it was excluded from factor analyses.

The first CFA model results generated a correlated three-latent factor model of the pre-treatment dynamic items with very good fit: CFI = .957, TLI = .949, RMSEA = .084, χ^2 (df = 101) =

244.408, $\chi^2/df = 2.42$. However, modification indices indicated that D11 (Release to High Risk Situations) was better represented as a cross loading on Sexual Deviance ($\chi^2 = 94.333$, $p < .001$) and Criminality ($\chi^2 = 87.857$, $p < .001$), rather than Treatment Responsivity. Changing these parameters resulted in an improved model with the following values indicative of acceptable model fit: CFI = .986, TLI = .983, RMSEA = .048, $\chi^2(df = 100) = 146.419$, $\chi^2/df = 1.464$.

Item loadings were substantial in both the initial (all > .40) and final (all > .45) models and both models were significant ($p < .001$). Factor 2 is significantly associated with Factor 1 and 3 ($ps < .001$), but Factor 1 and 3 are not significantly associated. The final model is presented in Figure 3.16. Given the close correspondence of the two models, factors from the original model (i.e., D11 loading on Treatment Responsivity rather than Sexual Deviance and Criminality; Olver et al., 2007) were employed for all other analyses.

Figure 3.16 *Confirmatory Factor Analysis of VRS-SO dynamic items. Final correlated three-factor model with standardized loadings shown*

84
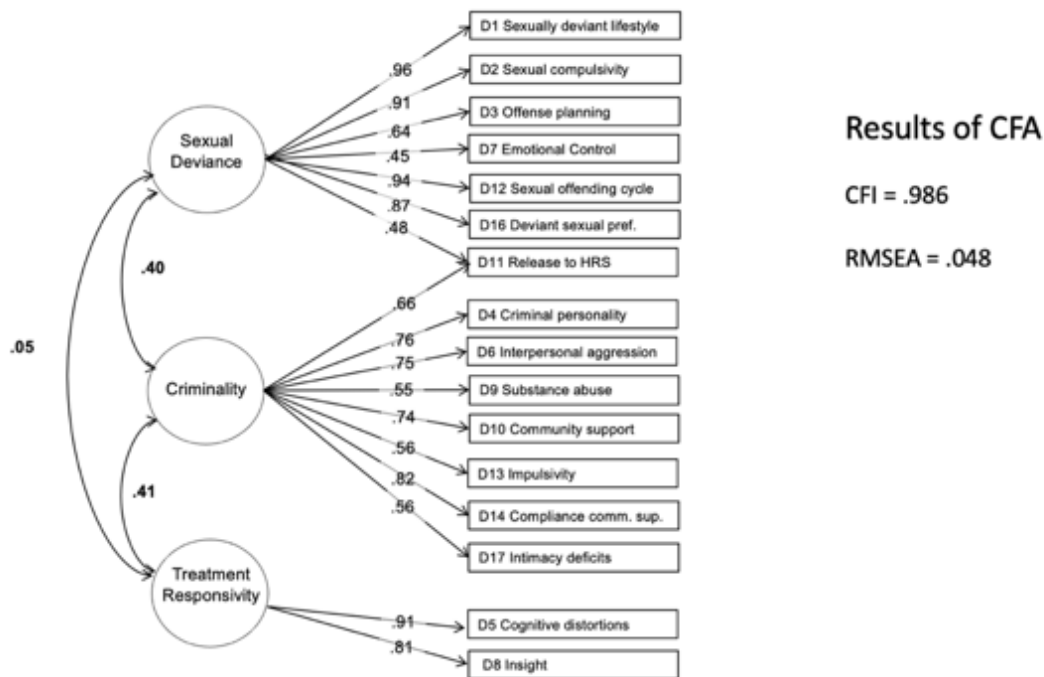
CHAPTER 4: DISCUSSION

4.1 General Discussion

The current study was a community-based investigation of several forensic measures used to assess sexual violence risk and protective factors (Static-99R, VRS-SO, STABLE-2007, SOTIPS, and SAPROF-SO) in a community-supervised sample of men with convictions for sexual offenses who were participating in sexual offense specific treatment. This study aimed to extend the literature by examining the risk-related, protection-related, and psychometric properties of these tools. While a large body of research exists for Static-99R, VRS-SO, and STABLE-2007, the VRS-SO has not previously been validated in a non-incarcerated sample. This study also aimed to bolster the research base for two more recently developed tools, the SOTIPS and SAPROF-SO. SOTIPS is a statistically derived risk assessment and treatment planning tool, while SAPROF-SO is specifically designed as an assessment and treatment planning tool that measures protective factors theoretically and empirically related to desistance from sexual offending.

The sample for this study comprised 200 files of men with convictions for sexual offenses who were assessed and often attended treatment programming at Forensic Assessment and Community Services (FACS) in Edmonton, Alberta between 2009 and 2015. FACS is a division of Alberta Health Services that provides court-mandated individuals with assessment and intervention services. Several forensic measures (VRS-SO, STABLE-2007, SOTIPS, and SAPROF-SO) were rated based on file information by the student researcher and a student research assistant. Roughly half of Static-99R and about one-third of STABLE-2007 measures were rated by the treating clinician and was extracted directly from the file when available. The mean age of the total sample at the time of FACS assessment was 39.6 ($SD = 14.4$), over half of the men (57%) were White or of European descent, 14.5% were Indigenous (First Nations, Métis, or Inuit), and the remaining 28.5% encompassed a diverse number of ethnocultural groups (South or Southeast Asian, Central American, West-Asian/Middle Eastern, African, and unidentified or mixed ancestry). The ethnic breakdown of the current sample was consistent with previous incarcerated and community-based samples regarding the proportion of White/European individuals but had a smaller proportion of Indigenous individuals and larger proportion of other ethnocultural groups (Jung & Wielinga, 2019; Olver et al., 2020).

In the current sample, 57% of men were employed at the time of the index offense, 64% had lived with an intimate partner for at least two years at some point in their lives, roughly half of the sample had completed grade 12 (49%), and approximately 14% had impaired cognitive abilities. About one-quarter (24%) of the men had a major mental illness, 16% were diagnosed with a substance use

disorder, 16.5% with a paraphilia, and 8.5% with a personality disorder. Over two-thirds of the men were assessed at a below average risk level (68%) according to the VRS-SO common language metric risk bins. About half of the individuals completed the treatment program (51%); the average attendance length in the program was 6.3 months.

In the present sample, 46% of the men were classified as having adult or teen victims, 23.5% as having extrafamilial child victims, 25.5% as having intrafamilial victims, and 4% had a mixed adult/teen and child victim profiles. According to VRS-SO static scale mean scores, individuals with a mixed victim profile were highest risk group, followed by those with extrafamilial child victims, extrafamilial adult/teen victims, and intrafamilial victims. Individuals with a mixed victim profile were the most sexually deviant group (as defined by the VRS-SO Sexual Deviance factor) and the most criminally oriented group (according to the VRS-SO Criminality factor).

Recidivism information was collected from federal criminal records (CPIC) in 2021; average follow-up time was over eight years. Compared to previous studies with community-supervised samples, the present study has a longer follow-up time (e.g., 41 months, Hanson et al., 2007; three years, McGrath et al., 2012), but shorter than studies that utilized an incarcerated sample (e.g., 12 years, Beggs & Grace, 2010; nine years; Sowden & Olver 2017).

During the follow-up time, 9% of the present sample acquired a new sexual offense. Individuals with sexual offense convictions who are being assessed at FACS are often a lower risk group than individuals who receive federal sentences or are assessed while in custody; this may contribute to lower recidivism base rates in the current study. However, sexual recidivism rates are comparable to previous research with a shorter follow-up time (Hanson et al., 2007), though Hanson et al.'s study used a broader definition of recidivism (i.e., included self-report of individual rather than just charges and convictions). Sexual recidivism base rates are also higher in the present study than McGrath et al. (2012). Regarding violent and general recidivism, about one-fifth (18.5%) of the present sample committed a new violent (including sexual) offense and roughly one-third (33.3%) acquired at least one new offense. These violent and general recidivism base rates are higher than Hanson et al. (2007) and McGrath (2012), perhaps owing to a longer follow-up time and varying recidivism inclusion criteria.

For interrater reliability, ratings for total scores were highly congruent across measures and significant at the p < .001 level. The same trend emerged for factor/domain scores, though significant varied between $p < .01$ and $p < .001$ depending on the factor/domain. However, the interrater reliability of the SOTIPS Criminality factor was low ($ICC = .39$) and non-significant, perhaps partially explained by less rigorous rater training on this tool. Regarding scale reliability, dynamic measures evidenced

higher internal consistency than static measures, likely owing to greater heterogeneity of item content in static scales. Reliability of the SAPROF-SO was not available due to inconsistent coding of certain items.

4.2 Validity of Risk/Protective Measures

4.2.1 Convergent Validity

Static-99R, VRS-SO, STABLE-2007, SOTIPS, and SAPROF-SO, and their factor/domain scores were almost all significantly correlated with one another, except the SAPROF-SO Professionally Provided Support factor, owing perhaps to small cell numbers. Tools that assessed risk factors demonstrated mostly strong and significant positive correlations with other tools designed to assess risk factors. The SAPROF-SO had moderate to strong and significant negative correlations with risk assessment tools. Total score correlations and factor/domain score correlations tended to be in the strong range, though there were some factor/domain scores in the weaker range and these correlations were generally not significant.

The direction and strength of correlations for Static-99R, STABLE-2007, and the VRS-SO are consistent with past research (Sowden & Olver, 2017). However, the SAPROF-SO showed larger, significant correlations with Static-99R and the VRS-SO dynamic items in the present study than in past research that investigated both routine and high-risk individuals; previous research showed small, nonsignificant correlations between these measures except between Static-99R and the SAPROF-SO Stability domain, which evidenced a moderate and significant negative correlation (Willis et al., 2020). At the time of this writing, no research could be found that investigated associations between the SOTIPS and other risk or protective measures.

This study provides additional evidence that well-established risk tools are strongly correlated and speaks to the relationship between risk factors and protective factors. Moderate to strong correlations indicate that the tool and its factor scores are measuring similar underlying constructs; though the strong correlations between risk and protective measures do not clarify whether these scores are capturing opposite ends of the same construct or two separate underlying constructs. The strength and significance of the negative correlations between the SOTIPS and SAPROF-SO total scores and some of the factor/domain scores is particularly striking when considering this question. Given that the factors contained in the SAPROF-SO do not merely indicate an absence of risk (Kelley et al., 2022), it is possible that the measure may be capturing a different underlying construct that is related to both risk and recidivism. Closer examination of the factors/domains suggests that each respective measure is assessing separate but overlapping constructs, a notion that is supported by previous research

(Coupland & Olver, 2020). Overall, the convergent correlations indicate that the measures encompass related constructs but were not so large as to suggest redundancy. However, as previously suggested, strong correlations do not simply imply redundancy; using more than one risk assessment tool may add information that is valuable for intervention and case management (Wong & Gordon, 2006) and may also incrementally predict recidivism (Babchishin et al., 2012b).

4.2.2 Predictive Validity

4.2.2.1 Discrimination. AUC values for the forensic measures and domain/factor scores generally fell in the small to moderate range, though the magnitude of effect was generally larger when the five-year fixed follow-up times were employed.

The predictive validity of the VRS-SO scores followed parallel patterns of significance regardless of whether a fixed or non-fixed follow-up time was used. More specifically, for both fixed and non-fixed follow-up, VRS-SO static scores were significantly predictive of violent and any recidivism, dynamic scores were significantly predictive of all recidivism outcomes, sexual deviance scores significantly predicted sexual recidivism, criminality scores significantly predicted violent and any recidivism, treatment responsivity scores were not significantly predictive of any outcome, and total scores were significantly predictive of all recidivism outcomes. Significant AUC values ranged from 0.61 to 0.72 when no fixed follow-up time was employed, and 0.60 to 0.76 when 5-year follow-up time was employed.

VRS-SO total, static, and dynamic score AUC values in the current study (AUCs = .62 to .72) with 5-year follow-up) are slightly lower than those found by Olver et al. (2018; AUCs = .68 to .73), and not all the AUC values in the current study were significant, though they were in Olver et al. (2018). VRS-SO factor score AUCs in the current study (AUCs = .49 to .76) were very similar to Olver et al. (2020; AUCs = .49 to .75), though the Sexual Deviance and Treatment Responsivity factors had markedly stronger and weaker predictive validity respectively in the current study. There are several potential explanations for the weak performance of the Treatment Responsivity factor. First, pre-treatment ratings on this factor may evidence less substantial relationships with recidivism than posttreatment or change ratings, owing to the intervention-focused nature of the factor items. Second, two of the items contained in this factor (Release to High Risk Situations and Treatment Compliance) were less reliably coded due to lack of file information, which may have impacted the robustness of this factor's relationship with recidivism outcomes. Despite not being specifically designed to assess risk of general recidivism, the VRS-SO was also a moderate to strong predictor of any recidivism, with significant AUC static, dynamic, and total values of .65 to .68, which is again slightly lower than Olver

et al. (2018). The results of the current study support the use of VRS-SO with community-supervised men with sexual offense convictions.

Static-99R was moderately predictive for violent and any, but not sexual recidivism when no fixed follow-up was employed (AUC values of .65, respectively), and moderately predictive for all recidivism outcomes when the 5-year follow up was used (AUC values ranging from .66 to .69). The present results for sexual recidivism (AUC = .66) are lower than Helmus et al.'s (2012) meta-analysis (AUC = .69), and Lee and Hanson's (2017) meta-analysis (AUC = .77).

STABLE-2007 was a moderate predictor of all recidivism outcomes with 5-year fixed follow-up time (AUCs from .65 to .71). STABLE-2007's predictive validity was higher in the present study (AUC = .71) than in the Brankley et al. (2017) meta-analysis (AUC = .67).

SOTIPS scores followed similar patterns of significance regardless of follow-up method. Specifically, sexual deviance scores predicted sexual but not violent or any recidivism, criminality scores predicted violent and any recidivism for no fixed follow-up, but only any recidivism with fixed follow-up time. Social stability and support scores and total scores significantly predicted all recidivism outcomes, with small to medium effects for no fixed follow-up time, and medium to large effects for five-year fixed follow-up time. Significant AUC values ranged from .67 to .71 when five-year fixed follow-up times were employed. SOTIPS demonstrated stronger predictive validity in the current study than in McGrath et al., (2012; AUCs = .62 to .63) and Hanson et al. (2020; AUCs = .62 to .67), though these two studies utilized either a shorter fixed follow-up or no fixed follow-up time.

When no fixed follow-up time was employed, SAPROF-SO domain and total scores evidenced variable patterns of significance and effect size magnitude. When 5-year fixed follow-up time was employed, internal capacity significantly predicted any recidivism, professionally provided support significantly predicted sexual recidivism, and professional identity, prosocial connection, stability, and total scores significantly predicted all outcomes. Significant AUC values ranged from .60 to .70 when 5-year fixed follow-up times were employed. One unexpected finding was the SAPROF-SO Professionally Provided Support domain significantly predicting sexual recidivism, but not violent or general recidivism. This may be influenced in part by both the nature of the items contained in the factor and service provider's attentiveness to capturing sexual recidivism-specific information. Specifically, the items that comprise the factor (sexual-offense specific treatment, medication, perception of therapeutic/supervisory alliance, and supervised living) specifically target sexual-offense relevant factors (e.g., medication includes libido-reducing medication). It is therefore less likely this factor is capturing factors that would evidence a relationship with desistance from violent and general

recidivism. Relatedly, health-care providers may be more likely to document information related to sexual offense risk and attend less to factors related to desistance from violent and general recidivism, which would in turn impact the Professionally Provided Support factor's predictive validity for these outcomes. The SAPROF-SO total score demonstrated weaker predictive accuracy in the present study for all recidivism outcomes (AUC = .65 to .66) than in previous research (AUCs = .63 to .81; Nolan et al., 2022), though Nolan and colleagues had a longer follow-up time of over 12 years.

4.2.2.2 Calibration. Calibration, or the examination of the association of risk bands and recidivism, was examined for each risk measure. The five-level risk classification was used for Static-99R and the VRS-SO, while low, medium, high risk/need band classification was used for STABLE-2007 and SOTIPS. Of note, Level IVa and IVb were combined for the VRS-SO given the small number of individuals in each group. For all risk tools, the survival curves were consistent with the anticipated recidivism trajectories. However, differences in recidivism between each risk bands on each measure were not always significant.

Specifically, for Static-99R and sexual recidivism; Level IVB was significantly different from each other risk group except for Level I; the remaining risk groups did not recidivate at significantly different rates. However, for violent recidivism, none of the Static-99R risk band groups significantly differed from one another, and for any recidivism, there was a general pattern of higher and faster recidivism as risk level increased.

For the VRS-SO total score and sexual recidivism, the only significant difference was between Level I and III; for violent recidivism, significant differences were evidenced between Level I and II and I and III. For any recidivism, significant differences were observed between Level I and II, III, and IVa/IVb but none of the remaining groups. The current study had more significant differences between VRS-SO total score risk groups for sexual recidivism than in Sowden's (2013) dissertation, though that sample utilized risk groups with different cut offs. For violent recidivism, the current study also showed different pattern of results from Sowden (2013), where higher risk groups evidenced significant differences from one another, unlike this study, where lower risk groups evidenced significant differences. However, this is likely because Sowden's sample was higher risk (VRS-SO total mean score of 42.2 (SD = 7.4) than in the current sample (*M* = 20.9, *SD* = 11.0).

VRS-SO dynamic risk bands significantly differed between most groups for sexual recidivism, while for violent recidivism, only the two highest risk groups (Level III and Level IVa/IVb) significantly differed from Level I; for any recidivism, Level III did not significantly differ from the two lowest risk groups.

For STABLE-2007 and all recidivism outcomes, risk/need bands significantly differed between low and high and medium and high, but not between low and medium. More significant differences between STABLE-2007 risk/need bands were present in the current study than in previous research, which found significant differences only between the moderate and high risk group for violent recidivism (Sowden, 2013). For SOTIPS and sexual recidivism, the high and low risk groups were significantly different, and for violent and any recidivism, all risk/need bands significantly differed except between low and medium.

Several potential explanations exist for the lack of significant differences between risk groups. First, as mentioned above, the current sample was relatively low risk, with low recidivism base rates, making it more difficult to differentiate between members of risk groups. Thus, lack of difference between lower risk groups for certain types of recidivism in this study could be partially explained by the uneven distribution of individuals in risk groups. Few individuals fell in the high range and when they did, they often tended to be on the lower end of it. Additionally, given that this sample was lower risk, it could be that the lack of significant differences between lower risk groups indicates the presence of a single risk category for some measures, such as STABLE-2007 and SOTIPS. Finally, previous findings indicate that posttreatment risk levels are better recidivism indicators than pre-treatment risk levels (Lewis et al., 2013). Pre-treatment calculation of risk levels in the current study may contribute to lack of differentiation between some groups.

When Chi-square and Cramer's V were used to examine calibration, risk bands for each risk tool were significantly associated with fixed 5-year recidivism rates for all recidivism outcomes, with small to moderate associations between risk measures and recidivism outcomes. There was one exception: the SOTIPS risk bands were not significantly associated with 5-year fixed rates of sexual recidivism; this may indicate that two risk bands are more appropriate for the SOTIPS, but more likely is due to the small number of individuals in each risk band. Overall, the chi-square findings are promising in light of the lack of significant differences in recidivism rates between risk bands for some risk tools found in the pairwise comparisons utilized in the Kaplan-Meier survival analyses.

The E/O index was also utilized to examine calibration, using both logistic regression generated estimates and observed numbers. Several noteworthy themes emerged. First, for the logistic regression generated estimates, all E/O index values were below 1.0, indicating that the expected sexual recidivism rates from the VRS-SO normative sample were lower than observed in the present sample. Second, no E/O index values in either analysis were statistically significant, indicating a relatively good fit between estimated and observed sexual recidivism rates. Third, E/O values approached 1.0 for

higher risk groups, with Level IVb (well above average risk) having extremely close expected and observed sexual recidivism rates in both analyses, indicating a very small difference between the present and normative samples. One exception to this is the Level IVa group when actual observed numbers were used; normative data overestimated recidivism in this category for the current sample, though not significantly so. Though there appears to be acceptable calibration for the VRS-SO in the current sample, the scale underestimates recidivism for individuals in lower risk categories (albeit not to a degree of statistical significance) and tends to be better calibrated for higher risk groups, except for Level IVa when actual observed numbers were employed. It is worth noting that the E/O index can amplify the magnitude of differences when the observed number of individuals who reoffended is very small, making it appear that calibration is worse than it really is (Olver, Beggs Christofferson et al., 2021). As noted, this disparity may be addressed by using a variation on the E/O index: the E-O index (David Thornton, as cited in Olver, Beggs Christofferson et al., 2021). The E-O index may be useful in replication research on VRS-SO calibration if the number of observed individuals who recidivated is also low, elucidating the true magnitude of differences between the groups.

Though calibration of the VRS-SO has not previously been examined in a community-supervised sample, Olver and Eher (2020) recently examined this question in a sample of incarcerated individuals convicted of sexual offenses and found lower estimated rates of 5-year sexual reoffense associated with VRS-SO score bands compared with observed rates from the normative sample. Mirroring the results of the current study, the closest correspondence was observed for the highest risk band using logistic regression-generated estimates (E/O index = 1.01).

4.2.3 Incremental Validity

The current study also examined the incremental predictive validity of specific tool pairings, and at the incremental validity of VRS-SO factors controlling for static risk. Most models resulted in one tool or factor being uniquely predictive of outcome.

For sexual recidivism, the VRS-SO Sexual Deviance total and SOTIPS were each uniquely predictive when entered with Static-99R, and the VRS-SO Dynamic total was incrementally predictive when entered with the VRS-SO static total, but Static-99R was predictive when added with the VRS-SO dynamic total. Olver and Eher (2019) found Static-99R and VRS-SO dynamic items both uniquely predicted sexual recidivism, while VRS-SO static items uniquely predicted this outcome over VRS-SO dynamic items. Research previously found the VRS-SO dynamic total to be incrementally predictive over Static-99R for sexual recidivism, (Sowden & Olver, 2017). Prior studies also found STABLE-2007 to be incrementally predictive of this outcome controlling for Static-99R (Brankley et al., 2019;

Sowden & Olver, 2017), while in the current study, STABLE-2007 was not uniquely predictive over Static-99R. SOTIPS was uniquely predictive over Static-99R in this study, mirroring previous research findings (McGrath et al., 2012). This unique contribution perhaps owes in part to the fact that SOTIPS was specifically designed for community supervised individuals. Unlike the present study, Nolan and colleagues (2022) found that the SAPROF-SO total significantly predicted sexual recidivism after controlling for Static-99R scores and significantly improved prediction model. Some evidence of dynamic measures uniquely predicting sexual recidivism utilized posttreatment scores (Beggs & Grace, 2010), though in Olver et al. (2013), Static-99R was still the dominant predictor of sexual recidivism after controlling for VRS-SO pre- and posttreatment dynamic risk. Additionally, Olver and Eher (2019) found Static-99R and VRS-SO dynamic scores were incrementally predictive of sexual recidivism.

For violent (including sexual) recidivism, Static-99R was the uniquely predictive variable in each model. When only VRS-SO factors were included in the model, the VRS-SO Criminality total was dominant, aligning with previous research (Olver & Eher, 2019). Besides the VRS-SO Criminality total, these results deviate from previous studies that have demonstrated the incremental predictive validity of dynamic measures (Hanson et al., 2007; McGrath et al., 2012; Olver et al., 2007; Olver et al., 2014. Specifically, Olver et al. (2014) found both pre- and posttreatment VRS-SO dynamic scores significantly and uniquely predicted violent and general recidivism after controlling for pretreatment static risk using Static-99R, while McGrath et al. (2012) found SOTIPS added incrementally to the prediction of all recidivism outcomes after controlling for Static-99R scores. Additionally, Olver and Eher (2019) found VRS-SO dynamic items and both Static-99R and VRS-SO static items to incrementally predict violent recidivism.

For any recidivism, results were considerably more variable: STABLE-2007 and SOTIPS each uniquely predicted any recidivism when controlling for static risk, both Static-99R and the SAPROF-SO were incrementally predictive when entered together, and the SAPROF-SO uniquely predicted any recidivism controlling for static and dynamic risk. VRS-SO Dynamic scores were not uniquely predictive of any recidivism when entered with Static-99R, or when entered with Static-99R and SAPROF-SO. Additionally, the VRS-SO Criminality total was uniquely predictive of any recidivism when entered with the other two VRS-SO factors, this remained true when Static-99R was added to the model, with both VRS-SO Criminality total and Static-99R uniquely predicting this outcome. Again, as noted above, previous research has found dynamic measures to be uniquely predictive of any recidivism after controlling for static risk (Brankley et al., 2019; McGrath et al., 2012; Olver et al., 2013).

Though dynamic tools at times did not add uniquely to the prediction of recidivism outcomes, it does not mean that the tools are not valuable. While capturing risk level is crucial, assessing need and responsivity factors is also important for treatment planning, to capture change related to treatment or other credible change agents (i.e., increasing age, secure employment, age-appropriate intimate relationships), and case management. While many measure pairings examined in the present study did not demonstrate incremental validity, this may reflect the strong relationship between many of the measures, indicating they are measuring similar underlying constructs.

It is also possible that the incremental validity results were influenced by low power. A simple rule that guides adequate sample size for Cox regression analyses is a minimum of ten events per variable, particularly with low prevalence binary predictors such as recidivism outcomes (Ogundimu et al., 2016); this helps eliminate bias in regression coefficients and improves predictive accuracy. Therefore, increasing sample size may lead to more clarity as to the incremental predictive validity of the forensic measures in the present study.

4.2.4 Construct Validity

4.2.4.1 Factor Structure of the VRS-SO Dynamic Items. The present study examined the structural properties of the VRS-SO dynamic items in relation to previous findings, which had supported a three-factor model. The current research excluded item D15 (Treatment compliance) as this item was not reliably coded. A CFA of the dynamic items broadly replicated the factor structure found in previous research (Olver et al., 2007; Olver & Eher, 2019; Olver, Neumann et al., 2018), apart from item D11 (Release to high risk situations), which cross-loaded across two factors in the current research and Olver and Eher (2019), instead of three factors as in Olver et al. (2007). The present study did not replicate Beggs and Grace's (2010) finding of a fourth factor (Self-Management).

The structural properties of the VRS-SO dynamic items in the present sample appear to broadly align with latent dimensions underpinning sexual violence risk found in other dynamic risk measures (e.g., STABLE-2007 and SOTIPS). With STABLE-2007, three factors were revealed: Antisociality, Sexual Deviance, and Hypersexuality; with Antisociality and Sexual Deviance being significant predictors of sexual recidivism. The SOTIPS generated three domains in the development sample: Sexual Deviance, Criminality/Antisociality, and Social Stability and Support (McGrath et al., 2012). More recently, Miner and colleagues (2022) revealed two factors: Sexual Risk and Antisocial Opposition. A similar pattern is seen when static measures are factor analysed. Brouillette-Alarie and colleagues (2015) examined factor structure of Static-99R and nonredundant items from Static-2002R and three domains emerged: Persistence/paraphilia, General Criminality, and Youthful/stranger

aggression. When both VRS-SO static and dynamic items were analysed, five factors emerged, with static and dynamic items largely loading on separate sets of factors: two static (Age, Sexual Criminality) and three dynamic (General Criminality, Sexual Deviance, and Treatment Responsivity; Olver et al., 2022).

While there is often a third factor that emerges, such as Treatment Responsivity with the VRS-SO, Hypersexuality with STABLE-2007, Social Stability and Support with SOTIPS, and Youthful/Stranger Aggression with the static measures, all of these studies agree on the presence of at least two factors: Sexual Criminality (i.e., sexual deviance) and General Criminality; these two factors constitute the broad domains most strongly associated with sexual recidivism (Doren, 2004; Hanson & Bussiére, 1998; Hanson & Morton-Bourgon, 2005).

The presence of these two factors across multiple dynamic sexual violence risk measures has various implications. First, they may serve to inform an individual's treatment recommendations. For example, an individual with high scores on the Sexual Deviance factor may benefit from skills training to manage sexual urges, whereas an individual with higher scores on the Criminality factor may benefit more from treatment that prioritizes substance abuse treatment, building prosocial supports, and skills to manage impulsivity and aggression. Second, and relatedly, varying levels of each risk domain according to offense type (i.e., victim profile) may yield different treatment recommendations, considering that individuals who score higher on the general criminality/antisociality factor generally have sexually offended against adults while individuals higher on sexual deviance are more likely to have offended against children (Eher et al., 2003; Harris et al., 2012; Rettenberger et al., 2010).

4.3 Protective Factors

Protective factors are those factors that are theoretically or empirically associated with reduced rates of sexual or violent recidivism in individuals with offense histories (de Vogel et al., 2012; Willis et al., 2017-2020). Some have asserted that protective factors merely represent the opposite pole of risk factors, can be subsumed within the Central Eight risk factors, and at best measure existing factors "better" than risk factors (Helmus, 2018). Others maintain that protective factors are theoretically distinct from risk factors and include not only factors that reflect the opposite pole of risk (e.g., self-control versus impulsivity), but factors independent of risk factors, such as goal-directed living, medication, and the therapeutic alliance (de Vogel et al., 2012; Willis et al., 2017-2020). Each side of the debate agrees that inclusion of protective factors may be valuable, even if they are not new constructs. Given that most risk tools, and more broadly, comprehensive risk assessments, tend to focus on deficits, including neutral and strength levels of these propensities may increase discrimination

accuracy of risk scales and encourage more balanced risk assessments that include a wholistic view of the individual. Most clinicians indicate they consider protective factors when completing assessments, even if a formal and validated measure is not used (Kelley et al., 2020). One aim of the present research was to examine the predictive validity of the SAPROF-SO with community-supervised men, and results provide preliminary support for including protective factor assessment in risk assessment, particularly based on its discriminative properties. A previous study suggests the SAPROF-SO has good predictive validity for sexual recidivism in a sample of men convicted of sexual offenses against children in New Zealand (Nolan et al., 2022), and recidivism outcome studies designed to provide information on predictive and incremental validity as well as provide base rate data are in progress (Kelley et al., 2022). Nolan (2021) notes that while it is premature to use SAPROF-SO to make predictions about sexual recidivism, it may be used in addition to validated risk tools. Bolstering the SAPROF-SO's research base investigating its use as an actuarial tool may strengthen the case for formal inclusion of protective factors in comprehensive risk assessments, service planning, and reintegration decisions.

There are several treatment and risk management implications related to inclusion of a formal protective tool such as SAPROF-SO in risk assessment. Regarding treatment implications, from the perspective of the therapeutic alliance, emphasizing strengths and gearing treatment goals towards positive outcomes could improve an individual's perception of fairness in the assessment process, set the stage for a stronger therapeutic relationship, and subsequently lead to greater treatment gains (Helmus, 2018). Fundamentally, a focus on increasing strengths is more motivating than decreasing deficits. It is intuitive that approach rather than avoidance interventions increases engagement, this remains true within sexual offense-specific treatment (Mann et al., 2004). While dynamic sexual violence risk instruments are concerned with problematic characteristics of an individual and encourage focus on removing or modifying deficits, the SAPROF-SO contains several items that are empirically and theoretically linked to desistance from reoffending, such as attention to the individual's current personal goals, leisure interests, and professional support and encourages treatment goals that are more engaging and less stigmatizing relative to those derived from dynamic sexual violence risk measures. The authors emphasize that a treatment focus on strengthening emotional connection to adults, sexual self-regulation, and self-control and adaptive schemas (all SAPROF-SO items) is likely more engaging than a focus on addressing intimacy deficits, sexual preoccupation, and cognitive distortions (as measured by dynamic risk assessment tools; e.g., Hanson et al., 2007; Wong et al., 2003-2017).

From a risk management perspective, the SAPROF-SO may be useful when formulating reintegration decisions. These decisions hold substantial weight, especially when working with individuals who score high on dynamic risk measures despite nearing warrant expiry or completing treatment while incarcerated. Kelley and colleagues (2022) note that for some individuals who have met treatment goals in secure settings, the only way to continue making further treatment progress is to generalize these learned skills to a community setting. To encourage successful reintegration, they suggest a balance must be struck between managing risk and allowing individuals opportunities to practice skills in more independent settings.

The SAPROF-SO may also have relevance when considering reintegration planning for individuals with responsivity concerns, such as intellectual disability, major mental illness, and traumatic brain injury. 41.5% of the individuals in the present sample met diagnostic criteria for a major mental illness at the time of assessment, and 28.2% had cognitive impairments or intellectual disabilities. Given that calibration of actuarial risk tools can be less accurate for these individuals (Hanson et al., 2013; Kelley et al., 2020), including a protective tool could provide clarity on risk management and treatment considerations as these individuals may also have slower treatment gains and require more professional support to achieve success in the community. Given this heavier reliance on environmental supports, Kelley and colleagues (2022) propose framing risk assessment as a question of whether protective factors within a proposed release setting combined with treatment progress will be sufficient to manage the individual's risk for sexual recidivism, rather than solely focusing on an individual's level of risk, which often heavily prioritizes internal risk factors.

A final implication of the present research concerns the lack of empirically tested guidance on how to systematically combine SAPROF-SO with actuarial risk measures, analogous to the work that has been done on how to combine static and dynamic actuarial measures to generate overall risk ratings (Brankley et al., 2017; Hogan & Sribney, 2019). Kelley and colleagues (2022) have produced a table that combines an individual's risk rating (according to static and dynamic risk as well as treatment change) with their level of protection (based on SAPROF-SO scores) to ascertain the required level of community supervision or case management. However, this is an empirically theorized table that has not been empirically tested.

Combining risk and protective ratings, as well as treatment change into comprehensive risk assessment requires employment of rigorous statistical methods to ensure integration of risk and protective factors is conducted in a systematic and nonarbitrary fashion. Past research has utilized

survival analytic techniques and logistic regression-based modelling in service of this goal (Coupland & Olver, 2020; Lovatt et al., 2022; Olver & Reimer, 2021).

Cox regression survival analysis-based modelling is one way to achieve this integration. Coupland and Olver (2020) examined the association between positive changes on SAPROF to possible reductions in recidivism by entering the pretreatment risk measure score followed by the risk measure change score in the first block, followed by the pretreatment SAPROF score SAPROF change score being entered in the second block. Olver and Reimer (2021) employed a similar method using Cox regression-based modelling to examine the incremental prediction of recidivism by PCL-R and SAPROF baseline and change scores.

One potentially more intuitive way to demonstrate the intersection of risk and protective factors is to use Kaplan-Meier survival analyses to initially examine trajectories of recidivism over time for each risk and protection subgroup based on their respective risk and protective measure scores (Lovatt et al., 2022). The potential risk mitigating properties of protective factors could then be demonstrated by comparing trajectories of recidivism for individuals who score high versus low on protection within a given risk group (i.e., high versus low sexual violence risk). Lovatt and colleagues (2022) then conducted logistic regression to use multivariate risk and protective measure models to generate recidivism rates that were associated with particular risk and protective factor score pairings. Olver and Reimer (2021) used logistic regression modelling to the same end, demonstrating that the level of risk posed by an individual who scored higher on the PCL-R and had several protective factors was not the same as an individual with similar scores on the PCL-R without those protective factors.

4.4 Strengths, Limitations, and Future Directions

The retrospective nature of this study was both an asset and a hindrance. While archival data allowed for a longer follow-up time (over eight years), data collection was limited to file information. This meant that certain variables (particularly protective factor variables from the SAPROF-SO) were consistently difficult to obtain; this impacted the integrity of some item ratings. Difficulty coding protective measures is well-documented and is likely influenced by the deficits-focused nature of files (de Vries Robbé, 2011; Willis et al., 2020). Still, relatively comprehensive pre-treatment information was available to code the risk measures and other variables, and strong inter-rater reliability lends to the credence of the present findings.

The sample of 200 individuals (172 with outcome information) was moderate in size, and limited statistical power reduced what was possible regarding categorical analyses, specifically regarding incremental validity of forensic measures. The smaller *n* in some groups may have

implications for the stability of findings in certain risk groups (e.g., there was only one observed individual who recidivated in the VRS-SO Level IVb risk group in E/O analyses). Limited statistical power also restricted investigation of differences in the forensic tool's predictive validity according to ethnocultural group. Previous research has found that the discrimination and calibration properties of some risk factors and risk tools varies according to ethnocultural group (Hanson et al., 2007; Lee & Hanson, 2017; Perley-Robertson et al., 2019), though a growing contingent of research also finds acceptable discrimination and calibration properties of sexual violence risk tools for sexual and general violence among ethnocultural minority groups (Babchishin et al., 2012a; Boccaccini et al., 2017; Lee et al., 2018; Olver, Sowden et al., 2018). In a Canadian context, this is particularly relevant to First Nations, Inuit, and Métis justice-involved individuals, who are overrepresented in the justice system. This makes the question of cultural bias and relatedly, psychometric equivalence in sexual violence risk tools between ethnocultural groups a critical concern. Specifically, do these risk tools predict recidivism with appropriately similar degrees of accuracy across ethnocultural groups?

Olver and Stockdale (2021) raise several important points regarding psychometric equivalence for ethnocultural minority groups. Specifically, they note there is currently no consensus on an acceptable scientific or clinical threshold for "good enough" discrimination or calibration and conversely, what level of disparity is cause for harm when utilizing sexual violence risk tools with ethnocultural minority groups. This lack of consensus has important implications given that it is inappropriate to simply stop using sexual violence risk measures with specific ethnocultural minority if there is not empirical support for their use with a particular group. Relatedly, the article discusses many factors that influence discrimination and calibration properties of the measures when used with ethnocultural minorities apart from the measures themselves, including systemic factors that may lead to increased recidivism (e.g., overpolicing, submission of a guilty plea) and examiner bias (e.g., conducting risk assessment in a culturally unresponsive manner and/or misapplying risk scores from one or more tools; Olver & Stockdale, 2021). Though it is beyond the scope of this program of research to fully explore concerns and proposed solutions, future research would do well to continue examining the psychometric properties of these risk and protective tools with ethnocultural minority groups and to engage in consistent consultation as part of the empirical process to ensure diverse voices and alternative ways of knowing are incorporated into best practices regarding comprehensive risk assessment. Because the question of psychometric equivalence with ethnocultural minority groups could not be addressed in the current sample; further research involving larger sample sizes would provide an opportunity to focus on this query.

Another limitation pertains to the present sample being relatively low risk, with mean VRS-SO scores more than a full standard deviation below the normative sample. While this may limit generalizability of findings, the risk level of this sample was expected, as most individuals who are managed in the community are likely to receive lower risk ratings than those who are incarcerated. Relatedly, while not a limitation in a practical sense, recidivism base rates were low, and this limited the power of analyses. Outcome data was also limited to federal criminal records; inclusion of provincial outcome data or even police records may provide a more comprehensive picture of recidivism for this sample, particularly as the lag in updating federal recidivism databases may lead to an underestimation of true recidivism rates.

A significant strength of this study is that it provided initial validations of several forensic measures. This study represents the first validation of the VRS-SO in a non-incarcerated sample, finding support for its predictive validity, and opening several potential avenues for future study. Specifically, recent research has found acceptable calibration for the VRS-SO for sexual recidivism projections with men over age 60, with E/O index demonstrating slight overprediction compared to actual observed rates for these older men (Olver, Beggs Christofferson et al., 2021). Investigation of the calibration properties of VRS-SO as a function of age in a community-supervised sample may further elucidate whether an age correction to account for age-related changes in risk that are not fully captured by changes on dynamic factors or on static age-item ratings is needed. (Olver, Beggs Christofferson et al., 2020). Another recent study (Olver et al., 2022) sought to develop empirical findings with dynamic risk factors contained in the VRS-SO that might guide further theory development, given that the empirical construct of dynamic risk is often considered largely atheoretical (Thornton, 2016; Ward & Beech, 2015). The study found three broad dimensions that run through pretreatment Long Term Vulnerabilities (LTVs), that similar factors occur in static and pretreatment dynamic items, and the structure of change in LTVs is different from the structure of LTVs. The authors suggest future theoretical development should seek to replicate these findings and explore the two change factors; future research may do well to explore these findings in a non-incarcerated sample.

Though future research with non-incarcerated populations should endeavour to measure change as a function of sexual offense-specific treatment and the relationship of these changes to recidivism using each dynamic risk tool included in this study, the VRS-SO may be a high priority for such community-based investigations given its unique structured mechanism to assess treatment-related change. Such research would add to a sparse but growing literature on treatment-related change and linkages between such change and recidivism. This highlights a major limitation of the present

research: This dissertation focuses heavily on dynamic measures but then proceeds to measure purportedly dynamic factors in a static fashion (i.e., assessing them at a single timepoint). An early goal of this study was to include posttreatment and consequently change data; this was not possible due to the paucity of treatment information contained on file. Previous research has linked positive changes in risk and protective factors to decreases in sexual and violent recidivism (de Vries Robbé et al., 2015; Olver et al., 2022), and changes on each dynamic sexual violence risk instrument included in the present study has shown significant associations with decreased sexual recidivism at follow-up (Olver, Mundt et al., 2018; Sowden & Olver, 2018; van den Berg et al., 2018).

In their 2020 review of the sexual offense treatment change literature, Olver and Stockdale highlight that integrating routine assessment of change into risk management interventions is not only a way to reduce sexual violence but represents an ethical responsibility for assessors and treatment providers; this stance is corroborated by guidelines put forth by the recent Canadian Psychological Association task force report on outcome and progress monitoring in psychotherapy (Tasca et al., 2018) Difficulty in obtaining risk ratings at multiple timepoints in the present study leads to consideration of how this data can be gathered in routine assessments. In addition to rating risk tools pre- and posttreatment, Olver and Stockdale (2020) suggest re-rating a tool at each time point (e.g., each probation meeting or therapy appointment) and using the most recent score as most indicative of current recidivism risk. They note this may be particularly useful in community supervision when risk changes from week to week may be small, or over longer time intervals when an obvious change agent may not be present. This stance aligns with research that the most recent score on a dynamic risk measure is more predictive of recidivism than earlier scores (Babchishin & Hanson, 2020; Olver, Kelley et al., 2021), and recommendations from the developers of the SAPROF-SO that the measure should be re-administered at least annually or when an environmental shift occurs (e.g., change of residence; Kelley et al., 2022). This recommendation regarding the SAPROF-SO leads to a final noted area for future research: just as evaluating changes in risk is important, so too is evaluating change is treatment change is relevant changes in protective factors and the relationship of such change with decreased recidivism.

To the author's knowledge, this is also the first Canadian validation of the SOTIPS, and given its performance, future research would do well to replicate predictive validity findings and investigate its ability to measure treatment-related change in additional community-supervised samples, given that a main goal of the tool is to monitor treatment progress for non-incarcerated individuals with sexual offense convictions.

As a final novel contribution, this study is the first Canadian validation of the SAPROF-SO in a community-supervised sample and one of the only studies (to the author's knowledge) to investigate how well the SAPROF-SO predicts recidivism outcomes. This measure appears to have promising psychometric properties linking it to assessment of risk and additional research could provide a unique and balancing role in the current approach to risk assessment, as well as standardizing the way protective factors for desistance from sexual offending are measured and presented in risk assessment. Future research should explore the relationship of specific protective factors with risk, and when treatment data is available, provide information about the benefits of focusing treatment on increasing the presence of protective factors to mitigate recidivism risk. Future investigations could also focus on determining the best methods for combining scores and producing normative data (Nolan et al., 2022), as recent research has examined how scores should be divided into protection categories and provided preliminary suggestions on how to integrate risk tool and SAPROF-SO ratings to determine an individual's required level of community supervision and/or case management (Kelley et al., 2022). Lastly, since commencement of this dissertation, findings from pilot research have informed the SAPROF-SO version 1 (Willis et al., 2021) which consists of 14 items in three domains: Resilience, Adaptive Sexuality, and Prosocial Connection and Reward, with six optional items captured in the Professional Risk Management domain. Future studies may wish to investigate the SAPROF-SO's structural properties in diverse samples (e.g., incarcerated or community-supervised, other jurisdictions, specific ethnocultural groups) to provide more information on latent dimensions underpinning desistance from sexual violence.

4.5 Conclusion

Assessing sexual violence risk and protective factors are important lines of research that have significant clinical and public safety implications. The present study provides evidence for the psychometric promise of the VRS-SO, SOTIPS, and SAPROF-SO in a community-supervised Canadian sample of men with sexual offense conviction histories and adds to the extant literature supporting the use of Static-99R and STABLE-2007 with this population. This study also provides preliminary support for the inclusion of protective measures in sexual violence risk assessment, indicating that protective factors may provide a more comprehensive understanding of the individual being assessed and that the SAPROF-SO demonstrated predictive validity for recidivism – or in alignment with the spirit of its creation, non-recidivism – when used alongside validated risk measures.

References

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

Alexander, M. A. (1999). Sexual offender treatment efficacy revisited. *Sexual Abuse: A Journal of Research and Treatment, 11*(2), 101-116. https://doi.org/10.1007/BF02658841

Allan, M., Grace, R.C., Rutherford, B., & Hudson, S.M. (2007). Psychometric assessment of dynamic risk factors for child molesters. *Sexual Abuse: A Journal of Research and Treatment, 19*(4), 347-367.http://doi.org/10.1007/s11195-007-9052-5

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

Andrews, D.A., Bonta, J., & Wormith, J.S. (2004). The Level of Service/Case Management Inventory. Multi-Health Systems. http://doi.org/10.1037/t05029-000

Armstrong, D., & Murdoch, K. (2015). Sex Offender Relapse Prevention Program Evaluation Plan Forensic Assessment and Community Services, Alberta Health Services

Andrews, D.A., & Bonta, J. (2006). *The psychology of criminal conduct* (4th edition). Newark, NJ: LexisNexis

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

Babchishin, K.M., Hanson, R.K., & Helmus, L.M. (2012b) Even highly correlated measures can add incrementally to predicting recidivism among sex offenders. *Assessment, 19*(4), 442-461. http://doi.org.10.1177/1073191112458312

Babchishin, K.M., & Hanson, R.K. (2020). Monitoring changes in risk of reoffending: A prospective study of 632 men on community supervision. *Journal of Consulting and Clinical Psychology, 88*(10), 886-898. http://doi.org/10.1037/ccp0000601

Barbaree, H.E., Langton, C.M., & Peacock, E.J. (2006). Different actuarial risk measures produce different risk rankings for sexual offenders. *Sexual Abuse*, *18*(4), *423*-40. http://doi.org/10.1177/107906320601800408

Baxstrom v. Herold. (1966). 383 U.S. 107 (86S. Ct. 760).

Beggs, S. (2010). Within-treatment outcome among sexual offenders: A review. *Aggression and Violent Behavior, 15,* 369-379.

Boccaccini, A.T., Rice, M.T., Helmus, L.M., Murrie, D.C., & Harris, P.B. (2017). Field validity of Static-99R scores in a statewide sample of 34,687 convicted sexual offenders. *Psychological Assessment, 29,* 611-623. http://doi.org.10.1037/pas0000377

Bonta, J. (2002). Offender risk assessment: Guidelines for selection and use. *Criminal Justice and Behavior, 29*(4), 355-379. http://doi.org/10.1177/0093854802029004002

Bonta, J., & Andrews, D.A. (2003). A commentary on Ward and Stewart's model of human needs. *Psychology, Crime & Law, 9,* 215-218.

Bouman, Y.H.A., de Ruiter, C., & Schene, A.H. (2010). Social ties and short-term self-reported delinquent behavior of personality disordered forensic outpatients. *Legal and Criminological Psychology, 15*(2), 357-371. http://doi.org/10.1348/135532509X444528

Bourgon, G., Mugford, R., Hanson, R.K., & Coligado, M. (2018). Offender risk assessment practices vary across Canada. *Canadian Journal of Criminology and Criminal Justice, 60*(2), 167-205. http://doi.org/10.3138/cjccj.2016-0024

Brown, T.A. (2015). *Confirmatory factor analysis for applied research, 2nd ed.* The Guilford Press.

Brankley, A.E., Babchishin, K.M., & Hanson, R.K. (2019). STABLE-2007 demonstrates predictive and incremental validity in assessing risk-relevant propensities for sexual offending: A meta-analysis. *Sexual Abuse* 1-29. http://doi.org/10.1177/1079063219871572

Brankley, A.E., Helmus, L.M., & Hanson, R.K. (2017). *STABLE-2007 evaluator workbook: Updated recidivism rates (includes combinations with Static-99R, Static-2002R, and Risk Matrix 2000).* Unpublished report, Public Safety Canada.

Briken, P., & Muller, J.L. (2014). Beurteilung der schuldfahigkeit bei paraphiler sorung; kann der schweregrad der storung mithilfe von kriterien aus prognoseinstrumenten erfasst warden? [Assessment of criminal responsibility in paraphilic disorder: Can the severity of the disorder be assessed with items of standardized prognostic instruments?]. *Der Nervenarzt, 85*(3), 304-311. http://doi.org/10.1007/s00115-013-3901-x

Brouillette-Alarie, S., & Proulx, J. (2013). Predictive validity of the Static-99R and its dimensions. *Journal of Sexual Aggression, 19*(3)*,* 311-328.

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 Behaviour, 36*(6), 567-590. http://doi.org/10.1177/0093854809333610

Cicchetti, D.V., & Sparrow, S.A. (1981). Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behaviour. *American Journal of Mental Deficiaency, 86*(1), 127-137.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum.

Conroy, S., & Cotter, A. (2017). Self-reported sexual assault in Canada, 2014. *Juristat, 42.* Statistics Canada Catalogue no. 85-002-X. Ottawa, ON: Statistics Canada.

Costello, A.B., & Osborne, J.W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most out of your analysis. *Practical Assessment, Research & Evaluation, 10*(7), 1-19. http://scholarworks.umass.edu.pare/vol10/iss1/7

Council of State Governments Justice Center. (2014). *A common language for risk assessment: Experts convene in Washington*. http://csgjusticecenter.org/a-common-language-for-riskassessments-experts-convene-in-washington/

Coupland, R.B.A., & Olver, M.E. (2020). Assessing protective factors in treated violent offenders: Associations with recidivism reduction and positive community outcomes. *Psychological Assessment, 32*(5), 493-508. http://doi.org/10.1037/pas0000807

Craig, L.A., Thornton, D., Beech, A., & Browne, K.D. (2007). The relationship of statistical and psychological risk markers to sexual reconviction in child molesters. *Criminal Justice and Behavior, 34*(3), 314-329. http://doi.org/10.1177/0093854806291416

de Vogel, V., de Ruiter, C., Bouman, Y., & de Vries, Robbé, M. (2012). *SAPROF. Guidelines for the assessment of protective factors for violence risk. 2nd Edition.* Van de Hoeven Stichting.

de Vogel, V., de Ruiter, C., van Beek, D., & Mead, G. (2004). Predictive validity of the SVR-20 and Static-99 in a Dutch sample of treated sex offenders. *Law and Human Behavior, 28*(3), 235-251. http://doi.org/10.1023/b:lahu.0000029137.41974.eb

de Vries Robbé, M., de Vogel, V., & de Spa, E. (2011). Protective factors for violence risk in forensic patients: A retrospective validation study of the SAPROF, *The International Journal of Forensic Mental Health, 10*, 178-186.

de Vries Robbé, M., de Vogel, V., Douglas, K.S., & Nijman, H.L.I. (2015). Changes in dynamic risk and protective factors for violence during inpatient forensic psychiatric treatment: predicting reductions in postdischarge community recidivism. *Law and Human Behavior, 39*(1), 53-61. http://doi:10.1037/lhb0000089

de Vries Robbé, M., de Vogel, V., Koster, K., & Bogaerts, S. (2015). Assessing protective factors for
    sexually violent offending with the SAPROF. *Sexual Abuse: A Journal of Research and*
    Treatment, 27, 51-70.

de Vries Robbé, M. & Olver, M.E. (2021, September). *Research and practice developments in the*
    assessment and enhancement of protective factors: Toward reducing risk and strengthening
    community safety and wellbeing. Plenary conducted at the 40th Annual ATSA Research and
    Treatment Conference, virtual

de Vries Robbé, M. & Willis, G.M. (2017). Assessment of protective factors in clinical practice.
    Aggression and Violent Behavior, 32, 55-63.

Douglas K.S., & Skeem, J.L. (2005). Violence risk assessment: Getting specific about being dynamic.
    Psychology, Public Policy, and Law, 11(3), 347-383. http://doi.org/10.1037/1076-
    8971.11.3.347

Dube, S.R., Anda, R.F., Whitfield, C.L., Brown, D.W., Felitti, V.J., Dong, M., & Giles, W.H. (2005).
    Long-term consequences of childhood sexual abuse by gender of victim. *American Journal of*
    Preventative Medicine, 28(5), 430-438. http://doi.org/10.1016/jamepre.2005.01.015

Eher, R., Neuwirth, W., Fruehwald, S., & Frottier, P. (2003). Sexualization and lifestyle impulsivity:
    Clinically valid discriminators in sexual offenders. *International Journal of Offender Therapy*
    and Comparative Criminology, 47(4), 452-467. http://doi.org/10.1177/0306624X03253844

Eher, R., Olver, M.E., Heurix, I., Schilling, F., & Rettenberger, M. (2015). Predicting reoffense in
    pedophilic child molesters by clinical diagnoses and risk assessment. *Law and Human*
    Behavior, 39(6), 571-580. http://doi.org/10.1037/lhb0000144

Etzler, S., Eher, R., & Rettenberger, M. (2020). Dynamic risk assessment of sexual offenders: Validity
    and dimensional structure of the STABLE-2007. *Assessment, 27*(4), 822-839.
    http://doi.org/10.1177/1073191118754705

Falshaw, L., Bates, A., Patel, V., Corbett, C., & Friendship, C. (2003). Assessing reconviction,
    reoffending, and recidivism in a sample of UK sexual offenders. *Legal and Criminological*
    Psychology, 8(2), 207-215. https://doi.org/10.1348/135532503322362979

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). SAGE Publications Ltd.

Finkelhor, D. (1984). *Child sexual abuse: New theory and research.* New York, NY: Free Press.

Gallagher, C.A., Wilson, D.B., Hirschfield, P., Coggeshall, M.B., & Mackenzie, D.L. (1999). A
    quantitative review of the effects of sex offender treatment on sexual reoffending. *Corrections*
    Management Quarterly, 3(4), 19-29.

Gannon, T.A., Olver, M.E., Mallion, J.S., & James, M. (2019). Does specialized psychological treatment for offending reduce recidivism? A meta-analysis examining staff and program variables as predictors of treatment effectiveness. *Clinical Psychology Review, 73,* 1-18. https://doi.org/10.1016/j.cpr.2019.101752

Gottfredson, M.R., & Hirschi, T. (1990). *A general theory of crime.* Palo Alto, CA: Stanford University Press.

Hall, G.C.N. (1995). Sexual offender recidivism revisited: A meta-analysis of recent treatment studies. *Journal of Consulting and Clinical Psychology, 63*(5), 802-809. https://doi.org/10.1037/0022-006x.5.802

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. https://doi.org/10.1177/0093854816683956

Hanson, R.K. (2018). Long-term recidivism studies show that desistance is the norm. *Criminal Justice and Behavior, 45*(9), 1340-1346. http://doi.org/10.1177/0093854818793382

Hanson, R. K., & Bussière, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology*, *66*(2), 348-362.

Hanson, R.K., Bourgon, G., Helmus, L.M., & Hodgson, S. (2009). The principles of effective correctional treatment also apply to sexual offenders: A meta-analysis. *Criminal Justice and Behavior, 36*(9), 865-891. http://doi.org/10.1177/0093854809338545

Hanson, R.K., Gordon, A., Harris, A.J.R., Marques, J.K., Murphy, W., Quinsey, V.L., & Seto, M.C. (2002). First report of the collaborative outcome data project on the effectiveness of psychological treatment for sex offenders. *Sexual Abuse: A Journal of Research and Treatment, 14*(2), 169-194. http://doi.org/10.1177/107906320201400207

Hanson, R.K., Harris, A.J.R., Scott, T., & Helmus, L. (2007). *Assessing the risk of sexual offenders on community supervision: The Dynamic Supervision Project* (User Report no.2007-05). http://vawnet.org/material/assessing-risk-sexual-offenders-community-supervision-dynamic-supervision-project

Hanson, R.K., Helmus, L.M., & Harris, A.J.R. (2015). Assessing thee 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. http://doi.org/10.1177/0093854815602094

Hanson, R.K., Helmus, L. & Hodgson, S. (2009). The principles of effective correctional treatment can also apply to sexual offenders: a meta-analysis. *Criminal Justice and Behavior, 36*(9)*,* 865-891.

Hanson, R.K., Lunetta, A., Phenix, A., Needley, J., & Epperson, D. (2014). The field validity of Static-99/R sex offender risk assessment tool in California. *Journal of Threat Assessment and Management, 1*(2), 102-117. http://doi.org/10.1037/tam0000014

Hanson, R.K., & Morton-Bourgon, K.E. (2005). The characteristics of persistent sexual offenders: A meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology, 73*(6)*, 1154-1163. http://doi.org/10.1037/0022-006X.73.6.1154

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

Hanson, R.K., Newstrom, N., Brouilette-Alarie, S., Thornton, D., Robinson, B.E., & Miner, M.H. (2020). Does reassessment improve prediction? A prospective study of the Sexual Offender Treatment Intervention and Progress Scale (SOTIPS). *International Journal of Offender Therapy and Comparative Criminology,* 1-29. http://doi.org/10.1177/0306624X20978204

Hanson. R.K., & Sheahan, C.L., & VanZuylen, H. (2013). Static-99 and RRASOR predict recidivism among developmentally delayed sexual offenders: A cumulative meta-analysis. *Sex Offender Treatment, 8*(1), 1-14.

Hanson, R.K. & Thornton, D. (2000). Improving risk assessment for sexual offenders: A comparison of three actuarial scales. *Law and Human Behavior, 24*(1), 119-136. http://doi.org/10.1023/A:1005482921333

Hanson, R.K., & Thornton, D. (2003). *Notes on the development of the Static-2002* (User report 2003-01). Department of the Solicitor General of Canada. http://www.publicsafety.gc.ca/cnt/rscs/pblctns/nts-dvlpmnt-sttc/index-en.aspx

Harris, A.J.R., & Hanson, R.K. (2004). *Sex offender recidivism: A simple question* (Corrections User Report No. 2004-03). Public Safety and Emergency Preparedness Canada. https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/sx-ffndr-rcdvsm/sx-ffndr-rcdvsm-eng.pdf

Harris, G.T., Lalumiére, M.L., Seto, M.C., Rice, M.E., & Chaplin, T.C. (2012) Explaining the erectile responses of rapists to rape stories: The contributions of sexual activity, non-consent, and violence with injury. *Archives of Sexual Behavior, 41*(1), 221-229. http://doi.org/10.1007/s10508-012-9940-8

Harris, G.T., & Rice, M.E. (2015). Progress in violence risk assessment and communication: Hypothesis versus evidence. *Behavioral Sciences and the Law, 33*(1), 128-145. http://doi.org/10.1002/bsl.2157

Harris, A. & Socia (2016). What's in a name? Evaluating the effects of the "sex offender" label on public opinions and beliefs. *Sexual Abuse: A Journal of Research and Treatment, 28*(7), 660-678. http://doi.org/10.1177.1079063214564391

Helmus, L.M. (2018). Sex offender risk assessment: Where are we and where are we going? *Current Psychiatry Reports, 20*(6), 46. http://doi.org/10.1007/s11920-018-0909-8

Helmus, L.M., 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. http://doi:10.1177/1079063211409951

Hilton, N. Z., Harris, G. T., & Rice, M. E. (2006). Sixty-six years of research on the clinical versus actuarial prediction of violence. *The Counseling Psychologist*, *34*(3), 400–409. http://doi.org/10.1177/0011000005285877

Judges, R.C. (2016). *An exploration into the value of protective factors in violence risk assessment of psychiatric inpatients.* (Doctoral Thesis) University of Nottingham http://eprints.nottingham.ac.uk/32959/

Jung, S., Pham, A., & Ennis, L. (2013). Measuring the disparity of categorical risk among various sex offender risk assessment measures. *Journal of Forensic Psychiatry & Psychology, 24*(3), 353-370. https://doi.org/10/1080/14789949.2013.806567

Kahn, R.E., Ambroziak, G., Hanson, R.K., & Thornton, D. (2017). Release from the "sex offender" label. *Archives of Sexual Behavior, 46*(4), 861-864. http://doi.org/10.1007/x10508-017-0972-y

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. http://doi.org/10.1177/1079063218800474

Kelley, S.M., Willis, G.M., & Thornton, D. (2022). Clinical applications of the structured assessment of protective factors against sexual offending (SAPROF-SO), Version 1. *Journal of Sexual Aggression* (Advance Online Publication). http://doi.org/10.1080/13552600.2021.2020345

Knight, R.A. (1999). Validation of a typology for rapists. *Journal of Interpersonal Violence, 14*(3), 297-323. http://doi.org/10.1177/088626099014003006

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*. 337-343. http://doi.org/10.1001/archpsych.199701830160065009

Kroner, D.G., & Derrick, B. (2020). The Council of State Governments Justice Center approach to increasing risk-level consistency in the application of risk assessment instruments. *Assessment,* 1-12. https://doi.10.1177/1073191120958066

Lasher, M.P., & McGrath, R. (2017). Desistance from sexual and other violence offending among child sexual abusers: Observations using the Sex Offender Treatment Intervention and Progress Scale. *Criminal Justice and Behavior, 44*(3), 416-431. http://doi.org/10.1177/0093854816670194

Lasher, M.P., McGrath, R.J., Wilson, D., & Cummings, G.F. (2015). Collaborative treatment planning using the Sex Offender Treatment Intervention and Progress Scale (SOTIPS): Concordance of therapist evaluation and client self-evaluation. *International Journal of Forensic Mental Health, 14*(1), 1-9. http://doi.org/10.1080/14999013/2014.974087

Heffernan, R., & Ward, T. (2019). Dynamic risk factors, protective factors, and value-laden practices. *Psychiatry, Psychology, and Law, 26*(2), 312-328. http://doi.org/10.1080/13218719.2018.1506721

Lee, S.C., & Hanson, R.K., (2017). Similar predictive accuracy of the Static-99R risk tool for White, Black, and Hispanic sex offenders in California. *Criminal Justice and Behavior, 44*(9), 1125-1140. http://doi.org/10.1177/0093854817711477

Lee, S.C., Hanson, R.K., & Zabarauckas, C.L. (2018). Sex offenders of East Asian heritage resemble other Canadian sex offenders. *Asian Criminology, 13*(1), 1-15. http://doi.org/10.1007/s11417-017-9251-y

Lodewijks, H.P.B., de Ruiter, C., & Doreleijers, T.A.H. (2010). The impact of protective factors in desistance from violent reoffending: A study in three samples of adolescent offenders. *Journal of Interpersonal Violence, 25*(3), http://doi.org/10.1177/088620509334403

Logie, C., Alaggia, R., & Rwigema, M.J. (2014). A social ecological approach to understanding correlates of lifetime sexual assault among sexual minority women in Toronto, Canada: Results from a cross-sectional internet-based survey. *Health Education Research, 29*(4), 671-182. http://doi.org/10.1093/her/cyt119

Lösel, F. & Schmucker, M. (2005). The effectiveness of treatment for sexual offenders: A comprehensive meta-analysis. *Journal of Experimental Criminology, 1*(1), 117-146. http://doi.org/10.1007/s11292-004-6466-7

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. http://doi.org/10.1037/PAS0001071

Luce, H., Schrager, S., & Gilchrist, V. (2010). Sexual assault of women. *American Family Physician, 81*(4)*,* 489-495.

Mann, R.E., Hanson, R.K., & Thornton, D. (2010). Assessing risk for sexual recidivism: Some proposals on the nature of psychologically meaningful risk factors. *Sexual Abuse: A Journal of Research and Treatment, 22*(2), 191-217. http://doi.org/10.1177/1079063210366039

Marsh, H.W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A.J.S., Trautwein, U., & Nagengest, B. (2010). A new look at the Big Five factor structure through exploratory structural equation modelling. *Psychological Assessment, 22*(3), 471-491. http://doi.org/10.1037/a0019227

Marshall, W.L., & Barbaree, H.E. (1988). The long-term evaluation of a behavioural treatment program for child molesters. *Behavior Research and Therapy, 26*(6), 499-511. http://doi.org/10.1016/0005-7967(88)90146-5

Marshall, W.L., Marshall, L.E., Serran, G.A., & O'Brien, M.D. (2011). *Rehabilitating sexual offenders: A strength-based approach.* American Psychological Association.

McGrath, R.J., Cumming, G.L., & Livingston, J.A. (2005). Predictive validity of the Sex Offender Treatment Needs and Progress Scale. Paper presented at the annual conference of the Association for the Treatment of Sexual Abusers, Salt Lake City, UT.

McGrath, R.J., & Hoke, S.E. (2001). *Vermont Assessment for Sex Offender Risk Manual.* Middlebury, VT: Author.

McGrath, R.J., Lasher, M.P., & Cummings, G.F. (2012). The Sex Offender Treatment Intervention and Progress Scale (SOTIPS): Psychometric properties and incremental predictive validity with the Static-99R. *Sexual Abuse: A Journal of Research and Treatment, 24*(5), 431-458. http://doi.org/10.1177/1079063211432475

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

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

Mills, J.F. (2017). Violence risk assessment: A brief review, current issues, and future directions. *Canadian Psychology, 58*(1)*,* 40-49. http://doi.org/10.1037/cap0000100

Monahan, J. (1981). *The prediction of violent behavior.* U.S. Department of Health and Human Services.

Mossman, D. (1994). Assessing predictions of violence: Being accurate about accuracy. *Journal of Consulting and Clinical Psychology, 62*(4), 783-792.

Muthén, B.O., & Muthén, L.K. (2013). *Mplus Version 7.11.* Author.

Myer, A.J. (2019). Examining the predictive validity of the Static-99R on Native American sex offenders. *Justice Evaluation Journal, 2*(2), 181-195. http://doi.org/10.1080/24751979.2019.1636614

Nolan, T. (2021). Attending to the positive: *Retrospective validation of the SAPROF-SO* [Masters thesis, University of Canterbury]. University of Canterbury Research Repository. https://ir.canterbury.ac.nz/handle/10092/101871

Nolan, R., Willis, G.M., Thornton, D., Kelley, S.M., & Beggs Christofferson, S. (2022). Attending to the positive: A retrospective validation of the Structured Assessment of Protective Factors – Sexual Offence Version. *Sexual Abuse, Online First,* http://doi/org/10.1177/10790632221098354

Northcott, M. (2012). *Intimate partner violence risk assessment tools: A review* (User report: rr12-08e). Ottawa, Ontario: Department of Justice Canada.

Ogundimu, E.O., Altman, D.G., & Collins, G.S. (2016). Adequate sample size for developing prediction models is not simply related to events per variable. *Journal of Clinical Epidemiology, 76*(X), 175-182. http://doi.org.10.1016/j.jclinepi.2016.02.031

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. http://doi.org/10.1037/tam0000065

Olver, M.E., Beggs Christofferson, S.M., Grace, R.C., & Wong, S.C.P. (2014). Incorporating change information into sexual offender risk assessments using the Violence Risk Scale- Sexual Offender Version. *Sexual Abuse: A Journal of Research and Treatment, 26*(5)*,* 472-499. http://doi.org/10.1177/1079063213502679

Olver, M.E., Beggs Christofferson, S.M., Nicholaichuk, T.P., & Wong, S.C.P (2021). Predictive properties of the Violence Risk Scale – Sexual Offense Version as a function of age. *Assessment, 28*(6), 1671-1693. http://doi.org/10.1177/1073191120914405

Olver, M.E., & Eher, R. (2019). Predictive properties and factor structure of the VRS-SO in an
    Austrian sample. *European Journal of Psychological Assessment, 36*(5), 748-757.
    https://doi.org/10.1027/1015-5759/a000551

Olver, M.E., Kelley, S.M., Kingston, D.A., Beggs Christofferson, S.M., Thornton, D., & Wong, S.C.P.
    (2021). Incremental contributions of static and dynamic sexual violence risk assessment.
    Integrating Static-99R and VRS-SO common language risk levels. *Criminal Justice and
    Behavior, 48*(8), 1091-1110. http://doi.org.10/1177/0093854820974400

Olver, M.E., Mundt, J.C., Thornton, D., Beggs Christofferson, S.M., Kingston, D.A., Sowden,
    J.N.,…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.
    http://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. http://doi.org/10.1037/pas0000533

Olver, M.E., Nicholaichuk, T.P., Kingston, D.A., & Wong, S.C.P. (2014). A multisite examination of
    sexual violence risk and therapeutic change. *Journal of Consulting and Clinical Psychology,
    82*(2), 312-324. http://doi.org/10.1037/a003540

Olver, M.E., Nicholaichuk, T.P., Kingston, D.A., & Wong, S.C.P. (2020). A prospective multisite
    examination of dynamic sexual violence risk: Extension and update to Olver, Nicholaichuk,
    Kingston, & Wong (2014). (2020). *Journal of Consulting and Clinical Psychology, 88*(4), 362-
    371. http://doi.org/10.1037/ccp0000478

Olver, M.E., Sowden, J.N., Kingston, D.A., Nicholaichuk, T.P., Gordon, A., Beggs Christofferson,
    S.M., & Wong, S.C.P. (2018). Predictive accuracy of Violence Risk Scale – Sexual Offender
    Version risk and change scores in treated Canadian Aboriginal and non-aboriginal sexual
    offenders. *Sexual Abuse, 30*(3), 254-275. http://doi.org/10.1177/1079063216649594

Olver, M.E., & Stockdale, K.E. (2020). Evaluating change in men who have sexually offended:
    Linkages to risk assessment and management. *Current Psychiatry Reports, 22*(22), 1-10.
    http:/doi.org/10.1007/s11920-020-01146-3

Olver, M.E., & Stockdale, K.E. (2021). Sexual violence risk assessment with Indigenous men: Context, controversies, and current findings. *Current Psychiatric Reports, 23*(8), 1-10. http://doi.org/10.1007/s11920-021-01261-9

Olver, M.E., & Reimer, E.K. (2021). 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. http://doi.org/10.1037/ccp0000638

Olver, M.E., Thornton, D., & Beggs Christofferson, S.M. (2022). Understanding the latent structure of dynamic risk: Seeking empirical constraints on theory development using the VRS-SO and the Theory of Dynamic Risk. *Sexual Abuse, 34*(2), 227-254. https://doi.org.19.1177/10790632211002858

Olver, M.E., Wong, S.C.P., Nicholaichuk, T., & Gordon, A. (2007). The validity and reliability of the Violence Risk Scale – Sexual Offender version: Assessing sex offender risk and evaluation of therapeutic change. *Psychological Assessment, 19,* 318-329.

Perley-Robertson, B., Helmus, M.L., & 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, and Law, 25*(3), 248-278. https://doi.org/10/1080/1068316X.2018.1519827.

Pettersen, C. (2013). *Factor structure of the Aggression Questionnaire with sexual offenders* [Master's thesis, Carleton University].

Pettersen, C., Nunes, K.L., Cortoni, F. (2016). Does the factor structure of the Aggression Questionnaire hold for sexual offenders? *Criminal Justice and Behavior, 43*6), 811-829. http://doi.org10/1177/0093854815618692

Prochaska, J. O., DiClemente, C. C., & Norcross, J. C. (1992). In search of how people change: Applications to addictive behaviors. *American Psychologist, 47,* 1102–1114.

Rennie, C., & Dolan, M. (2010). Predictive validity of the Youth Level of Service/Case Management Inventory in custody sample in England. *The Journal of Forensic Psychiatry and Psychology, 21*(3), 407-425. http://doi.org/10.1080/14789940903452311

Rettenberger, M., Matthes, A., Boer, D.P., & Eher, R. (2010). Prospective actuarial risk assessment: A comparison of five risk assessment instruments in different sexual offender subtypes. *International Journal of Offender Therapy and Comparative Criminology, 54*(X), 169-186. http://doi.org/10.1177/0306624X08328755

Roberts, C.F., Doren, D.M., & Thornton, D. (2002). Dimensions associated with assessments of sex offender recidivism risk. *Criminal Justice and Behavior, 29*(5), 569-589. http://10.1177/009385402236733

Rockhill, B., Byrne, C., Rosner, B., Louie, M.M., & Colditz, G. (2003). Breast cancer risk prediction with a log-incidence model: Evaluation of accuracy. *Journal of Clinical Epidemiology, 56*(9), 856-861. https://doi:10.1016/S0895-4356(03)00124-0

Ryan, J., Wilson, N., Kilgour, G., & Reynolds, N. (2014). Balancing efficiency and accuracy: Guidelines for the psychological assessment of extended supervision orders for child sex offenders in New Zealand. *Psychiatry, Psychology and Law, 21*(1), 139-157. http://doi.org:10.1080/13218719.2013.793154

Sample, L.L., & Kadleck, C. (2008). Sex offender laws: Legislators' accounts of the need for policy. *Criminal Justice Policy Review, 19*(1), 40-62. http://doi.org/10.1177/0887403407308292

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: test of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research – Online, 8*(8)*,* 23-74.

Schmucker, M. & Lösel, F. (2015). The effects of sexual offender treatment on recidivism: An international meta-analysis of sound quality evaluations. *Journal of Experimental Criminology, 11*(4), 597-630. http://doi.org/10.1007/s11292-015-9241-z

Sinha, (2013). Measuring violence against women: Statistical trends. *Juristat, 33.* Statistics Canada Catalogue no. 85-002-X. Ottawa ON: Statistics Canada.

Smallbone, S. & Rallings, M. (2013). Short-term predictive validity of the Satic-99 and Static-99R for Indigenous and nonindigenous Australian sexual offenders. *Sexual Abuse: A Journal or Research and Treatment, 25*(3), 302-316. http://doi.org/10.1177/1079063212472937

Smid, W.J., Kamphuis, J.H., Wever, E.C., & Van Beek, D.J. (2014). A comparison of the predictive properties of nine sex offender risk instruments. *Psychological Assessment, 26*(3)*,* 691-703. http://doi.org/10.1037/a0036616

Sowden, J.N., & Olver, M.E. (2017). 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. http://doi.org/10.1037/pas0000345

Stinson, J.D., & Becker, J.V. (2013). *Treating sex offenders: An evidence-based manual.* Guilford Press.

Stinson, J.D., Becker, J.V., & McVay, L. (2017). Treatment progress and behavior following 2 years of inpatient sex offender treatment: A pilot investigation of Safe Offender Strategies. *Sexual Abuse: A Journal of Research and Treatment, 29*(1), 3-27. http://doi.org/10.1177/1079063215570756

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 violent young offenders. *Criminal Justice and Behavior, 41*(1), 114-138. http://doi.org/10.1177/0093854813496999

Sun, J. (2005). Assessing goodness of fit in confirmatory factor analysis. *Measurement and Evaluation in Counseling and Development, 37*(4), 240-256. http://doi.org/10.1080/07481756.2005.11909764

Tabachnick, B.G., & Fidell, L.S. (2013) *Using multivariate statistics (6th ed).* Pearson.

Tasca, G.A., Angus, L., Bonli, R., Drapeau, M., Fitzpatrick, M., & Hunsely J. (2018). Outcomes and progress monitoring in psychotherapy. Ottawa, ON: Task force report prepared for the Canadian Psychological Association.

Tharp, A.T., DeGue, S., Valle, L.A., Brookmeyer, K.A., Massetti, G.M., & Matjasko, J.L. (2013). A systematic qualitative review of risk and protective factors for violence perpetration. *Trauma, Violence, and Abuse, 14*(2), 133-167. http://doi.org/10.1177/1524838012470031

Thornton, D. (2002). Constructing and testing a framework for dynamic risk assessment. *Sexual Abuse: A Journal of Research and Treatment, 14*(), 139-153. http://doi.org/10.1023/A:1014620214905

Thornton, D. (2016). Developing a theory of dynamic risk. *Psychology, Crime & Law, 22*(1-2), 138-150. http://doi.org/10.1080/1068316X.2015.1109092

Townsend, C., & Rheingold, A.A. (2013). Estimating child abuse prevalence rates for practitioners: A review of child sexual abuse prevalence studies. Charleston, S.C., Darkness to Light. https://www.d2l.org/wp-content/uploads/2017/02/PREVALENCE-RATE-WHITE-PAPER-D2L.pdf

Tully, R.J., Chou, S., & Browne, K.D. (2013). A systematic review on the effectiveness of sex offender risk assessment tools in predicting sexual recidivism of adult male sex offenders. *Clinical Psychology Review, 33*(2), 287-316. http://doi.org/10.1016/j.cpr.2012.12.002

Ullrich, S., & Coid, J. (2011). Protective factors for violence among released prisoners – Effects over time and interactions with static risk. *Journal of Consulting and Clinical Psychology, 79*(), 381-390. http://doi.org/10.1037/a0023613

van den Berg, J.W., Smid, W., Schepers, K., Wever, E., van Beek, D., Janssen, E., & Gijs, L. (2018). The predictive properties of dynamic sex offender risk assessment instruments: A meta-analysis. *Psychological Assessment, 30*(2), 179-191. http://doi.org/10/1037/pas0000454

Veith, A. (2018). *The predictive validity of the Static-99R and Stable-2007 in a community sample of sex offenders.* (Publication No.2372) [Doctoral dissertation, University of North Dakota]. UND Scholarly Commons.

Wakeling, H., Beech, A.R., & Freemantle, N. (2013). Investigating treatment change and its relationship to recidivism in a sample of 3773 sex offenders in the UK. *Psychology, Crime & Law, 19*(3), 233-252. http://doi.org/10.1080/1068316X.2011.626413

Walker, M., & O'Rourke, M. (2013). Probation officers' experience using Risk Matrix 2000 and Stable & Acute 2007 when supervising sex offenders living in the community. *Irish Probation Journal, 10,* 162-176. Available from http://www.probation.ie

Ward, T. (2010). The good lives model of offender rehabilitation: Basic assumptions, aetiological commitments, and practice implications. In F. Neil, P. Raynor, & C. Trotter (Eds.), *Offender supervision: New directions in theory, research, and practice* (pp. 41-64). New York, NY: Willan.

Ward, T. & Steward, C. (2003). Criminogenic needs and human needs: A theoretical model. *Psychology, Crime, & Law, 9,* 125-143. http://doi:10.1080/1068316031000116247

Ward, T., Yates, P.M., & Willis, G.M. (2012). The good lives model and the risk-need-responsivity model: A critical response to Andrews, Bonta, and Wormith (2011). *Criminal Justice and Behavior, 39*(1), 94-110. http://doi.org.10.1177/0093854811426085

Webster, C.D., Douglas, K.S., Eaves, D., & Hart, S.D. (1997). *HCR-20. Assessing the risk of violence. Version 2.* Burnaby, Canada: Simon Fraser University and Forensic Psychiatric Services Commission of British Columbia.

Webster, S.D., Mann, R.E., Carter, A.J., Long, J., Milner, R.J., OBrien, M.D., Wakeling, H.C., & Ray, N.L. (2006). Inter-rater reliability of dynamic risk assessment with sexual offenders. *Psychology, Crime, and Law, 12*(4), http://doi.org/10.1080/10683160500036889

Willis, G.M., Kelley, S.M., & Thornton, D. (2020). Interrater reliability and construct validity of proposed protective factors against sexual reoffending. *Criminal Justice & Behavior, 47*(11), 1448-1467. http://doi.org/10.1177/0093854820941039

Willis, G.M., & Thornton, D. (2021, May). *Introducing the SAPROF-SO version 1.* Symposium conducted at NYS ATSA & The Alliance Virtual Joint Conference.

Willis, G.M., Thornton, D., Kelley, S.M., & de Vries Robbé, M. (2021). *Forging a new reality in sexual recidivism risk assessment: From the Pilot Version to Version 1 of the Structured Assessment of Protective Factors against Sexual Offending.* Symposium conducted at the 40th Annual ATSA Research and Treatment Conference, virtual.

Willis, G.M, Thornton, D., Kelley, S.M., & de Vries Robbé, M. (2017-2020). *The Structured Assessment of Protective Factors for Violence Risk – Sex Offence version (SAPROF-SO) Pilot manual.*

Willis, G.M., Thornton, D., Kelley, S.M., & de Vries Robbé, M. (2021, September) Forging a new reality in sexual recidivism risk assessment: From the Pilot Version to Version 1 of the Structured Assessment of Protective Factors against Sexual Offending. Symposium conducted at the 40th Annual ATSA Research and Treatment Conference, virtual

Willis, G.M., Yates, P.M., Gannon, T.A., & Ward, T. (2013). How to integrate the Good Lives Model into treatment programs for sexual offending: An introduction and overview. *Sexual Abuse: A Journal of Research and Treatment, 25*(2), 123-142. http://doi.org.10.1177/1079063212452618

Wilson, R., Cortoni, F., Picheca, J., Stirpe, T., & Nunes, K. (2009). *Community-Based Sexual Offender Maintenance Treatment Programming: An Evaluation.* Research Report R-188, Ottawa: Correctional Service Canada.

Wong, S.C.P., & Gordon, A. (2006). 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.C.P., Olver, M.E., Nicholaichuk, T.P., & Gordon, A. (2003, 2017). *The Violence Risk Scale: Sexual Offense version (VRS-SO).* Regional Psychiatric Centre and University of Saskatchewan.

Wormith, J.S., Gendreau, P., & Bonta, J. (2012). Deferring to clarity, parsimony, and evidence in reply to Ward, Yates, and Willis. *Criminal Justice and Behaviour, 39*(1), 111-124, http://doi.org/10.1177/0093854811426087

Wormith, J.S., Hogg, S.M., & Guzzo, L. (2012). The predictive validity of a general risk/needs assessment inventory on sexual offender recidivism and an exploration of professional override. *Criminal Justice and Behavior, 39*(12), 1511-1538. http://doi.org/10.1177/009385481255741

Zevitz, R.G., & Farkas, M.A. (2000). Sex offender community notification: Managing high risk criminals or exacting further vengeance? *Behavioral Science & the Law, 18*(2-3), 375-391. http://doi.org/10.1002/1099-0798

Appendix A

University of Saskatchewan Behavioural Research Ethics Board Certificate

UNIVERSITY OF SASKATCHEWAN

Behavioural Research Ethics Board (Beh-REB)   21-Mar-2022

Certificate of Re-Approval

Application ID: 907

Principal Investigator: Mark Olver   Department: Department of Psychology

Locations Where Research Activities are Conducted: Alberta Health Services Forensic Assessment and Community Services 10225 106th Street, Edmonton Alberta., Canada

Student(s): Farron Wielinga
Martina Faitakis

Funder(s): Centre for Forensic Behavioural Science and Justice Studies

Sponsor:

Title: Use of Dynamic Risk Instruments to Assess Sexual Violence Risk and Evaluate Change in a Community Based Sexual Offence Program

Approval Effective Date: 02-Mar-2022

Expiry Date: 02-Mar-2023

Acknowledgment Of: N/A

Review Type: Delegated Review

* This study, inclusive of all previously approved documents, has been re-approved until the expiry date noted above

CERTIFICATION

The University of Saskatchewan Behavioural Research Ethics Board (Beh-REB) is constituted and operates in accordance with the current version of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2 2014). The University of Saskatchewan Behavioural Research Ethics Board has reviewed the above-named project. The proposal was found to be acceptable on ethical grounds. The principal investigator has the responsibility for any other administrative or regulatory approvals that may pertain to this project, and for ensuring that the authorized project is carried out according to the conditions outlined in the original protocol submitted for ethics review. This Certificate of Approval is valid for the above time period provided there is no change in experimental protocol or consent process or documents.

ONGOING REVIEW REQUIREMENTS

In order to receive annual renewal, a status report must be submitted to the REB Chair for Board consideration within one month prior to the current expiry date each year the project remains open, and upon project completion. Please refer to the following website for further instructions: https://vpresearch.usask.ca/researchers/forms.php.

Digitally Approved on behalf of the Chair
Behavioural Research Ethics Board
University of Saskatchewan

1 / 1

Dynamic Risk Assessment in a Community Setting

BASIC INFORMATION

1. _ _ / _ _ / _ _ _ _ **DATECODE:** Date the file was coded (mm/dd/yyyy)
2. _ _ **CODER:** The initials of the coder
3. _ _ _ _ _ _ **FILENO:** FACS file number – 5 digits
4. _ _ / _ _ / _ _ _ _ **AXDATE:** Date of Assessment (mm/dd/yyyy)
5. _ _ / _ _ / _ _ _ _ **INDEXDAT:** Date of Index Offense (mm/dd/yyyy)
   _ _ -_ _ _ _ _ _ _ _ _ **POLICE.OCC:** Police occurrence # corresponding to the index offense

DEMOGRAPHIC INFORMATION

6. _ _ / _ _ / _ _ _ _ **ODOB:** Offender date of birth (mm/dd/yyyy)
7. __ **ETHNICITY:** Offender's Ethnicity (1 = White/European; 2 = First Nations/Métis/Inuit 3 = Other; specify: _____)
8. __ **EDUCATE:** Highest grade completed (elementary, junior high, or high school) (#; use "at least" number)
9. __ **HIGHSCHOOL**: Completed high school (high school diploma or equivalent) *
   ☐ High School Diploma [1]   ☐ Equivalent (i.e., GED) [2]
10. __ **STARTCOLLEGE:** Started college or university training
11. __ **COMPLETECOLLEGE:** Completed college or university program/degree
12. __ **EMPLOYED:** Employed at time of index? (specify: _____)
13. __ **SOTIPS14:** Employment
    0 = Minimal or no problems
    1 = Some problems
    2 = Considerable problems
    3 = Serious problems

14. __ **SAPROF_SO13:** Work (consider the past six months)
    4 = A stable and suitable work situation is clearly present
    2 = A stable and suitable work situation is present to some extent OR consistent evidence of job preparedness activities are present
    0 = A stable and suitable work situation is not present, no consistent evidence of job preparedness activities

RELATIONSHIPS/SOCIAL SUPPORT

15. __ **SOTIPS16:** Social Influences (consider previous six months)
    0 = Associates primarily with and values the opinions of friends, family, and associates who are positive influences.
    1 = Associates more with and values the opinions of friends, family, and associates who are positive influences than those who are negative influences.
    2 = Associates more with and values the opinions of friends, family, and associates who are negative influences than those who are positive influences or is very socially isolated and does not associate with others.

3 = Associates primarily with friends, family, and associates who are negative influences.

16. __ **SAPROF_SO15:** Social network (consider the past six months)
    4 = A prosocial and supportive network is clearly present
    2 = A prosocial and supportive social network is present to some extent
    0 = A prosocial and supportive social network is not present

17. __ **SAPROF_SO16:** Emotional connection to adults (consider the past six months)
    4 = Emotional connections to adults are clearly present
    2 = Emotional connections to adults are present to some extent.
    0 = Emotional connections to adults are not present.

18. __ **SAPROF_SO17:** Intimate relationships (consider the past 6 months)
    4 = A stable intimate relationship of good quality is clearly present
    2 = A stable intimate relationship of good quality is present to some extent
    0 = A stable intimate relationship of good quality is not present

19. __ **SAPROF_SO2:** Secure attachment in childhood (consider ages 0-18)
    4 = Secure attachment with at least one prosocial adult was clearly present in childhood.
    2 = Secure attachment with at least one prosocial adult was present in childhood to some extent.
    0 = Secure attachment with at least one prosocial adult was not present in childhood.

SEXUAL DEVIANCE

20. __ **SAPROF_SO7:** Sexual self-regulation (consider past 12 months)
    4 = Sexual self-regulation is clearly present
    2 = Sexual self-regulation is present to some extent
    0 = Sexual self-regulation is not present

21. __ **SAPROF_SO8:** Prosocial sexual interests (consider past 24 months)
    4 = Enduring pro-social sexual interests are clearly present.
    2 = Enduring pro-social sexual interests are present to some extent
    0 = Enduring pro-social sexual interests are not present

22. __ **SAPROF_SO9:** Prosocial sexual identity (consider past 12 months)
    4 = Enduring pro-social sexual identity is clearly present
    2 = Enduring pro-social sexual identity is present to some extent
    0 = Enduring pro-social sexual identity is not present

23. __ **SAPROF_SO21:** Medication (consider past 6 months)
    4 = Motivation for and effective use of medication is clearly present
    2 = Motivation for and effective use of medication is present to some extent.
    0 = Motivation for and effective use of medication is not present
    n/a = no medication has been prescribed or recommended

24. __ **SOTIPS2:** Sexual Behaviour (consider past 6 months)
0 = No problems evident. Sexual behaviour is limited to legal and non-compulsive sexual activity. If behaviour involves a fetish, it is not illegal, and has no known relationship to the individual's sexual offending pattern.
1 = Some problems evident. These include occasional promiscuous behaviour and pornography use against probation conditions, treatment requirement, or facility rules.
2 = Considerable problems evident. These include masturbating to the extent that it interferes with life activities or causes physical harm, masturbating to offense-related sexual fantasies, frequenting strip bars, using phone sex lines, using pornography habitually, being promiscuous, and consistently engaging in sexual activity leading to orgasm more than seven times a week during the previous six months.
3 = Engaged in illegal sexual behaviour. This includes child molesting, rape, exhibitionism, child pornography, and prostitution. If in a residential setting, engaged in sexual behaviour that is against facility rules.

25. __ **SOTIPS3:** Sexual Attitudes (consider past 6 months)
0 = Has no or minimal difficulty recognizing and self-correcting attitudes and thoughts that support sexual offending.
1 = Has some difficulty recognizing or self-correcting attitudes and thoughts that support sexual offending. Is open to examining and changing these attitudes and thoughts.
2 = Has considerable difficulty recognizing or self-correcting attitudes and thoughts that support sexual offending. Has some openness to examining and changing these attitudes and thoughts.
3 = Does not recognize or self-correct attitudes and thoughts that support sexual offending. Is not open to examining and changing these attitudes and thoughts.

26. __ **SOTIPS4:** Sexual interests (consider past 6 months)
0 = All sexual interests are appropriate themes. (SSPI scores typically = 0 or 1).
1 = Most sexual interests are appropriate themes. (SSPI scores typically = 1, 2, or 3).
2 = Most sexual interests in offense-related themes. (SSPI scores typically = 3, 4, or 5).
3 = All sexual interests in offense-related themes. (SSPI scores typically = 4 or 5).

27. __ **SOTIPS5:** Sexual risk management (consider past 6 months)
0 = Good understanding of sexual offense risk factors and risk management strategies and uses effective risk management strategies on a very consistent basis
1 = Good understanding of sexual offense risk factors and risk management strategies and uses effective risk management strategies on a relatively consistent basis with occasional minor lapses.
2 = Partial understanding of sexual offense risk factors and risk management strategies or inconsistently uses effective risk management strategies with several lapses.
3 = Poor understanding of sexual offending risk factors and risk management strategies or intermittently or rarely uses effective risk management strategies or has had a serious lapse.

122

COGNITIONS/EMOTIONS/FUNCTIONING

28. __ **SAPROF_SO1:** Intact cognitive functioning
    4 = Intact cognitive functioning including at least average intelligence (FSIQ $\geq$ 90) and no evidence of other cognitive impairments
    2 = Low average IQ (FSIQ $\geq$ 80-89), or average IQ with some cognitive impairments.
    0 = Borderline or extremely low IQ and/or marked cognitive impairment. Individuals diagnosed with Intellectual Disability are included in this category.

29. __ **SAPROF_SO3:** Adaptive schema (consider past 12 months)
    4 = Adaptive schema dominate across the domains of self and others/world
    2 = Adaptive schema dominate most of the time across the domains of self and others/world
    0 = Maladaptive schema dominate across the domains of self and others/world.

30. __ **SAPROF_SO4:** Empathy (consider past 6 months)
    4 = Empathic perspective-taking leading to helpful behaviours clearly present.
    2 = Empathic perspective-taking leading to helpful behaviours is present to some extent.
    0 = Empathic perspective-taking leading to helpful behaviours is not present.

31. __ **SAPROF_SO5**: Coping (consider past 6 months)
    4 = Effective coping with daily life stressors is clearly present.
    2 = Effective coping with daily life stressors is present to some extent.
    0 = Effective coping with daily life stressors is not present.

32. __ **SAPROF_SO6:** Self-control (consider past 12 months)
    4 = Self-control is clearly present.
    2 = Self-control is present to some extent.
    0 = Self-control is not present.

33. __ **SOTIPS11:** Emotional Management (consider past 6 months)
    0 = No emotion management problems.
    1 = Minor emotional management problems. They are:
    - Relatively infrequent, and
    - Managed relatively effectively
    2 = moderate emotional management problems. They are:
    - Relatively frequent, or
    - Managed relatively ineffectively.
    3 = Serious emotional management problems. They are:
    - Frequent and intense, and
    - Managed very ineffectively

34. __ **SOTIPS13:** Impulsivity (consider past 6 months)
   0 = Behaviour is planned, thoughtful, and purposeful. Rarely or never does things that are impulsive, unplanned, and lack deliberation.
   1 = Occasionally does things that are impulsive, unplanned, and lack deliberation
   2 = Frequently does things that are impulsive, unplanned, and lack deliberation
   3 = Regularly does things that are impulsive, unplanned, and lack deliberation

35. __ **SOTIPS12:** Problem solving (consider past 6 months)
   0 = Successful at identifying and addressing typical life problems
   1 = Some problem-solving deficits:
   - Occasionally makes poorly considered decision, but
   - is able to self-correct when difficulties are pointed out.
   2 = Considerable problem-solving deficits:
   - Occasionally makes poorly considered decisions, and
   - Has trouble correcting even when difficulties are pointed out
   3 = Serious impairment:
   - Fails to identify obvious life problems
   - Frequently makes poorly considered decisions, and
   - Has difficulty recognizing negative consequences of decisions and self-correcting even when consequences are pointed out.

36. __ **SAPROF_SO10:** Life goals (describe)
   4 = Positive life goals are clearly present.
   2 = Positive life goals are present to some extent.
   0 = Positive life goals are not present.

COMMUNITY LIVING

37. __ **SAPROF_SO14:** Leisure activities (consider past 6 months)
   4 = Structured leisure activities are clearly present
   2 = Structured leisure activities are present to some extent
   0 = Structured leisure activities are not present

38. __ **SAPROF_SO18:** Housing Stability (consider past 6 months)
   4 = Stable housing is present.
   2 = Stable housing is present to some extent.
   0 = Stable housing is not present.

39. __ **SAPROF_SO19:** Financial management (consider past 6 months)
   4 = Steady income and sound financial management are clearly present.
   2 = Steady income and sound financial management are present to some extent.
   0 = Steady income and sound financial management are not present.

40. __ **SOTIPS15:** Residence (in the last six months)
   0 = Not more than one address change and satisfied with accommodation.
   1 = Two address changes or somewhat dissatisfied with accommodation.
   2 = Three or more address changes or very dissatisfied with accommodation.
   3 = No fixed address.

41. __ **SAPROF_SO23:** Supervised living (consider past 6 months)
    4 = Living circumstances are intensively supervised by professionals.
    2 = Living circumstances are supervised by professionals, or the individual lives with prosocial others.
    0 = Living circumstances are not intensively supervised and the individual does not live with prosocial others.

SUBSTANCE ABUSE AND MENTAL HEALTH

42. __ **OSUBIDX:** Alcohol or drugs used by the perpetrator at index?
    0 = No
    1 = Yes

43. __ **SUBUSE:** Evidence of substance use/dependence (consider lifespan, but note recency)
    0 = No evidence of substance abuse or dependence
    1 = Less serious substance abuse
    2 = Serious (impair health, social, legal) substance abuse or dependence

44. **SUBLIST:** List substances

_____
_____
_____
_____

45. __ **MHEALTH:** Indicate whether the perpetrator experienced mental health problems (consider lifespan but prioritize recency)
    0 = No evidence of problems with mental health
    1 = Possible or partial evidence of problems with mental health
    2 = Evidence of problems with mental health
                       ☐ Definite [1]   ☐ Provisional [2]
46. **MHDX:** List psychiatric diagnoses on file

_____
_____
_____
_____

VICTIM VARIABLES
47. _ _ / _ _ / _ _ _ _ **VDOB:** Index victim's date of birth (mm/dd/year)

48. __ **VETH:** Victim's Ethnicity
  1 = White/European
  2 = First Nations/Métis/Inuit
  3 = Other
  If other is selected, please specify ethnicity. If victim's ethnicity is of mixed origin, code as "3" and specify the ethnicity that the victim most identifies with (where possible).

49. __ **Number of victims** (Index and historical)

Indicate number of victims	Index	History
Male		
Female		
Child (< 12 years) *		
Adolescent (12-18 years) *		
Related/intrafamilial (<18 years) *		
Unrelated but known (i.e., friends, acquaintances)		
Stranger		
Total		

*Indicate number of each victim

CRIMINAL ATTITUDES/ LIFESTYLE

50. __ **SOTIPS1:** Sexual offense responsibility (current level of functioning)
     0 = Takes full responsibility for sexual offense behaviour.
     1 = Takes most of the responsibility for sexual offense behaviour. Places some blame elsewhere
     2 = Takes some of the responsibility for sexual offense behaviour. Places considerable blame elsewhere.
     3 = Is in categorical denial, or otherwise takes no responsibility for offense behaviour.

51. __ **SOTIPS6:** Criminal and rule breaking behaviour (previous six months)
     0 = No criminal or rule breaking behaviour evident
     1= Minor non-sexual problems evident, such as:
     - A minor motor vehicle charge,
     - A minor residential/correctional facility or program rule-breaking incident, or
     - A minor manipulative behaviour
     2= Moderate non-sexual problems evident, such as behaviour that has or could reasonably lead to:
     - Two or more minor motor vehicle charges
     - Two or more minor residential/correctional facility rule infractions,
     - One or more major residential/correctional facility rule infractions,
     - One or more major treatment program rule-breaking behaviours,
     - One or more misdemeanor offense charges
     - A technical violation of probation, parole, or other community supervision status, or
     - Multiple single incidents of any of the above.
     3 = Serious non-sexual problems evident, such as behaviour that has or could reasonably lead to:
     - A felony offense charge, or
     - A residential/correctional facility rule infraction that could lead to a felony offense charge.

52. __ **SAPROF_SO12:** Attitudes towards rules and regulations (consider past 6 months)
4 = Positive attitudes towards rules and regulations are clearly present.
2 = Positive attitudes towards rules and regulations are generally present but with isolated exceptions.
0 = Positive attitudes towards rules and regulations are not present.

53. __ **SAPROF_SO11:** Motivation for managing risk (consider past 12 months)
4 = Motivation for managing risk is clearly present (i.e., evidence the individual is in a Maintenance stage of change)
2 = Motivation for managing risk is present to some extent (i.e., evidence the individual is in a Preparation stage of change)
0 = Motivation for managing risk is not present (i.e., evidence the individual is in a Precontemplation stage of change)
n/a = Individual is at a low risk of sexual recidivism

54. __ **SOTIPS10:** Cooperation with community supervision (previous six months)
0 = Has no or minimal problems
1= Some problems – compliance less than 90% of the time with unexcused absences, minor supervision compliance problems or closed channel of communication
2= Considerable problems
3 = Severe problems

55. __ **SOTIPS7:** Criminal and rule breaking attitudes (previous six months)
0 = Has no or minimal difficulty recognizing and self-correcting attitudes and thoughts that support criminal or rule-breaking behaviour.
1= Has some difficulty recognizing or self-correcting attitudes and thoughts that support criminal or rule-breaking behaviour. Is open to examining and changing these attitudes and thoughts.
2= Has considerable difficulty recognizing or self-correcting attitudes and thoughts that support criminal or rule-breaking behaviour. Has some openness to examining and changing these attitudes and thoughts.
3 = Does not recognize or self-correct attitudes and thoughts that support criminal or rule-breaking behaviour. Is not open to examining and changing these attitudes and thoughts.

56. __ **SAPROF_SO24:** External control: Court-ordered or mandatory supervision and/or treatment (consider current sentence conditions/court order)
4 = Intensive external control is present
2 = External control is present to some extent.
0 = External control is absent.

57. __ **FCOND:** Failure on prior conditional release or community supervision
0 = No
1 = Yes

58. __ **PCUST:** Prior custodial sentence of 30 days or more
0 = No
1 = Yes

TREATMENT RELATED

59. _ _ / _ _ / _ _ _ _ **TXTIME:** Date that perpetrator entered treatment (mm/dd/yyyy)

60. __ **TXCOMPLETE:** Did the perpetrator complete treatment?
    0 = No
    1 = Yes

61. _ _ / _ _ / _ _ _ _ **TXCOMP:** Date that perpetrator completed treatment (mm/dd/yyyy)

62. __ **TXNONREASON:** If no, provide reason that perpetrator did not complete treatment
    1 = Disruptive behaviour
    2 = Low motivation/ poor effort
    3 = Program infractions
    4 = Security concerns
    5 = Client requested
    6 = Add as needed (specify reason)

63. __ **TXINITIATE:** Who initiated discharge from treatment?
    1 = Staff-initiated
    2 = Client-initiated
    3 = Mutually initiated
    4 = System-initiated

64. __ **TXSUSPEND:** Was the perpetrator suspended from treatment?
    0 = No
    1 = Yes

65. __ **TXSUSREASON**: If yes, provide reason that the perpetrator was suspended from treatment.
    1 = Disruptive behaviour
    2 = Low motivation/ poor effort
    3 = Program infractions
    4 = Security concerns
    5 = Client requested
    6 = Add as needed (specify reason)

66. __ **SAPROF_SO20:** Sexual offense-specific treatment: Availability of appropriate treatment services
    4 = Appropriate sexual offense-specific treatment is clearly present
    2 = Appropriate sexual offense-specific treatment is present to some extent
    0 = Appropriate sexual offense-specific treatment is not present
    n/r = Appropriate sexual offense-specific treatment is not present AND the individual does not require treatment

67. __ **SAPROF_SO22:** Perception of the therapeutic/supervisory alliance
    4 = A positive therapeutic/supervisory alliance is clearly present
    2 = A positive therapeutic/supervisory alliance is present to some extent
    0 = A positive therapeutic/supervisory alliance is not present

n/a = The individual is not in therapy or under supervisory, or the individual has been in therapy or under supervision for less than one month

68. __ **SOTIPS9**: Cooperation with treatment (previous six months)
    0 = No or minimal problems
    1 = Some problems (compliance less than 90% of the time with absences, failure to complete assignments, poor participation)
    2 = Considerable problems (compliance less than 90% of the time with absences, failure to complete assignments, poor participation <u>or</u> incidents of serious disruptive behaviour)
    3 = Severe problems

69. __ **SOTIPS8**: Stage of change (previous six months)
    0 = Maintenance stage
    1 = Action stage
    2 = Ambivalent stage
    3 = Pre-contemplation stage

70. __ **TXINDIVID:** Is the perpetrator seeing an individual therapist?
    0 = No
    1 = Yes

71. __ **TXOTHER:** Has the perpetrator participated in any previous treatment? (e.g., substance abuse, anger management etc.)
    0 = No
    1 = Yes

72. __ **TXOTHERSO:** Has the perpetrator participated in past <u>sex-offense specific</u> treatment?
    0 = No
    1 = Yes

CONDITIONS ON PERPETATOR (restrictions and/or treatment requirements) 0 = No, 1 = Yes

73. __ **ConSubs:** Drug/alcohol use (e.g., prohibition, referral for treatment, etc.)
74. __ **ConCouns:** Anger management and/or other treatment (e.g., referral for treatment/counselling, etc.)
75. __ **ConPsych**: Psychiatric and/or forensic mental health (e.g., subject referred for forensic psychiatric evaluation, consult with specialized mental health unit(s) within referring police service, etc.)
76. __ **ConWeapon**: Firearms and/or weapons (e.g., firearms/weapon prohibition requested/maintained)
77. __ **ConRes:** Residency and/or area restriction
78. __ **ConEmploy:** Conditions related to employment (e.g., must notify if changes, vocational assistance, etc.)
79. __ **ConRelat:** Conditions related to relationships (e.g., must advise of new intimate partner, etc.)

80. __ **ConChild:** Conditions related to children (e.g., consultation with AHS or other organization, review of child custody access and agreement, restriction of child visitation until other conditions met, etc.)

CONFIDENCE EVALUATION

81. __ **FILEQUAL:** What was the quality of the information contained in this file? Consider the number of reports, presence of police narrative, presence of completed risk assessments etc.
    1 = Poor quality
    2 = Moderate quality
    3 = High quality

82. __ **RATERCONF:** How confident are you scoring these measures with the information contained in this file?
    1 = Not confident
    2 = Somewhat confident
    3 = Quite confident
    4 = Very confident

83. __ **OTHERTOOL:** Was there another fully completed tool on file to assess risk (e.g., HCR-20)?
    (specify: _____)

84. __ **INDEXOFFENSE:** What was the index offense
    1 = Sexual Assault
    2 = Sexual Interference
    3 = Indecent Act/Exhibitionism
    4 = Voyeurism
    5 = Child Pornography
    6 = Sexual Exploitation
    7 = Add as needed

STATIC-99R

85. __ **Static0:** Researcher or clinician scored?
    0 = Researcher
    1 = Clinician
86. __ **Static1:** Age at release
    1 = Age 18 to 34.9
    0 = Age 35 to 39.9
    -1 = Age 40 to 59.9
    -3 = Age 60 or older
87. __ **Static2:** Ever lived with a lover for at least two years?
    0 = Yes
    1 = No
88. __ **Static3:** Index non-sexual violence – any convictions
    0 = No
    1 = Yes
89. __ **Static4:** Prior non-sexual violence – any convictions
    0 = No
    1 = Yes
90. __ **Static5:** Prior sex offenses

	Charges	**Convictions**
0 =	0	0
1 =	1, 2	1
2 =	3-5	2, 3
3 =	6+	4+

91. __ **Static6:** Prior sentencing dates (excluding index)
    0 = 3 or less
    1 = 4 or more
92. __ **Static7:** Any convictions for non-contact sex offenses
    0 = No
    1 = Yes
93. __ **Static8:** Any unrelated victims
    0 = No
    1 = Yes
94. __ **Static9:** Any stranger victims
    0 = No
    1 = Yes
95. __ **Static10:** Any male victims
    0 = No
    1 = Yes

STABLE-2007

96. __ **Stable0:** Researcher or clinician scored?
    0 = Researcher
    1 = Clinician

97. __ **Stable1:** Significant social influences
    0 = no problem evident
    1 = some problem evident
    2 = significant problem evident

98. __ **Stable2:** Capacity for stable relationships
    0 = no problem evident
    1 = some problem evident
    2 = significant problem evident

99. __ **Stable3:** Emotional identification with children
    0 = no problem evident
    1 = some problem evident
    2 = significant problem evident

100. __ **Stable4:** Hostility toward women
    0 = no problem evident
    1 = some problem evident
    2 = significant problem evident

101. __ **Stable5:** General social rejection
    0 = no problem evident
    1 = some problem evident
    2 = significant problem evident

102. __ **Stable6:** Lack of concerns for others
    0 = no problem evident
    1 = some problem evident
    2 = significant problem evident

103. __ **Stable7:** Impulsive
    0 = no problem evident
    1 = some problem evident
    2 = significant problem evident

104. __ **Stable8:** Poor problem-solving skills
    0 = no problem evident
    1 = some problem evident
    2 = significant problem evident

105. __ **Stable9:** Negative emotionality
    0 = no problem evident
    1 = some problem evident
    2 = significant problem evident

106. __ **Stable10:** Sex drive/ sex preoccupation
    0 = no problem evident
    1 = some problem evident
    2 = significant problem evident

107. __ **Stable11:** Sex as coping
    0 = no problem evident
    1 = some problem evident
    2 = significant problem evident

108. __ **Stable12:** Deviant sexual preference
    0 = no problem evident
    1 = some problem evident
    2 = significant problem evident

109. __ **Stable13**: Cooperation with supervision
    0 = no problem evident
    1 = some problem evident
    2 = significant problem evident

VRS-SO Score Sheet

Pre-Treatment Rater: Pre-Treatment Rating Date:
Post-Treatment Rater: Post-Treatment Rating Date:

Static Factors

Risk Factor[1]		Codes	Score	I or N
S1	**Age at Time of Release**	Under 25 years 25 to 34 years 35 to 44 years 45 years or older	3 2 1 0	
S2	**Age at First Sexual Offense**	Under 20 years 20 to 24 years 25 to 34 years 35 years or older	3 2 1 0	
S3	**Sexual Offense Victim Profile**	Mixed adult/teen and child victim profile Extrafamilial child victim profile Extrafamilial adult/teen victim profile Intrafamilial victim profile	3 2 1 0	
S4	**Prior Sexual Offenses**	4-4+ prior arrests/charges/convictions for a sexual offense 2-3 prior arrests/charges/convictions for a sexual offense 1 prior arrest/charge/conviction for a sexual offense No prior arrests/charges/convictions for a sexual offense	3 2 1 0	
S5	**Unrelated Victims**	4 or more unrelated victims 2-3 unrelated victims 1 unrelated victim No unrelated victims (related victims only)	3 2 1 0	
S6	**Number and Gender of Victims**	2 or more male victims & any number of female victims 2 or more female victims or 1 female and 1 male victim 1 male victim only 1 female victim only	3 2 1 0	
S7	**Prior Sentencing Dates**	11 or more prior sentencing occasions 5-10 prior sentencing occasions 2-4 prior sentencing occasions 0-1 prior sentencing occasions	3 2 1 0	
Total Static Factor Score		Before Treatment		
		After Treatment		

[1]   If it is necessary to omit a Static or Dynamic Factor, the rater should indicate whether the omission is because there is insufficient information (I) or because the item is not applicable (N).

For Stage Of Change: Use these symbols to indicate the Stage of Change: # of Stages changed:

P/C = Precontemplation/Contemplation   O = Pre-treatment   No change = 0.   1 stage = .5

P = Preparation. A = Action M = Maintenance   X = Post-treatment   2 stage = 1.0.   3 stages = 1.5

DYNAMIC FACTORS AND TOTAL SCORES

	RATINGS											
	Pre-Tx (a)	F 1†	F 2	F 3	Stage of Change††	# of Stages changed x .5 (b)	Post-Tx (a-b)†††	F 1	F 2	F 3	I or N
D1 Sexually Deviant Lifestyle	0 1 2 3				P/C P A M	1.5 1 .5 0					
D2 Sexual Compulsivity	0 1 2 3				P/C P A M	1.5 1 .5 0					
D3 Offence Planning	0 1 2 3				P/C P A M	1.5 1 .5 0					
D4 Criminal Personality	0 1 2 3				P/C P A M	1.5 1 .5 0					
D5 Cognitive Distortions	0 1 2 3				P/C P A M	1.5 1 .5 0					
D6 Interpersonal Aggression	0 1 2 3				P/C P A M	1.5 1 .5 0					
D7 Emotional Control	0 1 2 3				P/C P A M	1.5 1 .5 0					
D8 Insight	0 1 2 3				P/C P A M	1.5 1 .5 0					
D9 Substance Abuse	0 1 2 3				P/C P A M	1.5 1 .5 0					
D10 Community Support	0 1 2 3				P/C P A M	1.5 1 .5 0					
D11 Release to High Risk Situations	0 1 2 3				P/C P A M	1.5 1 .5 0					
D12 Sexual Offending Cycle	0 1 2 3				P/C P A M	1.5 1 .5 0					
D13 Impulsivity	0 1 2 3				P/C P A M	1.5 1 .5 0					
D14 Compliance with Community Supervision	0 1 2 3				P/C P A M	1.5 1 .5 0					
D15 Treatment Compliance	0 1 2 3				P/C P A M	1.5 1 .5 0					
D16 Deviant Sexual Preference	0 1 2 3				P/C P A M	1.5 1 .5 0					
D17 Intimacy Deficits	0 1 2 3				P/C P A M	1.5 1 .5 0					
	Pre-Tx:	Factors: 1 2 3			Change Total →		Post-Tx:	Factors: 1 2 3			
Total Dynamic Factor Score →					Total Dynamic Factor Score →						
Total Static Factor Score from Previous Page →					Total Static Factor Score from Previous Page →						
Total Static + Total Dynamic Factor Score →					Total Static + Total Dynamic Factor Score →						

Indicate if Clinical Override was used: Yes ❏   No ❏

†   To calculate scores for Factors 1 (Sexual Deviance), 2 (Criminality), & 3 (Treatment Responsivity): Place Pre-Tx score in the corresponding shaded box to the right (Note: D7 and D17 are excluded). Tally each column (F1, F2, F3) and enter total score in appropriate box.

††   For treatment purposes, specify whether the client is in Precontemplation or Contemplation stage by circling (O) or marking (X) the 'P' or 'C' stage for pre- and post treatment, respectively.

†††   If there is a deterioration during treatment, 'b' score is added to 'a' score for the corresponding Dynamic Factor