


Attending to the Positive: A Retrospective Validation of the Structured Assessment of Protective Factors-Sexual Offence Version

Sexual Abuse
2022, Vol. 0(0) 1–20
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/10790632221098354
journals.sagepub.com/home/sax


Thomas Nolan¹, Gwenda M. Willis² , David Thornton³ , Sharon M. Kelley⁴, and Sarah Beggs Christofferson¹

Abstract

Sexual recidivism risk assessment tools focus almost exclusively on risk factors associated with increased rates of recidivism and do not attend to protective factors that might mitigate reoffense risk. The present study investigated the predictive validity of the Structured Assessment of Protective Factors - Sexual Offence version (SAPROF-SO), developed to assess hypothesised protective factors against sexual recidivism in adult males. The SAPROF-SO pilot version contains 24 items across two domains: Personal and Professionally Provided Support. SAPROF-SO scores were rated retrospectively from a review of archived case files of 210 men with convictions for child sexual offenses, using the SAPROF-SO pilot manual and a supplementary retrospective scoring guide developed for the current study. SAPROF-SO Total and Personal domain scores were significantly predictive of sexual recidivism after an average follow-up period of 12.24 years (AUC = .81), and to a lesser extent, violent and general recidivism. SAPROF-SO Total and Personal scores additionally provided significant incremental validity over Static-99R scores in the prediction of sexual recidivism. Results support the predictive validity of protective factors for reduced sexual recidivism and

¹School of Psychology, Speech and Hearing, University of Canterbury, Christchurch, New Zealand

²School of Psychology, University of Auckland, Auckland, New Zealand

³FASTR, LLC, Madison, WI, USA

⁴Sand Ridge Secure Treatment Centre, Madison WI, USA

Corresponding Author:

Gwenda M. Willis, School of Psychology, University of Auckland, Private Bag, Auckland 92019, New Zealand.
Email: g.willis@auckland.ac.nz

invite future research examining how to integrate the SAPROF-SO alongside contemporary sexual recidivism risk assessment tools.

Keywords

SAPROF-SO, protective factors, risk assessment, sexual recidivism

Assessment of risk for sexual recidivism influences important decisions that can have significant implications for the individual assessed and wider society. Risk assessments may inform charging and sentencing decisions, the availability and intensity of treatment services offered, whether or when to release, conditions for supervision in the community, or even removal from “sexual offender” registries (Thornton et al., 2021). Contemporary instruments for assessing sexual recidivism risk are largely focused on static and dynamic risk factors (e.g., Brankley et al., 2021; Helmus, Thornton et al., 2012), both of which can be regarded as indexing the operation of long-term vulnerabilities for offending (Olver et al., 2021; Thornton, 2016; Ward & Beech, 2004).

This focus on negative aspects of the individuals being assessed likely contributes to a number of limitations of current risk assessment technology. Rogers (2000) has argued that by solely focusing on the negative, important information is missed, leaving risk assessment practices unbalanced. As a consequence, risk assessment may be experienced as an oppressive and adversarial process in which the person being assessed seeks to conceal risk factors while the assessor seeks to uncover them. Relatedly, framing treatment as solely concerned with reducing or eliminating risk factors may be demotivating for treatment participants and may make it harder to develop a rapport between them and the clinicians providing treatment (Cording & Beggs Christofferson, 2017; de Vries Robbé & Willis, 2017; Kelley et al., 2022). Recognition of such limitations has led to the increasing popularity of strengths-based approaches to treatment as epitomised by the Good Lives Model (GLM; Ward & Maruna, 2007; Ward & Stewart, 2003), with preliminary research finding greater client motivation and engagement in GLM consistent programs compared to relapse prevention oriented programs (Mallion et al., 2020). There is thus an increasing divergence between the negative focus of risk assessment and the positive focus of treatment.

One way to reconcile risk assessment and strengths-based treatment is to incorporate the assessment of protective factors, which encourages evaluators to look for evidence of an individual’s achievements and strengths, leading to a more balanced perspective of risk. There is debate in forensic psychology literature about the definition of protective factors (Cording & Beggs Christofferson, 2017). The current study utilises the definition put forward by Willis et al. (2020) which was influenced by the works of (de Vogel et al (2012); de Vries Robbé, Mann, et al., 2015). They define protective factors as follows:

Those factors that are theoretically or empirically associated with reduced rates of sexual or violent recidivism in individuals with at least one apprehension for a sexual offence as an adult. They must signal the presence of a strength, not merely the absence of a risk factor or deficit (Willis et al., 2020, p. 2).

Willis et al. go on to clarify that some protective factors may represent a protective pole of a dimension that has a risk factor at its opposite end. A person may have related protective and risk factors present simultaneously either because the risk factor exists as a long-term vulnerability while the protective pole of the dimension dominates current functioning, or because the poles are related but not incompatible, as when someone has both offense-related and normophilic sexual interests. Other protective factors are theorised to operate independently of known risk factors, including life goals and medication (de Vogel et al., 2012).

The Structured Assessment of PROtective Factors against violence risk (SAPROF; de Vogel et al., 2012) is one of the few validated tools to assess protective factors against violence in adults. Studies examining the predictive validity of the SAPROF have generally yielded favorable findings for the prediction of violent recidivism (for a review of SAPROF research, see de Vries Robb  et al., 2020). Moreover, both Coupland and Olver (2020) and de Vries Robb , de Vogel, Douglas et al. (2015) found that strengthening of protective factors during treatment was associated with a reduction in community recidivism after short and long term (10+ years) follow up periods, lending support to the use of strengths-based assessment to inform strengths-based treatment. Results for the prediction of sexual recidivism have been variable. de Vries Robb , de Vogel, et al. (2015) found that SAPROF Total scores were significantly predictive of sexual reoffending at follow-up times of 3 years (AUC = .76) and 15 years (AUC = .71), but Yoon et al. (2018) found SAPROF Total scores were not predictive of sexual reoffending (AUC = .53) as did Turner et al. (2016) who found an AUC of 0.52. One possible explanation for the variation in results could be that the latter two studies both carried out assessment at prison intake while the positive results of the first study were for assessments carried out at the end of inpatient treatment just prior to release. Alternatively (or in addition), the SAPROF does not include protective factors that are relevant to sex offense specific risk factors. Of note, de Vries Robb  et al.'s sample contained few individuals with paraphilias, meaning the absence of protective factors relevant to sex offense specific risk factors would have held less relevance in that sample compared to the Yoon et al. and Turner et al. samples.

The Structured Assessment of Protective Factors – Sexual Offence version (SAPROF-SO) was developed in response to the perception that the original SAPROF was insufficiently tailored to issues specific to sexual offending. This instrument was developed from the original SAPROF but existing items were tailored for relevance to sexual offending and new items were written to tap protective factors that were considered specific to sexual offending (Willis et al., 2017). Items were initially grouped into five domains aligning with contemporary rehabilitation theory (namely the Good Lives Model; see Ward & Maruna, 2007) and desistance research (see Harris,

2021): Internal Capacity, Prosocial Identity, Prosocial Connection, Stability, and Professionally Provided Support. Preliminary findings from a factor analysis suggested two overarching factors, whereby items in the first four domains separated from items in the Professionally Provided Support domain and were labelled Personal protective factors (Thornton & Kelley, 2020). Willis et al. (2020) have reported research supporting the SAPROF-SO's interrater reliability and construct validity but its predictive validity relative to sexual recidivism had not been tested prior to the present study.

The Current Study

The current study continues the research program initiated by Willis et al. (2020) to develop and validate the SAPROF-SO for use in sexual recidivism risk assessment practice. The primary aim was to evaluate the predictive validity of the SAPROF-SO for sexual recidivism, utilising file information for men who participated in a New Zealand prison-based sexual offending treatment program between 1993 and 2000 and for whom 12-year recidivism data were available from earlier research (Beggs & Grace, 2010, 2011). Predictive validity for general and violent recidivism were also examined. Assuming predictive validity for sexual recidivism was supported, an additional aim was to assess incremental predictive validity over and above the Static-99R (Helmus, Thornton, et al., 2012). The degree to which the SAPROF-SO offers incremental prediction beyond that provided by Static-99R is of particular importance since the Static-99R is much easier to score and static actuarial instruments similar to Static-99R typically form a base component of any risk assessment.

It was hypothesised that SAPROF-SO scores would show predictive validity for sexual offending. In addition, it was hypothesised that SAPROF-SO Future context ratings (based on the proposed release environment) would show superior predictive validity to Current context ratings (prison), given Future context ratings may better reflect the environment in which desistance or recidivism occurred compared to Current context ratings.

Method

Participants

Participants were adult males who had participated in the Kia Marama treatment program for men who have sexually offended against children between 1993 and 2000. Attempts were made to include all 218 cases used in Beggs and Grace's (2010, 2011) sample. Of the 218 files requested from the New Zealand Department of Corrections, four files were in use and unavailable at the time of data collection, and four case files were irretrievable following the 2011 Christchurch earthquake, resulting in a sample size for the current study of $N = 210$. At the time of treatment commencement, men were aged between 18 and 74, with an average age of 41.2 years ($SD = 12$ years). Ethnically, 77.6% were of Pākehā (New Zealand European) descent, 20.0% were of

New Zealand Māori descent, and the remaining 2.4% were from other ethnic groups including those from the Pacific Islands. On program entry, participants consented to their file information being used for future research purposes. This research was reviewed and approved by the New Zealand Department of Corrections and the University of Canterbury Human Ethics committee.

Measures and Data Collection Procedures

SAPROF-SO – Pilot Version. The SAPROF-SO pilot version (Willis et al., 2017) is a 24-item clinician-rated tool that exclusively measures factors hypothesised to protect against sexual reoffending in adult biologically born males. The full list of items is presented in Table 1. In the current study, because the nature of the therapeutic alliance was not routinely documented in case files, the *Therapeutic Alliance* item was omitted for all cases. SAPROF-SO ratings are made based on all available information (e.g., behavioral observations, interview data, file review) and the SAPROF-SO manual stipulates that the coder should only consider information from within a specific time window, which varies from item to item. Typically, only information that pertains to the previous 6–12 months is analysed, but earlier information is considered when coding certain items that are thought to be more stable, including *Prosocial Sexual Interests* and *Adaptive Schemas* (Willis et al., 2020). Higher scores are indicative of stronger protection against sexual recidivism. Across all items, a score of 4 is assigned when the protective factor is *clearly present*, a score of 2 when it is *somewhat present* and a score of 0 when it is *not present*. These are referred to as anchor points. In the event the rater cannot decide between two anchor points, they are to assign a middle score of 1 or 3. If a rater cannot decide between a middle score and an anchor point, the anchor point is recorded. Selected items (*Motivation for Managing Risk*, *Sexual Offence-Specific Treatment*, *Medication*, and *Therapeutic Alliance*) have “not applicable” (N/A) or “not required” (N/R) response options along with numerical coding options.

Separate scores can be assigned for a Future context if the individual’s context is expected to change (e.g., release from prison). The first 12 items (previously comprising the Internal Capacity and Prosocial Identity domains) are thought to be reasonably stable across time and contexts, and the SAPROF-SO pilot manual therefore guides raters to copy Current context scores for these items over to the Future context column (unless the rater has reason to believe these may change). Future context scores are coded separately for the remaining items. In the present study, the current context was prison and the future context was the individual’s proposed release environment. Current context scores were copied over to Future context scores for the first 12 items and *Medication*, but were coded separately for the remaining 10 items (*Therapeutic Alliance* was omitted). Future context scores were based on documentation about post-release plans (which did not include information about *Medication*). One participant was convicted and sentenced to further prison time while participating in treatment, therefore, his Future context scores were informed by information pertaining to the prison he was being transferred to. In calculating Personal and Professionally Provided

Table 1. SAPROF-SO Descriptive Statistics.

SAPROF-SO Items and Domains	Current Context			Future Context		
	<i>n</i> valid ratings	Mean	SD	<i>n</i> valid ratings	Mean	SD
Personal						
1 Intact cognitive functioning	210	3.25	1.19			
2 Secure attachment in childhood	201	1.44	1.38			
3 Adaptive schemas	210	0.90	0.88			
4 Empathy	208	0.94	0.92			
5 Coping	210	1.25	0.95			
6 Self-control	210	1.44	1.17			
7 Sexual self-regulation	200	0.45	0.85			
8 Prosocial sexual interests	210	1.02	0.93			
9 Prosocial sexual identity	210	0.45	1.25			
10 Goal-directed living	207	0.70	0.87			
11 Motivation for managing risk	210	2.06	0.75			
12 Attitudes towards rules and regulations	210	1.85	1.16			
13 Work	206	0.35	0.77	205	0.64	1.02
14 Leisure activities	35	1.66	0.98	170	1.38	1.17
15 Social network	208	2.13	1.01	207	2.29	1.01
16 Emotional connection to adults	179	0.50	0.83	179	0.70	0.97
17 Intimate relationship	210	0.52	0.85	208	0.75	1.31
18 Housing stability	210	4.00	0.00	203	3.26	1.12
19 Financial management	209	2.01	0.83	206	2.18	1.09
Subtotal (possible range: 0–76)	210	26.91	8.99	210	27.00	10.10
		(Obtained range: 8.44–55.81)			(Obtained range: 6.00–55.94)	
Professionally provided support						
20 Sexual offence-specific treatment	210	4.00	0.00	206	3.79	0.78
21 Medication	210	0.18	0.78	210	0.18	0.78
22 Therapeutic alliance	a	a	a	a	a	a
23 Supervised living	210	4.00	0.00	203	1.48	0.76

(continued)

Table 1. (continued)

SAPROF-SO Items and Domains	Current Context			Future Context		
	<i>n</i> valid ratings	Mean	SD	<i>n</i> valid ratings	Mean	SD
24 External control	210	4.00	0.00	208	2.00	0.25
Subtotal (possible range: 0–16)	210 (Obtained range: 12–16)	12.18	0.78	210 (Obtained range: 0–12)	7.38	1.62
Total score (possible range: 0–92)	210 (Obtained range: 20.44–67.81)	39.08	8.94	210 (Obtained range: 11.40–63.94)	34.38	10.38

Note. SAPROF-SO = Structured Assessment of Protective Factors – Sexual Offence version (Willis et al., 2017). SAPROF-SO items are rated 0–4. Missing scores were replaced with case mean domain scores in calculation of subtotal and total scores. Not applicable ratings were replaced with scores of 0. For Sexual self-regulation scores, cases were deemed valid when at least two elements could be rated (without which a score above zero is not possible) and omitted elements replaced with 0 in the calculation of subtotal and total scores.

^aTherapeutic alliance was not coded in the current study due to a lack of relevant information.

Support domain scores and SAPROF-SO Total scores for each participant and context, omitted item scores were replaced with the mean domain score for that participant, and consistent with Willis et al. (2020), scores of “N/A” were replaced with 0.

Retrospective Scoring Guide. A Retrospective Scoring Guide was developed for the current study to supplement the SAPROF-SO pilot version coding manual (available in the Online Supplemental Materials). The SAPROF-SO was developed to align with strengths-based approaches to treatment; however, the period participants completed treatment predated strengths-based rehabilitation frameworks and information relevant to the SAPROF-SO items was therefore often difficult to locate. The Retrospective Scoring Guide was developed to aid coders in finding pertinent information, maximise interrater reliability, and to minimise item omissions. It did so through orienting coders to use information beyond that recommended in the manual when making decisions regarding the presence or absence of protective factors. For example, when necessary, the Guide allowed raters to consider information that predated the time window stipulated in the SAPROF-SO pilot manual when rating specific items, while giving more weight to recent functioning. For example, when coding the *Empathy* item, coders were to consider evidence of empathic behaviors that predated the stipulated time window (6 months), such as previous involvement in charity groups or caring for a vulnerable family member. In addition, the Guide oriented raters to consider specific psychometric test scores readily accessible in participants’ files when coding specific items; for example, specific subscale or item scores on the State-Trait Anger Expression Inventory (STAXI; Spielberger, 1988), State-Trait Anxiety Inventory (STAI; Spielberger, 1983), and Psychopathy

Checklist Revised (PCLR; Hare, 1991) were consulted when rating *Coping*. Minor deviations from the pilot manual were made for scoring *Self-control* and *Sexual Self-regulation* on the basis of information available on participants' files. Specifically, a global *Self-control* rating aligning with the SAPROF-SO response scale (0 = poor, 1 = below average, 2 = average, 3 = above average, 4 = good) was assigned, based on all relevant information (detailed in the Retrospective Scoring Guide). For the *Sexual Self-regulation* item, each of the four elements of sexual self-regulation were coded as either present, absent, or unclear/omit: (i) a lifestyle that deliberately avoids triggering offense-related impulses, (ii) strategies to negotiate high risk situations; (iii) offense-related sexual impulses arising rarely and are effectively interrupted, and (iv) healthy expression of sexual drive. Elements present were summed, and total scores aligned with the pilot manual scoring (see Willis et al., 2017): a score of four was given when all four elements were present and had been present for at least 12 months in the community before the individual was imprisoned (e.g., in cases of historic offending), a score of 3 was given when all four elements were present in prison, a score of 2 was given when three elements were present in prison, a score of 1 was given when 2 elements were present in prison, and a score of 0 was given when zero or one elements were present in prison. To ensure guidance provided in the Retrospective Scoring Guide was congruent with the intent of each item, the Guide was critiqued, amended, and approved by the SAPROF-SO authors.

SAPROF-SO Coding Procedure. SAPROF-SO training was provided to the first author (a postgraduate clinical psychology student) by the SAPROF-SO authors via webinar. Coding was informed by a comprehensive review of relevant material which included each participant's final psychological treatment report (written posttreatment and usually shortly before release from prison) and documentation predating the final report such as court documents, treatment exercises, and psychometric test results. The full list of documents that were examined are detailed in the Retrospective Scoring Guide (see [Online Supplemental Materials](#)). As detailed in the Retrospective Scoring Guide, coders were instructed to record evidence associated with each SAPROF-SO item as they reviewed relevant documents. The SAPROF-SO pilot manual and Retrospective Scoring Guide were then consulted to score each item.

A randomly selected 10.5% of cases ($n = 22$) were rated independently by both the first author and a SAPROF-SO author. Files were rated in batches of two or three; both raters then met to review scores, discuss discrepancies and resolve discrepancies by consensus scoring. The process of discussion and consensus scoring provided the first author with the opportunity to resolve misunderstandings and clarify scoring criteria which led to some minor changes to the Retrospective Scoring Guide. Remaining files were then coded by the first author alone. Coders were blind to all recidivism outcomes at the time of coding.

Static-99R. The Static-99R (Helmus, Thornton, et al., 2012) consists of 10 static risk items, and total scores range between -3 and 12 with higher scores indicative of higher

risk, as follows: very low risk (scores of -3 and -2), below average risk (scores of -1 or 0), average risk (scores of 1 , 2 , or 3), above average risk (scores of 4 or 5) and well above average risk (scores ≥ 6 ; Hanson et al., 2017). The Static-99R was identified as the most frequently used sexual recidivism risk assessment tool in Kelley et al.'s (2020) survey of predominantly US-based evaluators and treatment providers; it is used in at least 31 countries and has been translated into 10 languages (Helmus et al.). The Static-99R demonstrates a moderate ability to discriminate recidivists from nonrecidivists (Helmus, Hanson, et al., 2012). Static-99R scores for the current sample were calculated by applying age weightings (detailed in Helmus, Hanson, et al., 2012) to Static-99 scores previously rated and available for all cases. Raters were blind to Static-99 and Static-99R scores at the time of coding the SAPROF-SO. The mean Static-99R score for the current sample was $M = 1.70$ ($SD = 2.46$), which falls in the Average risk category.

Recidivism. Recidivism data used in the current study were extracted from Beggs and Grace's (2010) dataset, which they sourced from criminal history information maintained by the New Zealand Department of Corrections. Beggs and Grace (2010) recorded whether each participant received convictions for sexual, violent, or general offenses between release and the end of the follow-up period. As reported in their study:

Sexual recidivism was defined according to the Static-99 scoring criteria for Category "A" offenses (Harris et al., 2003), that is, an offense with an identifiable victim (e.g., incest, sexual assault, exhibitionism). Category "B" offenses (i.e., no identifiable victim) were excluded, except for possession of child pornography. Violent recidivism was recorded when the offender had been convicted for a non-sexual offense against a person (e.g., assault, robbery, kidnapping). General recidivism was recorded for offenses that were neither sexual nor violent (e.g., possession of cannabis; p. 239).

Time at large prior to each reconviction, or to the end of the follow-up period (July 1, 2008) was calculated. The average follow-up time was 12.24 years ($SD = 1.86$, range = 7.92–14.88 years).

Planned Analyses

ICCs were calculated to assess the interrater reliability of SAPROF-SO scores. ROC analyses were conducted to assess the validity of the SAPROF-SO and Static-99R in the prediction of sexual, violent, and general recidivism. Cox regression analyses were performed to evaluate whether the SAPROF-SO contributed incremental validity over and above Static-99R scores in the prediction of sexual recidivism. Statistical analyses were conducted with SPSS (Version 25.0).

Results

SAPROF-SO Descriptive Statistics

The number of valid ratings, along with the means and standard deviations for item, domain and Total scores are presented in [Table 1](#). The mean number of omitted items per participant was 1.57 ($SD = 0.83$) for Current context ratings (range = 0–4), and 1.06 ($SD = 1.35$) for Future context ratings (range = 0–9). Only three participants (1.4%) were missing more than three Current context scores and seven participants (3.3%) were missing more than three Future context scores. Missing scores were replaced with the individual's mean domain score in the calculation of subtotal and total scores. The most frequently omitted item was *Leisure Activities* which could only be scored for 35 participants (16.6%) for the Current context and 170 cases (81%) for the Future context. Sometimes there was insufficient information to rate all four sexual self-regulation elements; all four elements could be rated for 51% of cases and at least two elements could be rated for 95% of cases. In scoring *Sexual Self-regulation*, omitted elements were replaced with zero.

Across all participants, the mean Total score was 39.08 ($SD = 8.94$) for the Current context and 34.38 ($SD = 10.38$) for the Future context. Limited variability was observed across specific items, especially in the Professionally Provided Support domain given that all participants were receiving very similar Professionally Provided Support. For Current context ratings, all participants were assigned scores of 4 for *Housing Stability*, *Supervised Living*, *External Control*, and *Sexual Offence-Specific Treatment* because they were incarcerated and participating in the Kia Marama programme. Similarly, 97.1% of participants were due to be released on parole and subject to parole conditions, and therefore received a score of 2 for *External Control* for the Future context.

Interrater Reliability

Using the subset of files independently coded by two raters ($n = 22$), single-rater ICCs (two-way mixed, absolute agreement) were calculated. Decisions regarding the types of ICC analyses used and the interpretations of these statistics (e.g., poor, moderate, good, or excellent) were informed by the widely cited paper by [Koo and Li \(2016\)](#). The single-rater ICCs for the Total Current and Total Future scores were .98, 95% CI [.93, .99], and .98, 95% CI [.95, .99], respectively, indicating excellent interrater reliability. The single-rater ICCs for the Personal domain were .98, 95% CI [.93, .99] for Current context ratings and .98, 95% CI [.95, .99] for Future context ratings, indicating excellent interrater reliability. The ICC for the Professionally Provided Support domain for the Future context was .90, 95% CI [.77, .96], indicating excellent reliability (a Current context ICC could not be computed due to a lack of variance). Detailed interrater reliability results for items, domains and Total scores across both contexts are available in the Online Supplemental Material (see [Supplemental Tables S1 and S2](#)).

Predictive Validity of the SAPROF-SO

Over the follow-up period ($M = 12.24$ years; range = 7.92–14.88 years), 12.9% ($n = 27$) of participants received convictions for a new sexual offense, 12.9% ($n = 27$) for a new violent offense, and 36.7% ($n = 77$) for a new general (nonsexual and nonviolent) offense. Of cases that reoffended, the average time between release and reoffending was 4.03 years for sexual ($SD = 2.83$; range = 36 days–10.15 years), 4.07 years for violent ($SD = 2.67$; range = 33 days–11.69 years), and 4.48 years for general recidivism ($SD = 3.85$; range = 7 days–14.14 years).

Results of ROC analyses assessing the predictive validity of SAPROF-SO domain and Total scores (Current and Future) and the Static-99R for sexual, violent, and general recidivism are presented in Table 2. SAPROF-SO Total scores for both contexts were significantly predictive of sexual (AUC Current = .81, $p < .001$; AUC Future = .78, $p < .001$), violent (AUC Current = .66, $p < .01$; AUC Future = .65, $p < .05$), and general recidivism (AUC Current = .63, $p < .01$; AUC Future = .63, $p < .01$). A similar pattern of results to that obtained with the SAPROF-SO Total scores was observed for the predictive validity of the SAPROF-SO Personal domain whereas the Professionally Provided Support domain was not predictive of any type of recidivism. In the prediction of sexual recidivism, SAPROF-SO Personal domain and Total scores for the Current context produced marginally higher AUCs than Future context scores. There was almost no difference between Current and Future context ratings in the prediction of violent and general recidivism. In the prediction of sexual recidivism, AUCs for

Table 2. Predictive Validity of SAPROF-SO Scale Scores (Current and Future) and Static-99R Scores for Sexual, Violent, and General Recidivism.

Scale	Sexual Recidivism		Violent Recidivism		General Recidivism	
	AUC	AUC 95% CI	AUC	AUC 95% CI	AUC	AUC 95% CI
Personal – current	.81***	[.72,.91]	.66**	[.54,.78]	.63**	[.55,.71]
Personal - future	.80***	[.70,.90]	.65*	[.53,.78]	.63**	[.55,.71]
Professionally provided support – current	.45	[.32,.57]	.51	[.40,.63]	.51	[.43,.59]
Professionally provided support – future	.48	[.36,.60]	.45	[.35,.56]	.49	[.41,.57]
Total – current	.81***	[.72,.90]	.66**	[.54,.79]	.63**	[.56,.71]
Total - future	.78***	[.68,.87]	.65*	[.52,.77]	.63**	[.55,.70]
Static-99R	.74***	[.64,.85]	.66**	[.56,.77]	.68***	[.61,.76]

Note. SAPROF-SO = Structured Assessment of Protective Factors – Sexual Offence version (Willis et al., 2017). Static-99 R (Helmus, Thornton, et al., 2012). AUC = Area Under the Curve, CI = Confidence Interval. Missing scores were replaced with case mean domain scores in calculation of domain and total scores. Not applicable ratings were replaced with scores of 0.

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

SAPROF-SO Personal domain and Total scores were larger than for Static-99R scores (AUC = .74, $p < .001$), although confidence intervals were overlapping.

Incremental Predictive Validity of the SAPROF-SO

A series of hierarchical Cox regressions (proportional hazards model) were conducted to investigate whether and to what extent SAPROF-SO scales provided incremental validity in predicting sexual recidivism after controlling for Static-99R scores. Results of the Cox regression analyses are presented in Table 3. Static-99R scores were entered into the prediction model in Step 1 (Model 1) and SAPROF-SO Total and Personal scores for both Current and Future contexts were entered in Step 2, in four separate analyses (Models 2–5). Hierarchical results showed that SAPROF-SO Total scores for Current and Future contexts significantly predicted sexual recidivism after controlling for Static-99R scores and significantly improved the prediction model ($\Delta\chi^2(1) = 15.91$, $p < .001$, $\Delta\chi^2(1) = 12.99$, $p < .001$, respectively). Similarly, SAPROF-SO Personal scores for Current and Future contexts significantly predicted sexual recidivism after controlling for Static-99R scores, and significantly improved the prediction model ($\Delta\chi^2(1) = 17.25$, $p < .001$, $\Delta\chi^2(1) = 16.02$, $p < .001$, respectively). Static-99R scores retained significant unique predictive validity for sexual recidivism in each model. In summary, results supported the incremental predictive validity for sexual recidivism of the SAPROF-SO Total and Personal domain scores over and above the Static-99R.

Discussion

The current study represents the first evaluation of the predictive validity of the SAPROF-SO. Utilising a retrospective design and a sample of adult males convicted for sexual offenses against children, findings supported the predictive validity of the SAPROF-SO for sexual recidivism, with higher levels of protective factors associated with lower rates of sexual recidivism. The inverse relationship between protective factors and recidivism held after Static-99R scores were statistically controlled. Although the primary intent of the SAPROF-SO was to measure factors that are protective relative to sexual recidivism, some were expected to be protective relative to other kinds of recidivism. Indeed, SAPROF-SO scores were also inversely associated with violent and general recidivism. Results demonstrated clearly that the SAPROF-SO Personal domain predicted recidivism whereas the Professionally Provided Support domain was not predictive of recidivism. Such a finding is somewhat unsurprising considering that personal protective factors are relatively enduring over time whereas professionally provided protective factors are often time-limited (i.e., linked to an individual's prison sentence or parole/probation conditions). Indeed, as indicated by Future context scores for the External control item, the overwhelming majority of participants were subject to time-limited standard parole conditions on release which would have lapsed early in the follow-up period.

Table 3. Incremental Validity of SAPROF-SO Total and Personal Scores for both Contexts after Controlling for Static-99R Scores for Sexual Recidivism.

Model	Regression Coefficient				Exp(B)	Hazard Ratio 95% CI	Model Fit and Improvement	
	B	SE	Wald	Exp(B)			χ^2	$\Delta \chi^2 (1)$
Model 1								
Step 1								
Static-99R	0.34	.07	24.26***	1.41		[1.23,1.61]	25.42***	22.22***
Model 2								
Step 2								
Static-99R	0.19	.07	6.19*	1.20		[1.04,1.40]		
SAPROF-SO total current	-0.12	.03	14.01***	0.89		[0.83,0.94]		
Model 3								
Step 2								
Static-99R	0.23	.07	10.38**	1.26		[1.09,1.44]	36.16***	12.97***
SAPROF-SO total future	-0.09	.03	11.95**	0.92		[0.87,0.96]		
Model 4								
Step 2								
Static-99R	0.18	.07	5.70*	1.20		[1.03,1.38]	39.14***	17.25***
SAPROF-SO personal current	-0.13	.03	14.99***	0.88		[0.83,0.94]		
Model 5								
Step 2								
Static-99R	0.20	.07	7.70**	1.22		[1.06,1.41]	37.95***	16.02***
SAPROF-SO personal future	-0.11	.03	14.06***	0.90		[0.85,0.95]		

Note. SAPROF-SO = Structured Assessment of Protective Factors – Sexual Offence version (Willis et al., 2017). Static-99 R (Helmus, Thornton, et al., 2012). $\Delta \chi^2$ = change in chi square from previous step. CI = Confidence interval. Step 1 χ^2 df = 1. Step 2 χ^2 df = 2. Static-99 R scores were entered into the prediction model in Step 1 (Model 1), and SAPROF-SO Total and Personal scores for both Current and Future contexts were entered in Step 2, in four separate analyses (Models 2–5). * $p < .05$, ** $p < .01$, *** $p < .001$.

Contrary to hypotheses, the predictive validity for SAPROF-SO Future context ratings (based on the proposed release environment) were not superior to Current context ratings (prison). Rather, there was almost no difference in AUCs between Current and Future contexts across all three types of reoffending. Future context ratings (with the exception of one participant who remained in prison) were informed by postrelease plans pertaining to the upcoming 6–12 months, but for individuals who did reoffend, the average time between release and reoffending was over 4 years for each type of offending. Thus, Future context ratings may not have reflected the environment in which recidivism occurred. Also, the release plans on which Future context ratings were based may or may not have eventuated, whereas Current context scores were based on the individual's presentation in prison and observations from psychologists. Moreover, the potential for meaningful variation in Future context scores was limited, especially for the Personal domain, where different Future and Current context ratings were only possible for seven of the 19 items. Thus, it is possible the predictive validity of the Future context ratings was hampered by the reliability of the information used to inform them, and the greater number of items considered generally stable across contexts (e.g., *Empathy, Coping, Prosocial Sexual Interests*) relative to those that are more context dependent (e.g., *Work, Leisure Activities*).

AUCs for SAPROF-SO Total scores in the prediction of sexual recidivism were larger than AUCs reported for SAPROF Total scores in research described in the introduction (de Vries Robbé, de Vogel, Koster, et al., 2015; Turner et al., 2016; Yoon et al., 2018). The difference between AUCs may be explained by multiple factors, including sampling and methodological differences between studies. Alternatively (or additionally), the larger AUCs obtained in the current study may provide further support that the SAPROF-SO was successful in its intent to incorporate protective factors specific to sexual offending.

Obtained AUCs were equivalent to those found by Beggs and Grace (2010) in their retrospective validation of the VRS-SO using the same sample and recidivism data. They reported AUCs of .79 and .80 for VRS-SO pre and posttreatment Total scores, respectively. The equivalent AUCs suggest that the SAPROF-SO and VRS-SO may predict sexual recidivism with comparable accuracy in samples of men who have sexually offended against children. Given the strong correlations between the SAPROF-SO and VRS-SO change score found in Willis et al. (2020), and their findings suggesting that the SAPROF-SO was capturing variance independent of the VRS-SO change score, an important question for future research is whether that variance contributes incremental predictive validity over and above the VRS-SO. Irrespective of whether incremental predictive validity is demonstrated, there are inherent merits in assessing risk through a strengths-based lens. Briefly, risk assessment findings help inform treatment targets, and a focus on strengthening an individual's internal, social, and environmental resources is more engaging than a focus on reducing dynamic risk factors (e.g., Mann et al., 2004). The SAPROF-SO offers a strengths-based assessment tool that can be used to help identify treatment goals and inform ongoing risk management in ways that align with strengths-based approaches to treatment. For further

discussion of the clinical implications of the SAPROF-SO, we refer readers to [Kelley et al. \(2022\)](#).

Limitations and Implications for Future Research

Several limitations of the current study must be acknowledged. The retrospective design afforded the opportunity to efficiently assess the predictive validity of the SAPROF-SO over a long follow-up period; however, the choice of design meant that coders were limited to information found in archived case files. Future research utilising a prospective design whereby coders could interview participants face to face, would allow researchers to gather all of the required information to score the SAPROF-SO, which would minimise omissions and provide predictive validity results that would more closely emulate the accuracy that could be expected if the SAPROF-SO was used in practice. Being restricted to case files presented a considerable challenge in the current study because protective factors, being a newer concept to emerge in the field ([Rogers, 2000](#)), were not as commonly explicitly referred to in treatment documents for the sample as they are in more recent psychological practice. Using the Retrospective Scoring Guide was reasonably effective in minimising omissions and overcoming this limitation. However, the mean number of item omissions per participant was higher in the current study than what was found by [Willis et al. \(2020\)](#) in their high risk sample where scores were informed by far more recent file information (G. M. Willis, personal communication, February 18, 2021). In the present study, relevant information was particularly sparse in relation to *Leisure Activities*, leading to many item omissions. Further, information relating to the *Therapeutic Alliance* item was so minimal that this item was not included in the current study.

The use of the Retrospective Scoring Guide represents a threat to the external validity of the present study for two main reasons. First, it likely increased interrater reliability. Indeed, the Retrospective Scoring Guide was developed to maximise interrater reliability by providing guidance for scoring each item from archival records. Therefore, the reported ICCs likely represent an overestimation of reliability relative to what might be expected in clinical settings or in file-based studies that do not utilise a supplementary scoring guide. However, [Willis et al. \(2020\)](#) found good interrater reliability using retrospective files of more recent information without the use of a supplemental scoring guide. Second, the Retrospective Scoring Guide instructed coders to consider information not routinely considered when scoring the SAPROF-SO such as historical and psychometric information. Future studies that are prospective or at least utilise more recent file information (and therefore do not necessitate the use of a supplementary scoring guide), may yield more externally valid results. Also, the homogenous nature of the sample limits the generalisability of findings. The current sample was comprised of men imprisoned for child sexual offenses who had volunteered for treatment. Whether findings generalise to nontreatment samples, females with sexual offense convictions, individuals with convictions for sexual offenses against adults, subpopulations including individuals with a history of sexual offending

and major mental health problems or cognitive impairment, and individuals without official convictions is unknown and replication studies utilising heterogeneous samples are needed.

SAPROF-SO Total and Personal scores were significantly predictive of sexual reoffending after a long average follow-up time of 12.24 years, implying that protective factors in the SAPROF-SO may be relatively stable over long time periods. Future research is needed to assess the extent to which protective factors are malleable, including through intervention, and the relationship between SAPROF-SO changes and recidivism outcomes. Another important question for future research concerns whether Professionally Provided Support (which was time-limited in the current study) predicts reduced recidivism when professional sources of support or sentence conditions/case management extend for many years postrelease. Additional avenues for future research include (i) comparing psychometric properties for institutional versus community ratings, (ii) developing models for using the SAPROF-SO to help inform treatment goals and exploring client motivation and engagement in such an approach to treatment planning, and (iii) determine the training and supervised practice requirements needed to support reliable scoring of the SAPROF-SO.

Findings from the current study make a small but important step towards the SAPROF-SO being considered actuarial and meeting admissibility standards for use in risk assessment practice. In our view, for the SAPROF-SO to be considered “actuarial,” it will need to have empirically developed estimates of sexual offense recidivism rates associated with its scores. These estimates will need to be generated from multiple samples, which is currently a task in progress. Other tasks for ongoing validation of the SAPROF-SO include (i) examining relationships (including incremental predictive validity) with dynamic risk assessment tools and (ii) producing normative data to help interpret SAPROF-SO Total and domain scores. The SAPROF-SO can be currently used in practice as a structured professional judgment measure to guide assessment and clinical practice (see [Kelley et al., 2022](#)). Whether it can be considered admissible in a court setting will vary by jurisdiction. For example, in the USA, those states operating under the standard set by [Frye v. United States \(1923\)](#) emphasize scientific consensus to determine admissibility whereas states operating under [Daubert v. Merrell Dow Pharmaceuticals \(1993\)](#) require professional consensus as well as evidence of peer reviewed publications, validity and reliability, a known error rate, and a test manual. Thus, the current study may be sufficient for those using the SAPROF-SO in some settings, but more validity studies are necessary before other professionals are likely to begin utilizing it within court reports in other settings.

Conclusion

Results of the current study are consistent with the growing body of empirical research supporting the value of protective factors in assessing risk for recidivism (e.g., [Abbiati et al., 2017](#); [de Vries Robbé et al., 2011](#); [de Vries Robbé, de Vogel, Douglas, et al., 2015](#); [de Vries Robbé, de Vogel, Koster, et al., 2015](#); [Yoon et al., 2018](#)). The current

study is the first to both investigate and find evidence for the predictive validity of the SAPROF-SO. In summary, it was found that SAPROF-SO Total and Personal scores were significantly predictive of sexual, violent, and general offending after an average follow-up period of 12.24 years. Further, in the prediction of sexual recidivism, the obtained AUC values for Total and Personal scores were equivalent to or higher than values obtained from research evaluating well-validated and widely used assessment tools that rely on risk factors (e.g., Brankley et al., 2021; Helmus, Hanson, et al., 2012; Oliver et al., 2018). Throughout the history of risk assessment it has been common practice to focus only on negative aspects of individuals assessed, likely leading to unbalanced evaluations. Results from the current study together with the burgeoning research on protective factors support a rationale for attending to the positive.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by a Rutherford Discovery Fellowship awarded to GMW.

Author's Note

This article is based on the Master of Science thesis completed by Nolan (2021). GMW, DT, and SMK are authors of the SAPROF-SO and may benefit financially from providing trainings related to the instrument. We have no further known conflict of interest to disclose. The opinions are those of the authors and not necessarily those of the New Zealand Department of Corrections.

ORCID iDs

Gwenda M. Willis  <https://orcid.org/0000-0001-9827-3397>

David Thornton  <https://orcid.org/0000-0003-0127-1104>

Supplemental Material

Supplemental material for this article is available online.

References

- Abbiati, M., Azzola, A., Palix, J., Gasser, J., & Moulin, V. (2017). Validity and predictive accuracy of the structured assessment of protective factors for violence risk in criminal forensic evaluations. *Criminal Justice and Behavior*, 44(4), 493–510. <https://doi.org/10.1177/0093854816677565>

- Beggs, S. M., & Grace, R. C. (2010). Assessment of dynamic risk factors: An independent validation study of the violence risk scale: Sexual offender version. *Sexual Abuse, 22*(2), 234–251. <https://doi.org/10.1177/1079063210369014>
- Beggs, S. M., & Grace, R. C. (2011). Treatment gain for sexual offenders against children predicts reduced recidivism: A comparative validity study. *Journal of Consulting and Clinical Psychology, 79*(2), 182–192. <https://doi.org/10.1037/a0022900>
- Brankley, A. E., Babchishin, K. M., & Hanson, R. K. (2021). STABLE-2007 demonstrates predictive and incremental validity in assessing risk-relevant propensities for sexual offending: A meta-analysis. *Sexual Abuse, 33*(1), 34–62. <https://doi.org/10.1177/1079063219871572>
- Cording, J. R., & Beggs Christofferson, S. M. (2017). Theoretical and practical issues for the measurement of protective factors. *Aggression and Violent Behavior, 32*(January-February 2017), 45-54. <https://doi.org/10.1016/j.avb.2016.12.007>.
- 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. <https://doi.org/10.1037/pas0000807>
- Daubert, v Merrell Dow Pharmaceuticals, Inc (1993). *509 U.S. 579, 113 S. Ct. 2786, 125 L. Ed. 2d 469.*
- 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 ed.). Van der Hoeven Stichting.
- de Vries Robbé, M., de Vogel, V., & de Spa, E. (2011). Protective factors for violence risk in forensic psychiatric patients: A retrospective validation study of the SAPROF. *The International Journal of Forensic Mental Health, 10*(3), 178-186. <https://doi.org/10.1080/14999013.2011.600232>.
- 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. <https://doi.org/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, 27*(1), 51–70. <https://doi.org/10.1177/1079063214550168>
- de Vries Robbé, M., Mann, R. E., Maruna, S., & Thornton, D. (2015). An exploration of protective factors supporting desistance from sexual offending. *Sexual Abuse, 27*(1), 16–33. <https://doi.org/10.1177/1079063214547582>
- de Vries Robbé, M., Vivienne de, V., & Agnes, V. (2020). The structured assessment of protective factors for violence risk (SAPROF). In K. S. Douglas, & R. K. Otto (Eds), *Handbook of violence risk assessment* (pp. 410–437). Routledge. <https://doi.org/10.4324/9781315518374-21>
- de Vries Robbé, M., & Willis, G. M. (2017). Assessment of protective factors in clinical practice. *Aggression and Violent Behavior, 32*(January-February 2017), 55-63. <https://doi.org/10.1016/j.avb.2016.12.006>.
- Frye v. United States (1923). *293 F. 1013, 34 A.L.R. 145. D.C.Cir.*

- Hanson, R. K., Babchishin, K. M., Helmus, L. M., Thornton, D., & Phenix, A. (2017). Communicating the results of criterion referenced prediction measures: Risk categories for the Static-99R and Static-2002R sexual offender risk assessment tools. *Psychological Assessment, 29*(5), 582–597. <https://doi.org/10.1037/pas0000371>
- Hare, R. D. (1991). *The psychopathy checklist-revised*. Multi-Health Systems.
- Harris, D. A. (2021). Desistance from sexual offending. *Current Psychiatry Reports, 23*(2), 1–7. <https://doi.org/10.1007/s11920-020-01219-3>
- Harris, A., Phenix, A., Hanson, R. K., & Thornton, D. (2003). *STATIC-99 coding rules revised - 2003*.
- Helmus, L. M., Hanson, R. K., Thornton, D., Babchishin, K. M., & Harris, A. J. R. (2012). Absolute recidivism rates predicted by Static-99R and Static-2002R sex offender risk assessment tools vary across samples: A meta-analysis. *Criminal Justice and Behavior, 39*(9), 1148–1171. <https://doi.org/10.1177/0093854812443648>
- Helmus, L. M., Kelley, S. M., Frazier, A., Fernandez, Y. M., Lee, S. C., Rettenberger, M., & Boccaccini, M. T. (in press). Static-99R: Strengths, limitations, predictive accuracy meta-analysis, and legal admissibility review. *Psychology, Public Policy and Law*. <https://doi.org/10.1037/law0000351>.
- 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, 24*(1), 64–101. <https://doi.org/10.1177/1079063211409951>
- Kelley, S. M., Ambroziak, G., Thornton, D., & Barahal, R. M. (2020). How do professionals assess sexual recidivism risk? An updated survey of practices. *Sexual Abuse, 32*(1), 3–29. <https://doi.org/10.1177/1079063218800474>
- 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. <https://doi.org/10.1080/13552600.2021.2020345>.
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Mallion, J. S., Wood, J. L., & Mallion, A. (2020). Systematic review of ‘Good Lives’ assumptions and interventions. *Aggression and Violent Behavior, 55*(November-December 2020), 101510. <https://doi.org/10.1016/j.avb.2020.101510>.
- Mann, R. E., Webster, S. D., Schofield, C., & Marshall, W. L. (2004). Approach versus avoidance goals in relapse prevention with sexual offenders. *Sexual Abuse, 16*(1), 65–75. <https://doi.org/10.1177/107906320401600105>
- Olver, M. E., Mundt, J. C., Thornton, D., Beggs Christofferson, S. M., Kingston, D. A., Sowden, J. N., Nicholaichuk, T. P., Gordon, A., & Wong, S. C. P. (2018). Using the violence risk scale-sexual offense version in sexual violence risk assessments: Updated risk categories and recidivism estimates from a multisite sample of treated sexual offenders. *Psychological Assessment, 30*(7), 941–955. <https://doi.org/10.1037/pas0000538>
- Olver, M. E., Thornton, D., & Christofferson, S. M. B. (2021). 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/10.1177/10790632211002858>
- Rogers, R. (2000). The uncritical acceptance of risk assessment in forensic practice. *Law and Human Behavior*, 24(5), 595–605. <https://doi.org/10.1023/a:1005575113507>
- Spielberger, C. D. (1983). *State trait anxiety inventory*. Consulting Psychologists.
- Spielberger, C. D. (1988). *State-trait anger expression inventory (STAXI) professional manual*. Psychological Assessment Resources.
- Thornton, D. (2016). Structured risk assessment. In A. Phenix, & H. M. Hoberman (Eds.), *Sexual Offending: Predisposing Antecedents, Assessments and Management* (pp. 503–520). Springer. https://doi.org/10.1007/978-1-4939-2416-5_23
- Thornton, D., Hanson, R. K., Kelley, S. M., & Mundt, J. C. (2021). Estimating lifetime and residual risk for individuals who remain sexual offense free in the community: Practical applications. *Sexual Abuse*, 33(1), 3–33. <https://doi.org/10.1177/1079063219871573>
- Thornton, D., & Kelley, S. M. (2020) Developing protective factors responsive to different criminogenic need profiles. In Association for the Treatment Of Sexual Abusers (ATSA) Annual Research And Treatment Conference, 26–29 October 2020. https://www.atsa.com/Public/Conference/2020/ATSA_2020_Conference_Program.pdf
- Turner, D., Rettenberger, M., Yoon, D., Klein, V., Eher, R., & Briken, P. (2016). Risk assessment in child sexual abusers working with children. *Sexual Abuse*, 28(6), 572–596. <https://doi.org/10.1177/1079063214564390>
- Ward, T., & Beech, A. R. (2004). The etiology of risk: A preliminary model. *Sexual Abuse*, 16(4), 271–284. <https://doi.org/10.1177/107906320401600402>
- Ward, T., & Maruna, S. (2007). *Rehabilitation: Beyond the risk assessment paradigm*. Routledge.
- Ward, T., & Stewart, C. A. (2003). The treatment of sex offenders: Risk management and good lives. *Professional Psychology: Research and Practice*, 34(4), 353–360. <https://doi.org/10.1037/0735-7028.34.4.353>
- Willis, G. M., Kelley, S. M., & Thornton, D. (2020). Are protective factors valid constructs? Inter-rater reliability and construct validity of proposed protective factors against sexual reoffending. *Criminal Justice and Behavior*, 47(11), 1448–1467. <https://doi.org/10.1177/0093854820941039>
- Willis, G. M., Thornton, D. T., Kelley, S. M., & de Vries Robbé, M. (2017). *The structured assessment of protective factors - sexual offence version (SAPROF-SO) pilot manual*. [Unpublished manual].
- Yoon, D., Turner, D., Klein, V., Rettenberger, M., Eher, R., & Briken, P. (2018). Factors predicting desistance from reoffending: A validation study of the SAPROF in sexual offenders. *International Journal of Offender Therapy and Comparative Criminology*, 62(3), 697–716. <https://doi.org/10.1177/0306624x16664379>