



Kent Academic Repository

McEwan, Troy E., Slikboer, Reneta, Spivak, Benjamin, Fullam, Rachael, Darjee, Rajan, Papalia, Nina, Dunne, Ashley, Simmons, Melanie and Ogloff, James R. P. (2026) *A real-world validation of the Risk for Sexual Violence Protocol (RSVP)*. Psychology, Crime & Law . ISSN 1068-316X.

Downloaded from

<https://kar.kent.ac.uk/114560/> The University of Kent's Academic Repository KAR

The version of record is available from

<https://doi.org/10.1080/1068316x.2026.2660082>

This document version

Publisher pdf

DOI for this version

Licence for this version

CC BY (Attribution)

Additional information

Versions of research works

Versions of Record

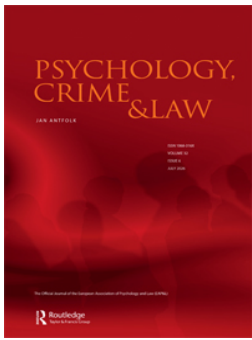
If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in **Title of Journal** , Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).



A real-world validation of the Risk for Sexual Violence Protocol (RSVP)

Troy McEwan, Reneta Slikboer, Benjamin Spivak, Rachael Fullam, Rajan Darjee, Nina Papalia, Ashley Dunne, Melanie Simmons & James R. P. Ogloff

To cite this article: Troy McEwan, Reneta Slikboer, Benjamin Spivak, Rachael Fullam, Rajan Darjee, Nina Papalia, Ashley Dunne, Melanie Simmons & James R. P. Ogloff (28 Apr 2026): A real-world validation of the Risk for Sexual Violence Protocol (RSVP), *Psychology, Crime & Law*, DOI: [10.1080/1068316X.2026.2660082](https://doi.org/10.1080/1068316X.2026.2660082)

To link to this article: <https://doi.org/10.1080/1068316X.2026.2660082>



© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Published online: 28 Apr 2026.



[Submit your article to this journal](#)



Article views: 843



[View related articles](#)



[View Crossmark data](#)

A real-world validation of the Risk for Sexual Violence Protocol (RSVP)

Troy McEwan^{a,b}, Reneta Slikboer^a, Benjamin Spivak^a, Rachael Fullam^a, Rajan Darjee^a, Nina Papalia^a, Ashley Dunne^a, Melanie Simmons^c and James R. P. Ogloff^a

^aCentre for Forensic Behavioural Science, Swinburne University of Technology and Forensicare, Melbourne, Australia; ^bCentre of Research and Education in Forensic Psychology, University of Kent, Canterbury, UK; ^cDepartment of Justice and Community Safety, Melbourne, Australia

ABSTRACT

This study assessed the concurrent and predictive validity of clinician judgements made using the Risk for Sexual Violence Protocol (RSVP) in routine practice, and examined how they used the guideline. Risk assessments were accessed from the files of 100 men referred to an Australian community forensic mental health service due to sexual violence. Predictive validity was examined at four time points based on police charges, though only 5 men reoffended during the follow-up period. RSVP Case Prioritization judgements were strongly correlated with the Static-99 and had excellent predictive validity at all follow-up times. Those judged High priority reoffended significantly more often than cases judged Moderate or Low priority. Generally, clinician judgements had better sensitivity than specificity, though the low number and rate of recidivism mean these findings must be interpreted cautiously and require replication. In practice clinicians rarely completed the Future Relevance ratings, Psychopathy risk factor, or the Serious Physical Harm and Immediate Action Required Summary Judgements. Aspects of the practice setting in which the RSVP was used may have helped to maintain predictive validity despite clinicians' departure from its instructions.

ARTICLE HISTORY

Received 12 November 2025
Accepted 12 March 2026

KEYWORDS

Sexual offending; risk assessment; offender management; structured professional judgement; sexual violence

Over the past two decades many guidelines and tools have been developed to aid assessment of the risk of future harmful sexual behaviour among men with previous sexual offending (Kelley et al., 2020). While differing in important ways, these tools and guidelines share the common goal that risk outcomes derived from the guideline or tool should be associated with relative differences in the probability of subsequent sexual offending by men classified into different risk groups (Singh et al., 2013). On this criterion, a 2009 meta-analysis suggested better predictive performance for actuarial tools, though

CONTACT Troy McEwan  tmcewan@swin.edu.au  Centre for Forensic Behavioural Science, Swinburne University of Technology, John St, Hawthorn, Victoria 3122, Australia

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/1068316X.2026.2660082>.

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

there were few published evaluations of structured professional judgement (SPJ) guidelines at the time (Hanson & Morton-Bourgon, 2009). Subsequent research has produced inconsistent results with some studies showing stronger performance for SPJ instruments, and others for actuarial (Hart & Boer, 2021; Wertz et al., 2023), with considerable variability in predictive performance of the SPJ guidelines across studies (Thornton, 2024). Despite this, advocates for SPJ guidelines argue that they are useful as they to help clinicians inform a detailed and individualized risk management plan (Hart et al., 2007). Nonetheless, demonstrating that clinician judgements about risk and case priority demonstrate predictive validity is essential for any instrument calling itself a risk assessment.

There are two SPJ guidelines specifically designed to structure assessment of sexual violence risk and risk management: the Sexual Violence Risk-20 (Boer et al., 1997) and the Risk for Sexual Violence Protocol (Hart et al., 2003), with a second edition of both being published in the past decade (Boer et al., 2017; Hart et al., 2022). The SVR-20 and RSVP are closely related, have good concurrent validity, and have been considered parallel forms of risk assessment (Hart & Boer, 2021). There is a sizeable body of research into the predictive validity of risk judgements made using the SVR-20, with studies showing medium to large effect sizes discriminating between men with and without subsequent sexual offending (e.g. Kanters et al., 2017; Rettenberger et al., 2011; Tsao & Chu, 2021). The SVR-20 has also demonstrated the ability to predict non-sexual violence (Blacker et al., 2011; Smid et al., 2014; Stadtland et al., 2005). The same is not true of the RSVP, with only three published validation studies (discussed further below). The RSVP is used by a sizeable minority of practitioners (Kelley et al., 2020) and although SVR-20 validation research provides some evidence-base, there is need for more research demonstrating that clinicians' conclusory opinions guided by the RSVP are both reliable between raters and meaningfully associated with differences in sexual recidivism.

The RSVP purpose and structure

Like all SPJ guidelines, the RSVP guides the qualified user through a series of steps to come to an overall opinion about the priority that should be given to risk management in the case. In the RSVP, sexual violence is defined as attempted, threatened or actual non-consensual sexual contact. The RSVP requires a relatively high degree of knowledge of the nature, causes and management of sexual violence, and training and expertise in the diagnoses of mental disorders. Additionally, a high level of skill in individual assessment, interviewing, information review and interpretation of standardized assessments is required to use the RSVP (Hart & Boer, 2010).

Assessors are instructed to rate the presence of 22 factors in the 12 months prior to assessment and over the lifetime. Each factor is, according to the authors, associated with various facets of risk for sexual violence including its 'nature, severity, imminence, frequency and likelihood'. Assessors can consider additional case-specific factors that are not already captured by the guidelines when forming judgements about risk and case priority (e.g. access to weapons, specific victim vulnerability, etc.). The potential relevance of each risk factor to future risk management strategies is also rated as yes, possible/partial or no. After rating the presence and relevance of all risk factors, the assessor is instructed to develop 'risk scenarios' that encapsulate plausible ways and contexts in

which a person may commit sexual violence in the future. Each scenario should consider the nature, seriousness, imminence, frequency and relative likelihood of the sexual violence. Risk management plans should then be developed to address each scenario and risk factors thought to play key causal roles. Finally, the assessor records a Case Prioritization judgement capturing the overall need for management to reduce sexual violence risk, and specific judgements about the risk for serious physical harm and whether there is perceived to be any need for immediate action. Each judgement may be 'Low', 'Moderate', or 'High' (Hart et al., 2003).

The reliability and validity of the RSVP

Interrater reliability of the RSVP has been reported in several studies. When case vignettes are used rather than information originating from clinical cases, intraclass correlations between multiple assessors tend to be lower for individual risk factors and summary judgements (e.g. correlations ranging from .43 to .69 in Sutherland et al., 2012). In contrast, in studies using clinical cases (file review and field data) the range in correlations are much stronger, .81 to 1.00 (Darjee et al., 2016; Sea & Hart, 2021; Vargen et al., 2020). Overall, the limited evidence indicates a satisfactory level of inter-rater reliability when used in real-world settings.

Concurrent validity of the RSVP with other risk assessments for sexual violence has been reported by Vargen and colleagues (2020) and Raymond and colleagues (2024). In the earlier study, RSVP total scores (lifetime presence and relevance ratings) were strongly and significantly correlated with total scores from the Static-99R (Phenix et al. 2016a) and the Sex Offender Risk Appraisal Guide (SORAG; Quinsey et al., 1998) ($r = .43-.68$). Similarly, RSVP Case Prioritization judgements were significantly and strongly correlated with Static-99R and SORAG categories ($r = .65-.66$). In Raymond and colleagues' Australian study, correlations between RSVP, Static-99 (Hanson & Thornton, 1999) and Static-99R total scores were smaller ($r = .32-.37$), but still significant. The association between RSVP Case Prioritization categories and Static-99 and 99R risk categories was smaller again ($r = .19-.21$), and not significant for the Static-99. As discussed previously, Raymond and colleagues demonstrated strong concurrent validity between the RSVP and SVR-20, with a .92 correlation between clinician judgements made using the two guidelines, and a .93 correlation between total scores.

To date, the predictive validity of the RSVP when used in its intended form has been evaluated in three studies, all of which used samples from community forensic settings. Two of these studies (Raymond et al., 2024; Vargen et al., 2020) used pseudo-prospective designs in which researchers completed the RSVP from information recorded at the time of the original assessment, then each case was followed up using administrative data capturing reoffending until the time of the research study. Vargen and colleagues operationalized recidivism as 'police contact', meaning convictions, charges, and allegations of sexual violence, while Raymond and colleagues used convictions only. Darjee et al. (2016), in contrast, used a truly prospective design in a sample of serious sexual offenders from Scotland. They drew on RSVP data recorded by clinicians in their practice and followed each assessed individual over an average of 3.75 years using any police contact for sexual violence, differentiating between any sexual offending and 'serious' sexual violence (contact sexual offences and sexual homicide).

Vargen and colleagues (2020) found that the RSVP Case Prioritization judgement was associated with increased odds of reoffending as the risk level increased, with a significant difference in odds of reoffending between each risk level. Conversely, Raymond and colleagues found that RSVP Case Prioritization judgements were not associated with significant differences in the odds of reoffending in their sample of older men who had used sexual violence (AUC = .54–.60 depending on follow-up time, all not significant). Darjee et al. (2016) obtained inconsistent results, concluding that only the Immediate Action Required judgement could significantly discriminate between those with and without serious sexual offending with a moderate to large effect size (AUC = 0.68), although higher Case Prioritization judgements were associated with quicker time to reoffending. Interestingly, while clinician judgements in Darjee and colleagues' study did not discriminate between those with and without serious sexual reoffending, total scores compiled from the Past ratings or Future Relevance ratings of risk factors did so with a strong effect (AUC = .71 and .72 respectively). Given the prospective nature of the study, these results were confounded by the fact that more risk management was directed towards those judged to be at higher risk of sexual reoffending. The difference in performance between the total score and summary judgements could indicate that the structured professional judgement element of the guidelines produced less predictively accurate risk assessment than did the total score. Alternatively, these results could be because those judged higher risk (regardless of total score) received more risk management, which was effective in reducing reoffending and so invalidated any predictive effect of the judgements. Darjee and colleagues concluded that their findings were more consistent with the latter hypothesis. Being undermanaged – i.e. those considered high risk but managed at a moderate level – tended to be associated with quicker and more frequent sexual offending in their samples whereas Low risk individuals who received low levels of management had low rates of reoffending.

A fourth study of the RSVP (Bengtson et al., 2025) created a composite RSVP from relevant risk factors included in the HCR-20 and SVR-20, excluding item 5 Psychological Coercion, which did not exist in the other guidelines. As the RSVP did not exist at the time the data was collected, no clinicians responsible for ratings were trained in the RSVP (though they were trained in the HCR-20 and SVR-20). Because they were coded for different guidelines, risk factors were rated only as 'ever present', and the methodology suggests that a total score rather than a clinical summary risk rating was used in predictive validity analyses. The sample of 96 men was followed for convictions for a further sexual offence, including use of child sexual exploitation material. Recidivism rates were 10.4% at 2 years, 16.7% at 5 years and 20.8% at 10 years. Total scores produced AUCs ranging between .64 (non-significant at 10 years) and .72 (significant at 5 years). Given the substantial departures from intended clinical use of the RSVP, it is difficult to know whether these results are comparable to those of the earlier studies, or how they would generalize to clinical practice.

The value of real-world investigation of SPJ guidelines

Darjee and colleagues' findings highlight the importance of exploring how clinicians use the RSVP in the real world to make case prioritization judgements and risk management decisions. In addition to providing insight into how clinicians use SPJ guidelines, real-world validations are important because it is possible that researcher-simulated judgements may not generalize into how clinicians apply the guidelines to their clinical practice

(Edens & Boccaccini, 2017). For example, the HCR-20 has been shown repeatedly to provide a predictively valid SPJ assessment of violence risk (Douglas & Shaffer, 2021). However, many studies do not examine the predictive validity of structured judgements about future violence, instead using total scores (see Challinor et al., 2021 for a review). Moreover, some authors have questioned whether the predictive validity of HCR-20 risk judgments made in carefully planned and conducted research trials generalizes to its use in routine practice by practitioners in the field (see Tully, 2017 for review). This does not mean that tightly controlled research into predictive validity is not essential, only that it is also important to evaluate performance in conditions that more closely reflect day-to-day clinical practice.

In the SPJ sexual offending risk assessment field, there are only two studies based on real world data. In addition to Darjee and colleagues' (2016) study described above, Rettenberger and colleagues' (2011) investigated SVR-20 ratings made by Austrian clinicians. While judgements produced good predictive validity, the authors reported that clinicians departed from the instructions in the guidelines and used total scores rather than risk categories to draw conclusions. Although real world or field trial designs are subject to several limitations (perhaps most importantly, the confounding effects of intervention based on risk judgments), research evaluating the predictive effects of clinician judgements when used in routine practice is necessary to understand how clinicians use the guidelines and whether they can achieve results consistent with those achieved under tightly controlled conditions.

Aims of the current study

The current study adds to the limited literature on the predictive validity of clinician judgements about the risk of sexual violence guided by the RSVP. The study addresses three aims. The first was to explore the concurrent validity of the RSVP with the Static-99 (Hanson & Thornton, 1999) and the second was to investigate the predictive validity of RSVP outcomes for sexual violence recidivism over time. The following questions were of interest in relation to the two aims:

- (a) What is the overall ability of Summary Judgements and RSVP total scores (Total Past, Total Recent and Total Lifetime) to discriminate between those who do and do not reoffend over time?
- (b) How effectively do Summary Judgements correctly classify individuals as reoffending or not reoffending over time?
- (c) How well does the Case Prioritization Summary Judgement predict future sexual violence over time?
- (d) Does Case Prioritization improve predictive validity over and above the RSVP Total Lifetime score?

The third aim was to examine how clinicians use the RSVP in a real-world setting. We intentionally chose to use the phrase 'real world' in this study rather than 'field trial' to make it clear that the data was drawn from clinicians' routine practice rather than during a controlled research trial conducted in a field setting. Specifically, the following questions were of interest:

- (a) What is the relationship between Past, Recent and Lifetime risk factors and clinicians' Case Prioritization Summary Judgements?
- (b) What is the relationship between risk factor total and domain total scores and clinicians' Case Prioritization Summary Judgements?
- (c) Are there any patterns evident in how clinicians use the RSVP that provide insight into its application in real-world settings?

Method

Study design

This study used a pseudo-prospective design. Forensic psychologists and/or psychiatrists employed at a state-wide community forensic mental health service in Victoria, Australia, used the RSVP in assessments conducted between 2012 and 2016. The service provided assessments based on the presence of past harmful behaviour (violence, harmful sexual behaviour, stalking, firesetting, etc.) and did not require a diagnosis or presentation of mental illness to accept referrals. The RSVP was the risk assessment guideline used in all cases where the presenting behaviour meets the RSVP's definition of sexual violence and recommendations about risk assessment and management are needed. Clinicians could choose to use additional risk assessment guidelines if they thought it necessary and helpful. Both assessments and the subsequent police record of individuals assessed were collected in 2019 and sexual violence recidivism was identified from the presence of charges for relevant offences during the follow-up period.

Procedure

All procedures were in accordance with the ethical standards of the Swinburne University of Technology (project 2019/069) and the Victorian Department of Justice (project CF1822498) Human Research Ethics Committees, and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. The files of every client presenting to the service between 24 January 2012 and 5 December 2016 ($n = 615$) were accessed by researchers with waiver of informed consent. Risk assessment, offending, and demographic information was extracted and coded. The name, sex and date of birth of each person assessed was linked to records held on the Victoria Police Law Enforcement Assistance Program (LEAP) data base using SOUNDEX matching. LEAP is a live database maintained by Victoria Police officers who are required to record the presence of a crime as soon as they are satisfied that, on the balance of probabilities, a crime has occurred.¹ Even though a charge is recorded in LEAP, it may not proceed to court, may be withdrawn at court, or may not result in a conviction. Five hundred and one of the original sample were able to be linked (96%). For each matched individual, all charges recorded by police between 1 June 1967 and 9 March 2020 were extracted from LEAP on 31 August 2023, with details of the date each alleged crime was committed. The same identifying information was provided to Corrections Victoria, the state correctional service, who identified individuals with periods of incarceration between 17 June 1997 and 12 February 2021 ($n = 264$).

Time-at-risk was calculated from the date the individual was assessed to 9 March 2020, controlling for periods of incarceration. Time-at-risk ranged from 1043 days to 2967 days (median = 1983 days, IQR = 845) but was truncated to 1460 days (4 years) to maintain maximum sample size to allow for meaningful analyses. Of the 100 individuals in the sample 21 had spent less than four years in the community during the follow-up period due to incarceration.

Sample

Of the 615 individuals assessed at the service in the collection period, sexual violence was relevant to the referral in 116 cases, meaning the RSVP was used by the assessing clinician. In the remaining 499 cases assessment was of risk of general violence, stalking, firesetting, etc., meaning the RSVP was not used and cases were excluded. Sixteen of the 116 sexual violence cases were also excluded, including one woman, 10 cases that did not have a Case Prioritization Summary Judgement and five cases for which all risk factor ratings were missing from the file, leaving 100 cases for analyses (see Supplementary Figures 1 and 2 for visual representation of missing data).

RSVP judgements were made by 21 clinicians. Clinical forensic psychologists were responsible for 84 assessments, forensic psychiatrists responsible for 4, and both a psychologist and psychiatrist involved in 12 assessments. All assessors were registered health professionals with forensic expertise who had completed training and supervised practice in the use of the RSVP. Of the 100 cases included in the study, clinicians chose to use both the RSVP and the Static-99 in 29 cases. In addition, 24 cases were assessed using the RSVP and another risk assessment instrument or guideline for non-sexual offending behaviour (one case with two additional risk instruments). In three cases this was the Level of Service/Risk-Need-Responsivity (Andrews et al., 2008), in eight the HCR-20 was used (Douglas et al., 2014), in two the Stalking Risk Profile (MacKenzie et al., 2009), and in 15 the Static-99R (Phenix et al., 2016b) was used. Static-99 data was used to assess concurrent validity in a non-random sample, while data from guidelines or tools other than the Static-99 are not reported in this study.

As a tertiary service, referrals came from community correctional services (i.e. probation services; 50%), general mental health services or practitioners (29%), child protection (12%), self-referrals (4%) and 1% was missing referral source. Of the 100 individuals in the sample, 88% were matched to police records and had no evidence of sexual reoffending. Of the remaining 12 cases, which were treated as not re-offending, four were not referred from a criminal justice source and so might not be expected to have a criminal record. Eight were referred by probation or court meaning a match would be expected and it is possible reoffending occurred but was not detected. Forty-one percent had a record of being incarcerated in Victoria prior to the assessment.

Measures

Risk for sexual violence protocol (RSVP)

The RSVP is a SPJ guideline intended to assist practitioners' judgements about the risk for future sexual violence among adult men with a history of sexual violence. The RSVP (Hart et al., 2003) comprises 22 risk factors divided into five domains: History of Sexual Violence

(five items); Psychological Adjustment (five items); Mental Disorder (five items); Social Adjustment (four items); and Manageability (three items) (see Table 1). Each item is rated twice (past and recent) according to a three-point ordinal scale (N = *no evidence*, ? = *possible/partial evidence*, and Y = *definite presence*). Past presence refers to the individual's life up until 12 months before the assessment while recent presence is the 12 months immediately prior to assessment. Factors can be omitted if there is insufficient information and case-specific factors can be added if the assessor deems it relevant.

Finally, the combination of presence and relevance of risk factors, risk scenarios and risk management considerations are used to inform conclusory opinions about the level of Case Prioritization, Risk of Serious Physical Harm and Need for Immediate Action. In this study, these opinions are referred to as Summary Judgements. Case Prioritization is judged as being Low/Routine, Moderate/Elevated or High/Urgent, and has very similar meaning to risk judgements in contemporaneous guidelines such as the HCR-20 version 2 (i.e. the higher the priority, the greater the need for immediate and more extensive intervention (Webster et al., 1997)). The Risk of Serious Physical Harm is made on a similar three-point scale while the Need for Immediate Action (e.g. immediate preventive steps) can be recorded using a 'yes' or 'no' categorization.

The raw data contained 16 cases (16%) in which intermediate Case Prioritization judgements were recorded by clinicians (e.g. low-moderate, moderate-high). These judgements were recoded to the higher adjacent category prior to analysis. This decision reflects the conceptual and applied meaning of such ratings in structured professional judgment (SPJ) approaches, where an in-between category indicates that risk is at least at the lower level but that there is credible concern it may fall within the higher category

Table 1. The five domains and twenty-two risk factors of the RSVP.

Domain and item number	Item description
Sexual violence history	
1	Chronicity of sexual violence
2	Diversity of sexual violence
3	Escalation of sexual violence
4	Physical coercion in sexual violence
5	Psychological coercion in sexual violence
Psychological adjustment	
6	Extreme minimization of denial of sexual violence
7	Attitudes that support or condone sexual violence
8	Problems with self-awareness
9	Problems with stress or coping
10	Problems resulting from child abuse
Mental disorder	
11	Sexual deviance
12	Psychopathic personality disorder
13	Major mental illness
14	Problems with substance use
15	Violent or suicidal ideation
Social adjustment	
16	Problems with intimate relationships
17	Problems with non-intimate relationships
18	Problems with employment
19	Non-sexual criminality
Manageability	
20	Problems with planning
21	Problems with treatment
22	Problems with supervision

(Hart et al., 2017; Douglas et al., 2014). Collapsing these ratings upward preserves the evaluator's expressed concern and avoids misclassifying cases with potentially elevated risk as lower-risk. This approach is also consistent with how risk categories are used in applied decision-making, where management strategies are typically guided by the highest plausible level of assessed risk rather than the lowest (Douglas & Skeem, 2005).

The summary judgements of 'Serious Physical Harm' and 'Need for Immediate Action' were rarely used, with 'Omit' responses recorded in 85% and 89% of cases respectively, and were therefore not considered in predictive validity analyses.

RSVP total scores were calculated by converting risk factor presence ratings to numerical scores ($N = 0$, $? = 1$, $Y = 2$) and tallying a total score based on the highest of either past or recent ratings for each factor. Total scores ranged from 0 to 44. Any missing item ratings were replaced with a '0' rating (i.e. no evidence). Future Relevance ratings i.e. how relevant the risk factor is to future offending were calculated following the same method. We did not investigate whether case specific risk factors were identified by clinicians.

Static-99

The Static-99 (Hanson & Thornton, 1999) is a 10-item actuarial risk assessment tool used to classify adult males with a history of contact sexual offending based on their probability of a further sexual offence. All items are scored based on historical events, meaning scores cannot reduce over time. Scores can range from 0 to 12, with one item scored 0–3 (history of sexual offences) and all others scored either 0 or 1. Item scores are totalled and converted into categories based upon sexual reoffending rates in the sample upon which the measure was developed. The risk categories range from Low (0 to 1), Moderate-Low (2 to 3), Moderate-High (4 to 5) and High (6+). An Australian validation study involving 621 men assessed at the same clinic as the current research showed that the Static-99 had good inter-rater reliability, a large discriminatory effect for sexual offending recidivism ($AUC = .67-.68$), and there were significant differences in relative risk between Low, Moderate and High risk categories over follow-up from 5 to 15 years (Reeves et al., 2018).

Sexual violence recidivism

Charges recorded in LEAP denoting sexual offending were those categorized within the ANZSOC Division 03 Sexual assault and related offences (Australian Bureau of Statistics, 2011) with Bestiality and Child Sexual Abuse Material charges removed for consistency with the RSVP sexual violence definition (see Supplementary Table 1 for details of charge names). Using this definition, 5 of the 100 (5%) participants had charges coded as sexual violence recidivism during the four-year follow-up period.

Analyses

All analyses were conducted with RStudio v2023.12, scripts were written in R-Project v4.3.2 and the compiler Rtools43 were used. Time at risk consisted of the number of days an individual remained in the community with days incarcerated removed. This means that each follow-up period used (i.e. 365 days [one year], 730 days [two years], 1095 days [three years] and 1460 days [four years]) does not represent a continuous

period of time if an individual spent some of those days incarcerated. One case had a missing value for RSVP assessment date. Rather than excluding the individual from the study, the date was estimated based on the average number of days between RSVP assessment and the final assessment report date in the sample (Tabachnick & Fidell, 2007, p. 66).

To investigate any potential bias arising from excluding cases, exploratory analyses (with uncorrected p values for multiple tests) compared cases that were included and excluded on demographics, Case Prioritization and data matching. A further check was conducted by re-running all analyses without cases in which the Case Prioritization Summary Judgements was amended (from five to three categories). Only minor changes to the results were found in response to this latter analysis, which did not influence interpretations, and we retained these cases in reported analyses.

Concurrent validity was assessed using a sub-sample of cases ($n = 29$) who had both RSVP and Static-99 risk assessments. Spearman's rho correlations were used to explore the relationship between the Static-99 total score and Static-99 risk categories with RSVP Case Prioritization Summary Judgement. Given the Static-99 was likely to have been applied non-randomly, a series of comparisons were conducted to determine whether cases where the Static-99 was applied differed meaningfully from those where it did not, allowing for clearer interpretation of IRR analyses.

When assessing the predictive validity of any SPJ guideline, the focus is usually on whether case prioritization or risk judgements are meaningfully associated with future behaviour. This is examined in various ways, the first being whether judgements meaningfully discriminate between those who do and do not reoffend, usually assessed using a receiver operating characteristic (ROC) curve. The area under the curve (AUC) is equivalent to the probability that a randomly selected person who reoffends will be assessed as higher risk than a randomly selected person who doesn't reoffend. The ability of Case Prioritization Summary Judgements to discriminate between those who sexually reoffend or not was evaluated at one, two, three, and four years post RSVP assessment using time-dependent AUC analyses and associated 95% confidence intervals (per Singh et al., 2014). Additionally, predictive validity was assessed by determining whether Case Prioritization Summary Judgements correctly classified sexual violence reoffending using sensitivity and specificity statistics over the four periods of follow-up time (95% confidence intervals calculated using bootstrapping). As these statistics require a 2×2 contingency table, classification accuracy was calculated for Low/Moderate versus High. The same metrics were used to assess predictive validity of the SPJ guidelines' total and domain scores. These scores cannot be used to communicate about risk in practice, though of course they underpin summary judgements given the presence of more risk factors should generally be associated with higher judgments across an entire sample.

To evaluate how well the Case Prioritization Summary Judgement predicted future sexual violence, a Kaplan-Meier survival curve was calculated, stratified by Case Prioritization level. The log of the negative log of survival was used to calculate the 95% confidence interval to avoid truncating the confidence intervals (those more than 100%) at the beginning of the survival curve. Log rank tests were performed to assess differences in the pattern of sexual violence reoffending between Low, Moderate and High Case Prioritization, with a Bonferroni corrected $\alpha = 0.016$ used for significance testing.

We intended to report positive and negative predictive values to ascertain whether elevated Case Prioritization judgements were associated with improvements in prediction of sexual violence recidivism over the base rate. We also intended to use Cox regression and hazard ratios to evaluate whether the judgements provided incremental improvements in prediction of sexual violence over RSVP total scores. However, given the low base rate of reoffending, results of these base rate-dependent analyses would be highly uncertain and we chose not to pursue them.

The small number of recidivists raised other concerns about the potential for uncertain results that could not be generalized to other samples. Sensitivity, specificity, and ROC AUC are not dependent on base rate of recidivism, however the small *number* of recidivists could lead to problems as a difference of one case is a sizeable proportional difference that could influence results. We managed this issue by reporting 95% confidence intervals for all statistics, providing information about the uncertainty around any effects given the sample size and number of recidivists. To further demonstrate the impact that the small number of recidivists could have on the results, we conducted a robustness analysis (also known as a sensitivity analysis, Greenland, 1996). This involved shifting the risk classification of one recidivist from a 'high' to a 'moderate' classification and reporting the impact on sensitivity, specificity and AUC in each analysis.

To address the third aim, descriptive statistics were calculated for all risk factor ratings, including the presence of Omitted factors. To explore which risk factors were most closely related to clinicians' Case Prioritization Summary Judgement, Kendall's Tau-c correlations were used. Sixty-six bivariate correlations were performed between the 22 risk factors for Past, Recent and Lifetime durations with Case Prioritization summary judgements. Kendall's Tau-c correlations were also used to explore the relationship between total scores (lifetime, Total Past and Total Recent), domain scores and Case Prioritization Summary Judgement. A Bonferroni correction was used for statistical significance cross all correlations ($\alpha = 0.0006$).

Results

Sample inclusion and exclusion

There were no statistically significant differences in age between those included ($n = 100$, $M = 39.49$, $SD = 13.29$) and excluded ($n = 16$, $M = 36.89$, $SD = 13.05$) from the sample ($t = -0.73692$, $df = 20.309$, $p = 0.470$). Fisher's exact tests indicated no statistically significant difference between those included and excluded on having a record of offending (Included 81%, Excluded 75%, $p = 0.195$), having a record of incarceration (Included 41%, Excluded 12%, $p = 0.106$), presence of a mental health, intellectual disability or acquired brain injury diagnosis (Included 61%, Excluded 81%, $p = 0.162$) or recommendation for treatment at PROGRAM (Included 48%, Excluded 25%, $p = 0.106$).

Descriptive information

The sample were aged between 18.36 and 79.92 years ($M = 39.49$, $SD = 13.29$). Most were born in Australia (86%), with the remaining 14% being born across 13 countries. Sixty-one men had a mental illness diagnosis (M number of diagnoses = 1.06, $SD = 1.14$) and nine had an identified intellectual disability or acquired brain injury. Thirty percent of the sample were in a

romantic relationship at the time of assessment. Four percent had completed an undergraduate degree, 17% completed high school, and 55% had completed less than grade 12. Seventy-nine men had a history of offending prior to assessment, with a wide range in the number and types of offending present: sexual violence (Median number of charges = 2, range = 0–33); child-victim sexual violence (Median = 1, range = 0–9); non-sexual physical violence (Median = 1, range = 0–19); all offending (Median = 8, range = 1–86).

The raw and corrected frequencies of RSVP Summary Judgements are presented in Table 2. Serious Physical Harm and Immediate Action Required Summary Judgements were very frequently omitted by the assessing clinicians. Based on the corrected Case Prioritization Summary Judgement, among those who reoffended, four were from the 'High' group (20% of that group), one was from the 'Low' group (2.7% of that group) and there were no reoffenders in the Moderate group.

As shown in Table 3, on average the sample tended to have higher total scores in the Past (prior to a year before assessment) compared to Recently (within the last year), with varied total scores across the five domains. The domain with the highest average lifetime score was Psychological Adjustment and the domain with fewest endorsed risk factors was Sexual Violence.

Concurrent validity between the RSVP and Static-99

Comparison of the 29 cases with a Static-99 with the remainder of the RSVP sample showed that there was no significant difference in the age of the two groups, though the former was more likely to have prior convictions for any sexual offending and/or

Table 2. Frequencies of clinician summary judgements and corrected summary judgements.

Judgement	Low	Low/moderate	Moderate	Moderate/high	High	Omit	Missing
Clinician judgement							
Serious physical harm	9	1	1	0	1	85	3
Immediate action	6	0	1	1	0	89	3
Case prioritization	37	10	33	6	14	0	NA
Corrected judgment							
Serious physical harm	9	–	2	–	1	85	3
Immediate action	6	–	1	–	1	89	3
Case prioritization	37	–	43	–	20	0	NA

Note: Numbers represent both frequencies and percentages as $n = 100$; NA = not applicable i.e. cases that did not have Case Prioritization Summary Judgement were removed from the study, dashes indicate where data has been recoded.

Table 3. Descriptives for risk factor total scores and domain total scores.

Calculated risk factor and domain total scores	Descriptives			
	<i>M</i>	<i>SD</i>	<i>Median</i>	<i>Range</i>
Risk factor total scores				
Total Past (sum of past items)	20.52	8.21	21	0–41
Total Recent (sum of present items)	14.78	7.78	14	1–35
Total Lifetime (only the highest of past or present summed; $n = 100$)	21.79	7.15	22	6–41
Domain total scores				
Sexual violence history	4.64	4.17	4	0–20
Psychological adjustment	10.15	3.93	10	0–20
Mental disorder	6.93	3.94	7	0–16
Social adjustment	8.60	3.75	9	0–16
Manageability	5.02	3.46	5	0–12

Note: Participants with any missing values for past and recent risk factor scores were removed $n = 96$.

child sexual offending (potentially reflecting this requirement of using the Static-99). There was no observable clinician effect on whether the Static-99 was used (i.e. particular clinicians were not choosing to use the Static-99 as part of their assessment), or time effect (i.e. the Static-99 wasn't used more often at different points in the data collection period). Those who received the Static-99 were more likely to have been referred by Child Protective Services than from other sources (though the Static-99 was used in cases from other sources too). Subsequent consultation with program staff suggested that this might reflect a program procedure that specified the approach to risk assessment in child protection referrals in a way that could have resulted in more frequent use of the Static-99. Results of these comparative analyses are available on request.

Total Static-99 scores ranged from 0–10 (median = 3, $q_1 = 1$, $q_3 = 7$). When categorized, 28% were Low, 24% were Moderate-Low, 14% were Moderate-High and 34% were in the High category. Four of the cases with a Static-99 had sexual violence recidivism; all were categorized as either High or Moderate-High risk. Both the Static-99 total score ($\rho = .61$, $p < .001$) and the Static-99 risk categorization ($\rho = .60$, $p < 0.0001$) were significantly and strongly correlated with the RSVP Case Prioritization Summary Judgement.

Predictive validity of RSVP case prioritization summary judgment and total score

Table 4 shows the ability of the Case Prioritization Summary Judgement to distinguish between those with and without subsequent sexual violence at various follow up

Table 4. AUC of sexual reoffending post assessment over four time points for case prioritization rating and risk factor totals.

Days at risk	AUC	95%CI
Case prioritization summary judgement		
365 (1 year)	0.90	0.87–0.94
	<i>0.91</i>	<i>0.87–0.94</i>
730 (2 years)	0.92	0.88–0.95
	<i>0.84</i>	<i>0.65–0.94</i>
1095 (3 years)	0.77	0.44–0.94
	<i>0.71</i>	<i>0.40–0.93</i>
1460 (4 years)	0.78	0.45–0.95
	<i>0.72</i>	<i>0.40–0.94</i>
Total past (sum of past items) $n = 96$		
365	0.82	0.73–0.88
730	0.89	0.78–1.00
1095	0.91	0.82–0.98
1460	0.92	0.83–0.99
Total recent (sum of present items) $n = 96$		
365	0.87	0.80–0.93
730	0.89	0.69–0.99
1095	0.73	0.32–0.99
1460	0.73	0.31–0.99
Total lifetime (only highest of past or present summed; $n = 100$)		
365	0.84	0.77–0.91
730	0.89	0.78–1.00
1095	0.90	0.81–0.98
1460	0.92	0.82–0.99

Note: Total Past and Recent $n = 96$ due to missing values. Results in *italics* are the results of analyses testing the robustness of results to changes in the data given the small number of recidivists. In this row, one recidivist was moved from the High to Moderate risk category.

times. There was a strong discriminatory effect at each point, though this became non-significant after two years (95%CI range incorporated .50). This finding was robust to a change in the classification of one recidivist from High to Moderate Case Prioritization. Total scores similarly showed a strong discriminatory effect at most time points. Notably, the ability of Total Past and Total Lifetime scores to predict sexual violence improved as days at risk increased, whereas the discriminatory performance of Total Recent scores reduced over time and, like the Case Prioritization Summary Judgements, became non-significant after two years.

The High Case Prioritization Summary Judgement demonstrated excellent sensitivity and specificity, correctly classifying all individuals who reoffended out to two years (see Table 5), then dropping slightly out to four years. Robustness testing showed that this result was somewhat unstable over time, with a dramatic reduction in sensitivity after 12 months, and wide confidence intervals around the estimates. There was no change in specificity over time (see Table 5).

Time to reoffending was between 257 and 933 days at risk ($M = 578$, $SD = 256.44$). Figure 1 shows the survival curves for the 100 participants categorized as Low ($n = 37$), Moderate ($n = 43$) and High ($n = 20$) case priority. A statistically significant difference in survival was found for those deemed High compared to Moderate priority ($\chi^2 = 9.3$, $df = 1$, $p = 0.002$), but not High compared to Low (log rank $\chi^2 = 5.1$, $df = 1$, $p = 0.02$) or Moderate compared to Low (log rank $\chi^2 = 1.2$, $df = 1$, $p = 0.3$). Sensitivity analyses showed no impact on the pattern of results.

The presence and use of risk factors in clinicians' practice

As shown in Table 6, the Past and Recent risk factors most frequently recorded as present were Stress and Intimate Relationships. Psychopathy and Physical Coercion were least frequently recorded as present across both periods. The Psychopathy risk factor attracted the most 'Omit' responses from clinicians; being rated only once in the sample of 100. Notably, Future Relevance ratings were omitted by clinicians in at least 80% of cases for all risk factors.

Correlations between risk factor scores (Past, Recent and Lifetime) and Case Prioritization Summary Judgements are displayed in Table 7. Significant positive correlations indicate that the presence of the risk factor was associated with a higher Case Prioritization

Table 5. Prediction related indices for sexual reoffending from case prioritization judgements at four time points, comparing high ($N = 20$) to moderate/low ($N = 80$) case prioritization.

Days at risk	Censored	Reoffended	Cutoff	Sensitivity % (95% CI)	Specificity % (95% CI)
365	0	1	High	100 (100–100) <i>100 (100–100)</i>	80.81 (72.73–88.00) <i>81.82 (73.99–89.69)</i>
730	0	4	High	100 (100–100) <i>75 (20.00–100)</i>	83.33 (75.25–90.83) <i>83.33 (75.25–90.83)</i>
1095	2	5	High	80 (32.29–100) <i>60 (0–100)</i>	82.80 (74.74–90.63) <i>82.80 (74.74–90.63)</i>
1460	21	5	High	80 (32.29–100) <i>60 (0–100)</i>	83.78 (74.65–92.19) <i>83.78 (74.65–92.19)</i>

Note: Results in *italics* are the results of analyses testing the robustness of results to changes in the data given the small number of recidivists. In this row, one recidivist was moved from the High to Moderate risk category. As this individual did not reoffend until the second year of follow-up, sensitivity and specificity did not change in year 1.

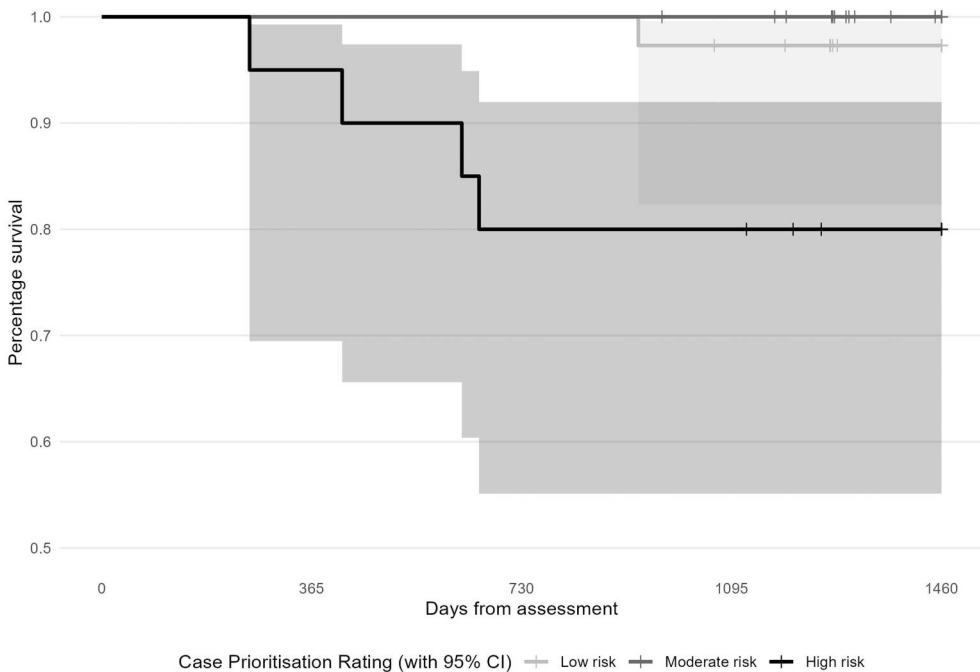


Figure 1. Kaplan–Meier survival curves by case prioritization summary judgement.

Summary Judgement. The correlations suggest a small to moderate positive relationship between Case Prioritization judgements and Past and Recent presence of Chronicity of Sexual Violence, Attitudes Condoning sexual violence, Sexual Deviance, and Problems with Planning and Treatment. Table 8 shows strong positive correlations between RSVP total scores (Past, Recent and Lifetime) and Case Prioritization Summary Judgements, indicating that the number of risk factors rated as present was strongly related to clinicians' judgements. Domain total scores were all correlated to a moderate degree, with the strongest correlations observed for Sexual Violence History and Manageability.

Discussion

The aims of this study were to evaluate the concurrent and predictive validity of Summary Judgements made by clinicians using the RSVP, and to explore how clinicians used the RSVP in routine practice, with a view to understanding the predictive validity of risk assessments guided by the RSVP in a real world environment.

Consistent with Vargen and colleagues (2020), RSVP Case Prioritization Summary Judgements were strongly correlated with Static-99 total scores and risk categories recorded by the same rater in a sub-sample of 29 cases. This sub-sample is not random, being more likely to have a formal sexual offending history, more likely to be referred from child protective services, and containing four of the five recidivists. This likely limits the generalizability of this result to those at higher risk of sexual violence. We also had no information about whether raters coded the Static-99 or RSVP first, though anecdotally clinicians reported that they typically completed the actuarial tool and then the SPJ guideline.

Table 6. Frequencies of past, recent and lifetime risk factor presence and future relevance for each risk factor.

Risk factor	Past presence				Recent presence				Lifetime presence				Future relevance			
	No	P/P	Yes	Omit	Miss	No	P/P	Yes	Omit	Miss	No	P/P	Yes	Omit	Miss	Miss
Chronicity	34	19	44	3	0	67	10	19	4	0	35	20	45	9	<3	<3
Diversity	55	21	19	≈4	<3	78	10	8	4	0	59	22	19	11	<3	≈3
Escalation	56	20	18	5	<3	82	7	5	6	0	59	22	19	12	0	<3
Physical coercion	72	8	15	≈4	<3	87	<3	5	≈6	0	76	8	16	11	0	<3
Psychological coercion	50	17	29	≈3	<3	81	8	4	7	0	53	18	29	9	<3	<3
Extreme minimization	33	24	35	≈5	<3	35	37	22	≈4	<3	32	31	37	5	<3	88
Attitudes	41	18	34	≈5	<3	54	26	16	≈4	<3	43	23	34	8	<3	<3
Self awareness	7	27	61	≈4	<3	16	38	43	<3	<3	9	28	63	4	<3	86
Stress	8	14	74	4	0	16	27	54	3	0	≈8	16	76	<3	4	≈3
Child abuse	51	16	28	5	0	56	24	16	4	0	55	17	28	14	<3	<3
Sexual deviance	35	11	46	5	3	42	17	34	≈5	<3	38	13	49	9	0	84
Psychopathy	74	6	<3	18	<3	77	5	<3	16	<3	93	≈6	<3	13	<3	≈3
Major mental illness	30	17	49	4	0	37	21	40	<3	<3	29	20	51	7	3	83
Substance use	45	12	41	<3	<3	73	11	16	0	0	45	13	42	10	<3	<3
Violent or suicidal	43	21	32	4	0	61	21	17	<3	<3	42	23	35	9	4	84
Intimate relationships	9	13	73	≈4	<3	17	26	54	<3	<3	7	14	79	<3	4	90
Non-intimate rel'ships	22	18	54	≈5	<3	21	29	48	<3	<3	18	19	63	3	<3	≈3
Employment	30	19	48	3	0	26	25	47	<3	<3	21	25	54	5	3	84
Non sexual criminality	50	23	22	5	0	83	9	5	3	0	52	25	23	11	<3	<3
Problem with planning	28	24	43	5	0	33	33	31	4	0	28	26	46	6	<3	87
Problem with t'ment	37	27	32	4	0	37	32	27	4	0	33	31	36	4	<3	≈3
Problem with s'vision	49	18	28	6	0	63	25	10	<3	<3	51	21	28	11	<3	<3

Note: as $N = 100$ the number represented is both count and percent. Cell sizes <3 but greater than 0 are obscured to maintain individual privacy, with approximations in neighbouring cells where necessary to avoid arithmetic identification of small cell size values. For lifetime risk factor frequencies, the highest of Past and Recent was summed, the Omit and missing values were recoded as 0 'no evidence'; Miss = missing data; P/P = Possible/Partial; Future relevance scores are presented here to explore how the RSVP is used in practice. Past presence = prior to one year before assessment; Recent presence = within one year from assessment.

Table 7. Correlation of case prioritization summary judgement with risk factor presence.

Risk factor	Past presence risk factors			Recent presence risk factors			Lifetime risk factors		
	<i>df</i>	τ_c	<i>p</i>	<i>df</i>	τ_c	<i>p</i>	<i>df</i>	τ_c	<i>p</i>
Chronicity	98	0.38	<0.001*	98	0.25	<0.001*	98	0.38	<0.001*
Diversity	97	0.36	<0.001*	98	0.13	0.050	98	0.36	<0.001*
Escalation	97	0.42	<0.001*	98	0.14	0.010	98	0.43	<0.001*
Physical coercion	97	0.02	0.812	98	0.02	0.646	98	0.02	0.763
Psychological coercion	97	0.15	0.092	98	0.06	0.266	98	0.15	0.089
Extreme minimization	96	0.14	0.102	96	0.19	0.021	98	0.17	0.054
Attitudes	96	0.36	<0.001*	97	0.39	<0.001*	98	0.37	<0.001*
Self awareness	97	0.12	0.114	97	0.23	0.006	98	0.12	0.103
Stress	98	0.01	0.935	98	0.17	0.036	98	0.02	0.813
Child abuse	98	0.13	0.099	98	0.13	0.122	98	0.13	0.091
Sexual deviance	95	0.28	<0.001*	96	0.39	<0.001*	98	0.28	<0.001*
Psychopathy	97	0.09	0.070	97	0.08	0.090	98	0.09	0.071
Major mental illness	98	0.14	0.104	98	0.19	0.030	98	0.16	0.068
Substance use	98	-0.01	0.882	98	0.08	0.271	98	0.01	0.937
Violent or suicidal	98	0.19	0.030	98	0.20	0.012	98	0.19	0.030
Intimate relationships	96	0.19	0.006	97	0.28	<0.001*	98	0.19	0.004
Non-intimate rel'ships	97	0.29	<0.001*	97	0.39	<0.001*	98	0.29	<0.001*
Employment	98	0.29	<0.001*	97	0.20	0.014	98	0.26	0.001
Non sexual criminality	98	0.07	0.407	98	0.03	0.533	98	0.10	0.229
Problem with planning	98	0.33	<0.001*	98	0.36	<0.001*	98	0.34	<0.001*
Problem with treatment	98	0.33	<0.001*	98	0.33	<0.001*	98	0.32	<0.001*
Problem with S'vision	98	0.27	0.002	98	0.18	0.017	98	0.33	<0.001*

Note: $N = 100$; $df = n-2$. Missing data in the Past and Recent risk factors remained in the data set, with four individuals missing 8, 3, 10 and 4 data points leading to differing degrees of freedom for each correlation between risk factor with Case Prioritization summary judgement. Lifetime risk factors are created by selecting the highest Past or Recent value, missing values were coded as 0. Case Prioritization was entered into the correlations as Low, Moderate or High.

*Significant.

Table 8. Correlations of case prioritization level with total score and domain total scores.

Risk factor totals	Correlation with case prioritization		
	<i>df</i>	τ_c	<i>p</i>
Total past (sum of past presence scores)	94	0.56	<0.001*
Total recent (sum of Recent presence scores)	94	0.54	<0.001*
Total lifetime	98	0.62	<0.001*
Domain totals			
Sexual violence history	94	0.47	<0.001*
Psychological adjustment	94	0.42	<0.001*
Mental disorder	94	0.40	<0.001*
Social adjustment	94	0.43	<0.001*
Manageability	94	0.46	<0.001*

Note: $N = 100$; $df = n-2$; cases with any missing values for Past and Recent risk factor scores were removed; all correlations $p < 0.001$; τ_c = Kendall's Tau-c; Total Lifetime included $n = 100$ is the highest of Past or Recent for each individual, summed.

*Significant.

Nonetheless, these results suggest that when rated by the same clinician in their practice in higher risk cases, the Static-99 and RSVP produce consistent results.

Consistent with prior Canadian and Scottish research, and contrary to the only previous Australian study of the RSVP that was limited to men aged over 50 (Raymond et al., 2024), Case Prioritization Summary Judgements showed a moderate to strong ability to discriminate between those with and without further sexual violence in the first two years post-assessment. Classification analyses showed that this was largely due to the sensitivity of

High priority judgements in this sample, which captured all but one of the participants who sexually reoffended. However, changing the classification of one recidivist to moderate resulted in a substantial drop in sensitivity, suggesting this strong classification finding may not generalize to a larger sample with more recidivists.

It is important to emphasize that while High Case Priority Summary Judgements did provide effective classification in this sample, only 20% of High priority cases actually reoffended over the four-year follow-up. This group was clearly at increased relative risk, given fewer than 5% of those in the Moderate or Low priority groups reoffended, but in absolute terms, sexual reoffending remained unlikely in the High priority group in this sample, even over four years. The low base rate of reoffending does not necessarily mean that the 80% of the High priority group who did not reoffend were not appropriately categorized. Non-offending individuals may have had a greater propensity to reoffend than those in lower priority groups but were either not identified as reoffending (i.e. the behaviour was present without charge), were not presented with circumstances conducive to reoffending, and/or received appropriate treatment and management to mitigate risk. In relation to this latter point, a proportion of the sample would have been on probation orders during some of the follow-up period, and others would have received forensic psychological treatment.

The overall pattern of results suggests that, despite substantial limitations associated with the small number of recidivists, clinician judgements informed by the RSVP do demonstrate predictive validity. Results are similar to Vargen and colleagues' (2020), despite the far lower rate of sexual violence recidivism (5% vs 29%). The higher rates of sexual reoffending across all priority categories in the Canadian study (62% of High priority cases, 32% of Moderate priority cases, and 17% of Low priority cases reoffended over nearly 10 years of follow-up) could reflect broader definitions of sexual violence recidivism, an overall higher risk sample, the pseudo-prospective research design, and/or less effective interventions.

Unfortunately, we were unable to address the question of whether Case Prioritization Summary Judgements provided incremental improvement in predictive validity over total scores, with the low base rate of sexual reoffending in the sample preventing regression analyses. Examination of the AUC values for RSVP total score, and the Past and Current scores suggested that a good level of discrimination could be achieved by applying the tool actuarially, broadly similar to that provided by the Case Prioritization Summary Judgement. Notably, these analyses showed that the Recent Total score achieved better discrimination between those with and without sexual recidivism in the first two years of follow-up (similar to the Case Prioritization Summary Judgement), while Past Total score achieved better discrimination in years three and four. This is consistent with the idea that historical risk factors provide an important indicator of longer-term risk, while the recent presence of risk factors is more indicative of increased risk in the short to medium term (even if those risk factors are not truly dynamic, such as a recent history of sexual violence).

The stated purpose of an SPJ approach to risk assessment is to help structure the assessor's thinking about why the person may present at higher risk at particular times, allowing for the fact that different risk factors may be more or less relevant for different people (Hart et al., 2022). Examination of the relationships between Past and Recent ratings of risk factors and Case Prioritization Summary Judgements suggested that the presence of a

more substantial history of sexual violence was associated with clinician perception that the individual was of higher case priority, as was historical evidence of sexual deviance and attitudes supportive of sexual violence. The two latter factors were also associated with High Case Prioritization Summary Judgements if they were present in the 12 months prior to assessment, as were recent relationship problems (intimate or non-intimate), and the Manageability domain items capturing problems with planning and treatment. Raymond and colleagues (2024) identified that the Problems with Planning risk factor was significantly associated with sexual recidivism in their sample of older Australian men (as was the Future Plans domain of the SVR-20). The Problems with Planning item has been removed from the RSVP-V2, incorporated into Problems with Stress and Coping and a new risk factor introduced (Problems with Living Circumstances). It will be interesting to see whether and how these changes influence clinicians use of the guidelines to make Summary Judgements, and the overall validity of the RSVP-V2.

In this sample, clinicians' judgments about the need for intervention to prevent sexual reoffending appeared to be broadly predictively valid (though with a high rate of false positives). However, it is essential to interpret these findings carefully given the small sample size and high level of uncertainty around the sensitivity of High Case Prioritization Judgements. Although all four individuals with further charges in the first two years post-assessment were identified as being of High Case Priority by clinicians (100% sensitivity), because there were only five recidivists, moving one into the Moderate Case Priority group led to a dramatic drop in sensitivity (to 75%) and a confidence interval ranging from 20%-100%. This means that the findings related to correct classification rate of recidivists should not be generalized until replicated in samples with larger numbers of recidivists and less uncertain results.

Analyses examining how clinicians used the RSVP in the real world showed that they did not adhere to all the user instructions. Clinicians typically did not use Future Relevance ratings to inform their Case Prioritization Summary Judgements, they routinely omitted the Psychopathy risk factor, and they did not record Summary Judgements about the potential for Serious Physical Violence or Immediate Action Required. It is unclear why these two Summary Judgements were not recorded by clinicians – possibly they did not find them helpful in their particular setting or there may have been some other reason that would require qualitative research to establish. The omission of the Future Relevance items is perhaps the most notable of the departures from the RSVP guidance as they are considered central to the purpose of an SPJ risk assessment, being viewed as a primary way of individualizing the assessment and allowing clinician judgement to be expressed (Douglas et al., 2014; Logan, 2016). It is unclear why clinicians did not rate the Future Relevance Items given the absence of contemporaneous qualitative data about clinicians' use of the RSVP. Given the very high rates of missing data, it seems that omitting these ratings was routine among the entire group of clinicians working in the program, rather than being the result of individuals occasionally not following instructions. This may be unique to this sample and setting, but future research should be alert to this pattern of missing data in research conducted in real-world settings.

It is notable that, despite the omission of Future Relevance ratings, clinicians' Case Prioritization Summary Judgements were still able to discriminate between those with and without further charges for sexual violence in the two years post-assessment, equivalent to discrimination achieved in previous studies using either file review or field trial

methodology (Darjee et al., 2016; Vargen et al., 2020). It is possible that the specific setting of this study could explain this result. The program from which data were collected uses a case formulation model of assessment (see McEwan & Darjee, 2021; Warren et al., 2005, for a program description). For every client, all available information is integrated into an explanatory narrative of past harmful sexual behaviour that draws on relevant theory to propose potential causal mechanisms for their problem behaviour (i.e. a formulation of the client's past sexual violence). This assessment approach is routine, regardless of which or whether an SPJ guideline is used. This formulation of past behaviour is then projected forward to inform judgements about the imminence and severity of future sexual violence and factors that could be associated with improvement in the individual's circumstances and behaviour. The formulation informs clinicians' risk management recommendations, which may in some cases include psychological treatment within the program.

It is possible that clinicians' departures from the highly structured approach outlined in the RSVP did not meaningfully affect the validity of their judgements because their standard practice means that they routinely work in ways consistent with how most current SPJ guidelines conceptualize the tasks of risk assessment. Future relevance of various risk factors was likely to have been considered implicitly while formulating, and imminence and serious harm may have featured in clinicians' consideration of future plausible events, even though the full set of conclusory judgements were not used. It may be that in a different context, with less emphasis on the importance of case formulation, departing from the RSVP instructions could undermine the validity of clinician judgements.

These findings do raise some unforeseen questions about the role of formulation in SPJ risk assessment. Over the past decade, many authors have argued that formulation is an integral part of SPJ risk assessment (see Tarpey et al., 2023, for review). This is reflected in current versions of SPJ guidance where clinicians are required to formulate before making risk judgements (Douglas et al., 2014; Hart et al., 2022; Kropp & Hart, 2015). However, the RSVP does not require clinicians to develop a formulation prior to making Case Prioritization judgements and it is unclear whether clinicians in this study made judgements before or after developing an explanatory formulation of the clients' behaviour. It is possible that the RSVP judgement performed well because the clinicians developed a formulation prior to determining case priority. Alternatively, it is possible that they formed their judgements based on the presence of risk factors, and possibly risk scenarios, before formulating. If the latter, this suggests that formulation may not be a necessary step in developing predictively valid risk judgements when using an SPJ approach. Good formulation would seem central to developing individualized risk management plans (McMurran & Bruford, 2016; Logan & Johnson, 2023; McEwan et al., 2023) and should be part of a thorough forensic assessment. However, whether developing a formulation prior has any effect (positive or negative) on the predictive validity of an SPJ risk judgement deserves further research.

There were several significant limitations associated with this study. The small sample size and limited number of recidivists means that there is substantial uncertainty about whether results would generalize to other samples. Moreover, eight cases were not matched to reoffending data where a match might be expected, and they were treated as not reoffending and retained in the sample due to the limited sample size. Given

these sample limitations, results of this study must be interpreted cautiously. Nonetheless, as only the third peer-reviewed study of the RSVP, they still add meaningfully to the literature on this under-evaluated SPJ guideline. A further limitation is the lack of inter-rater reliability data. The pseudo-prospective real-world research design meant that we were unable to gather contemporaneous inter-rater reliability data directly from the clinicians' interviews (as exemplified by Sea & Hart, 2021). While it is possible to calculate IRR based on file review ratings, we did not have access to trained research assistants who could complete such ratings. Moreover, it is questionable whether file-review ratings would be consistent with the aims of the study as a real-world investigation. A further limitation is the lack of risk scenario data, meaning we were unable to examine whether these were used by the clinicians and how risk scenarios might relate to Case Prioritization Summary Judgements and future sexual violence. Risk scenarios are a unique element of the RSVP that are not part of the parallel guideline, the SVR-20, and their contribution to the validity of risk judgements made using SPJ guidelines is an area in need of further research (Peikarnegar, 2021).

The potential impact of risk management interventions on reoffending in this pseudo-prospective study is a further limitation. All individuals were assessed in the community, however some would have been subject to probation orders and others would have received psychological treatment following the assessment to try to change a pattern of harmful sexual behaviour. The effects of these interventions on sexual recidivism is unknown and may be part of the reason for the low reoffending rate in this sample. Another limitation is the use of formal police charges to establish the presence of sexual recidivism. While the use of charges rather than convictions means that we potentially captured a wider range of sexual violence, it is possible that we included cases where the behaviour did not actually involve sexual violence. Police charges also obviously underestimate the true prevalence of harmful sexual behaviour in the community, with more than 90% of sexual violence in Australia going unreported to police (Australian Bureau of Statistics, 2021-22). Further, we only accessed data from Victoria Police, meaning any offending that occurred outside of this state was not included. A similar issue affects incarceration data, which was only available for Victoria, meaning those incarcerated in other states may not have been at risk for the full follow-up period.

As noted throughout this article, a new version of the RSVP was published in 2022. The RSVP-V2 differs from the original RSVP in several ways – there is no Future Relevance item for the five Nature of Sexual Violence risk factors, there is an additional risk factor (Sexual Health Problems), three RSVP risk factors have been replaced or amended (Suicidal/Violent Ideation, Nonsexual Criminality, Problems with Planning), an additional summary rating (Likelihood of Future Sexual Violence) has been added, and the Case Prioritization judgement is focused more on the perceived effort required to prevent sexual violence. There is also now an explicit Formulation step between making Future Relevance ratings and considering risk scenarios, reflecting other SPJ guidelines that share similar authorship. Given these differences between editions the findings of this study may or may not generalize to the RSVP-V2 and further research is required to investigate the validity and clinical utility of assessments using the more recent guideline.

This study suggests that Case Prioritization Summary Judgements informed by the RSVP can provide a predictively valid assessment of the future risk of sexual violence

over two years, when used by trained clinicians in their routine practice. This result requires replication in a larger sample with more reoffending, given the very low rates in this study, and the substantial limitations described above. Predictive validity was unaffected by clinicians' departure from aspects of the instructions in the RSVP guidelines. That may be due to specific characteristics of the program in which this research was conducted and highlights the important role of professional decision making and knowledge when using an SPJ approach to risk assessment. This study provides evidence that some clinicians do depart from the strictures of SPJ guidelines, meaning studies that investigate how and why this happens will be important, both to the development of future guidelines and to inform the content of user training.

Note

1. Further information on LEAP data recording practices can be found at <https://www.crimestatistics.vic.gov.au/about-the-data/explanatory-notes#Victoria%20Police%20recorded%20crime%20standards%20and%20practices>.

Acknowledgements

The authors wish to thank the Victorian Institute for Forensic Mental Health (Forensicare) Victoria Police and Corrections Victoria for assistance with data collection. The views expressed herein are solely those of the authors and do not necessary reflect the views or policies of Forensicare, Victoria Police or Corrections Victoria.

Disclosure statement

Rajan Darjee is a co-author of the RSVP-V2 and receives financial benefit from training facilitation. Other authors have no conflicts of interest to disclose.

Funding

This project was supported by the Victorian Department of Corrections and the Victorian Institute of Forensic Mental Health through the Catalyst Consortium.

References

- Andrews, D. A., Bonta, J., & Wormith, S. J. (2008). *The level of service/risk need responsivity inventory (LS/RNR): Scoring guide*. Multi-Health Systems.
- Australian Bureau of Statistics. (2011). *03 Sexual assault and related offences*. ABS. <https://www.abs.gov.au/statistics/classifications/australian-and-new-zealand-standard-offence-classification-anzsoc/2011/03>
- Australian Bureau of Statistics. (2021-22). *Sexual violence*. ABS. <https://www.abs.gov.au/statistics/people/crime-and-justice/sexual-violence/2021-22>
- Bengtson, S., Kristensen, E., Ibsen, M., Lund, J., & Långström, N. (2025). Predictive validity of SVR-20, RSVP, HCR-20, PCL-R, and clinician summary risk ratings among adult males convicted of sexual offending: Nationwide prospective field study. *Journal of Sexual Aggression*, <https://doi.org/10.1080/13552600.2025.2533172>
- Blacker, J., Beech, A. R., Wilcox, D. T., & Boer, D. P. (2011). The assessment of dynamic risk and recidivism in a sample of special needs sexual offenders. *Psychology, Crime & Law*, 17(1), 75–92. <https://doi.org/10.1080/10683160903392376>

- Boer, D. P., Hart, S., Kropp, P. R., & Webster, C. D. (2017). *Manual for version 2 of the Sexual Violence Risk-20: Structured professional judgment guidelines for assessing and managing risk of sexual violence*. Protect International Risk and Safety Services Inc.
- Boer, D. P., Hart, S. D., Kropp, P. R., & Webster, C. D. (1997). *Manual for the Sexual Violence Risk-20: Professional guidelines for assessing risk of sexual violence*. Mental Health, Law & Policy Institute, Simon Fraser University.
- Challinor, A., Ogundalu, A., McIntyre, J. C., & Bramwell, V. (2021). The empirical evidence base for the use of the HCR-20: A narrative review of study designs and transferability of results to clinical practice. *International Journal of Law and Psychiatry*, 78(September-October), 101729. <https://doi.org/10.1016/j.ijlp.2021.101729>
- Darjee, R., Russell, K., Forrest, L., Milton, E., Savoie, V., Baron, E., Kirkland, J., & Stobie, S. (2016). Risk for Sexual Violence Protocol (RSVP): A real world study of the reliability, validity and utility of a structured professional judgement instrument in the assessment and management of sexual offenders in South East Scotland. *Risk Management Authority and NHS Scotland Forensic Network*. https://www.rma.scot/wpcontent/uploads/2018/04/Risk_for_Sexual_Violence_Protocol_-_RSVP.pdf
- Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2014). *HCR-20V3: Assessing risk for violence: User guide*. Mental Health, Law, and Policy Institute, Simon Fraser University.
- Douglas, K. S., & Shaffer, C. S. (2021). The science of and practice with the HCR-20 V3 (Historical-Clinical-Risk Management-20, Version 3). In K. S. Douglas, P. Zapf, & R. K. Otto (Eds.), *Handbook of violence risk assessment*. (2nd ed., pp. 253–293). Routledge. <https://doi.org/10.4324/9781315518374-16>
- Douglas, K. S., & Skeem, J. L. (2005). Violence risk assessment: Getting specific about being dynamic. *Psychology, Public Policy, and Law*, 11(3), 347–383. <https://doi.org/10.1037/1076-8971.11.3.347>
- Edens, J. F., & Boccaccini, M. T. (2017). Taking forensic mental health assessment “out of the lab” and into “the real world”: Introduction to the special issue on the field utility of forensic assessment instruments and procedures. *Psychological Assessment*, 29(6), 599–610. <https://doi.org/10.1037/pas0000475>
- Greenland, S. (1996). Basic methods for sensitivity analysis of biases. *International Journal of Epidemiology*, 25(6), 1107–1116. <https://doi.org/10.1093/ije/25.6.1107>
- 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.
- Hanson, R. K., & Thornton, D. (1999). *Static-99: Improving actuarial risk assessments for sex offenders (User Report 99-02)*. Department of the Solicitor General of Canada.
- Hart, S., Kropp, P. R., Laws, D. R., Klaver, J., Logan, C., & Watt, K. A. (2003). *The Risk for Sexual Violence Protocol (RSVP): Structured professional guidelines for assessing risk of sexual violence*. Mental Health, Law, and Policy Institute, Simon Fraser University.
- Hart, S. D., & Boer, D. P. (2010). Structured professional judgment guidelines for sexual violence risk assessment: The sexual violence risk-20 (SVR-20) and risk for sexual violence protocol (RSVP). In R. K. Otto, & K. S. Douglas (Eds.), *Handbook of violence risk assessment* (pp. 269–294). Taylor & Francis Group.
- Hart, S. D., & Boer, D. P. (2021). Structured professional judgment guidelines for sexual violence risk assessment. The sexual violence risk-20 (SVR-20) versions 1 and 2 and risk for sexual violence protocol (RSVP). In K. S. Douglas, & R. K. Otto (Eds.), *Handbook of violence risk assessment* (pp. 312–358). Taylor and Francis.
- Hart, S. D., Douglas, K. S., & Guy, L. S. (2017). The structured professional judgment approach to violence risk assessment. In A. R. Felthous, & H. Saß (Eds.), *The Wiley handbook on the psychologies of violence* (pp. 643–666). Wiley.
- Hart, S. D., Kropp, P. R., Watt, K. A., Darjee, R., Davis, M. R., Laws, R. & Logan, C. (2022). *RSVP-V2 version 2 of the risk for sexual violence protocol*. Protect International.
- Hart, S. D., Michie, C., & Cooke, D. J. (2007). Precision of actuarial risk assessment instruments. Evaluating the ‘margins of error’ of group v. individual predictions of violence. *British Journal of Psychiatry*, 190, s60–s65. <https://doi.org/10.1192/bjp.190.5.s60>

- Kanters, T., Hornsveld, R. H. J., Nunes, K. L., Zwets, A. J., Muris, P., & Marle, H. J. C. V. (2017). The sexual violence risk-20: Factor structure and psychometric properties. *The Journal of Forensic Psychiatry & Psychology*, 28(3), 368–387. <https://doi.org/10.1080/14789949.2017.1284887>
- 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>
- Kropp, P. R., & Hart, S. D. (2015). *SARA-V3: User manual for version 3 of the spousal assault risk assessment guide*. ProActive Resolutions.
- Logan, C. (2016). Risk formulation: The new frontier in risk assessment and management. In D. Richard Laws, & W. O'Donohue (Eds.), *Treatment of sex offenders strengths and weaknesses in assessment and intervention* (pp. 83–105). Springer International Publishing. <https://doi.org/10.1007/978-3-319-25868-3>
- Logan, C., & Johnstone, L. (2023). *Managing clinical risk: A guide to effective practice* (2nd ed.). Routledge.
- MacKenzie, R. D., McEwan, T. E., Pathé, M. T., James, D. V., Ogloff, J. R. P., & Mullen, P. E. (2009). *Stalking risk profile: Guidelines for the assessment and management of stalkers*. StalkInc. & Monash University.
- McEwan, T., Galletta, M., & Underwood, A. (2023). *Treating stalking: A practical guide for clinicians*. John Wiley & Sons.
- McEwan, T. E., & Darjee, R. (2021). The problem behaviour program: Threat assessment and management in a community forensic mental health context. In J. R. Meloy, & J. Hoffmann (Eds.), *International handbook of threat assessment* (pp. 536–553). Oxford University Press. <https://doi.org/10.1093/med-psych/9780190940164.003.0031>
- McMurran, M., & Bruford, S. (2016). Case formulation quality checklist: A revision based upon clinicians' views. *Journal of Forensic Practice*, 18(1), 31–38. <https://doi.org/10.1108/JFP-05-2015-0027>
- Peikarnegar, M. (2021). *A pilot investigation of the utility of case formulation and scenario planning in structured professional judgment using the Spousal Assault Risk Assessment Guide–Version 3 (SARA-V3)*. [Master's thesis, Simon Fraser University]. Summit Research Repository. https://summit.sfu.ca/_flysystem/fedora/2022-08/input_data/21612/etd21545.pdf
- Phenix, A., Fernandez, Y., Harris, A. J. R., Helmus, L., Hanson, R. K., & Thornton, D. (2016a). *Static-99R coding rules revised – 2016*. Public Safety Canada. https://www.publicsafety.gc.ca/cnt/rsrscs/pb_lctns/sttc-2016/index-en.aspx
- Phenix, A., Helmus, L. M., & Hanson, R. K. (2016b). *Static-99R & Static-2002R evaluators' workbook*. Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (1998). *Violent offenders: Appraising and managing risk*. American Psychological Association. <https://doi.org/10.1037/10304-000>
- Raymond, B., Davis, M. R., Ogloff, J. R. P., & McEwan, T. E. (2024). Convergent and predictive validity of the Sexual Violence Risk-20 and Risk for Sexual Violence Protocol with older sexual offenders. *Psychology, Crime & Law*, 1–22. Advance online publication. <https://doi.org/10.1080/1068316X.2024.2401489>
- Reeves, S. G., Ogloff, J. R. P., & Simmons, M. (2018). The predictive validity of the Static-99, Static-99R, and Static-2002/R: Which one to use? *Sexual Abuse*, 30(8), 887–907. <https://doi.org/10.1177/1079063217712216>
- Rettenberger, M., Boer, D. P., & Eher, R. (2011). The predictive accuracy of risk factors in the sexual violence risk-20 (SVR-20). *Criminal Justice and Behavior*, 38(10), 1009–1027. <https://doi.org/10.1177/0093854811416908>
- Sea, J., & Hart, S. D. (2021). Interrater reliability and concurrent validity of the risk for sexual violence protocol for Korean sexual offenders: A field study. *International Journal of Offender Therapy and Comparative Criminology*, 65(13-14), 1423–1445. <https://doi.org/10.1177/0306624X20967969>
- Singh, J. P., Desmarais, S. L., & Dorn, R. A. V. (2013). Measurement of predictive validity in violence risk assessment studies: A second-order systematic review. *Behavioral Sciences and the Law*, 31(1), 55–73. <https://doi.org/10.1002/bsl.2053>
- Singh, J. P., Yang, S., & Mulvey, E. P. (2014). Reporting guidance for violence risk assessment predictive validity studies: The RAGEE statement. *Law and Human Behavior*, 38(1), 15–22. <https://doi.org/10.1037/lhb0000034>

- Smid, W., Kamphuis, J. H., Wever, E. C., & Beek, D. J. V. (2014). A comparison of the predictive properties of nine sex offender risk assessment instruments. *Psychological Assessment*, 26(3), 691–703. <https://doi.org/10.1037/a0036616>
- Stadtland, C., Hollweg, M., Kleindienst, N., Dietl, J., Reich, U., & Nedopil, N. (2005). Risk assessment and prediction of violent and sexual recidivism in sex offenders: Long-term predictive validity of four risk assessment instruments. *Journal of Forensic Psychiatry*, 16(1), 92–108. <https://doi.org/10.1080/1478994042000270247>
- Sutherland, A., Johnstone, L., Davidson, K. M., Hart, S. D., Cooke, D. J., Kropp, R., Logan, C., Michie, C., & Stocks, R. (2012). Sexual violence risk assessment: an investigation of the interrater reliability of professional judgments made using the risk for sexual violence protocol. *International Journal of Forensic Mental Health*, 11(2), 119–133. <https://doi.org/10.1080/14999013.2012.690020>
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Pearson Education, Inc.
- Tarpey, E., Stephenson, Z., & Barker, R. (2023). Risk formulation in forensic practice: A review of the evidence. *Journal of Forensic Practice*, 25(3), 243–262. <https://doi.org/10.1108/JFP-12-2022-0064>
- Thornton, D. (2024). What works in assessing recidivism risk in men convicted of contact sexual offenses. In I. L. A. Craig, L. Dixon, & T. A. Gannon (Eds.), *The Wiley handbook of what works in correctional rehabilitation: An evidence-based approach to theory, assessment and treatment* (pp. 93–104). Wiley-Blackwell.
- Tsao, I. T., & Chu, C. M. (2021). An exploratory study of recidivism risk assessment instruments for individuals convicted of sexual offenses in Singapore. *Sexual Abuse*, 33(2), 157–175. <https://doi.org/10.1177/1079063219884575>
- Tully, J. (2017). HCR-20 shows poor field validity in clinical forensic psychiatry settings. *Evidence Based Mental Health*, 20(3), 95–96. <https://doi.org/10.1136/eb-2017-102745>
- Vargen, L. M., Jackson, K. J., & Hart, S. D. (2020). Interrater reliability, concurrent validity, and predictive validity of the Risk for Sexual Violence Protocol. *Law and Human Behavior*, 44(1), 37–50. <https://doi.org/10.1037/lhb0000353>
- Warren, L. J., MacKenzie, R., Mullen, P. E., & Ogloff, J. R. P. (2005). The problem behavior model: The development of a stalkers clinic and a threateners clinic. *Behavioral Sciences & the Law*, 23(3), 387–397. <https://doi.org/10.1002/bsl.593>
- Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20: Assessing risk for violence (version 2)*. Simon Fraser University.
- Wertz, M., Schobel, S., Schiltz, K., & Rettenberger, M. (2023). A comparison of the predictive accuracy of structured and unstructured risk assessment methods for the prediction of recidivism in individuals convicted of sexual and violent offense. *Psychological Assessment*, 35(2), 152–164. <https://doi.org/10.1037/pas0001192>