



THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e. g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Risk Assessing Sexual Deviance, including:

**The Predictive Validity of the Violence Risk Scale for Sexual
Offenders (VRS-SO): A Systematic Review**

and

**Understanding Sexual Deviance in Risk Assessment Practice: A
Criminal Justice Social Worker Perspective**

Claire Forbes

Doctorate of Clinical Psychology

The University of Edinburgh

August 2022

Acknowledgements

This thesis is the culmination of more than a 3-year doctorate. It is the final step of a 12-year journey towards a career which I feel gives me so much more than it takes. I am fortunate enough to have been supported by innumerable people along the way and no acknowledgement section will ever be enough to do justice to the extent of my gratitude. Nonetheless I shall try...

I'd like to thank those who participated in this research, for their time, enthusiasm, and openness to the process. I'd also like to thank my supervisors, Dr David Gillanders and Dr Louise Tansey, for staying solution-focused and helping guide the way out of periods of uncertainty. You have helped foster my confidence, as well as my competence. Research is a less scary place now which I am endlessly grateful for.

I'd like to thank my second-reviewer Alice Marten. Both for her time and thoughtfulness in her contribution to my thesis and for her selflessness in housing me during her own thesis turmoil. Thank you for your honesty in our friendship and for being my Susan.

I'd like to thank Pernille and Hannah for leading the way and helping me see the light at the end of the tunnel. For taking the lessons they learned, sharing them with me and offering endless compassion and encouragement. You continue to empower and inspire me in all facets of life. I'd also like to thank Sarah for being in this last stretch with me. I am selfishly so glad we were in the same boat. I don't think I could have done this without you. Thank you for helping me to laugh as well as cry.

I'd like to thank Toni for keeping me fed and Léni for keeping me grounded. You helped me to keep perspective, to connect with the present moment and most importantly, to connect to myself, when things were hard. In a similar vein, thank you to Skye and Quinn for bringing joy into my days and reminding me of life outside of psychology.

I'd like to thank my parents, who have encouraged me from a young age that I could do anything. Regardless of circumstances, you have never doubted my capabilities. You have never shied away from shouldering any burden in order to make my life easier and to help me keep my eye on the prize.

Finally, to all my friends who have supported me through the biggest and most turbulent year of my life. Thank you to Rach, Sean, Rocky, Stu, Eidann, Hannah, Martin, Spratley, Maisie, Steve, Iris, Briony, Sumner, Woody, Anna and everyone mentioned above, for letting me draw on your support when I didn't have the capacity to return the favor. At a time when I felt the ground was continuously shifting under my feet, you made me feel secure enough to put life on hold and to take care of myself in order to get this done. To be able to trust these friendships to endure has been the most incredible gift. Each of you has shaped me, the person I am and the person I want to be. In submitting this thesis, I am excited to be able to return to reality and give back a little of what I have got. Sincerely, I thank you.

Table of Contents

| | |
|--|-----------|
| Lay Summary | 1 |
| Thesis Portfolio Abstract | 3 |
| The Predictive Validity of the Violence Risk Scale for Sexual Offenders (VRS-SO): A Systematic Review | 4 |
| Abstract..... | 4 |
| Introduction..... | 5 |
| <i>Developing and validating risk assessment measures</i> | 6 |
| <i>The Violence Risk Scale – Sexual Offender Version (VRS-SO)</i> | 7 |
| <i>Possible Clinical Contributions</i> | 8 |
| Methods | 10 |
| <i>Search Strategy</i> | 10 |
| <i>Eligibility</i> | 11 |
| <i>Data Extraction</i> | 12 |
| <i>Synthesis of Results</i> | 12 |
| <i>Risk of Bias</i> | 13 |
| <i>Inter-rater Reliability</i> | 14 |
| Results | 15 |
| <i>Study Characteristics</i> | 15 |
| <i>Risk of Bias in Studies</i> | 22 |
| <i>What is the predictive validity for the VRS-SO Static Total Score?</i> | 27 |
| <i>What is the predictive validity for the VRS-SO Dynamic Total Score?</i> | 28 |
| <i>What is the predictive validity for the VRS-SO Dynamic Domain Scores?</i> | 29 |
| <i>What is the predictive validity for the VRS-SO Overall Scores?</i> | 32 |
| Discussion..... | 33 |
| <i>Summary of Results</i> | 33 |
| <i>Study Strengths and Limitations</i> | 35 |
| <i>Review Strength and Limitations</i> | 38 |
| Conclusion | 39 |
| References..... | 40 |
| Understanding Sexual Deviance in Risk Assessment Practice: A Criminal Justice Social Workers Perspective..... | 46 |
| Abstract..... | 46 |
| Introduction..... | 47 |
| <i>“What” are the Risk Factors for Sexual Recidivism?</i> | 47 |
| <i>“Why” does Sexual Deviance increase risk?</i> | 52 |
| <i>Sexual Deviance as a Risk Factor</i> | 53 |

| | |
|--|------------|
| <i>“How” is the risk of recidivism assessed?</i> | 54 |
| <i>Current Study</i> | 58 |
| Methods | 59 |
| <i>Ethical Considerations</i> | 59 |
| <i>Design</i> | 59 |
| <i>Context</i> | 59 |
| <i>Recruitment</i> | 60 |
| <i>Participant Demographics</i> | 61 |
| <i>Materials</i> | 61 |
| <i>Procedure</i> | 62 |
| <i>Data Analysis</i> | 63 |
| <i>Quality and Rigour</i> | 63 |
| Results | 65 |
| <i>Navigating initial Biases</i> | 67 |
| <i>Understanding Sexual Deviance</i> | 73 |
| <i>Gathering Evidence</i> | 76 |
| <i>Internet offending vs. Contact offending</i> | 82 |
| <i>Respondent Validation</i> | 83 |
| <i>Summary of the theoretical model</i> | 83 |
| Discussion..... | 83 |
| <i>Understanding Sexual Deviance in Practice</i> | 84 |
| <i>Assessing Sexual Deviance in Practice</i> | 85 |
| <i>The Motivation-Facilitation Model</i> | 88 |
| <i>Difficulties in the literature</i> | 88 |
| <i>Considerations for Internet Offenders</i> | 90 |
| <i>Strengths and Limitations</i> | 91 |
| <i>Clinical implications and Future Directions</i> | 92 |
| Conclusion | 93 |
| References..... | 94 |
| Appendices..... | 103 |
| Appendix 1: Sage Journal of Assessment - Submission Guidelines | 103 |
| Appendix 2: Prediction model Risk of Bias Assessment Tool (PROBAST)..... | 106 |
| Appendix 3: Ethical Approval..... | 110 |
| Appendix 4: Participant Information Handout | 111 |
| Appendix 5: Example of Line-by-Line Coding..... | 114 |
| Appendix 6: Samples from Researchers Memo’s | 115 |

Lay Summary

The aim of this research portfolio is to contribute to the understanding of how sexual deviance is assessed in forensic practice. Risk assessment is an important process that allows professionals to consider the likelihood that someone will re-offend (recidivate), in what circumstances this is most likely to occur and whether services can manage that risk.

Research has shown that lots of different factors seem to be related to sexual recidivism (e.g., poorer problem solving is related to higher sexual recidivism; higher antisocial traits is related to higher recidivism). Factors that are related to recidivism are interpreted as risk factors. The presence of these is assessed and monitored to inform risk and management decisions.

Sexual deviance has one of the strongest individual relationships with sexual recidivism and is therefore an important factor to risk assess. However, the way this relationship has been shown is through a measure of sexual arousal that is not commonly available in practice.

Other means of assessment try to capture sexual interest, preference, behaviours, and fantasies, but it is not clear whether these instruments allow practitioners to assess the same construct. The different tools and terms in research literature creates confusion as to what sexual deviance is. This confusion likely translates down into practice and makes risk assessment challenging.

The first half of this portfolio reviews the quality and quantity of evidence that a risk assessment instrument has predictive validity for sexual recidivism. Predictive validity is inferred when a relationship exists between instrument scores and sexual recidivism. The Violence Risk Scale for Sexual Offenders (VRS-SO) groups different types of risk factors together to provide total, scale and domain scores which aim to support practitioners to make objective judgements of risk. The primary reason to consider the VRS-SO for the focus of this review is that it groups certain factors into a sexual deviance domain, which could provide utility in practice. A systematic review allowed this researcher to explore the evidence for the predictive validity of the VRS-SO, to summarise the collective findings and interpret this in relation to the quality of the studies. Nine studies were included, and they indicated a promising relationship between sexual recidivism and the VRS-SO's representation of sexual deviance, as well as other VRS-SO scores. However, most of the included studies shared several methodological issues that mean these results have to be

interpreted cautiously. Most importantly, the author of the tool was involved in 7 of the 9 studies, which increases the risk that results will be biased to find positive results. There is a need for more independent studies of the VRS-SO's properties.

The second part of this portfolio explores a different aspect of assessing sexual deviance. To better understand how the research on sexual deviance was being translated into practice, this author conducted semi-structured interviews with 10 criminal justice social workers (CJSW's). CJSW's are responsible for a large portion of sex offender risk assessments in Scotland. An approach called grounded theory was used. This meant that the aim of the study was to create a theory explaining CJSW experience that was grounded in the data provided by them in the interviews. The theory suggests that CJSW's understand sexual deviance within the wider context of their risk assessment process. Interviews indicated that CJSW's most commonly thought about sexual deviance as an underlying, motivating sexual interest or preference that was considered sexually deviant due to its focus on an inappropriate being (e.g., child/animal) or an inappropriate act (e.g., violence and harm). In the process of risk assessment, they felt that presence of sexual deviance was best evidenced through repeating patterns of behaviour that could indicate an enduring preference/interest in a certain act or being. The difficulties in the literature around providing a clear, consistent definition of sexual deviance, were mirrored in participant interviews. Participants shared the view that defining and assessing sexual deviance was a difficult task, best understood within the context of the individual offender and their other risk and protective factors. They acknowledged that wider influences such as social norms, their dual role of assessment and management and their own experiences could create bias in their assessment. They felt this was important to recognise and manage. The theory presented does not claim to be fully inclusive, but in follow-up interviews CJSW's felt that that it represented their experience. It is hoped that representing the experiences of assessing sexual deviance in practice within the literature, will help direct future research in a clinically useful way.

Thesis Portfolio Abstract

Background: Sexual deviance is an important risk factor for sexual recidivism that is notoriously difficult to define and therefore to measure in practice.

Aims: The aim of this portfolio is to contribute to the literature regarding the assessment of sexual deviance in forensic practice. Firstly, by reviewing the evidence for an actuarial instrument which presents a sexual deviance domain that shows promising concurrent validity with phallometric measures. Secondly, by qualitatively exploring the experience of those most commonly charged with assessing sexual deviance in practice and capturing how they understand and approach this.

Methods: Systematic review methodology was used to search the literature for the Violence Risk Scale for Sexual Offenders (VRS-SO). The tool presents a total score, static and dynamic scale scores and three dynamic domain scores (Sexual deviance, criminality, and treatment responsivity). Nine studies were included. The risk of bias in these studies was assessed using the Prediction model Risk of Bias Assessment Tool (PROBAST) and their findings synthesised.

The researcher adopted a grounded theory approach to capture the experience of 10 Criminal Justice Social Workers (CJSW's). Data collection and analysis occurred in an iterative cycle to inform the development of a theoretical model representing how they understand and assess sexual deviance as a risk factor.

Results: The systematic review suggested that total, dynamic and sexual deviance domain scores had good predictive validity for sexual recidivism. However, these scores were interpreted cautiously within the context of the methodological quality.

Qualitative data analysis led to the formation of a theoretical model by which CJSW's primarily conceptualised sexual deviance as a deviant sexual interest or preference which could act as an underlying motivator for sexual recidivism. The process of assessing this is described and relied predominantly upon repeated patterns of behavioural evidence that spoke to an enduring interest in a victim-group or sexual act over time.

Conclusions: The results of these studies are discussed in the context of the existing literature on sexual deviance and risk assessment. Suggestions for future research are considered alongside implications for clinical practice.

The Predictive Validity of the Violence Risk Scale for Sexual Offenders (VRS-SO): A Systematic Review

Abstract

This review explores the evidence of predictive accuracy for the Violence Risk Scale for Sexual Offenders (VRS-SO). The VRS-SO is a risk assessment tool designed to assess the risk of sexual recidivism while also identifying areas of criminogenic need that could serve as intervention treatment targets. It incorporates static and dynamic items as well as a measure of change for dynamic items. This systematic review assessed the quality and quantity of evidence for the VRS-SO's predictive accuracy for sexual, violent and general recidivism. Nine studies were included. Results of data synthesis suggest the total, dynamic and sexual deviance domain scores had good predictive validity for sexual recidivism. Total score, criminality domain scores and treatment responsivity domain scores were promising for violent offending. These results are discussed in relation to methodological differences and their risk of bias (measured using the Prediction model Risk of Bias Assessment Tool; PROBAST).

Word Count: 13,821

Introduction

Risk assessment tools are used to guide decision-making about the lives of offenders and the safety of the general public. Their outcomes often determine the restrictions put in place to manage an offender's risk. As demonstrated by Andrews and Bonta (1998; 2010)'s Risk-Need-Responsivity principle, it is essential that the restrictions and interventions of professionals match the needs and risk of the offender, otherwise risk can be increased rather than mitigated.

Historically, risk assessments were based on unstructured clinical judgement. Concerns about accuracy and replicability motivated a search for evidence-based approaches to risk assessment (Menziez et al., 1994; Quinsey et al., 1998). Two types of risk assessment tool emerged from this, Actuarial Risk Assessment Instruments (ARAI) and Structured Professional Judgement (SPJ) Tools.

ARAI's endeavor to make objective estimates of an offender's likelihood of re-offending by rating a range of risk factors. These ratings form a weighted total score which is compared to the scores and associated recidivism rates of a normative sample. Advocates of ARAI's emphasize the ease of administration, cost-effectiveness, and statistical grounding of the tools, which they argue reduces subjective bias (Craig & Beech, 2009). However, the statistical underpinnings of these tools have been criticised. Some suggest that the margin of error in these tools is not appropriately understood and does not give rise to reliable estimates of an individual's likelihood of re-offending (Cooke & Michie, 2013). Additionally, an element of subjectivity likely remains in the selection process of appropriate tools and administration of them. Structured Professional Judgement (SPJ) tools also involve rating a series of risk factors but have increased focus on capturing narrative detail relevant to how these risk factors relate to the individual and each other. By nature, these are more subjective, however proponents argue that the structure and grounding in the evidence-base minimizes the risk of bias and offers scope for more practical utility (Cooke & Michie, 2013; Green, Carrol & Brett, 2010; Khiroya, Weaver & Madden, 2009). The evidence-base suggests ARAI's and SPJ's both demonstrate a significant relationship with sexual recidivism and therefore both are considered to have predictive validity (Hanson & Morton-Bourgon, 2009).

Hanson and Morton-Bourgon's (2009) meta-analysis indicated that different ARAI's were similarly efficacious when it comes to predicting risk of sexually violent recidivism. Nonetheless, this should not be interpreted to mean that any tool will do. An assessor should

consider the risk factors included, the evidence for different types of recidivism and the normative sample when selecting an ARAI. Risk factors are typically categorised as static (unchangeable) or dynamic (more malleable) risk factors (Andrews & Bonta, 1998). Static risk factors speak more to enduring risk and may be useful for informing restrictions. Dynamic risk factors consider how risk may change with time and environment. These may be more useful identifying targets for intervention. Sexual offenders are more likely to re-offend with non-sexual offences (McCann & Lussier, 2008; Rice, Harris, Lang, & Cormier, 2006; Zimring et al., 2007, 2009). Therefore, assessors may consider type of recidivism that a tool has a relationship with. Critics have highlighted that the predictive value of an ARAI reduces the more an individual differs from the normative sample. Therefore, assessors may consider the sample in the selection of a tool.

This highlights the importance of understanding the evidence-base for risk assessment tools to inform appropriate selection and use in practice. The development of accurate risk assessment tools facilitates the appropriate distribution of resources, in line with the risk-need-responsivity principle and thus reducing risk of recidivism. However, there are several methodological difficulties in this process.

Developing and validating risk assessment measures

First there are theoretical challenges to developing risk assessment tools. In essence each risk assessment tool is representative of a hypothesised model of risk. It presents a collection of risk factors, anticipated to interact in some way to influence an individual's propensity to re-offend. However, the process by which any one factor affects risk is not understood (Cooke & Michie, 2013; Ward & Fortune, 2016). A model, or measure, of risk is most likely to be effective if it has a strong theoretical underpinning (Olver, Neumann et al., 2018), but the current state of the literature limits this.

Second, a range of psychometric properties need to be demonstrated to allow an assessor to trust in a tool's ability to consistently measure what it aims to (reliability) and the accuracy with which it does that (validity). While both these components are important, studies tend to focus primarily on predictive validity (Van Calster, 2019). In this context, this is the ability of the tool to accurately predict recidivism. It can be evidenced by a tool's ability to discriminate between offenders who go on to recidivate and those who do not. As well as its ability to calibrate increasing risk effectively (e.g., higher risk scores should relate to increasing rates of recidivism; Olver & Eher, 2020; Wolff et al., 2019). As this is the primary

aim of these tools, the focus on this property is understandable. However, the ability to fully interpret the utility of any tool relies on a fuller understanding of its psychometric properties (Olver, Neumann et al., 2018; Wolff et al., 2019).

The ability to define and measure recidivism presents a further challenge. Most studies use official criminal record information as a measure of recidivism; however, the majority of sexual offences are not reported. Therefore, this is likely to be an underestimate of true recidivism (Andrews & Bonta, 1998; Conroy & Murrice, 2007). Tools validated in this way can only speak to the predictive accuracy for known recidivism.

Known sexual recidivism rates are low (Hanson & Morton-Bourgon, 2009; Harris et al., 2009). This creates difficulties in sampling, study design and statistical accuracy. The less commonly an outcome occurs, the more room for error when trying to predict it (Ozkan et al., 2020). Within validation studies, it is therefore important to have a significant proportion of the participants recidivate which requires larger sample sizes (Wolff et al., 2019).

The Violence Risk Scale – Sexual Offender Version (VRS-SO)

The VRS-SO (Olver and colleagues, 2003; Olver et al., 2007) is an ARAI designed to assess risk for sexual violence in sexual offenders, to assist in treatment planning and monitor changes in risk ratings over time. The first iteration was modelled on the Violence Risk Scale (Wong & Gordon, 2001). Olver and colleagues (2003) drew upon three theoretical models; the Psychology of Criminal Conduct (PCC; Andrews & Bonta, 1998), Relapse Prevention Theory (RPT; Pithers, 1990; Ward & Hudson, 1998) and the transtheoretical model of therapeutic change (TTM; Prochaska, DiClemente & Norcross, 1992). PCC and RPT contribute to the view that there are genetic and developmental contributions to offending behaviours but primarily it is an outcome of social learning, with recidivism occurring due to an interaction of dynamic factors. The TTM highlights the different changes which a person has to go through in order to change their behaviour (Pre-Contemplation, Contemplation, Preparation, Action, Maintenance, Relapse).

An iterative series of studies, led from the initial 16-item dynamic tool to the current 24-item version consisting of 7 static items and 17 dynamic items. The static component captures objective details regarding the frequency and severity of sexual offending behaviour, including the age at onset of offending, number of prior offences and victim demographics. The dynamic component consists of three domains – sexual deviance, criminality, and

treatment responsivity. The sexual deviance domain is most specific to sexual recidivism. It captures details of offenders' sexual interests and behaviours. The criminality domain captures a range of factors associated with violent/general offending, including offence supportive attitudes, interpersonal aggression, emotional control, and impulsivity. The treatment responsivity domain captures how the offender has engaged with treatment, supervision, and restrictions in the past and present.

Each item is rated using information gathered from a thorough file review and an optional semi-structured interview schedule. Items are rated on a 4-point Likert scale (0-3). Ratings are then combined into total scores. The VRS-SO presents 6 scores - total, static, dynamic and one for each dynamic domain. Higher ratings indicate a higher risk of sexual recidivism. As is common practice with actuarial tools, these scores are categorised into five risk categories which represent how much more/less likely the individual is to sexually re-offend compared to a normative sample of sexual offenders (Olver, Mundt et al., 2018; See Table 1). The normative sample for the VRS-SO consists of incarcerated, male, sex offenders who are predominantly white (approximately 60%) and have a relatively even distribution between child and adult victims (approximately 50% each).

Table 1
Risk Categories and Associated Recidivism Rates

| Risk Category | 5-year Sexual Recidivism Base Rates | | |
|---------------------------------------|-------------------------------------|---------------|--------------|
| | Total Score | Dynamic Score | Static Score |
| Category I: Very Low Risk | <1.0-2.0% | 1.3-3.1% | 1.7-2.1% |
| Category II: Below Average Risk | 2.1-4.1% | 3.2-5.2% | 2.5-4.6% |
| Category III: Average Risk | 4.3-13.8% | 5.4-12.5% | 5.6-17.4% |
| Category IVA: Above Average Risk | 14.3-26.7% | 13.0-21.0% | 20.7-28.4% |
| Category IVB: Well Above Average Risk | 27.5-60.4% | 21.7-53.1% | 32.8-53.1% |

Possible Clinical Contributions

An interesting aspect of the VRS-SO is the inclusion of an assessment of change for dynamic items based on the TTM. When a dynamic item scores for risk, it is considered a suitable target for change and is attributed one of the six stages of change. Progression across the

stages represents positive change. This is captured in a change score which has demonstrated predictive validity for recidivism (Beggs & Grace, 2011; Eher et al., 2020; Olver et al., 2020; Sowden & Olver, 2017). The authors have suggested that this component could add depth to repeat measurement and have clinical utility in the allocation of treatment resources.

The sexual deviance domain may also add to the clinical utility of this measure. Sexual deviance is one of the risk factors with the strongest relationship with sexual recidivism but is notoriously difficult to define (Laws & O'Donohue, 2008; Mann et al., 2010). Most of the research is based upon phallometric assessment of sexual arousal, but this is not readily available in practice (Stinson & Becker, 2008). This has led tool developers to try and capture sexual deviance by other means. An early example of this was the screening scale for paedophilic interests (SSPI, Seto & Lalumière, 2001), which consists of 4-items based on behavioural indicators of sexual deviance (e.g., number of victims, sex of victims). ARAI's tend to focus on offence behaviour to inform sexual deviance. However, the sexual deviance domain of the VRS-SO consists of 5-items which aim to look beyond the offending behaviour (e.g., sexually deviant lifestyle, deviant sexual preferences).

Studies have found that the sexual deviance domain is associated with sexual recidivism in samples of mixed sexual offenders (Olver & Wong, 2006). Canales, Olver and Wong (2009) found it correlated with phallometric measures of interest in paedophilic stimulus. Though not as strongly as the SSPI. Nonetheless, Hawes, Boccaccini and Murrie (2013) found it contributed to the predictive validity of the Psychopathy Checklist-Revised (PCL-R; Hare, 1991) comparably to the SSPI. This is promising evidence for use in practice and would benefit from further exploration.

The Current Review

The current review was motivated by the need to understand the strengths and weaknesses of available risk assessment tools. The majority of sexual offender risk assessments in the UK are conducted by Police, Social Workers and Probation Officers, who are primarily trained in the use of ARAI's. As the similarity of an offender to the normative sample of an ARAI affects the predictive accuracy of the tool, it is important to explore the psychometric properties of the range of tools developed on different populations. The VRS-SO offers utility for incarcerated samples. It also offers potential advantages over other available measures. The VRS-SO takes a broader approach to the construct of sexual deviance (i.e., Stable-2007,

Hanson et al., 2007; SVR-20, Boer et al., 1997), which shows promising construct validity (Canales et al., 2009). Additionally, the VRS-SO offers a novel approach to assessing change in the dynamic risk factors of offenders. While other tools in use in the UK such as the Stable-2007 (Hanson et al., 2007) are designed to assess change, there has been relatively little evidence exploring this (Brankley, Babchishin & Hanson, 2021).

Meta-analyses comparing risk assessment instruments have indicated that the VRS-SO presents as a potentially promising addition to the repertoire currently available, with an evidence base for incarcerated adult male sexual offenders (Hawes et al., 2013; van den Berg, 2018). At the time of these meta-analyses, limited studies were available for inclusion, but the research base for the VRS-SO has grown, facilitating a more in-depth evaluation of its evidence for psychometric properties.

This review began with the broad aim of exploring the predictive properties of static, dynamic and change scores of the VRS-SO, including the predictive accuracy and incremental validity. Scoping searches narrowed the focus to the following question: To what extent do the total, scale, and domain scores of the VRS-SO demonstrate predictive accuracy for recidivism in adult, male sex offenders?

Methods

The protocol for this review is registered with the International Register of Systematic Reviews (Registration number: CRD42022338129). It can be accessed online at: (https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42022338129).

Search Strategy

A literature search of the online databases (including PsychINFO, EMBASE and MEDLINE), was conducted on the 3rd of March 2022. This primary search was restricted to articles that had been published in peer-reviewed journals, available in English, online or in print. To reduce the possible impact of publication bias, ProQuest Dissertations and Theses Global Database were searched for relevant, unpublished doctoral theses. Additional studies were identified by reviewing citations of included papers, the reference sections of included papers and contacting key authors.

Table 2
Search Terms

| Psychometric Properties AND | Risk Assessment AND | Sexual Offenders |
|--|---|--|
| Psychometric Propert* OR Psychometric* OR Validity OR Reliability OR Statistic* OR Methodology OR Questionnaire* OR Measurement | Risk Assessment OR Risk Measure* OR Recidivism OR Risk Factor* OR Violence Risk Scale OR VRS | Sexual Offen* OR P?edophilia OR Rapist OR Child Molestor OR Exhibitioni* OR Bestiality OR Incest OR Indecent Exposure OR Sexual Sadis* OR Sexual Devian* OR Paraphilia |

Truncation () and wildcard (?) symbols were used to ease the searching of term-variant and international spellings.*

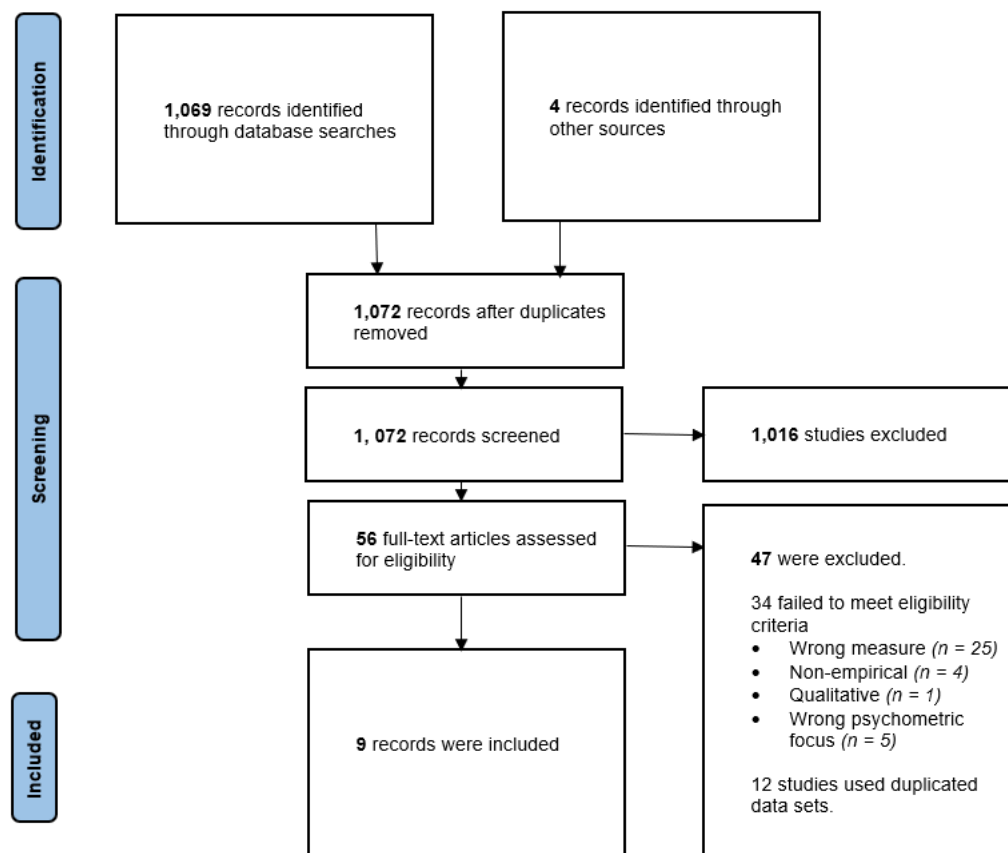
As shown in Table 2., the search terms were derived from the key concepts of the aim of this review which was to explore the “psychometric properties”, of the Violence Risk Scale - Sexual Offenders version (VRS-SO) “risk assessment” in its use with “sexual offenders”.

Eligibility

All relevant studies were included for review. Studies were considered relevant if they purported to conduct a quantitative analysis of the predictive properties of the VRS-SO total and/or domain scores with an adult (18+), male, sex offender population.

The selection process is detailed in Figure 1. Fifty-six potentially relevant studies were included for a full-text review. Thirty-four were excluded because they did not meet the eligibility criteria. Twelve were excluded due to sample duplication. In total, 17 studies used data that involved amalgamations of the same samples. To reduce the risk of over-representing positive findings, studies that repeated the same analyses on the same sample were excluded. The decision of which to include was made based on 1) length of follow-up for recidivism and 2) relevance to research questions.

Figure 1
Study Selection Process



Data Extraction

A data extraction form was piloted on 4 of the included papers and then adapted for use with all included papers. It was designed to capture descriptive and statistical information relevant to the research question. Descriptive information included details of the country of study, study design and participant demographics. To aid interpretation and comparison, information on how recidivism rates were operationalised and measured was extracted, along with the details of participant follow-up and details of statistical analyses.

Synthesis of Results

A narrative synthesis was used to compare and present the data relevant to this review. All included studies used area under the Receiver Operating Characteristics (ROC) curve, known as Area Under the Curve (AUC) analysis. AUC analyses are considered the gold standard for predictive accuracy of recidivism. They are robust to fluctuations in base rates and selection ratio's (Rice & Harris, 2005). This allows comparisons across measures and studies which consist of varying base rates.

The ROC curve is a probability curve which plots the true positive rate of a model (the sensitivity) against the false positive rate. This allows the predictive performance of a model to be assessed by the rate of increase in true positives compared to false positives. The AUC statistic summarises how capable the model is of distinguishing between binary outcomes. It provides equal weight to sensitivity (true positive rate) and specificity (true negative rate). This can be problematic with some predictive models where a skew in one direction has more significant consequences (e.g., health conditions). In the context of recidivism risk, over and under restriction have significant consequences, this equal balance is appropriate.

The AUC value represents the degree of probability that a person who has an outcome will be accurately predicted by the model to have that outcome. In this context, the AUC value speaks to how capable the VRS-SO is of correctly distinguishing between a recidivist and a non-recidivist. AUC analyses provide a value between 0 and 1. The higher the AUC value, the better it is at distinguishing between outcomes. Lower values indicate a higher likelihood of type I (false positive) and type II (false negative) errors. A value of 0.5 indicates that the model has a 50% chance of accurately predicting the binary outcome - equivalent to chance. Values below 0.5 indicate that the model is more likely to make inaccurate prediction than accurate (Rice & Harris, 2005).

Rice and Harris (2005) set the precedent of categorising AUC values as small (AUC = 0.56-0.63), medium (AUC = 0.64-0.70) and large effect sizes (AUC = 0.71-1.0). For the VRS-SO to be considered to have a large effect size, it would therefore need to correctly distinguish between recidivists and non-recidivists at least 71% of the time. The AUC, significance and confidence intervals were reported for consistency between studies. Where data was not available, study authors were contacted to request additional information, though none responded.

Risk of Bias

Quality appraisal was conducted using the Prediction model Risk of Bias Assessment Tool (PROBAST; Wolff et al., 2019; See Appendix 2). This is a relatively novel tool designed to assess the risk of bias within studies on the predictive properties of diagnostic and prognostic models of risk. According to the PROBAST a prognostic model is one which estimates the likelihood of a specific outcome within a certain time frame. Although it was initially developed for use with medical models, this definition is applicable to the VRS-SO. There is

some early precedent for the use of the PROBAST with forensic risk assessment instruments (Burghart et al., 2022; Fazel et al., 2021).

The PROBAST identifies problems with design and analysis as the most common sources of ROB in studies validating models of risk. The PROBAST for validation studies consists of 17-items which are categorised into four domains: 1) participants, 2) predictors, 3) outcome and 4) analysis. The participant domain explores sources of data, inclusion/exclusion criteria and facilitates consideration of study design. The predictor and outcome domain explores whether these variables have been clearly and consistently defined and assessed. The availability of the variables, blinding of researchers and timing of data collection is explored. The analysis domain explores the base-rate of the outcome, how data was handled and evaluated. Each item questions whether the decisions made by the researchers resulted in a low risk of bias. These can then be answered; yes; probably yes; probably no; no and no information. Each domain is then summed into a score of high, low, or unclear risk. Domains with more than one item scored as “no, probably no” are at higher risk of bias. While the developers feel this scoring system is useful, an emphasis is placed on the inclusion of qualitative information. This forces users of the PROBAST to justify their ratings and facilitates consideration of more subtle differences between studies.

The PROBAST also incorporates an assessment of applicability, to check that a study is appropriate for inclusion in a review. Due to the narrow focus of this review, there were no concerns with applicability of the included studies.

Inter-rater Reliability

To minimise risk of bias, a second reviewer screened 25% of the papers at each phase of study collection (abstract review and full-text review). There were no discrepancies regarding the inclusion/exclusion of studies reviewed. The same reviewer conducted quality appraisals using the PROBAST on all included studies. Initial comparison of PROBAST rating indicated a 66.7% agreement on items, with 51 items scored differently. Most of these differences were within the analysis domain. First round of discussions revealed that 46 of these items were related to differences in understanding of factors specifically related to recidivism as an outcome measure and the analytic methods used in these studies. Once these were clarified, agreement between raters was 96.7%, with only 5 items scored differently. These were primarily related to whether there was enough information in the papers to infer a quality rating (e.g., “Probably Yes” or “Probably No”) and were resolved collaboratively.

Results

Study Characteristics

The characteristics of each study and its participants are outlined in Table 2. Nine studies were included in this review. Studies 1 and 2 shared a participant sample, though looked at different domains of the VRS-SO and so did not repeat the same analysis (Olver et al., 2020; Olver, et al., 2016). Sample sizes ranged from 91-668 participants, with 2,453 participants in total. All were adult male sex offenders who participated in an assessment of risk between 1983 and 2015. For 6 studies, assessment took place as part of a treatment programme. Seven studies reported the mean scores on the VRS-SO to demonstrate the average level of risk amongst their participant population. Four reported the mean VRS-SO total scores as representative of “average risk” (Beggs & Grace, 2010; Goodman-Delahunty & O’Brien, 2014; Olver & Eher, 2020; Olver et al., 2007). The other 3 studies reported mean VRS-SO total scores as “above average risk” (Eher et al., 2020; Sowden & Olver, 2017; Todd, 2013). Four studies reported higher risk ratings on dynamic than static components (Eher et al., 2020; Goodman-Delahunty & O’Brien, 2014; Sowden & Olver, 2017; Todd, 2013). The rest were approximately equivalent (Beggs & Grace, 2010; Olver & Eher, 2020; Olver et al., 2007).

Participant demographics were described differently across studies. For example, participant age was described at the time of offence, of release, of treatment and of recidivism. This inhibited the ability to synthesise across the longitudinal timeframe of these studies. Of relevance to this review question is the offence characteristics of the participants. Two studies included only offenders with child victims (Beggs & Grace, 2010; Goodman-Delahunty & O’Brien, 2014), while the rest included offenders with a range of victims. Despite differing terminology, studies described the offence of the participants in relation to the age of their victims, which facilitated cross-cultural comparison. There was some variation in how offenders with mixed ages of victims were represented. As shown in Table 2., four studies used a “mixed offender” category, while others recorded this type of offender either under the two distinct categories or only coded the victim from the index offence. This creates some uncertainty as to the exact proportions of offences in these samples. It is clear that at least 63.2% of the participants had offences against children. At least 47.6% had offences against victims over the age of 14. At least 27% of participants had committed incest offences. Only one study explicitly included non-contact offenders in their population (Olver, et al., 2016).

One explicitly excluded non-contact offenders (Todd, 2013). The rest did not state whether non-contact offenders were represented.

Studies were conducted in Canada (n = 4), Austria (n = 2), Australia (n = 1), New Zealand (n = 1) and the USA (n = 1). Eight out of the 9 studies used a retrospective/archival cohort design (Dekkers et al., 2012), with only 1 prospective study. The archival nature of these designs prevented the use of the interview schedule of the VRS-SO. Results are based on file review alone. The prospective study (Olver et al., 2020) did not state whether the interview was used.

The included studies looked at the predictive validity of the static total scores (n = 8), dynamic total scores (n = 8), dynamic domain scores (n = 7) and overall VRS-SO score (n = 7). Five reported on the incremental validity of the VRS-SO dynamic scores (Beggs & Grace, 2010; Olver & Eher, 2020; Olver et al., 2007; Sowden & Olver, 2017; Todd, 2013) and four the predictive validity of change scores (Eher et al., 2020; Olver et al., 2020; Sowden & Olver, 2017; Todd, 2013). The evidence on change scores and incremental validity was insufficient to include in this review.

Predictive validity was evaluated using AUC analyses for scores with recidivism incidences across various follow up periods (range of 2-14 years). Four studies measured recidivism as any new conviction or charge following release into the community, while 5 relied on convictions alone. Sexual, violent, and other forms of recidivism were defined differently across studies, as demonstrated in Table 2.

Table 3
Descriptive Characteristics of Included Studies and their Samples

| Study No. | Authors | Country | Study Design | VRS-SO Score of interest | Sample | | | Recidivism | | | | RoB Rating* |
|-----------|--|---------|----------------------------|--|--------|--|---|------------|-------------------------|--|---|-------------|
| | | | | | N | Offending Demographics | Average VRS-SO Risk Scores of Overall Sample | N | Follow-up Period (yrs.) | Definition | Rates | |
| 1 | Olver, Nicholaichuk, Kingston & Wong (2020) | Canada | Prospective Cohort Study | Static Total Score | 570 | Offences against victims <14yrs old n = 260 (45.6%) Offences against victims >14yrs old n = 296 (51.9%) Unknown n = 14 (2.5%) ** | Average for sample not reported | 570 | 5 & 10 fixed | Convictions & Charges Sexual = sexually motivated offence (contact or non-contact) Violent = an offence against a person involving actual, attempted, or threatened physical or psychological harm, including sexual offences. General = any category of offence, including technical breaches. | Sexual Recidivism, n = 61 (10.7%) Violent Recidivism, n = 157 (27.5%) General Recidivism, n = 258 (45.5%) | Low |
| 2 | Olver, Klepfisz, Stockdale, Kingston, Nicholaichuk & Wong (2016) | Canada | Retrospective Cohort Study | Dynamic Total Score Dynamic Domain Scores | 668 | Offences against <u>unrelated</u> victims <14yrs old n = 195 (29.2%) Incest offences against victims <14yrs old n = 120 (17.9%) Offences against >14yrs old n = 261 (39.2%) Mixed Age Offenders n = 89 (13.3%) Non-contact n = 3 (0.4%) | Static M = 5.14 (SD = 1.8) Below Average Risk | 647 | 5 & 10 fixed | Conviction & Charges Sexual = sexually motivated offence Violent = an offence against the person (e.g., assault, robbery), including sexual offences. General = any criminal code conviction. | Sexual Recidivism, n = 67 (10.4%) Violent Recidivism, n = 150 (23.2%) General Recidivism, n = 269 (41.6%) | Unclear |

| | | | | | | | | | | | | |
|---|---|---------|---|--|-----|---|--|-----|-----------------------|---|--|------|
| 3 | Todd (2013) | USA | Retrospective Cohort Study Treatment | Static Total Score Dynamic Total Score Dynamic Domain Scores VRS-SO Total Score | 94 | <p>“Child Molesters” Offences against <u>unrelated</u> victims <13yrs old <i>n</i> = 36 (37.5%)</p> <p>“Rapists” Offences against unrelated adult victims <i>n</i> = 16 (16.7%)</p> <p>Incest Offenders (age not specified) <i>n</i> = 9 (9.4%)</p> <p>Mixed Age Offenders <i>n</i> = 32.3(34.4%)</p> <p>Non-contact offenders not included.</p> | Dynamic M = 35.74 (SD = 5.60) Well Above Average Risk Total M = 48.03 (SD = 6.86) Above Average Risk | 87 | m = 5.34 | Conviction & Charges Sexual = any charge or conviction for a sexual offence. General = any charge or conviction for a non-violent, non-sexual offence. Non-sexual charges or reconvictions = all non-violent and violent offences. | Sexual Recidivism, <i>n</i> = 5 (5.7%) Sexual Charge or Recidivism, <i>n</i> = 9 (10.3%) Non-sexual recidivism, <i>n</i> = 18, (20.7%) Non-sexual Charge or Recidivism, <i>n</i> = 19 (21.8%) | High |
| 4 | Olver & Eher (2020) | Austria | Retrospective Cohort Study | Static Total Score Dynamic Total Score Dynamic Domain Scores VRS-SO Total Score | 668 | <p>Offences against victims <14yrs old <i>n</i> = 361 (54%)</p> <p>Offences against victims >14yrs old <i>n</i> = 307 (46%) **</p> | Dynamic M = 26.8 (SD = 7.5) Average Risk Static M = 6.8 (SD = 4.0) Average Risk Total M = 33.4 (SD = 9.8) Average Risk | 335 | 5 fixed | Convictions only Sexual = sexually motivated offence (contact or non-contact) Violent = any contact sexual offence or nonsexual violent offence. | Sexual Recidivism, <i>n</i> = 52 (14.7%) Violent Recidivism, <i>n</i> = 97, (27.5%) | Low |
| 5 | Eher, Hofer, Buchgeher, Domany, Turner & Olver (2020) | Austria | Retrospective Cohort Study Treatment | Static Total Score Dynamic Total Score Dynamic Domain Scores VRS-SO Total Score | 91 | <p>Offences against victims <14yrs old <i>n</i> = 53 (58.2%)</p> <p>Offences against victims >14yrs old <i>n</i> = 38 (41.8%) **</p> | Sexual Deviance M = 11.71 (SD = 4.19) Criminality M = 11.40 (SD = 6.71) Treatment Responsivity | 70 | m = 7.17 SD = 2.47 | Conviction & Incarceration Sexual = sexually motivated offence (contact or non-contact) Other recidivism = any offence leading to a subsequent incarceration or psychiatric placement. | Sexual recidivism, <i>n</i> = 11 (12.1%) Other recidivism, <i>n</i> = 19 (20.9%) | Low |

| | | | | | | | | | | | | |
|---|--|--------|---|--|-----|--|--|-----|--------------------|--|--|------|
| | | | | | | | M = 9.54 (SD = 2.22) Dynamic M = 37.33 (SD = 6.71) Well Above Average Risk Total M = 46.93 (SD = 9.34) Above Average Risk | | | | | |
| 6 | Olver, Wong, Nicolaichuk & Gordon (2007) | Canada | Retrospective Cohort Study Treatment | Static Total Score Dynamic Total Score Dynamic Domain Scores VRS-SO Total Score | 321 | Offences against victims <14yrs old <i>n</i> = 56 (17.5%) Offences against victims >14yrs old <i>n</i> = 169 (52.7%) Mixed age offenders <i>n</i> = 45 (14%) Incest Offenders (age not specified) <i>n</i> = 51 (15.8%)** | Sexual Deviance M = 6.3 (SD = 4.0) Criminality M = 8.8 (SD = 3.6) Treatment Responsivity M = 6.8 (SD = 2.1) Dynamic M = 24.9 (SD = 7.5) Average Risk Static M = 10 (SD = 4) Average Risk Total M = 34.9 (SD = 10) Average Risk | 321 | m = 10.0 SD = 4 | Conviction only Sexual = sexually motivated offence Non-sexual Violent = an offence against a person that was not sexually motivated Recidivism – any conviction for a new sexual or non-sexual violent offence following release to community. | Sexual recidivism, <i>n</i> = 79 (24.6%) Violent recidivism, <i>n</i> = 115 (35.8%) Recidivism, <i>n</i> = 194 (60.4%) | High |

| | | | | | | | | | | | |
|---|------------------------------------|-----------|---|--|-----|---|--|---------------------|--|---|---------|
| 7 | Sowden & Olver (2017) | Canada | Retrospective Cohort Study Treatment | Static Total Score Dynamic Total Score Dynamic Domain Scores VRS-SO Total Score | 180 | <p>“Child Molesters” Offences against victims <14yrs old <i>n</i> = 40 (22%)</p> <p>“Rapists” Offences against victims >14yrs old <i>n</i> = 82 (45.5%)</p> <p>Mixed Age Offenders <i>n</i> = 36 (20%)</p> <p>Incest Offenders (age not specified) <i>n</i> = 22 (12.5%) **</p> | <p>Sexual Deviance M = 8.97 (SD = 3.72)</p> <p>Criminality M = 10.88 (SD = 3.64)</p> <p>Treatment Responsivity M = 7.34 (SD = 2.32)</p> <p>Dynamic M = 31.31 (SD = 5.36) Above Average Risk</p> <p>Static M = 11.24 (SD = 3.53) Average Risk</p> <p>Total M = 42.43 (SD = 7.26) Above Average Risk</p> | m = 9.3 SD = 3.0 | <p>Convictions Only</p> <p>Sexual = sexually motivated offence (contact or non-contact)</p> <p>Nonsexual violent = an offence against a person that was not sexually motivated</p> <p>Violent = any sexual or nonsexual violent offence</p> <p>General = any new criminal conviction, including breaches and technical violations.</p> | <p>Sexual recidivism, <i>n</i> = 36 (20%)</p> <p>Non-sexual Violent recidivism <i>n</i> = 59 (32.8%)</p> <p>Violent recidivism, <i>n</i> = 81 (45%)</p> <p>General recidivism, <i>n</i> = 109 (60.6%)</p> | High |
| 8 | Goodman-Delahunty & O'Brien (2014) | Australia | Retrospective Cohort Study Treatment | Static Total Score Dynamic Total Score Dynamic Domain Scores VRS-SO Total Score | 213 | <p>Offences against an <u>unrelated</u> child victim <i>n</i> = 118 (55.4%)</p> <p>Incest offences against a child victim <i>n</i> = 97 (45.5%) **</p> | <p>Sexual Deviance M = 13.10 (SD = 1.9)</p> <p>Criminality M = 11.64 (SD = 4.1)</p> <p>Treatment Responsivity M = 9.59 (SD = 2.8)</p> <p>Dynamic</p> | m = 9.1 SD = 4.5 | <p>Conviction & Charge</p> <p>Sexual = a new sexual offence.</p> <p>Violent = a new violent offence.</p> <p>General = all reoffences including violent and sexual</p> | <p>Sexual recidivism, <i>n</i> = 20 (11.6%)</p> <p>Violent recidivism, <i>n</i> = 17 (9.9%)</p> <p>General recidivism, <i>n</i> = 55 (32%)</p> | Unclear |

| | | | | | | | | | | | | |
|---|----------------------|-------------|---|---|-----|---|--|-----|-----------|--|---|-----|
| | | | | | | <p>M = 36.32 (SD = 8) Well Above Average Risk</p> <p>Static M = 2.72 (SD = 2.6) Below Average Risk</p> <p>Total M = 39.01 (SD = 9) Average Risk</p> | | | | | | |
| 9 | Beggs & Grace (2010) | New Zealand | Retrospective Cohort Study Treatment | <p>Static Total Score</p> <p>Dynamic Total Score</p> <p>Dynamic Domain Scores</p> <p>VRS-SO Total Score</p> | 218 | <p>Incest offences against a child victim <i>n</i> = 123 (56.4%)</p> <p>Mixed offences against <u>related</u> and <u>unrelated</u> child victims <i>n</i> = 95 (43.6%) **</p> | <p>Sexual Deviance M = 9.1 (SD = 2.7)</p> <p>Criminality M = 4.1 (SD = 3.0)</p> <p>Treatment Responsivity M = 4.3 (SD = 2.1)</p> <p>Dynamic M = 21.5 (SD = 5.9) Average Risk</p> <p>Static M = 7.6 (SD = 2) Average Risk</p> <p>Total M = 29.1 (SD = 9.2) Average Risk</p> | 218 | m = 12.24 | <p>Convictions only</p> <p>Sexual = a sexual offence with an identifiable victim (e.g. sexual assault, incest, exhibitionism) & possession of child pornography</p> <p>Violent = any nonsexual offence against a person (e.g., assault, robbery, kidnapping).</p> <p>General = offences that were neither sexual nor violent (e.g., possession of cannabis).</p> | <p>Sexual recidivism, <i>n</i> = 29 (13.3%)</p> <p>Violent recidivism, <i>n</i> = 30 (13.8%)</p> <p>General recidivism, <i>n</i> = 81 (37.2%)</p> | Low |

*Risk of Bias Rating as assessed by PROBAST (Prediction model Risk of Bias Assessment Tool)

** Inclusion of non-contact offenders not specified.

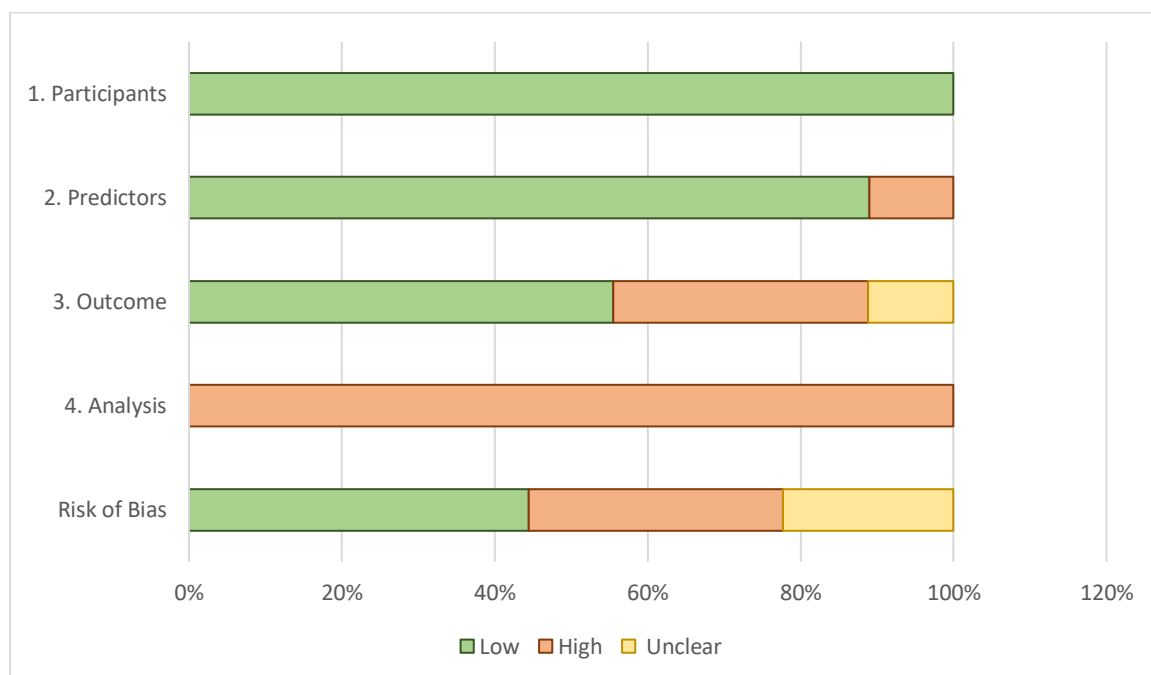
*** The average VRS-SO presented were all taken pre-treatment; Risk Categories as defined by Olver, Thornton et al., (2018).

Risk of Bias in Studies

The PROBAST assessment indicated that 4 studies were at low RoB, 3 at high RoB and 2 were unclear. As shown in Figure 2/Table 4, most of this risk derived from issues in outcome and analysis domains.

Figure 2

PROBAST Ratings - Percentage of studies rated in each domain



Participants. All studies scored low for risk of bias in this domain. All participants were reflective of the intended population for use of the VRS-SO with data drawn from appropriate sources. None of the studies clearly reported their inclusion/exclusion criteria, though convenience sampling of those in sexual offender services naturally contained the demographics of the sample.

Predictor Variables. All except Study 8 scored low for risk of bias in this domain (Goodman-Delahunty & O'Brien, 2008). The VRS-SO is manualised which facilitates consistency of item ratings. Six studies referenced training on the manual from the tool developers which further enhances consistency (Beggs & Grace, 2010; Goodman-Delahunty & O'Brien, 2014; Olver & Eher, 2020; Olver, et al., 2016; Olver et al., 2020 & Sowden & Olver, 2017). Five studies reported inter-rater reliability of the VRS-SO items, which ranged between moderate-good (0.5-0.99; Beggs & Grace, 2010; Goodman-Delahunty & O'Brien, 2014; Olver et al., 2007; Sowden & Olver, 2017 & Todd, 2013). Six studies reported that the

VRS-SO was scored blind to recidivism outcomes (Beggs & Grace, 2010; Eher et al., 2020; Olver & Eher, 2020; Olver et al., 2020; Olver et al., 2007 & Sowden & Olver, 2017). The remaining 3 did not reference this.

Despite a shared low RoB rating, there were some qualitative differences due to study design. Prospective, longitudinal and multisite cohort studies are arguably the strongest design for studies of predictive validity (Wolff et al., 2019). Only Study 1 (Olver et al., 2020) utilised this design. The rest were longitudinal but archival which may have limited the availability of information to allow accurate scoring of predictor variables. Studies 1 and 2 (Olver et al., 2016; Olver et al., 2020) were open to increased RoB because they initially prospectively used a 16-item version of the VRS-SO and then retrospectively rated the 17th dynamic item and the static items from archival information. The process of this in study 2 is protected by the more objective nature of the static items. RoB is increased by the subjective nature of dynamic items in Study 1. One researcher completed retrospective scoring, reducing the risk of inter-rater differences. However, they only scored 1/3rd of the items and then prorated the rest. The method of proration is not described which leaves the extent of RoB unclear. Study 4 (Olver & Eher, 2020) is at higher RoB as half the VRS-SO scores were rated routinely by frontline staff and half were rated by the researchers. While the manual and training may reduce RoB, different people collecting scores at different time periods increases RoB.

Study 8 (Goodman-Delahunty & O'Brien, 2014), received a rating of high RoB, because it made adaptations to the coding of the VRS-SO. These alterations were made to ensure that the archival data was being scored effectively and was done in collaboration with the tool developers to minimise deviation. Nonetheless, adaptations increase the risk of changing the underlying model and impact generalisability of findings.

Outcome Variables. Recidivism as an outcome creates RoB in a number of ways. Firstly, there is no gold standard specification for operationalising recidivism in risk research (Klinge, 2019). Five studies defined recidivism as any new conviction received during follow-up (Beggs & Grace, 2010; Eher et al., 2020; Olver & Eher, 2020 & Olver et al., 2007; Sowden & Olver, 2017). Four studies defined recidivism as any new conviction or charge during follow-up (Goodman-Delahunty & O'Brien, 2014; Olver et al., 2016; Olver et al., 2020 & Todd, 2013). The use of convictions alone is likely a significant under-representation of the true rate of recidivism as it only includes those which have been caught, charged and proven. Some authors therefore encourage the inclusion of charges as it increases the number

of offences represented (Olver et al., 2020). However, it likely also increases the inclusion of false positives in the statistic. Both these methods for capturing recidivism are considered acceptable in the research on risk assessments and so both have been scored as low risk in this appraisal. Nonetheless, it is important to remember that all formal recidivism rates are estimates. How this is defined will influence interpretation of study outcomes.

All studies accessed outcome data through governmental public records. This limits the control of the researcher over the classification and recording of data. Drawing data from a single database reduces risk of variability of information available across participants, which minimises risk of bias. However, three studies which drew upon a single database (Olver et al., 2016; Olver et al., 2007; Sowden & Olver, 2017), referenced additional documents were available for some participants. For example, those that received a federal sentence had more information available. This may mean there was variety within the outcome information available, and some recidivism data may have been wrongly categorised due to missing details (e.g., whether a violent offence was sexually motivated). Only studies 3 and 8 (Goodman-Delahunty & O'Brien, 2014; Todd, 2013), used multiple databases.

Five studies scored low for RoB (Beggs & Grace, 2010; Eher et al., 2020; Goodman-Delahunty & O'Brien, 2014; Olver & Eher, 2020 & Olver et al., 2020) 1 was unclear (Olver et al., 2016) and 3 were high RoB (Olver et al., 2007; Sowden & Olver, 2017 & Todd, 2013). However, the distinction between ratings in this domain is less clear. The higher ratings of bias for studies, 3, 6 and 7 were primarily attributed to lack of blinding of the researcher and more variable periods of follow-up. While these three studies explicitly stated that the researcher was not blind to predictor variables when accessing outcome data, no other studies reported that they protected blinding in their procedure – there was merely a lack of information which inhibits the ability to judge RoB. An element of blinding is embedded in collection of outcomes as it was initially recorded by government agencies un-associated with the research. However, this does not protect from bias at the point of data extraction from these records.

Included studies had follow-up times ranging from 2 to 14 years. There is no gold standard follow-up time for recidivism studies. Most recidivism occurs within the first few years following release from an incarceration, which leads some to argue 2-3years is a sufficient follow-up (Blumstein & Nakamura, 2012; Klingele, 2019). However, Blumstein and Nakamura (2012) have suggested that a person's risk does not reduce until they have desisted

from re-offending for 10-13years. In assessing risk of bias for these studies, 5 years was viewed as an acceptable follow-up period, with 10-15 years considered the gold standard. Five studies reported a follow-up range with a minimum below 5 years (Goodman-Delahunty & O'Brien, 2014; Eher et al., 2020; Olver et al., 2007; Sowden & Olver, 2017; Todd, 2013). The average follow-up time was 5years or more for all studies (as can be seen in Table 3). However, studies 3 and 6 (Olver et al., 2007; Todd, 2013) had a wider range of follow-up time and more participants followed-up for less than 5 years, increasing their RoB.

Table 4
PROBAST Ratings for each Study

| Study | Risk of Bias | | | | Overall |
|-------|--------------|------------|---------|----------|---------|
| | Participants | Predictors | Outcome | Analysis | |
| 1 | + | + | + | - | + |
| 2 | + | + | ? | - | ? |
| 3 | + | + | - | - | - |
| 4 | + | + | + | - | + |
| 5 | + | + | + | - | + |
| 6 | + | + | - | - | - |
| 7 | + | + | - | - | - |
| 8 | + | ? | + | - | ? |
| 9 | + | + | + | - | + |

PROBAST = Prediction model Risk of Bias Assessment Tool; ROB = Risk of Bias
+ indicates low risk of bias; - indicates high risk of bias; and ? Indicates unclear risk of bias affected by information missing from the study.

Analysis. All studies scored high for RoB in this domain. In prediction model studies, the number of participants who experience the observed outcome is important when considering the power of analysis. With a binary outcome, such as recidivism, there needs to be a reasonable proportion of both to draw meaningful conclusions regarding the accuracy and precision of the measure. The number or percentage of an outcome that constitutes a reasonable proportion varies, but the designers of PROBAST recommend at least 100 participants for a validation study (Wolff et al., 2019). Based on this, none of the included studies had sufficient participants to assess sexual recidivism, though three studies had more than 100 participants with sufficient violent recidivism rates. Two of these included sexual recidivism within violent recidivism (Olver et al., 2016; Olver et al., 2020) and one was just violent (Olver et al., 2007).

Another common source of bias within predictive model studies stems from the treatment of missing data. Only two studies indicated that there was no missing data throughout their studies (Beggs & Grace, 2010; Sowden & Olver, 2017). Five studies reported missing predictor data. This was only for the dynamic domain. Studies 3 and 4 removed participants with missing predictor items (Olver & Eher, 2020; Todd, 2013). This is a common yet

controversial approach. However, in both of these studies the number of participants removed was minimal. Both studies 6 and 8 (Goodman-Delahunty & O'Brien, 2014; Olver et al., 2007) had a high proportion of participants with missing items (>75%). Six studies excluded participants with missing data. For study 6 (Olver et al., 2007), participants with 4 or more missing items were excluded. The rest were estimated using a regression-based imputation procedure. Study 8 (Goodman-Delahunty, 2014) used a regression procedure to estimate missing values. Study 1 prorated missing predictor data, as described earlier (Olver et al., 2020). Five studies reported different numbers of participants included in analyses than in the overall study, though did not discuss the reasons for this (Eher et al., 2020; Olver & Eher, 2020; Olver et al., 2020; Olver et al., 2016; Todd, 2013). It is assumed that this reduction in numbers is due to missing outcome data which led to exclusion during statistical analysis.

The PROBAST has an item relating to “complexities within the data”, which includes censoring and accounting for multiple outcome events. Censoring is an inherent risk in studies which have an unnatural endpoint in order for analysis to take place. For example, participants are categorised as having re-offended or not within the time frame, but participants may have gone on to re-offend after the study. Therefore, some participants who are classed as “non-recidivists” in the study, may go on to recidivate. Fixed follow-up periods can help with interpretation of results as it can speak to the rate of recidivism within a specific timeframe. Only studies 1, 2 and 4 (Olver & Eher, 2020; Olver et al., 2016; Olver et al., 2020) used fixed follow-up times. Some statistical procedures can account for censoring. However, the analysis of interest in this review is more vulnerable to bias. Multiple outcome events per participant are also a potential source of bias. None of the studies explained how this was represented/controlled for (e.g., if a participant had multiple incidents of recidivism).

The final area explored within this domain was the extent to which different performance measures were evaluated and how appropriately. The quality and applicability of a predictive measure or model is typically assessed in terms of discrimination (to distinguish between those with and without an outcome) and calibration (to recognise increasing risk of an outcome with increasing rating on the measure). All included studies spoke to discrimination (using AUC analyses), which were the focus of this review. Only study 4 (Olver & Eher, 2020) specifically explored calibration, using E/O analysis (Hanson, 2017).

What is the predictive validity for the VRS-SO Static Total Score?

Eight of the 9 included studies reported on the total static score of the VRS-SO. Table 5 reports the AUC values for the prediction of recidivism using the total static score. For sexual recidivism, 4 out of 8 studies reported a significant predictive effect. Three showed large effect sizes for prediction of sexual recidivism (Olver et al., 2016, Olver & Eher, 2020, Olver et al., 2007) and one demonstrated a medium effect size (Beggs & Grace, 2010). Three out of the 7 studies that looked at violent recidivism reported a significant predictive effect size. Two studies showed a large effect size (Olver et al., 2016; Olver & Eher, 2020) and one study found a small effect size (Olver et al., 2007) for violent recidivism. Three of the 6 studies that looked at general recidivism reported a significant predictive effect size; 1 large (Olver et al., 2020), 1 medium (Beggs & Grace, 2010) and 1 small (Sowden & Olver, 2017). Studies 2 and 4 also conducted the same analysis with a fixed 5-year follow-up period, which provided results of the same significance. For study 2, the AUCs were weaker (sexual and violent) or equivalent (general; Olver et al., 2016). For study 4, the AUC was better for sexual (AUC = 0.80, $p < 0.001$) but worse for violent (AUC = 0.66, $p < 0.001$; Olver & Eher, 2020).

Of the three studies which reported no effect for any form of recidivism, two also found no significant predictive findings on other measures of static factors (Static-99; Eher et al., 2020; Todd, 2013). This suggests that the lack of effect may be reflective of a feature of the participant samples or study design as opposed to a representation of the VRS-SO's predictive capabilities.

Table 5
Static Score - AUC values for VRS-SO prediction of Sexual, Violent and General Recidivism

| No | Years Follow Up | Recidivism | | |
|----|---------------------------|--|--|--|
| | | Sexual | Violent | General |
| 2 | 10-year fixed | $n = 319, AUC = 0.71^{***}$ <i>95% CI = 0.64-0.79</i> | $n = 319, AUC = 0.72^{***}$ <i>95% CI = 0.66-0.78</i> | $n = 319, AUC = 0.74^{***}$ <i>95% CI = 0.68-0.79</i> |
| 3 | $m = 5.34$ | $n = 58, AUC = 0.58$ | $n = 58, AUC = 0.54$ | $n = 58, AUC = 0.63$ |
| 4 | $M = 11.4$ (SD = 1.9) | $n = 353, AUC = 0.76^{***}$ <i>95% CI = 0.69-0.82</i> | $n = 353, AUC = 0.65^{***}$ <i>95% CI = 0.59-0.70</i> | |
| 5 | $m = 7.17$ (SD = 2.47) | $n = 70, AUC = 0.65$ <i>95% CI = 0.53-0.70</i> | | $n = 70, AUC = 0.59$ <i>95% CI = 0.47-0.71</i> |
| 6 | $m = 10.0$ (SD = 4) | $n = 321, AUC = 0.74^{***}$ <i>95% CI = 0.68-0.80</i> | $n = 321, AUC = 0.60, **$ <i>95% CI = 0.53-0.66</i> | |
| 7 | $m = 9.3$ (SD = 3.0) | $n = 180, AUC = 0.53, p > 0.05$ <i>95% CI = 0.43-0.64</i> | $n = 180, AUC = 0.57, p > 0.05$ <i>95% CI = 0.49-0.66</i> | $n = 180, AUC = 0.59^*$ <i>95% CI = 0.51-0.68</i> |
| 8 | $m = 9.1$ (SD = 4.5) | $n = 172, AUC = 0.54, p > 0.05$ <i>95% CI = 0.28-0.80</i> | $n = 172, AUC = 0.77, p > 0.05$ <i>95% CI = 0.39-1.0</i> | $n = 172, AUC = 0.58$ <i>95% CI = 0.48-0.69</i> |
| 9 | $m = 12.24$ | $n = 218, AUC = 0.70^{**}$ <i>95% CI = 0.60-0.80</i> | $n = 218, AUC = 0.59, p > 0.05$ <i>95% CI = 0.49-0.69</i> | $n = 218, AUC = 0.67^{***}$ <i>95% CI = 0.60-0.74</i> |

AUC values (Rice & Harris, 2005); 0.56 = small effect size; 0.64 = medium effect size; 0.71 = large effect size

** significant at $p < 0.05$; ** significant at $p < 0.01$; *** significant at $p < 0.001$; non-significant results shaded in grey.*

What is the predictive validity for the VRS-SO Dynamic Total Score?

Eight of the 9 included studies reported on the total dynamic score of the VRS-SO. Table 6. reports the AUC values for the prediction of recidivism using the total dynamic score of the VRS-SO. Seven out of 8 studies measured dynamic items pre- and post- a treatment programme, which facilitated two sets of analyses.

Table 6
Dynamic Total Score - AUC values for VRS-SO prediction of Sexual, Violent and General Recidivism

| No. | Years Follow Up | Sexual | Violent | General |
|-----|-------------------------|--|---|---|
| | | | | |
| 1 | 5-year fixed | <i>n</i> = 564, AUC = 0.68 *** 95% CI = 0.60-0.76 | <i>n</i> = 564, AUC = 0.68*** 95% CI = 0.63-0.74 | <i>n</i> = 564, AUC = 0.70*** 95% CI = 0.65-0.74 |
| | | <i>n</i> = 564, AUC = 0.70*** 95% CI = 0.63-0.78 | <i>n</i> = 564, AUC = 0.70*** 95% CI = 0.65-0.76 | <i>n</i> = 564, AUC = 0.70*** 95% CI = 0.65-0.74 |
| 3 | m = 5.34 | <i>n</i> = 58, AUC = 0.53 | <i>n</i> = 58, AUC = 0.30 | <i>n</i> = 58, AUC = 0.42 |
| | | <i>n</i> = 58, AUC = 0.55 | <i>n</i> = 58, AUC = 0.39 | <i>n</i> = 58, AUC = 0.52 |
| 4 | 5-year fixed | <i>n</i> = 340, AUC = 0.70*** 95% CI = 0.62-0.79 | <i>n</i> = 340, AUC = 0.66*** 95% CI = 0.59-0.73 | |
| 5 | m = 7.17 (SD = 2.47) | <i>n</i> = 70, AUC = 0.72** 95% CI = 0.60-0.81 | | <i>n</i> = 70, AUC = 0.70, ** 95% CI = 0.58-0.80 |
| | | <i>n</i> = 57, AUC = 0.67, 95% CI = 0.49-0.85 | | <i>n</i> = 57, AUC = 0.66 95% CI = 0.51-0.81 |
| 6 | m = 10.0 (SD = 4) | <i>n</i> = 321, AUC = 0.66*** 95% CI = 0.59-0.73 | <i>n</i> = 321, AUC = 0.53 95% CI = 0.46-0.59 | |
| | | <i>n</i> = 321, AUC = 0.67*** 95% CI = 0.60-0.74 | <i>n</i> = 321, AUC = 0.55 95% CI = 0.48-0.61 | |
| 7 | m = 9.3 (SD = 3.0) | <i>n</i> = 180, AUC = 0.63* 95% CI = 0.53-0.73 | <i>n</i> = 180, AUC = 0.65*** 95% CI = 0.57-0.73 | <i>n</i> = 180, AUC = 0.66*** 95% CI = 0.58-0.74 |
| | | <i>n</i> = 180, AUC = 0.64* 95% CI = 0.53-0.74 | <i>n</i> = 180, AUC = 0.68*** 95% CI = 0.60-0.76 | <i>n</i> = 180, AUC = 0.66*** 95% CI = 0.58-0.74 |
| 8 | m = 9.1 (SD = 4.5) | <i>n</i> = 172, AUC = 0.49 95% CI = 0.30-0.69 | <i>n</i> = 172, AUC = 0.55 95% CI = 0.34-0.75 | <i>n</i> = 172, AUC = 0.64** 95% CI = 0.54-0.73 |
| | | <i>n</i> = 172, AUC = 0.41 95% CI = 0.16-0.65 | <i>n</i> = 172, AUC = 0.75* 95% CI = 0.61-0.89 | <i>n</i> = 172, AUC = 0.62 95% CI = 0.46-0.77 |
| 9 | m = 12.24 | <i>n</i> = 218, AUC = 0.78*** 95% CI = 0.69-0.88 | <i>n</i> = 218, AUC = 0.65* 95% CI = 0.53-0.77 | <i>n</i> = 218, AUC = 0.60* 95% CI = 0.52-0.68 |
| | | <i>n</i> = 218, AUC = 0.81*** 95% CI = 0.72-0.89 | <i>n</i> = 218, AUC = 0.64* 95% CI = 0.52-0.75 | <i>n</i> = 218, AUC = 0.58 95% CI = 0.50-0.66 |

AUC values (Rice & Harris, 2005); 0.56 = small effect size; 0.64 = medium effect size; 0.71 = large effect size

* significant at $p < 0.05$; ** significant at $p < 0.01$; *** significant at $p < 0.001$; non-significant results shaded in grey.

For sexual recidivism, 8 studies conducted 15 analyses and reported 10 significant predictive effects. Two studies reported 3 large effect sizes for prediction of sexual recidivism (Beggs & Grace, 2010; Eher et al., 2020). Four studies reported 7 medium effect sizes for prediction of sexual recidivism (Olver & Eher, 2020; Olver et al., 2020; Olver et al., 2007; Sowden & Olver, 2017). Only study 3 (Todd et al., 2013) and study 8 (Goodman-Delahunty & O'Brien, 2014) reported no significant effects for sexual recidivism. Study 3 had the highest RoB for all included studies and found no significant predictor relationships on any of the measures that they used, including the VRS-SO. It is therefore likely that the methodological

limitations of this study contributed to the non-significance of effects reported. Study 8 was the only study to report adapting their coding for scoring dynamic items, which reduces the comparability of their results to those of other studies included in this review. They reported no significant results for any scores related to sexual recidivism.

Of the 13 analyses across 7 studies looking at violent recidivism, there were 8 significant predictive effect sizes reported. For violent recidivism only one study reported a large effect size (Goodman-Delahunty & O'Brien, 2014). While 4 studies reported 7 medium effect sizes (Beggs & Grace, 2010; Olver et al., 2020; Olver & Eher, 2020; Sowden & Olver, 2017). Six studies conducted 12 analyses for general recidivism, of which 5 studies reported 6 medium and 1 small effect sizes (Beggs & Grace, 2010; Eher et al., 2020; Goodman-Delahunty & O'Brien, 2014; Olver et al., 2020; Sowden & Olver, 2017).

What is the predictive validity for the VRS-SO Dynamic Domain Scores?

Seven of the 9 included studies reported on the dynamic domain scores of the VRS-SO. Table 7. reports the AUC values for the prediction of recidivism using the three dynamic domain scores: sexual deviance, criminality and treatment responsivity. Six out of the 7 studies measured this pre- and post-treatment allowing for two sets of analyses.

For the sexual deviance domain, 6 studies had significant results for sexual, violent and general offending. However, the only effect sizes of note were for sexual recidivism, with 3 studies reporting 4 large effect sizes (Beggs & Grace, 2010; Eher et al., 2020; Olver & Eher, 2020), 1 study reporting a medium effect size (Eher et al., 2020) and 2 studies reporting 3 small effect sizes (Olver et al., 2020; Olver et al., 2007). This is not surprising as the content of this domain is specific to sexual offending. Study 7 (Sowden & Olver, 2017) and Study 8 (Goodman-Delahunty & O'Brien, 2014) were the only studies to not find a significant result for sexual recidivism in this domain. As noted, Study 8 reported no significant results relating to sexual recidivism. Interestingly their participants had the highest mean scores for sexual deviance and the smallest standardized deviation suggesting a tightly clustered population of scores. This may have limited the ability to analyse possible effects in this domain due to restricted variance. Study 7 explored their finding using a correlation analysis and found an inverse relationship between sexual deviance scores and recidivism, which is in direct contrast to the evidence base on this factor.

For the criminality domain scores, 7 studies had significant findings across the three recidivism categories. For sexual recidivism, 4 studies reported 7 medium effect sizes (Beggs

& Grace, 2010; Olver et al., 2020; Olver et al., 2007; Sowden & Olver, 2017;) and 1 study found a small effect size (Olver et al., 2007).

Table 7
Dynamic Domain Scores - AUC values for VRS-SO prediction of Sexual, Violent and General Recidivism

| Domain | | Follow-Up Period | Recidivism | | |
|--|-------------|-------------------------|--|--|--|
| | | | Sexual | Violent | General |
| Sexual Deviance | 1 | 5-year Fixed | $n = 563, AUC = 0.58$ $95\% CI = 0.50-0.66$ | $n = 563, AUC = 0.49$ $95\% CI = 0.43-0.55$ | $n = 563, AUC = 0.52$ $95\% CI = 0.47-0.57$ |
| | | | $n = 563, AUC = 0.61^*$ $95\% CI = 0.53-0.69$ | $n = 563, AUC = 0.52$ $95\% CI = 0.46-0.58$ | $n = 563, AUC = 0.54$ $95\% CI = 0.49-0.58$ |
| | 4 | M = 11.4 (SD = 1.9) | $n = 343, AUC = 0.71^{***}$ $95\% CI = 0.64-0.79$ | $n = 343, AUC = 0.44$ $95\% CI = 0.38-0.51$ | |
| | 5 | M = 7.17 (SD = 2.47) | $n = 70, AUC = 0.70^{**}$ $95\% CI = 0.58-0.80$ | | $n = 70, AUC = 0.54$ $95\% CI = 0.51-0.81$ |
| | | | $n = 57, AUC = 0.74^{***}$ $95\% CI = 0.23-0.60$ | | $n = 57, AUC = 0.51$ $95\% CI = 0.50-0.74$ |
| | 6 | M = 10.0 (SD = 4) | $n = 321, AUC = 0.59^*$ $95\% CI = 0.52-0.66$ | $n = 321, AUC = 0.35^{***}$ $95\% CI = 0.29-0.41$ | |
| | | | $n = 321, AUC = 0.61^{**}$ $95\% CI = 0.54-0.68$ | $n = 321, AUC = 0.35^{***}$ $95\% CI = 0.29-0.41$ | |
| | 7 | M = 9.3 (SD = 3.0) | $n = 180, AUC = 0.46$ $95\% CI = 0.34-0.57$ | $n = 180, AUC = 0.42^*$ $95\% CI = 0.33-0.50$ | $n = 180, AUC = 0.38^{**}$ $95\% CI = 0.29-0.46$ |
| | | | $n = 180, AUC = 0.48$ $95\% CI = 0.36-0.60$ | $n = 180, AUC = 0.44$ $95\% CI = 0.36-0.53$ | $n = 180, AUC = 0.40^{**}$ $95\% CI = 0.31-0.48$ |
| | 8 | M = 9.1 (SD = 4.5) | $n = 172, AUC = 0.52$ $95\% CI = 0.31-0.72$ | $n = 172, AUC = 0.57$ $95\% CI = 0.41-0.72$ | $n = 172, AUC = 0.57$ $95\% CI = 0.48-0.74$ |
| | | | $n = 172, AUC = 0.38$ $95\% CI = 0.04-0.72$ | $n = 172, AUC = 0.67$ $95\% CI = 0.36-0.99$ | $n = 172, AUC = 0.59$ $95\% CI = 0.43-0.74$ |
| | 9 | M = 12.24 | $n = 218, AUC = 0.72^{***}$ $95\% CI = 0.62-0.82$ | $n = 218, AUC = 0.60$ $95\% CI = 0.49-0.71$ | $n = 218, AUC = 0.51$ $95\% CI = 0.43-0.59$ |
| | | | $n = 218, AUC = 0.77^{***}$ $95\% CI = 0.62-0.82$ | $n = 218, AUC = 0.61$ $95\% CI = 0.50-0.72$ | $n = 218, AUC = 0.52$ $95\% CI = 0.44-0.60$ |
| | Criminality | 1 | 5-year Fixed | $n = 568, AUC = 0.67^{***}$ $95\% CI = 0.59-0.74$ | $n = 568, AUC = 0.75^{***}$ $95\% CI = 0.70-0.80$ |
| $n = 568, AUC = 0.69^{***}$ $95\% CI = 0.62-0.77$ | | | | $n = 568, AUC = 0.77^{***}$ $95\% CI = 0.72-0.82$ | $n = 568, AUC = 0.76^{***}$ $95\% CI = 0.72-0.80$ |
| 4 | | 5-year Fixed | $n = 323, AUC = 0.61$ $95\% CI = 0.52-0.70$ | $n = 323, AUC = 0.73^{***}$ $95\% CI = 0.66-0.80$ | |
| 5 | | M = 7.17 (SD = 2.47) | $n = 70, AUC = 0.51$ $95\% CI = 0.39-0.63$ | | $n = 70, AUC = 0.68^*$ $95\% CI = 0.56-0.79$ |
| | | | $n = 57, AUC = 0.51$ $95\% CI = 0.38-0.65$ | | $n = 57, AUC = 0.63$ $95\% CI = 0.49-0.76$ |
| 6 | | M = 10.0 (SD = 4) | $n = 321, AUC = 0.63^{***}$ $95\% CI = 0.56-0.70$ | $n = 321, AUC = 0.65^{***}$ $95\% CI = 0.59-0.71$ | |
| | | | $n = 321, AUC = 0.65^{***}$ $95\% CI = 0.58-0.72$ | $n = 321, AUC = 0.67^{***}$ $95\% CI = 0.61-0.73$ | |
| 7 | | M = 9.3 (SD = 3.0) | $n = 180, AUC = 0.65^{**}$ $95\% CI = 0.57-0.74$ | $n = 180, AUC = 0.72^{***}$ $95\% CI = 0.65-0.80$ | $n = 180, AUC = 0.78^{***}$ $95\% CI = 0.71-0.86$ |
| | | | $n = 180, AUC = 0.66^{**}$ $95\% CI = 0.57-0.76$ | $n = 180, AUC = 0.72^{***}$ $95\% CI = 0.53-0.70$ | $n = 180, AUC = 0.77^{***}$ $95\% CI = 0.69-0.84$ |
| 8 | | M = 9.1 (SD = 4.5) | $n = 172, AUC = 0.49$ $95\% CI = 0.29-0.68$ | $n = 172, AUC = 0.78^{**}$ $95\% CI = 0.66-0.90$ | $n = 172, AUC = 0.67^{**}$ |

| | | | | | |
|---|---|-------------------------|---|--|---|
| | | | <i>n</i> = 172, AUC = 0.43 95% CI = 0.19-0.67 | <i>n</i> = 172, AUC = 0.82* 95% CI = 0.60-1.0 | 95% CI = 0.58-0.76 <i>n</i> = 172, AUC = 0.67* 95% CI = 0.52-0.81 |
| | 9 | M = 12.24 | <i>n</i> = 218, AUC = 0.69** 95% CI = 0.59-0.79 | <i>n</i> = 218, AUC = 0.76*** 95% CI = 0.67-0.85 | <i>n</i> = 218, AUC = 0.71*** 95% CI = 0.64-0.77 |
| | | | <i>n</i> = 218, AUC = 0.70** 95% CI = 0.59-0.79 | <i>n</i> = 218, AUC = 0.77*** 95% CI = 0.67-0.86 | <i>n</i> = 218, AUC = 0.70*** 95% CI = 0.63-0.77 |
| Treatment Responsivity | 1 | 5-year Fixed | <i>n</i> = 566, AUC = 0.61* 95% CI = 0.51-0.70 | <i>n</i> = 566, AUC = 0.063*** 95% CI = 0.56-0.69 | <i>n</i> = 566, AUC = 0.59*** 95% CI = 0.54-0.64 |
| | | | <i>n</i> = 566, AUC = 0.64** 95% CI = 0.55-0.73 | <i>n</i> = 566, AUC = 0.64*** 95% CI = 0.58-0.70 | <i>n</i> = 566, AUC = 0.60*** 95% CI = 0.55-0.65 |
| | 4 | 5-year Fixed | <i>n</i> = 322, AUC = 0.59 95% CI = 0.46-0.71 | <i>n</i> = 322, AUC = 0.61** 95% CI = 0.53-0.69 | |
| | 5 | M = 7.17 (SD = 2.47) | <i>n</i> = 70, AUC = 0.51 95% CI = 0.38-0.63 | | <i>n</i> = 70, AUC = 0.6395% CI = 0.50-0.74 |
| | | | <i>n</i> = 57, AUC = 0.56 95% CI = 0.44-0.71 | | <i>n</i> = 57, AUC = 0.69** 95% CI = 0.56-0.81 |
| | 6 | M = 10.0 (SD = 4) | <i>n</i> = 321, AUC = 0.58* 95% CI = 0.51-0.65 | <i>n</i> = 321, AUC = 0.59* 95% CI = 0.53-0.65 | |
| | | | <i>n</i> = 321, AUC = 0.59** 95% CI = 0.52-0.66 | <i>n</i> = 321, AUC = 0.60** 95% CI = 0.53-0.65 | |
| | 7 | M = 9.3 (SD = 3.0) | <i>n</i> = 180, AUC = 0.59 95% CI = 0.49-0.69 | <i>n</i> = 180, AUC = 0.62** 95% CI = 0.54-0.70 | <i>n</i> = 180, AUC = 0.62** 95% CI = 0.54-0.71 |
| | | | <i>n</i> = 180, AUC = 0.62* 95% CI = 0.51-0.72 | <i>n</i> = 180, AUC = 0.65*** 95% CI = 0.57-0.73 | <i>n</i> = 180, AUC = 0.64** 95% CI = 0.56-0.73 |
| | 8 | M = 9.1 (SD = 4.5) | <i>n</i> = 172, AUC = 0.52 95% CI = 0.32-0.71 | <i>n</i> = 172, AUC = 0.70* 95% CI = 0.53-0.87 | <i>n</i> = 172, AUC = 0.55* 95% CI = 0.45-0.64 |
| | | | <i>n</i> = 172, AUC = 0.52, 95% CI = 0.30-0.75 | <i>n</i> = 172, AUC = 0.79* 95% CI = 0.00-1.0 | <i>n</i> = 172, AUC = 0.55* 95% CI = 0.47-0.70 |
| | 9 | M = 12.24 | <i>n</i> = 218, AUC = 0.73*** 95% CI = 0.65-0.82 | <i>n</i> = 218, AUC = 0.49, 95% CI = 0.38-0.60 | <i>n</i> = 218, AUC = 0.51, 95% CI = 0.43-0.59 |
| <i>n</i> = 218, AUC = 0.74*** 95% CI = 0.65-0.83 | | | <i>n</i> = 218, AUC = 0.47, 95% CI = 0.35-0.58 | <i>n</i> = 218, AUC = 0.49, 95% CI = 0.41-0.57 | |

AUC values (Rice & Harris, 2005); 0.56 = small effect size; 0.64 = medium effect size; 0.71 = large effect size

* significant at $p < 0.05$; ** significant at $p < 0.01$; *** significant at $p < 0.001$; non-significant results shaded in grey.

For violent recidivism, 5 studies reported 9 large effect sizes (Beggs & Grace, 2010; Goodman-Delahunty & O'Brien, 2014; Olver & Eher, 2020; Olver et al., 2020; Sowden & Olver, 2017) and 1 study reported 2 medium effect sizes (Olver et al., 2007). For general recidivism, 3 studies found 5 large effect sizes (Beggs & Grace, 2010; Olver et al., 2020; Sowden & Olver, 2017;) and 3 studies reported 4 medium effect sizes (Beggs & Grace, 2010; Goodman-Delahunty & O'Brien, 2014; Eher et al., 2020). The stronger predictive accuracy for violent and general recidivism respectively, is in line with the literature on the risk factors in this domain.

For the treatment responsivity domain, 7 studies produced significant findings for sexual, violent and general recidivism. For sexual recidivism, 1 study reported 2 large effect sizes

(Beggs & Grace, 2010), 1 study reported 1 medium effect size (Olver et al., 2020) and 3 studies reported 4 small effect sizes (Olver et al., 2020; Olver et al., 2007; Sowden & Olver, 2017).

For violent recidivism, 1 study reported a large effect size (Goodman-Delahunty & O'Brien, 2014). 3 studies reported medium effect sizes (Goodman-Delahunty & O'Brien, 2014; Olver et al., 2020; Sowden & Olver, 2017;) and 4 studies reported 5 small effect sizes (Olver & Eher, 2020; Olver et al., 2020; Olver et al., 2007; Sowden & Olver, 2017). For general recidivism 2 studies reported medium effect sizes (Eher et al., 2020; Sowden & Olver, 2017) and 3 studies reported 5 small effect sizes (Goodman-Delahunty & O'Brien, 2014; Olver et al., 2020; Sowden & Olver, 2017).

What is the predictive validity for the VRS-SO Overall Scores?

Seven of the 9 included studies reported on the overall score of the VRS-SO. Table 8. reports the AUC values for the prediction of recidivism using the total overall score. Four of the 7 studies measured this pre- and post-treatment allowing for two sets of analyses. For sexual recidivism, the 7 studies conducted 11 analyses.

Table 8
Total Scores - AUC values for VRS-SO prediction of Sexual, Violent and General Recidivism

| No. | Years Follow Up | Recidivism | | |
|-----|----------------------|---|---|---|
| | | Sexual | Violent | General |
| 3 | m = 5.34 | <i>n</i> = 58, AUC = 0.56 | <i>n</i> = 58, AUC = 0.40 | <i>n</i> = 58, AUC = 0.61 |
| 4 | 5-Year Fixed | <i>n</i> = 340, AUC = 0.79*** 95% CI = 0.71-0.87 | <i>n</i> = 340, AUC = 0.68*** 95% CI = 0.61-0.75 | |
| 5 | m = 7.17 (SD = 2.47) | <i>n</i> = 70, AUC = 0.71* 95% CI = 0.59-0.80 | | <i>n</i> = 70, AUC = 0.61 95% CI = 0.53-0.76 |
| 6 | m = 10.0 (SD = 4) | <i>n</i> = 321, AUC = 0.71*** 95% CI = 0.64-0.77 | <i>n</i> = 321, AUC = 0.56 95% CI = 0.50-0.62 | |
| | | <i>n</i> = 321, AUC = 0.72*** 95% CI = 0.66-0.78 | <i>n</i> = 321, AUC = 0.57* 95% CI = 0.51-0.64 | |
| 7 | m = 9.3 (SD = 3.0) | <i>n</i> = 180, AUC = 0.61* 95% CI = 0.51-0.71 | <i>n</i> = 180, AUC = 0.63** 95% CI = 0.55-0.71 | <i>n</i> = 180, AUC = 0.65*** 95% CI = 0.56-0.73 |
| | | <i>n</i> = 180, AUC = 0.62* 95% CI = 0.52-0.72 | <i>n</i> = 180, AUC = 0.66*** 95% CI = 0.58-0.74 | <i>n</i> = 180, AUC = 0.66*** 95% CI = 0.28-0.74 |
| 8 | m = 9.1 (SD = 4.5) | <i>n</i> = 172, AUC = 0.56 95% CI = 0.34-0.77 | <i>n</i> = 172, AUC = 0.73* 95% CI = 0.58-0.87 | <i>n</i> = 172, AUC = 0.65** 95% CI = 0.55-0.74 |
| | | <i>n</i> = 172, AUC = 0.44 95% CI = 0.12-0.76 | <i>n</i> = 172, AUC = 0.73* 95% CI = 0.61-0.89 | <i>n</i> = 172, AUC = 0.62 95% CI = 0.47-0.77 |
| 9 | m = 12.24 | <i>n</i> = 218, AUC = 0.79*** 95% CI = 0.69-0.88 | <i>n</i> = 218, AUC = 0.64* 95% CI = 0.52-0.75 | <i>n</i> = 218, AUC = 0.65*** 95% CI = 0.58-0.73 |
| | | <i>n</i> = 218, AUC = 0.80*** 95% CI = 0.71-0.89 | <i>n</i> = 218, AUC = 0.63* 95% CI = 0.52-0.74 | <i>n</i> = 218, AUC = 0.64** 95% CI = 0.56-0.71 |

AUC values (Rice & Harris, 2005); 0.56 = small effect size; 0.64 = medium effect size; 0.71 = large effect size

* significant at $p < 0.05$; ** significant at $p < 0.01$; *** significant at $p < 0.001$; non-significant results shaded in grey.

Four studies reported 6 large (Beggs & Grace, 2010; Eher et al., 2020; Olver & Eher, 2020; Olver et al., 2007;) and 1 study found 2 medium effect sizes (Sowden & Olver, 2017). For violent recidivism, 6 studies conducted 10 analyses. One study reported 2 large effect sizes (Goodman-Delahunty & O'Brien, 2014), 3 studies reported medium effect sizes (Beggs & Grace, 2010; Olver & Eher, 2020; Olver et al., 2007) and 3 studies reported 3 small effect sizes (Beggs & Grace, 2010; Olver et al., 2007; Sowden & Olver, 2017). For general recidivism 5 studies conducted 8 analyses. Three studies reported 5 medium effect sizes (Beggs & Grace, 2010; Goodman-Delahunty & O'Brien, 2014; Sowden & Olver, 2017).

Discussion

This paper used systematic review methodology to explore the predictive accuracy of the VRS-SO (i.e., the ability of the VRS-SO to predict recidivism outcomes). It presented narrative synthesis of the predictive accuracy of the total, static, dynamic and domain scores for recidivism. The VRS-SO was primarily designed as a tool for risk of sexual violence (Olver, 2003). However, in practice understanding the risk of different types of recidivism is important for risk management and public protection. Therefore, the validity of the VRS-SO is discussed for sexual, violent and general recidivism.

Summary of Results

Static scores are generally considered to be the most robust predictors of recidivism (Quinsey et al., 1998). However, this review found inconsistent evidence for the predictive validity of the static scores of the VRS-SO for all three types of recidivism. Approximately half of the studies which reported on the predictive validity of static scores found a significant effect for each type of recidivism. Slightly higher effect sizes were found for sexual recidivism than violent and general respectively. Six of the 8 studies that explored the validity of static scores compared the VRS-SO and the Static-99 and found similar results (Beggs & Grace, 2010; Eher et al., 2020; Olver & Eher, 2020; Olver et al., 2007; Sowden & Olver, 2014; Todd, 2013). This suggests that the differences between results are not specific to the static component of the VRS-SO but may reflect unclear differences in the samples and contexts of these studies.

For sexual recidivism, the scores of the dynamic, total and sexual deviance domain scores showed good predictive validity. The total score primarily demonstrated large effect sizes and the dynamic score medium-large effect sizes. Only two studies contradicted these findings (Goodman-Delahunty & O'Brien, 2014; Todd et al., 2013). It is worth noting that these were

two of the three studies conducted independent of the tool developers. Todd et al., (2013; Study 3) had the highest risk of bias rating and demonstrated no significant predictor relationships for the VRS-SO or any other measures used in their study. The lack of significant findings across included assessment tools indicates that this is reflective of the methodological limitations of this study rather than a true reflection of the VRS-SO's psychometric properties. Goodman-Delahunty and O'Brien (2014; Study 8) had an unclear risk of bias rating. Their findings are distinct from other included studies as they found no predictive results for sexual recidivism across their analyses. Their strongest outcomes were for violent and general recidivism. The authors attributed this to the small base-rate of sexual recidivism found in their sample. However, their rate was comparable to other studies included in this review and so it is unlikely to be the only explanation. Importantly, Goodman-Delahunty and O'Brien were the only included study to report making adaptations to the scoring of the VRS-SO to fit the retrospective nature of their data. Although they did this with input from the tool designers, these adaptations may have impacted the predictive validity of these particular domains. While it is important not to dismiss the results of this study, its comparability to the other studies is limited by the adaptation as they may have altered the underlying model of risk (Wolff et al., 2019).

The sexual deviance domain showed promising predictive validity for sexual recidivism. The majority of studies found significant predictive results, but the effect sizes were varied. The reasons for this are unclear, as there were no consistent patterns of findings based on participant demographics (e.g., mean scores for sexual deviance) or quality of the study. A surprising finding was presented by Sowden and Olver (2017), who reported an inverse relationship between sexual deviance scores and recidivism. This contradicts the wealth of evidence which suggests sexual deviance is one of the strongest predictors for sexual recidivism (Mann et al., 2010). Sowden and Olver's study received a higher risk of bias rating than other studies, but this was attributed to lack of blinding to recidivism outcomes which would not be expected to have this effect. There may be participant differences or other unclear reasons for this outcome. However, along with the varied effect sizes, it highlights the need to better understand the underlying relationship between domains and domain items. The criminality and treatment responsivity domains demonstrated mixed evidence for sexual recidivism, with about half the studies reporting significant findings. This is not surprising as these domains are more related to violent and general recidivism in the literature (Hanson & Bussiere, 1996; Mann et al., 2010).

For violent recidivism, the total scores and treatment responsivity domain scores showed promising predictive validity, with a strong number of studies reporting significant effect sizes, varying from small-large. The criminality domain presented the strongest evidence for violent recidivism with all studies demonstrating a significant medium-large effect size. A similar pattern with weaker effects emerged for general recidivism, with criminality and treatment responsivity showing promise. There were more mixed findings for total scores and dynamic scores. The increased variety in effect sizes for general recidivism may be reflective of the broader variety in definitions of this outcome, which makes it the least comparable across studies.

Considering the promising relationship found between the Sexual Deviance domain and phallometric measures (Canales et al., 2009) and with sexual recidivism (Hawes, Boccaccini & Murrie, 2013; Olver & Wong, 2006), the inconsistent results found in this review are surprising. Although some argue that the only true method of assessing sexual deviance is phallometry (Seto & Lalumière, 2001), there is a need for more accessible evidence-based, assessment tools in practice. Therefore, further studies better understanding the constructive, concurrent, and predictive validity of this domain could significantly contribute to the literature and add utility in forensic practice. The effect sizes for dynamic and total scores are comparable to effect sizes reported for other ARAI's. As is the finding that total scores are more predictive of sexual recidivism – the type of recidivism that the VRS-SO was primarily designed to measure (Hanson & Morton-Bourgon, 2009; Van den Berg et al., 2018). Nonetheless, these findings need to be interpreted cautiously in light of the methodological issues within the included studies.

Study Strengths and Limitations

The PROBAST indicated that 3 of the included studies were at high risk of bias, 2 were unclear and 4 had a low risk of bias. However, it is worth noting that with the exception of Todd (2013) there are fewer qualitative differences between included studies than implied by these overall ratings. The primary differences between quality ratings were related to researcher blinding. Sowden and Olver (2017; Study 7) along with Olver and colleagues (2007; study 6) both received a high rating of bias in the outcomes domain because their papers explicitly reported that the researchers were not blind to predictors when collecting outcome data. Other studies received a “no information” rating for this item. It is likely that at least some of the other studies also were not blind to predictor variables, however this could

not be inferred. The two studies which received an unclear risk of bias rating (Olver et al., 2016, Study 2; Goodman-Delahunty & O'Brien, 2014, Study 8) had provided "no information" on whether the researcher was blind to recidivism when collecting the predictor variables. While there were some other variations, methodological differences between the papers were minimal. This may explain why there were no consistent links between the quality rating of the studies and their findings.

All studies used an archival component in their design. This approach to data collection is useful in that it provides access to large samples and allows for longitudinal follow-up. Both these factors are essential when evaluating an outcome event that is rare, like sexual recidivism. For the studies which retrospectively scored the VRS-SO from file information, there is a risk that the scores will not accurately represent the participants. Key information might be lacking and result in someone receiving a lower risk score than they merit. This risk is mitigated in the studies which accessed scores of VRS-SO's that had been routinely collected. However, these are open to a different risk due to a lack of control over the administration of the measure. The manual of the VRS-SO should to some extent protect from variability in administration, but studies have shown difference in risk ratings when they are collected for clinical vs. research use (Boccaccini et al., 2009).

Although a strength of these studies is their large participant samples, a weakness is the small proportion of recidivists. Although all studies used AUC analyses which are considered robust to low base rates, this analysis is still open to Type II bias in practice (Babchishin & Helmus, 2016; Cooke & Michie, 2013; Rice & Harris, 1998). Sexual recidivism rates ranged from 5.7%-24.5% and violent recidivism ranged from 9.9%-45.5% of the sample. While there is not a clear guideline regarding how many participants should have an outcome for it to be considered an acceptable base rate, the PROBAST recommends a minimum number of 100 participants with an outcome event (Wolff et al., 2019). Based on this approach, none of the included studies had an acceptable base rate of sexual recidivism and only 3 had an acceptable rate of violent recidivism (Olver et al., 2016; Olver et al., 2020; Olver et al., 2007). The fact that six of these studies used a treatment sample, may have reduced the prevalence of recidivism, as treated samples may be less likely to re-offend (Piquero et al., 2012). The decision to use treatment samples was made to answer other research questions within these studies. However, future researchers may consider non-treated samples to enhance the base-rate. Though it is worth acknowledging that regardless of direct treatment

intervention, when risk is identified in offending populations it is inevitable that services will attempt to minimise that risk.

Difficulties were also introduced by the wide range of follow-up periods used. While most recidivism for incarcerated offenders occurs within the first few years after release, it has been estimated that risk of offending remains higher than for the general population at least 13 years after release (Blumstein & Nakamura, 2012). There is no gold standard of follow-up times for recidivism in research, 5 years is an accepted precedent (Conroy & Murrie, 2007). However, 13 years would likely reduce the risk of censoring biasing results. Although the studies all had acceptable mean follow-up times (between 5-12years), the range of follow-up times for individual participants was broad (2-13 years), increasing the risk of Type 1 errors, censoring and missing data. For example, a participant followed up for 2 years may recidivate at 5 years and thus be incorrectly recorded as a non-recidivist. The use of fixed follow-up times does not remove the risk of censoring but at least allows the researcher to state more concretely about the predictive accuracy of recidivism within a specific time frame.

Additionally, different studies within this review used different definitions of recidivism. Sexual recidivism was the most consistently defined and therefore the most comparable across studies. However, there were wider differences across how violent and general recidivism were conceptualised. Studies used a range of combinations of charge and conviction data to capture recidivism. The combination of charges and convictions is estimated to increase the base-rate of sexual recidivism, however that was not observed to be the case in this review. There were no clear patterns of differences between studies relating to the follow-up period or means of defining and measuring recidivism. However, it is possible that the small number of studies included within this review may be masking related effects.

PROBAST indicated that the biggest domain of difficulty for all included studies was their approach to analysis. There is a strong precedent for the use of AUCs to assess predictive validity of forensic risk assessment tools and to interpret them using Rice and Harris (1998)'s guidance. However, it is worth acknowledging that over-reliance on AUC's oversimplifies the concept of predictive validity (Cooke & Michie, 2013) and other less generous interpretations of effect sizes are available (Akonbeng, 2007). This over-reliance is further exemplified by the fact that all included studies reported on the discriminative properties of the VRS-SO scores but only Olver & Eher (2020; study 4) explored calibration. To fully understand the clinical utility of a risk assessment measure, more than one single component

(discrimination), must be understood. Van Calster (2019) highlights that a risk model with a lower AUC value, but better calibration could provide a more accurate prediction of risk than one with a higher AUC value but poor calibration. Olver and Eher (2020) observed good but non-significant calibration for the VRS-SO, describing a slight over-estimate of sexual recidivism risk. Given the importance of making accurate risk judgements, it is essential to better understand this property of the VRS-SO, before making claims of its clinical usefulness.

A key limitation for the evidence-base of the VRS-SO is introduced by the possible bias of the authors. The developers of the tool were involved in six out of the nine included studies which can increase the risk of confirmation bias. Sowden and Olver (2017; Study 7), highlighted that because of their connection to the tool and experience with it likely enhances their adherence and application of the scoring procedures. When considering clinical relevance, their administration may be above and beyond that which would be expected with less experienced professionals in practice and therefore be less generalisable to use in practice. Although the results need to be interpreted within the context of their methodological weaknesses, it cannot be ignored that the weakest results for the predictive validity were produced by two out of the three independent studies of the VRS-SO (Goodman-Delahunty & O'Brien, 2012; Todd, 2013).

The studies included in this study and the wider evidence-based for the VRS-SO are characterized by a lack of variety in terms of authors, participant samples and methodology. To better understand the efficacy of this assessment tool in practice, the literature would benefit from more independent explorations of the psychometric properties of the tool. The reduction of author bias would increase confidence generalising these promising findings to real life settings.

Review Strength and Limitations

This review started with the broad research aim to explore the predictive properties of the VRS-SO. Scoping searches iteratively narrowed the focus of this to the predictive validity of the VRS-SO for recidivism outcomes. A strength of this review was the inclusion/exclusion criteria which contained the review question and process, minimising disparities between included studies. An exception to this was the definition of recidivism. The review did not specify a means of defining recidivism. Although the definition and categorisation of sexual recidivism is comparable across studies, the differences in violent and general recidivism

impacted the interpretability of synthesised results. Additionally, the number of included studies was small, with an over-representation of studies conducted by the tool developer.

At the point of study selection, 12 studies were excluded due to sample duplication. The author decided to include data sets only once to reduce the risk of overrepresenting positive findings and based the decision of which studies to include on the length of follow up and relevance to the research question. However, some of the excluded studies may have offered a statistical advantage as they amalgamated several of the sample populations in order to create larger base-rates (Olver, Mundt et al., 2018; Olver, Sowden et al., 2018). As AUC's are less affected by base rates, the decision was made to include the samples individually. However, this may be considered a weakness of this review.

Conclusion

Overall, this review and the studies within it are representative of the literature available on the VRS-SO. The synthesis of results indicate that the dynamic components of this measure show good predictive validity for sexual recidivism, including the sexual deviance domain. This suggests that the VRS-SO could have a useful role in forensic practice as an ARAI specifically designed for incarcerated samples of sexual offenders. However, the lack of variety within the evidence-base suggests that there is a need for caution interpreting these results. More independent research, into the broader psychometric properties of this tool would increase confidence in understanding the utility of it in forensic practice.

References

- Akonbeng, A. K. (2007). Understanding diagnostic tests 3: receiver operating characteristic curves. *Acta paediatrica*, 96(5), 644.
<https://doi.org/10.1111/j.1651-2227.2006.00178.x>
- Andrews, D.A. & Bonta, J. (1998). *The Psychology of Criminal Conduct* (2nd ed.). Cincinnati, OH: Anderson.
- Andrews, D. A., & Bonta, J. (2010). Rehabilitating criminal justice policy and practice. *Psychology, Public Policy, and Law*, 16(1), 39.
<https://doi.org/10.1037/a0018362>
- Babchishin, K. M., & Helmus, L. M. (2016). The influence of base rates on correlations: An evaluation of proposed alternative effect sizes with real-world data. *Behavior research methods*, 48(3), 1021.
<https://doi.org/10.3758/s13428-015-0627-7>
- 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.
<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.
<https://doi.org/10.1037/a0022900>
- Blumstein, A., & Nakamura, K. (2012). *Extension of Current Estimates of Redemption Times: Robustness Testing, Out-of-State Arrests, and Racial Differences*. Report to the U.S. Department of Justice.
Retrieved from: <https://www.jrsa.org/webinars/presentations/cch-part2-nij-2012.pdf>
- Boccaccini, M. T., Murrie, D. C., Caperton, J. D., & Hawes, S. W. (2009). Field validity of the STATIC-99 and MnSOST-R among sex offenders evaluated for civil commitment as sexually violent predators. *Psychology, Public Policy, and Law*, 15(4), 278.
<https://doi.org/10.1037/a0017232>
- Boer, D. I., Hart, S. D., Kropp, P. R., & Webster, C. D. (1998). *SVR-20. Manual for the sexual violence risk-20. Professional Guidelines for Assessing Risk of Sexual Violence*. Burnaby: The Mental Health, Law, and Police Institute, Simon Fraser University.
- 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>
- Burghart, M., de Ruiter, C., Hynes, S. E., Krishnan, N., Levtova, Y., & Uyar, A. (2022). *The Structured Assessment of Protective Factors for Violence Risk (SAPROF): A Meta-Analysis of its Reliability and Predictive Validity*. Unpublished manuscript.
<https://doi.org/10.31234/osf.io/pjv6w>

Conroy, M. A., & Murrie, D. C. (2007). *Forensic assessment of violence risk: A guide for risk assessment and risk management*. John Wiley & Sons.

Cooke, D. J., & Michie, C. (2012). Violence risk assessment: from prediction to understanding—or from what? To why?. In C. Logan & L. Johnston (Eds.), *Managing Clinical Risk: A guide to effective practice* (pp. 22-44). Routledge.

<https://doi.org/10.4324/9780203106433>

Craig, L. A., & Beech, A. (2009). Best practice in conducting actuarial risk assessments with adult sexual offenders. *Journal of Sexual Aggression, 15*(2), 193.

<https://doi.org/10.1080/13552600902867847>

Dekkers, O. M., Egger, M., Altman, D. G., & Vandenbroucke, J. P. (2012). Distinguishing case series from cohort studies. *Annals of internal medicine, 156*(1), 37.

<https://doi.org/10.7326/0003-4819-156-1-201201030-00006>

Darjee, R., Russell, K., Forrest, L., Milton, E., Savoie, V., Baron, E., & 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 Southeast Scotland*. Report to Risk Management Authority and NHS Scotland Forensic Network. Retrieved from:

https://www.rma.scot/wpcontent/uploads/2018/04/Risk_for_Sexual_Violence_Protocol_-_RSVP.pdf.

Eher, R., Hofer, S., Buchgeher, A., Domany, S., Turner, D., & Olver, M. E. (2020). The predictive properties of psychiatric diagnoses, dynamic risk and dynamic risk change assessed by the VRS-SO in forensically admitted and released sexual offenders. *Frontiers in psychiatry, 10*, 922.

<https://doi.org/10.3389/fpsy.2019.00922>

Fazel, S., Burghart, M., Fanshawe, T., Gil, S. D., Monahan, J., & Yu, R. (2021). *The predictive performance of criminal risk assessment tools used at sentencing*. Unpublished Study.

<https://doi.org/10.1016/j.jcrimjus.2022.101902>

Goodman-Delahunty, J., & O'Brien, K. (2014). *Reoffence risk in intrafamilial child sexual offenders*. Report to the Criminology Research Advisory Council. Retrieved from: https://researchoutput.csu.edu.au/ws/portalfiles/portal/9387090/52805_Report

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

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

Hanson, R. K., & Bussière, M. T. (1996). *Predictors of sexual offender recidivism: A meta-analysis* (Vol. 4). Solicitor General Canada, Ministry Secretariat.

Hanson, R. K., Harris, A. J., Scott, T. L., & Helmus, L. (2007). *Assessing the risk of sexual offenders on community supervision: The Dynamic Supervision Project*. Ottawa, Ontario: Public Safety Canada.

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.

<https://doi.org/10.1037/a0014421>

Harris, A. J. R., & Hanson, R. K. (2010). Clinical, actuarial and dynamic risk assessment of sexual offenders: Why do things keep changing? *Journal of Sexual Aggression*, 16(3), 296.

<https://doi.org/10.1080/13552600.2010.494772>

Hawes, S. W., Boccaccini, M. T., & Murrie, D. C. (2013). Psychopathy and the combination of psychopathy and sexual deviance as predictors of sexual recidivism: Meta-analytic findings using the Psychopathy Checklist—Revised. *Psychological assessment*, 25(1), 233.

<https://doi.org/10.1037/a0030391>

Helmus, L., Hanson, R. K., & Morton-Bourgon, K. E. (2011). International comparisons of the validity of actuarial risk tools for sexual offenders, with a focus on Static-99. In D.P. Boe., R. Eher, L.A. Craig, M.H. Miner & F. Pfafflin (Eds). *International perspectives on the assessment and treatment of sexual offenders: Theory, practice, and research* (pp57-84). Wiley.

Klinge, C. (2019). Measuring change. *The Journal of Criminal Law and Criminology*, 109(4), 769.

Retrieved from: <https://www.jstor.org/stable/48572943>

Laws, D.R. & O'Donohue, W.T. (2008). *Sexual Deviance: Theory, Assessment and Treatment*. (2nd Ed.) The Guildford Press. London.

Mann, R. E., Hanson, R. K., & Thornton, D. (2010). Assessing risk for sexual recidivism: Some proposals on the nature of psychologically meaningful risk factors. *Sexual Abuse*, 22(2), 191.

<https://doi.org/10.1177/1079063210366039>

McCann, K., & Lussier, P. (2008). Antisociality, sexual deviance, and sexual reoffending in juvenile sex offenders: A meta-analytical investigation. *Youth Violence and Juvenile Justice*, 6(4), 363.

<https://doi.org/10.1177/1541204008320260>

Olver, M. (2003). *The development and validation of the Violence Risk Scale: Sexual Offender Version (VRS: SO) and its relationship to psychopathy and treatment attrition*.

Doctoral dissertation, University of Saskatchewan. Retrieved from:

<https://harvest.usask.ca/handle/10388/7922>

Olver, M. E., & Eher, R. (2020). Predictive properties and factor structure of the VRS-SO in an Austrian sample. *European Journal of Psychological Assessment*, 36(5), 748.

<https://doi.org/10.1027/1015-5759/a000551>

Olver, M. E., Klepfisz, G., Stockdale, K. C., Kingston, D. A., Nicholaichuk, T. P., & Wong, S. C. (2016). Some notes on the validation of VRS-SO static scores. *Journal of Sexual Aggression*, 22(2), 147.

<https://doi.org/10.1080/13552600.2015.1116625>

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. (2018). Using the Violence Risk Scale-Sexual Offence 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.

<https://doi.org/10.1037/pas0000538>

Olver, M. E., Neumann, C. S., Sewall, L. A., Lewis, K., Hare, R. D., & Wong, S. C. (2018). A comprehensive examination of the psychometric properties of the Hare Psychopathy Checklist-Revised in a Canadian multisite sample of indigenous and non-indigenous offenders. *Psychological Assessment*, 30(6), 779.

<https://doi.org/10.1037/pas0000533>

Olver, M. E., Nicholaichuk, T. P., Kingston, D. A., & Wong, S. C. (2020). A prospective multisite examination of dynamic sexual violence risk: Extension and update to Olver, Nicholaichuk, Kingston, and Wong (2014). *Journal of consulting and clinical psychology*, 88(4), 362.

<https://doi.org/10.1037/ccp0000478>

Olver, M.E., Sowden, J.N., Kingston, D.A., icholaichuk, T.P., Gordon, A., Christofferson, S.M.B. & Wong, S.C.P. (2018). Predictive Accuracy of Violence Risk Scale – Sexual Offender Version Risk and Change Scores in Treated Candaian Aboriginal and Non-Aboriginal Sexual Offenders. *Sexual Abuse*, 30(3), 254.

<https://doi.org/10.1177/1079063216649594>

Olver, M. E., Wong, S. C., Nicholaichuk, T., & Gordon, A. (2007). The validity and reliability of the Violence Risk Scale-Sexual Offender version: assessing sex offender risk and evaluating therapeutic change. *Psychological assessment*, 19(3), 318.

<https://doi.org/10.1037/1040-3590.19.3.318>

Ozkan, T., Clipper, S. J., Piquero, A. R., Baglivio, M., & Wolff, K. (2020). Predicting sexual recidivism. *Sexual Abuse*, 32(4), 375.

<https://doi.org/10.1177/1079063219852944>

Pithers, W. D. (1990). Relapse prevention with sexual aggressors. In W.L. Marshall, D.R. Laws & H.E. Barbaree (Eds.) *Handbook of sexual assault* (pp. 343-361). Springer, Boston, MA.

Piquero, A. R., Farrington, D. P., Jennings, W. G., Diamond, B., & Craig, J. (2012). Sex offenders and sex offending in the Cambridge study in delinquent development: Prevalence, frequency, specialization, recidivism, and (dis) continuity over the life-course. *Journal of Crime and Justice*, 35(3), 412.

<https://doi.org/10.1080/0735648X.2012.688527>

Prochaska, J. O., DiClemente, C. C., & Norcross, J. C. (1992). In search of the structure of change. In Y. Klar, J.D. Fisher, J.M. Chinsky & A. Nadler (Eds.) *Self change* (pp. 87-114). Springer, New York, NY.

Rice, M. E., Harris, G. T., Lang, C., & Cormier, C. (2006). Violent sex offences: How are they best measured from official records? *Law and human behavior*, 30(4), 525.

<https://doi.org/10.1007/s10979-006-9022-3>

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. *Law and human behavior*, 29(5), 615.

<https://doi.org/10.1007/s10979-005-6832-7>

Singh, J. P., & Fazel, S. (2010). Forensic risk assessment: A metareview. *Criminal Justice and Behavior*, 37(9), 965.

<https://doi.org/10.1177/0093854810374274>

Sample, L. L., & Bray, T. M. (2003). Are sex offenders dangerous? *Criminology & Public Policy*, 3(1), 59.

<https://doi.org/10.1111/j.1745-9133.2003.tb00024.x>

Sowden, J. N., & Olver, M. E. (2017). Use of the Violence Risk Scale—Sexual Offender Version and the Stable 2007 to assess dynamic sexual violence risk in a sample of treated sexual offenders. *Psychological Assessment*, 29(3), 293.

<https://doi.org/10.1037/pas0000345>

Stinson, J. D., Becker, J. V., & Sales, B. D. (2008). Self-regulation and the etiology of sexual deviance: Evaluating causal theory. *Violence and Victims*, 23(1), 35.

<https://doi.org/10.1891/0886-6708.23.1.35>

Todd, L. (2013). *An Independent Validation of the Violence Risk Scale: Sexual Offender Version (VRS:SO): Assessing Therapeutic Change and its Relationship to Recidivism in a Sample of SVP and SVP considered Sex Offenders*. Doctoral dissertation, University of Chicago. Retrieved from <https://www.proquest.com/dissertations-theses/independent-validation-violence-risk-scale-sexual/docview/1698250886/se-2>

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

<https://doi.org/10.1037/pas0000454>

Ward, T., & Fortune, C. A. (2016). The role of dynamic risk factors in the explanation of offending. *Aggression and violent behavior*, 29, 79.

<https://doi.org/10.1016/j.avb.2016.06.007>

Ward, T., & Hudson, S. M. (1998). The construction and development of theory in the sexual offending area: A metatheoretical framework. *Sexual Abuse*, 10(1), 47.

<https://doi.org/10.1177/107906329801000106>

Wolff, R.F., Moons, K.G.M., Riley, R.D., Whiting, P.F., Westwood, M., Collins, G.S., Reirnsma, J.B., Kleijnen, J. & Mallet, S. (2019). PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. *Research and Reporting Methods*, 170(1), 51.

<https://doi.org/10.7326/M18-1376>

Zimring, F. E., Piquero, A. R., & Jennings, W. G. (2007). Sexual delinquency in Racine: Does early sex offending predict later sex offending in youth and young adulthood? *Criminology & Public Policy*, 6(3), 507.
<https://doi.org/10.1111/j.1745-9133.2007.00451.x>

Zimring, F. E., Jennings, W. G., Piquero, A. R., & Hays, S. (2009). Investigating the continuity of sex offending: Evidence from the second Philadelphia birth cohort. *Justice Quarterly*, 26(1), 58.
<https://doi.org/10.1080/07418820801989734>

Understanding Sexual Deviance in Risk Assessment Practice: A Criminal Justice Social Workers Perspective

Abstract

Sexual deviance is a well cited and well-evidence risk factor for sexual recidivism. However, there is a lack of a clear, consistent operational definition in the literature. This likely affects the ability of assessment tools and assessors to accurately and consistently assess this in forensic practice. Using social constructivist grounded theory, this study explored the experience of criminal justice professionals assessing and understanding sexual deviance as a risk factor for sexual recidivism. Participant interviews led to the construction of a model representing how they do this. They primarily understood sexual deviance as a deviant sexual interest or preference that could act as an underlying motivator for sexual offending when facilitated by other risk factors and the opportunity to offend. This study explores how their understanding is shaped and interacts with the wider risk assessment process. Implications for future research and forensic practice are discussed.

Word Count: 16,165

Introduction

In the last 10-years, recorded sexual offences have increased by 96% in Scotland (Scottish Government, 2022). Although sexual offenders recidivate less frequently than other offenders (Hanson & Bussiere, 1998) and are more likely to recidivate non-sexually (McCann & Lussier, 2008; Rice, Harris, Lang, & Cormier, 2006; Zimring et al., 2007, 2009), the potential level of harm associated with sexual offending incites fear and demands prevention (Masters & Kebbell, 2019). Not all sexual offenders are equally likely to reoffend (Mann et al., 2010). The challenge of identifying those at higher risk of recidivism falls to criminal justice professionals (e.g., police, social work). Assessment of risk informs the allocation of resources and decision-making on intervention and management strategies (Darjee et al., 2016). Effective risk assessment requires understanding “what” factors are associated with increased risk, “why” they increase risk for an individual and “how” to assess their presence (Cooke & Michie, 2013; Monahan, 2007).

“What” are the Risk Factors for Sexual Recidivism?

The evidence on risk factors is primarily correlational, stating whether a relationship exists between a factor and sexual recidivism. Studies primarily focus on recidivism relating to in-person sexual offences, with higher representation of contact offences. There is little evidence as to how relationships function to increase risk, meaning risk factors are proxies for underlying risk processes that are not currently understood (Cooke & Michie, 2013; Ward & Fortune, 2016). Many risk factors have been explored but no one factor has demonstrated a consistent, high correlation with sexual recidivism. This has led researchers to postulate that sexual recidivism is determined by an interaction of multiple risk factors (Gannon, 2021; Mann et al., 2010; Seto, 2018; Smid & Wever, 2019).

The risk factors with the strongest evidence-base for sexual recidivism are sexual deviance, antisocial personality disorder and psychopathy (Doren & Elwood, 2009; Hanson & Morton-Bourgon, 2005; Mann et al., 2010). They are considered to have predictive value due to the strength of their correlations. Although it is well cited, sexual deviance is notoriously difficult to define. The literature refers to it as behaviour, fantasy, thought, interest, preference, orientation and/or arousal (Bartels & Gannon, 2011; Gross, 2014; Hart et al., 2003; Joyal & Carpentier, 2016; Laws & O’Donohue, 2008).

Defining Sexual Deviance. Considering the variety of terms used, often interchangeably in the literature (Smit et al., 2011), a logical starting point is to clarify terminology.

“Deviance” means to “differ from the accepted standard” (Oxford English Dictionary, 2012). The law sets a standard of what sexual behaviours are considered acceptable in society. A sexual offence may be considered a sexually deviant behaviour as it differs from the accepted standard. However, if this was a sufficient definition of sexual deviance as a risk factor, then all sexual offenders would have sexual deviance, which is not the case (Seto, 2008). Therefore, while sexual offending may be considered a sexually deviant act, sexual deviance as a risk factor must be more than committing a sexual offence.

The meta-analyses of Hanson and Morton-Bourgon (2005) and that of Mann and colleagues (2010), summarise the correlational evidence for sexual deviance in the following terms. Hanson and Morton-Bourgon (2005) referred to sexual deviance as any “deviant sexual interests, such as children, rape and other paraphilia’s, as well as sexual preoccupations”. Mann and colleagues (2010) categorised sexual preoccupations as a separate risk factor and presented three categories of evidence for sexual deviance: paedophilic preference, sexualised violence and multiple paraphilia’s. An important distinction is that between preference and interest.

A sexual interest is defined as an enduring sexual attraction to a stimulus (Hanson & Morton-Bourgon, 2005; Moser, 2016). Sexual preference is when a person’s interest in that stimulus is stronger than for other stimuli. It is their dominant, sometimes sole interest, with some likening it more to sexual orientation (Moser, 2016; Seto, 2012; Quinsey, 2003). Paraphilia is a medicalised term, referring to sexual arousal that is deviant to the societal norms. The DSM-5 (Diagnostic and Statistical Manual of Mental Disorders; APA, 2013) refers to a paraphilia as a “persistent, intense, atypical sexual arousal pattern” which can become a paraphilic disorder when associated with distress and/or offending.

The suggestion that a paraphilia is a deviation from a societal norm implies that there is an accepted understanding of what is “normal” sexual arousal. However, this assumption has been challenged by the findings that a statistically normative percentage of non-offending samples report interests, fantasy and behaviour that has been considered “deviant from the

norm”. This includes sexual interest, fantasy and behaviour involving sado-masochism, exposure, and voyeurism as well as interest and fantasy about children (Green, 2002; Joyal & Carpentier, 2017; Seto, 2008).

The most common methodology for assessing sexual deviance in research is phallometric measurement of sexual arousal to visual and auditory stimuli (Hanson & Morton-Bourgon, 2005; Mann et al., 2010). Arousal has been interpreted to indicate underlying sexual interest or preference (dependent on study design; Lalumière & Quinsey, 1994).

Paedophilia. The strongest evidence for sexual deviance as a risk factor stems from studies demonstrating that phallometrically measured sexual arousal to stimuli depicting children is the highest predictor of child sexual recidivism (Hanson & Bussiere, 1998; Hanson & Morton-Bourgon, 2005; Mann et al., 2010).

There is not a consensus on how to define paedophilia in research (Gannon, 2021). Seto (2010) defined paedophilia as a “sexual attraction to pre-pubescent children, reflected in a person’s sexual thoughts, fantasies, urges, sexual arousal or behaviour regarding children” that reflects a persistent “sexual preference”. Early studies categorised all child sexual offenders as paedophiles (Gannon, 2021; Seto, 2018). This is a common misconception held in society. Research has indicated that there are numerous pathways to sexual offending against children, not all involving a deviant sexual preference (Smallbone & Cale; Ward & Siegert, 2002).

Phallometric studies have demonstrated that sexual offenders against children generally show higher arousal to child stimuli than offending and non-offending control groups, evidencing a greater prevalence of sexual interest in children (Freund & Watson, 1991; Harris et al., 1992; Marshall, Barbaree, & Christophe, 1986; Quinsey, Steinman, Bergersen, & Holmes, 1975). However, less than half of child sexual offenders demonstrate a preference for child stimuli over adult (Seto & Lalumière, 2001; Seto, 2010). Studies on both interest and preference for children have demonstrated a relationship with sexual recidivism (Hanson & Bussiere, 1998; Hanson & Bourgon-Morton, 2005; Mann et al., 2010).

Sexualised Violence. Mann and colleagues (2010) defined sexualised violence as a sexual interest in sadism (deriving pleasure from infliction of pain, suffering or humiliation; Oxford English Dictionary, 2012) or a preference for coercive sex (persuading someone to do

something using force or threats; Oxford English Dictionary, 2012). The relationship between sexualised violence and contact sexual recidivism is less consistent than for sexual arousal for children (Kingston & Yates, 2008; Lalumière et al., 2017; Stinson & Becker, 2008).

There is some evidence that sexual arousal to depictions of sexual violence on phallometric measures is related to sexual and violent recidivism (Hanson & Morton-Bourgon, 2004; Knight & Thornton, 2007; Seto & Kuban, 1996). Lalumière and colleagues (2003) meta-analysis found rapists showed more phallometric responding to rape stimuli than non-sexual offenders in 20 out of 22 studies. This indicates higher prevalence of sexual interest. However, there has been no consistent evidence that rapists prefer violent or non-consensual stimuli over non-violent, consensual stimuli (Lalumière et al., 2003; Stinson & Becker, 2008).

Phallometric measures are less able to discriminate between rapists and non-rapists, than between child sexual offenders and non-child sex offenders (Hanson & Bussiere, 1998; Marshall & Fernandez, 2000; Stinson & Becker, 2008). Therefore, this may be a less effective assessment in this population. Some have taken the level of violence used in an offence to indicate greater arousal for violence and demonstrated a relationship to sexual recidivism (Hanson & Morton-Bourgon, 2004). However, this should be cautiously interpreted as there are multiple reasons why an offence could involve increased violence (e.g., anger, victim-resistance).

Rapists are a heterogenous group and there may be variety in the specific source of arousal in sadistic or coercive material. Heterosexual males may show some arousal to all sexual stimuli involving adult females, which could mask preference. This may contribute to the comparably weaker evidence base. Additionally, studies have not consistently used appropriate samples. Kingston and colleagues (2010) found on average, that 586 sexual offenders responded more to adult consenting stimuli than that depicting coercive sex or violence. However, only 15% of their sample were rapists (the rest were child sexual offenders).

Multiple Paraphilia's. Mann and colleagues (2010) found that having multiple paraphilias was related to sexual recidivism. Abel and colleagues (1988a; 1988b) provided the first evidence of this relationship, proposing high rates of paraphilia comorbidity. These

rates have never been replicated, potentially due to methodological issues. The term paraphilia specifically refers to arousal but instead of measuring this directly, Abel used self-report. They deviated from the DSM definition and assumed behaviour was reflective of underlying interest. Recent studies have suggested a smaller proportion, typically of higher risk sexual offenders, demonstrate multiple paraphilias (Dunsetih et al., 2004; Levenson & Morin, 2006; Smallbone & Wortley, 2004) and that this is related to increased sexual recidivism (Knight & Thornton, 2007; Prenky, Knight & Lee, 1997).

Measurement Difficulties. The evidence for sexual deviance as a risk factor is predominantly based on phallometric measures of sexual arousal. As with all assessment means, this is open to error. A portion of phallometric data tends to be invalid due to technical problems, patterns of arousal that cannot be interpreted and false responding (Seto & Lalumière, 2001; Stinson & Becker, 2008). There is a social pressure not to demonstrate arousal to sexual offending stimuli (Lalumière & Quinsey, 1994). Some men may be able to suppress their penile response when looking at stimuli that arouses them (therefore masking interest) or generate a penile response when looking at stimuli that does not arouse them (therefore masking preference). Renaud and colleagues (2009) explored this and found that 80% of participants could voluntarily control their erectile responses during phallometric measurement. As discussed, falsifying results may be easier in studies using rape stimuli as there are likely to be broader differences in interest and a shared gender/age across depictions. A lack of standardisation across procedures has opened phallometric measures up to criticism (Conroy & Murrie, 2007; Smid et al., 2011).

Paraphilia diagnoses are intended to capture deviant patterns of sexual arousal. However, paraphilic diagnoses do not typically rely on objective assessment tools. They are based upon criteria presented in the DSM-V (APA, 2013). This was not developed for forensic practice but has been used for the purpose of risk assessment as relates to sexual deviance (especially in the USA; First, 2014; Harris, Boccaccini & Rice, 2017). This practice has been criticised for demonstrating poor inter-rater reliability as DSM criteria are vulnerable to subjective interpretation (Doren & Elwood, 2009; Laws & O'Donohue, 2008; Smid et al., 2011). The primary difficulty with this use of the DSM is that criteria significantly change across editions (First, 2014). The evidence based on one edition of the DSM, can only be cautiously

applied to other editions. This creates significant implications for using this to assess sexual deviance in forensic research and practice.

“Why” does Sexual Deviance increase risk?

The research has made substantial progress since the 1980’s to understand “what” factors are associated with increased risk of sexual offending. However, the answer to the question “why” these factors increase risk for an individual remains elusive (Cooke & Michie, 2013; Ward & Fortune, 2016). Despite widespread agreement that sexual recidivism is multiply determined, many theories of sexual offending have been single factor theories (Ward, 2014). Those that attempt a comprehensive explanation of sexual offending tend to emphasise the role of sexual deviance and inhibition (Smid & Wever, 2019).

Seto (2008; 2018) presented the motivation-facilitation model of sexual offending. This presents the hypothesis that there is an underlying motivator to sexually offend which takes the form of sexual deviance (defined as a paraphilic interest/preference) or hypersexuality. Acting on this motivation is facilitated by reduced self-control or behavioural inhibition in a situation where there is an opportunity to offend.

Smid and Wevers (2019) also considered how sexual deviance and behaviour disinhibition could interact to lead to offending behaviour in the incentive-motivational model of sexual deviance. They conceptualise sexual deviance as a sexual interest or preference that can motivate an action towards offending behaviour. However, when paired with adequate behavioural inhibition, they believed an individual will not act on this. This could explain the prevalence of interest within the general population (Green, 2002; Joyal & Carpentier, 2017). They further indicate that sexual preference may provide a stronger drive towards action as phallometric studies of preference not only demonstrate a high arousal response to sexually deviant stimuli, but a low arousal towards stimuli that is normative or acceptable. Therefore, they may be less able to experience sexual gratification in other areas.

Smid and Wevers (2019) further posited that sexual deviance may develop through the process of excitation transfer (Zillmann, 1996), where a strong emotion coinciding with sexual arousal may enhance the feeling of arousal and the source of that emotion is attributed as the source of arousal (e.g., having one’s genitals seen could be associated with a sense of shame that could enhance arousal which would then be attributed to exposure/exhibitionism). The sexually deviant interest is thereby formed and reinforced through operant conditioning.

While this theory is primarily rooted in social learning theories, they acknowledge a likely genetic/biological component and hypothesise that those with lower arousal to normative/acceptable sexual stimuli are more susceptible to this process.

Gannon (2021) more specifically considered how sexual deviance in the form of paedophilia could develop and be maintained in the compositional explanatory theory of paedophilia (CEToP). Interestingly, she suggested that sexually deviant interests and sexually deviant preferences could form via different developmental pathways with interest formed more by environmental factors and preference more by biological processes. She indicated that both could be motivators for sexual offending behaviour in the context of other risk factors and opportunity. Importantly, her theory suggests sexually deviant interests are more malleable and could be more receptive to intervention. The research on sexual interests and preferences suggests implies that they are part of the same construct. However, considering them as distinct underlying process that result in similar observable behaviours could explain some of the inconsistencies in the evidence-base.

These theoretical conceptualisations of the role of sexual deviance in sexual offending suggest that sexual deviance (as a sexual interest or preference) serves as a motivator towards sexual offending. However, other factors must be present for this motivator to be acted upon. Specifically, disinhibition and opportunity to offend. These theories consider sexual interests as less likely to lead to action as individuals are likely to have other, legal sexual outlets. Additionally, it has been hypothesised that sexual preferences have a greater biological/genetic component which means they are less malleable to change than sexual interests. These theoretical accounts present hypothesised causal relationships in the onset and maintenance of sexual offending that can be generalised across offence types (contact, non-contact, internet). However, these have been based upon correlational evidence, for sexual recidivism (not onset) almost exclusively in contact-offending samples.

Sexual Deviance as a Risk Factor

The evidence of sexual deviance as a risk factor for sexual recidivism is based on arousal to stimuli that are considered unacceptable by society. This arousal is taken to indicate a sexual interest or preference (when the arousal is stronger to one stimulus over others). The deviant interest or preference can involve an unacceptable focus (e.g., a child, or object) or an unacceptable act (e.g., infliction of pain, coercion). Sexual deviance by this definition is associated with higher rates of sexual recidivism in sexual offenders.

While there is no clear consensus regarding the focus of sexual deviance, the DSM and other risk assessment measures suggest that a sexually deviant interest or preference is one that relates to causing harm or distress to others or oneself (APA, 2013; Hart et al., 2003). Research indicates there is not a consistent relationship between unusual, harmless/legal sexual interests (e.g., shoe fetish, transvestism) and sexual offending, which offers some support for this (Mann et al., 2010).

While phallometrically measured deviant sexual interests have been found to relate to increased sexual recidivism, the evidence into paedophilia suggests that a sexual preference is a stronger predictor. This fits with hypothesised explanations of sexual deviance as a motivating factor for sexual offending (Gannon, 2021; Seto, 2018; Smid & Wever, 2019). However, theories highlight that the presence of sexually deviant interests/preferences does not guarantee that someone will offend. Instead, they hypothesise the need for other facilitating and situational risk factors.

Those with a sexual preference may find it harder to gain that reward through other non-deviant stimuli and thus be more likely to act. Due to the stronger relationship between deviant sexual preference and sexual recidivism, Smid and colleagues (2011) recommended considering sexual deviance as referring to sexual preference. However, the research into multiple paraphilia's may contradict this as it suggests co-existing deviant sexual interests can also be a risk factor with a relationship to sexual recidivism. In order to interpret these two streams of evidence, it would be beneficial for further studies into the role of multiple paraphilia's using comparable methodological approaches (e.g., phallometry).

The theories which suggest sexual deviance is shaped by environmental factors and theories of learning (Gannon, 2021; Smid & Wever, 2019), indicate that sexual deviance may be a malleable risk factor open to change. Those which suggest a biological underpinning, indicate that sexual deviance is a stable and enduring factor (Seto, 2018). In practice, it is predominantly viewed as a stable and enduring factor, which is unchangeable (Dempster & Hart, 2002; Hanson & Harris, 2000; Seto 2010). However, associated risk may be managed indirectly through other risk and protective factors (Mann et al., 2010; First, 2014).

“How” is the risk of recidivism assessed?

Many means of assessing sexual deviance have been developed, a full summary of which is beyond the scope of this paper (see Akerman & Beech, 2012; Smit et al., 2011). This section

focuses on those most used in UK forensic practice. Aside from phallometric assessment and paraphilia diagnoses, the meta-analyses (Hanson & Bussiere, 1998; Hanson & Morton-Bourgon, 2005; Mann et al., 2010) report evidence based on self-report, offence history and risk assessment tools.

Self-Report. Abel and colleagues (1988) highlighted that the best source of information about a person's sexual deviance is the person themselves. However, societal pressures discourage discussion of both normative and deviant sex (Abel et al., 1998; Seto, 2010). Quality of self-report may be affected by cognition, insight, self-awareness, and the way in which it is gathered (Stinson & Becker, 2008). In an interview context better quality self-report was associated with a positive rapport and an experienced interviewer (Abel et al., 1998). However, this takes time, skill and is open to subjective differences. Although numerous tools to support self-report have been developed, it is generally not recommended due to the difficulty validating any information gathered (Smid et al., 2011).

Offence History and Behaviour. Research has suggested that previous sexual and violent offending is correlated with sexual recidivism (Conroy & Murrie, 2007; Dempster & Hart, 2002; Seto, 2010; Quinsey et al., 2006). After phallometry, a history of sexual offending is one of the strongest factors correlated with sexual recidivism (Hanson & Bussiere, 1998). Accessibility and affordability make this a more appealing assessment approach. Records of past offending can be less subject to falsification than self-report (e.g., criminal record). though there are likely limits to the amount of information available (Stinson & Becker, 2008). Despite its common use in practice, relying upon historical information as an assessment of sexual deviance requires the inference of underlying sexual interest/preference from behaviour. While there is evidence for the use of past behaviour to predict sexual recidivism, sexual offending does not always have underlying sexual deviance and therefore there is risk of incorrectly inferring sexual deviance (Seto, 2008; Stinson & Becker, 2008).

Risk Assessment Tools. There is a consensus that the best approach to risk assessment should be evidence-based and structured to minimise subjective biases (Cooke & Michie, 2013; Craig & Beech, 2009, Monohan, 2007). However, the best method for this is debated. Historically, risk assessment was based on unstructured clinical judgements. Assessments were guided by individual/team experience and intuition which meant they lacked transparency, replicability and were little better than chance (Menzies et al., 1994). Two types of instruments have been developed in the endeavor to facilitate a structured, evidence-based process: actuarial risk assessment instruments and structured professional judgement tools. In assessment, risk factors for sexual recidivism are commonly dichotomized as either static or dynamic (Andrews & Bonata, 1998). Static factors are considered unchangeable enduring factors (e.g., age, number of past offences). Dynamic factors are considered more open to change, with acute dynamic factors considered the most fluid (e.g., intoxication) and stable dynamic factors more resistant (e.g., offence-supportive attitudes).

Actuarial Risk Assessment Instruments. Actuarial risk assessment instruments (ARAI's) aim to predict the likelihood of an individual re-offending by comparing them to a normative group. Assessors rate the presence of a number of empirically based risk factors. An algorithm combines these ratings to produce a total score and estimated level of risk (Darjee et al., 2016; Harris & Hanson, 2010), typically presented as a percentage likelihood of reoffending within a time period. ARAI's often rely upon static factors (e.g., Static-99, Harris et al., 2003), though some dynamic tools have been developed (e.g., Stable-2007, Hanson et al., 2007; Dempster & Hart, 2002). The scores of actuarial assessments have demonstrated a significant relationship with sexual recidivism (Eher et al. 2011; Hanson & Morton-Bourgon, 2009; Seto, 2005). This is taken to indicate the predictive accuracy of these measures and led to their wide-spread use in forensic practice (Khiroya, Weaver, Maden, 2009). However, there are several points of controversy.

ARAI's rely on the assumption that an individual offender will behave in a similar way to other offenders with similar characteristics (Hanson, 2000; Judge et al., 2014). The predictive accuracy of ARAI's reduces the more an individual deviates from the normative sample. This has been evidenced for offence-type (especially non-contact offences; Bartosh et al., 2003; Rettenberger & Eher, 2007) and ethnicity (Långström, 2004). While group data provides an indication of the likelihood of recidivism occurring within a similar group, it provides little

information about the risk of an individual (Berlin et al., 2003; Darjee et al., 2016). The presentation of risk as a numerical score can give a false sense of confidence in the accuracy of actuarial tools (Cooke & Michie, 2013). Several researchers have suggested that ARAI's do not clearly present the true margins of error associated with these instruments (Hart, Michie & Cooke, 2007). Using different statistical methods to analyse the development data, these researchers found that confidence intervals of the static-99 were too wide to allow meaningful interpretation from an individual score (e.g., 95% confidence-prediction interval of 0.00-0.99% on one item). Proponents of ARAI's have criticised their methodology (Craig & Beech, 2009; Harris, Rice and Quinsey, 2007). Nonetheless, they raise important points about the sources of uncertainty in attempts to predict individual behaviour from group data.

The attraction of ARAI's lies in their apparent ease of administration, interpretation, cost-effectiveness, and minimisation of bias (Craig & Beech, 2009). However, studies and anecdotal accounts have found that different tools can provide conflicting outcomes, which can be difficult to interpret (Cooke & Michie, 2013).

Structured Professional Judgement Tools. Structured Professional Judgement (SPJ) tools also include the objective rating of a series of evidence-based risk factors but add a more qualitative focus on the individual. The priority is to formulate risk of recidivism and provide clinically useful information for risk management (Darjee et al., 2016; Harris & Hanson, 2010). Meta-analyses have indicated that when comparing the evidence for predictive accuracy for sexual recidivism, these tools are similarly efficacious (Hanson & Morton-Bourgon, 2009). Some have suggested that the evidence of ARAI's is more robust than for SPJ tools (Craig & Beech, 2009). This is likely because the literature tends to focus on predictive accuracy (Hart 1998; Douglas & Kropp, 2002). ARAI's provision of total scores lends themselves more naturally to this form of analysis (Darjee et al., 2016; Heilbrun, Douglas, & Yasuhara, 2009). Nonetheless, SPJ tools are widely incorporated into forensic practice and have been reported to provide more practical utility than ARAI's (Green, Carrol & Brett, 2010; Khiroya, Weaver & Madden, 2009).

The proponents of ARAI's have suggested that the qualitative nature of SPJ tools increases the risk of subjective bias. However, moderate interrater reliability has been found for SPJ tools in forensic practice (Darjee et al., 2016; Hart, 2003). While research studies find higher inter-rater reliability for ARAI's, there seems to be greater variability in practice. A Scottish found agreement of item ratings on the Stable-2007 varied from 40-100% (Risk Management

Association; RMA, 2013). This study indicates that there is room for subjective interpretation in the use of both tools. The difference is that SPJ's have been designed specifically to manage this subjective process. All assessment contains scope for error. The importance is to understand and estimate the size of that error in order to consider its influence when making decisions (Craig & Beech, 2009; Laws & O'Donohue, 2008).

With regards to sexual deviance, ARAI's such as the Stable-2007 (SA07; Hanson et al., 2007) define sexual deviance as a sexual interest. Its presence is assessed through observable items that correlate to sexual deviance (e.g., relationship to victim; victim age; victim gender). This approach provides an estimate of the likelihood that an offender has sexual deviance based on traits they share with individuals who have been deemed to have sexual deviance. It does not assess the sexually deviant interest of the individual (Smid et al., 2011). SJP's, such as the Risk of Sexual Violence Protocol (RSVP; Hart et al., 2003) tend to adopt a broader conceptualisation of sexual deviance as evidenced by behaviour, interest, preference, fantasy, arousal (etc.). It encourages a formulation of the different facets of the offenders sexual and offending behaviour to build an understanding of how they may indicate sexual deviance, how this affects risk of recidivism and how this could be managed.

Current Study

Sexual Deviance is an important risk factor for sexual recidivism. However, it is inconsistently defined in the literature and there is not a gold-standard assessment tool available in forensic practice. The lack of a clear, consistent, operational definition creates a gap between research and practice (Harris, Boccaccini & Rice, 2017) and poses a challenge to those responsible for assessing sexual deviance. Given the significant individual and societal consequences associated with inaccurate risk assessment (Andrews & Bonata, 1998), it is important to understand and optimise practice in this area.

The present study aimed to explore how the research on sexual deviance is being translated into practice and identify any discrepancies. To achieve this, the researcher used grounded theory to explore how Criminal Justice Social Workers (CJSW's) understand and assess sexual deviance as a risk factor for sexual offending. Grounded theory is a qualitative methodology useful for exploring constructs which are poorly understood. It facilitates a detailed description of participants experience while also allowing a deeper exploration of the processes which shape this experience. Therefore, this methodology allowed the researcher to identify not just if there are discrepancies between literature and practice, but to identify how

they are formed and maintained. The overarching goal of this study was to introduce the voices and experiences of those tasked with assessing sexual deviance into the literature and direct future research towards meeting the needs of those in practice.

Methods

Ethical Considerations

This study was approved by the Ethical Committee of the University of Edinburgh and the Local Authority Research and Development Offices for Edinburgh Council (see Appendix 3). Participants volunteered for the study and received written information in advance as to the planned procedure, data management and their right to withdraw from the study (see Appendix 4).

Design

Grounded theory is an inductive, qualitative methodology. It involves an iterative process of data collection and analysis to inform the development of a theory or model intended to enhance understanding of a poorly understood area (Charmaz, 2006). Since its first conception (Glaser & Strauss, 1967), there have been many adaptations of grounded theory (Kenny & Fourie, 2015; Lauridsen & Higginbottom, 2014). This study adopted the Social Constructivist Grounded Theory method (Charmaz, 2006) which recognises reality as constructed through interactions within a context. It explores subjective meanings through interaction, rather than seeking an objective truth. The researcher is recognised as an active participant in the construction of meaning. Therefore, the methodology incorporates reflexive practice and explicative methods which enhance the quality of the study.

Context

Participants. Recruitment for this study took place during 2020 when the UK was in lockdown due the Covid-19 pandemic. Therefore, the participants (and researcher) were living in a time of uncertainty and change, affecting work and personal lives. The participants were working from home for the first time which was accompanied by practical and emotional challenges. Aspects about working in this context were explored in the interviews and facilitated reflection on the processes and practices they valued. Several participants spoke to changes within their field over the last 10 years. They referred to changes in the way sex offenders were approached and managed by their team. They referenced a reduction in co-working for sex offender cases, except when especially complex. This change was generally attributed to a reduction in available resources. However, it was accompanied by

their own increase in experience and confidence lone working. They reflected on the role of social norms and public perception of their work, which will be discussed further in the results section.

Researcher. The lead researcher was a Trainee Clinical Psychologist, undergoing a 3-year training programme. Their research skills and clinical experienced increased over the period of this project. The researcher had a long-standing interest in forensic psychology. However, at the time of interviews, had no previous experience working with or researching sexual deviance, sex offenders or risk assessment. This provided some naivety which is desirable in a grounded theory study, allowing the participants to shape the researchers view. However, the researcher held some implicit biases and assumptions about sexual deviance and the risk assessment process which were revealed through reflexive memos.

The researchers' view of the constructs and processes involved in risk assessment reflected the view of a novice, in that they lacked nuance. Their initial assumptions were relatively black and white. For example, they initially viewed the participants as a homogenous group and did not anticipate the breadth of individual differences that would exist within the sample. From the initial literature review, they had developed the expectation that participants would be either unsure about how to define and assess sexual deviance or would describe following a clear, rigid procedure of what to assess and how (i.e., a locally used definition and interview protocol). There was a limited expectation that this could vary with experience, confidence, or interest.

Similarly, the researcher's concept of risk assessing sexual offenders considered risk factors as independent constructs - each individually assessed and individually associated with risk. This mirrors how the evidence for risk factors is presented in research but does not reflect the interactive process of risk assessment practice. The initial literature review shaped the expectation that participants would most commonly talk about behaviour, with a specific focus on paedophilia and child sexual offences. This is reflective of a broader assumption that the focus of these interviews would be on risk assessing contact sexual offenders.

Recruitment

Criminal Justice Social Workers (CJSW)'s were recruited through their local authority areas. Department managers facilitated recruitment by circulating emails containing study information. The researcher attended team meetings (held on Microsoft Teams) to increase

awareness and answer questions. Participants could opt-in to the study via email or phone. Participants were included if they were a CJSW, with a minimum of 1-year work experience and able to participate in a 1-hour interview. The only exclusion criterion was that the participants had to be English speaking. Sample sizes were not predetermined as the aim was to saturate theoretical concepts, not stop at an arbitrarily determined standard.

Grounded theory typically uses theoretical sampling, where participants can be selected for interview based on certain traits identified as relevant to interpretation of the data (Breckenridge & Jones, 2009). To facilitate this, participants were asked to provide information on their demographics and work experience when they opted in.

Participant Demographics

The participant sample of this study consisted of CJSWs working in a Scottish Urban area. Ten participants (6 male, 4 female) took part in initial interviews. Four participants (2 male, 2 female) participated in a second respondent validation interview. The average age was 46.3 years old (range of 38-60 years old). On average participants had 10 years of experience in the role (range of 2.5-13 years).

Materials

A semi-structured interview was used to explore the participants' understanding and assessment of sexual deviance. The interview schedule consisted of 4-6 planned, open ended questions which allowed the researcher to fully explore their responses. Table 1. presents examples of interview questions.

Table 1
Examples of Interview Questions

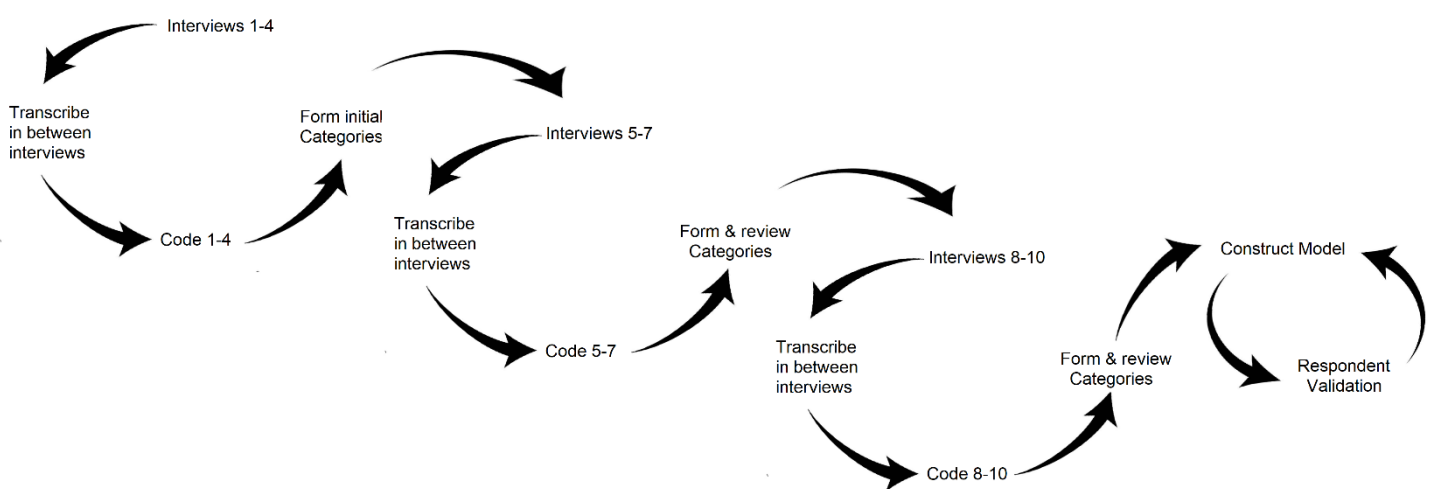
| Concept | Interview Question |
|-----------------------|---|
| Overt definition | <i>"If you had a new colleague who has never worked in this area before, how would you explain sexual deviance to them?"</i> |
| ARAI's | <i>"You mentioned the tools you use give you something but not necessarily the whole picture, what do you feel like they give you when it comes to risk assessing?"</i> |
| Process of assessment | <i>"I'm wondering, what helps you make those judgements?"</i> <i>"Can you try to talk me through how you would try and make sense of that?"</i> |
| Ending Question | <i>"Has anything else come up for you that we haven't had time to talk about?"</i> |

The interviewer also asked participants to expand on their answers (“Can you tell me more about that?”, “How do you make sense of that?”) or to provide summary questions with the aim of checking their understanding of responses (“So from what you have said...”, “Just to check I’ve picked you up right...”). This facilitated elaboration of concepts from participants. Due to the iterative nature of grounded theory, the interview questions were able to be added and removed in response to the developing interpretation of participant data.

Procedure

When participants opt-ed in, they were sent information regarding the purpose, procedure and ethical precautions taken in the study (Appendix 4). They were offered a choice of interview time and method. All participants chose to engage via Microsoft Teams - a secure, video-conferencing system, which was familiar to them. Before the interview, the researcher discussed the procedural and ethical information with the participants and provided the opportunity for questions. Participants were aware they could request the recording to be stopped, they could take a break or withdraw from the interview at any time. Interviews lasted between 45-90minutes and were recorded on an encrypted Dictaphone. After the interview participants were offered a debrief and the opportunity to ask further questions. They were also invited to opt-in to be contacted to participate in a further interview as a means of respondent validation.

Figure 1
Iterative process of data collection and analysis.



In grounded theory, data collection and analysis occur in an iterative process. Each interview was transcribed within 3 days of recording by the researcher. As shown in Figure 1, analysis

occurred in 3 chunks which allowed the interpretation of the data and the development of the theory to inform the interview process, which further fed into the construction of the theory.

Data Analysis

Following the protocol for analysis described by Charmaz (2006), the first author immersed themselves in the data by conducting and transcribing each interview. Line-by-line coding was used for the first 7 interviews (See Appendix 5). These initial in-depth codes were grouped into focused codes, which informed broader categories of constructs. Interviews 8-10 were coded first using the focused codes and then revisited to ensure that no new, relevant concepts were missed. In line with the constant comparative method (Glaser & Strauss, 1967), the first author revisited each transcript and set of codes after each new round of coding, to ensure relevant data was not lost as understanding developed. The process of comparison was facilitated by use of reflexive memos as well as the iterative cycle of tasks. Categories and models were cross validated in three ways, 1) by a literature review conducted after the coding process, 2) by a second coder and 3) by respondent validation.

Quality and Rigour

Table 2. demonstrates the quality guidelines reviewed by the researcher and their incorporation into study design. Fit and credibility was protected by immersion in the data (familiarisation) and the constant comparative method. A second qualitative reviewer (with no experience in forensic psychology or risk assessment practice) coded 10% of participant interviews to allow the researcher to explore whether their codes were true to the participant data. The resonance of the final model was explored through respondent validation interviews, lasting approximately 30-minutes. The interviewer explained the purpose of the interviews to dispel socially desirable responding. They then presented the theoretical model in four parts which corresponded with the categories formed. Participants had the opportunity to comment throughout and then some follow-up questions were asked to allow a deeper understanding of less saturated codes.

Table 2
Guidelines for Quality and Rigour in Qualitative Studies

| Glaser (1978) | Charmaz (2006) | Hammersley (2012) |
|---|--|--|
| Fit How well does the theory fit the codes? | | |
| | Credibility How well does the theory fit the data? | Consistency of Theoretical Claims to Data |
| | Resonance How well does the theory fit the experience of participants? | Credibility to those studied |
| | | Reflexive Is there reflection on the researcher's interaction with the data? |
| Workability How well does the theory explain/predict behaviour? | | Degree of Substantive theory produced |
| Relevance Does the theory have a clear focus? | | |
| | Usefulness Does the theory have an impact in practice? | |
| | | Transferable Is the theory transferable to other samples? |
| Modifiability Is the theory open to further modifications as new data is collected? | | |
| | Originality Has the theory added new insights to the area? | Novelty of Claims |

The researcher kept detailed, reflexive memos which documented the development of the researchers understanding and decision-making. These memos facilitated reflection on changes in understanding and interpretation as new data was collected and analysed. They also kept detailed notes of supervision meetings and other related discussions to explore how different sources of information might have an influence. Memos and supervision notes were reviewed at each stage of analysis (See Appendix 6).

Memos additionally assisted the research to hold the relevance and scope of data collection in mind. The researcher was able to identify when they were being drawn to gather data which was less relevant to the research aim and developing theory. Keeping a narrower focus was felt to be important given the time-period of this project, as it increased the chances of saturation of concepts and therefore, the substantive nature of the theory produced. A more

substantive theory is typically more able to explain behaviour in the participant sample, to be generalisable to other samples and to therefore have clinical utility.

The researcher has aimed for transparency in their presentation of results to allow these markers of quality and rigor to be assessed by readers of this paper.

Results

Participant interviews led to the development of a theoretical model which represented two intertwined processes, as depicted in Figure 2. These were 1) the process of understanding what sexual deviance was (the smallest circle) and 2) the process of assessing for sexual deviance (the larger circle).

As is demonstrated in Figure 2, the process of understanding sexual deviance occurred within the wider processes of assessment. In planning this study, how participants understood sexual deviance and how they assessed it were viewed as two separate questions. However, the understanding of each of these processes emerged iteratively. Therefore, to view these as separate would not provide a true representation of the process in practice.

An important caveat when reviewing any grounded theory model is the level of saturation and completeness. Ideally, interviews continue until every focused code is fully saturated and therefore the model can be viewed as complete. However, in the reality of research practice, saturation is often stated not evidenced (Mason, 2010). The intertwined nature of the processes in this model indicates that they likely take place within the context of other processes associated with risk assessment which were beyond the scope of this study to explore. Thus, the researcher does not claim this theoretical model is an inclusive representation of every facet of these processes. However, it does present key aspects of these processes that were given a shared weight by participants.

Figure 2

A Diagrammatic Representation of How Criminal Justice Social Workers Understand and Assess Sexual Deviance as a Risk Factor for Sexual Recidivism

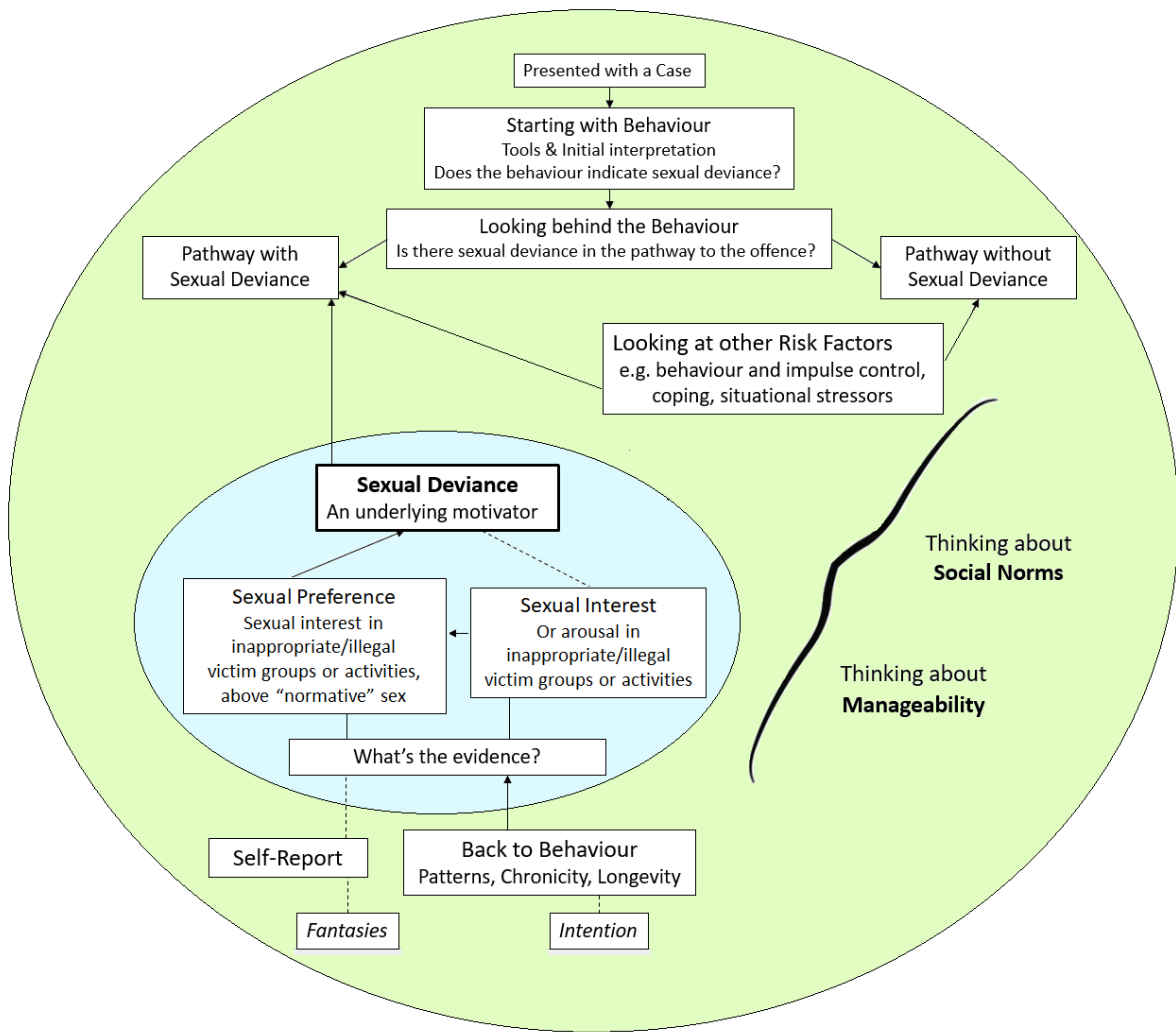
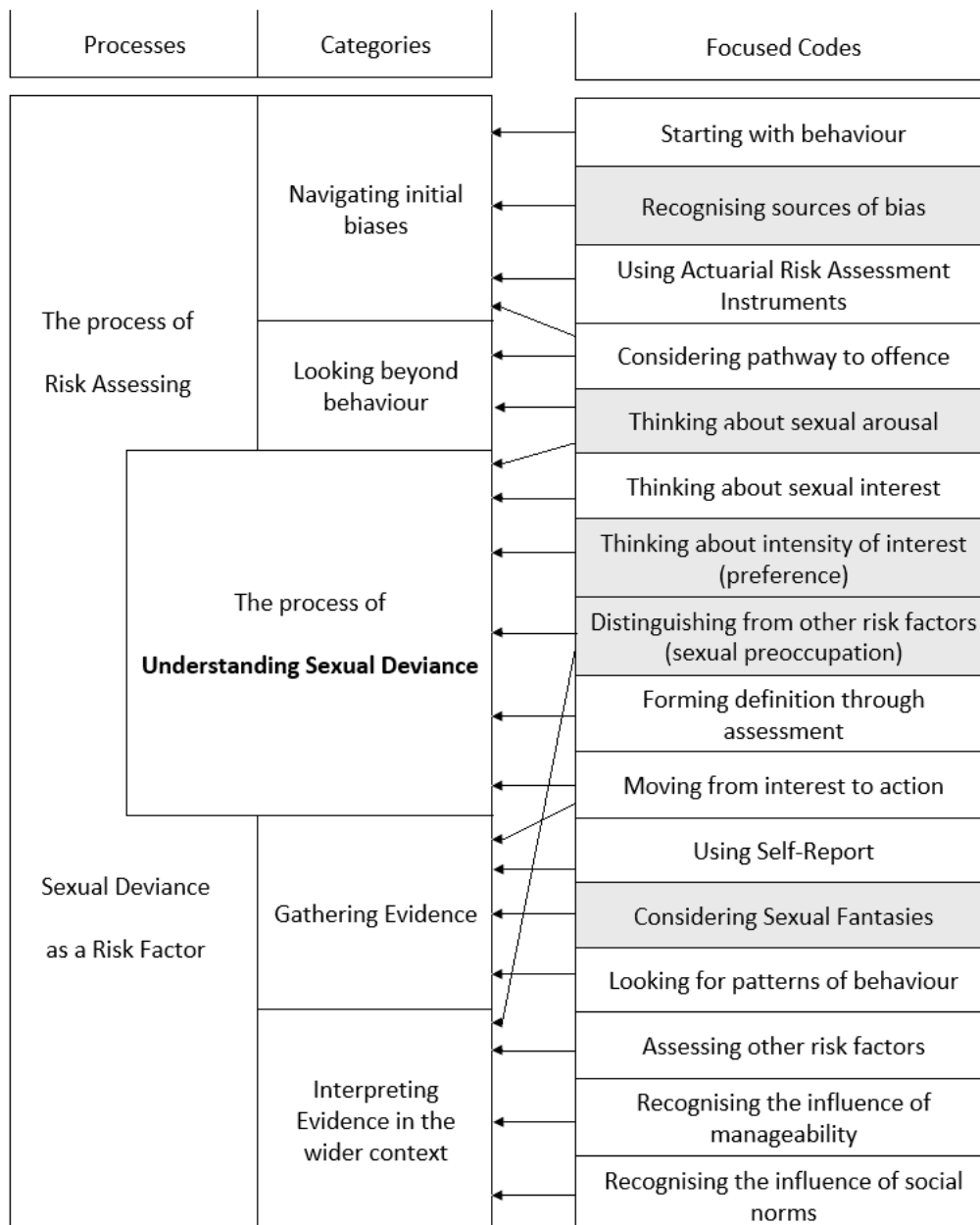


Figure 3 demonstrates the focused codes and categories which informed the development of this model. Promising but not fully saturated codes are shaded in grey. The evidence of each of these focused codes and categories will be presented in this section before a summary of the overarching model.

Figure 3
Focused Codes and Categories underlying the Model



Navigating initial Biases

Starting with behaviour. Behaviour was explicitly presented as the starting point for a risk assessment and consideration of sexual deviance. This is intuitive as the first information presented to the participants relates to the offender's behaviour (I.e., their offence). As Quinn said, “*you’re starting with an offence that suggests deviance.*”

Participants expressed that the offender's behaviour provided the first indications of the presence of sexual deviance. Especially when this behaviour was illegal and contravening social norms.

Billie: *"I think in terms of bestiality, thing that we're saying is this is just so far out of our social norms, em, or incest I suppose is the other one where you're saying, this is just completely outside of social norms... it's the behaviour that is the issue there."*

The participants shared that their assessments often begin with the use of the actuarial tools Risk Matrix 2000 (RM2000) and the Stable-2007 (SA07). The SA07 specifically considers sexual deviance in terms of behavioural indicators (e.g., number of times they have offended, selection of victims).

Alex: *"I suppose the way we're trained to think about it with the SA07 is in terms of sexual offending against prepubescent children.... [Using the SA07] you also might take into account their sexual interests... whether [they are normal/abnormal]. But it's not as great a concern with assessing deviance in terms of child victims."*

Participants identified an important distinction between illegal behaviour and unusual behaviour. There was a broad consensus that behaviour had to be considered as more than just "unusual" to qualify as sexual deviance. Davi highlighted, *"there are things we're going to be scoring people on [in the tools] that potentially lots of adults do in their own time and you know, it isn't necessarily considered deviant because it's healthy, consensual"*.

Participants felt that consent played an important role in determining what behaviours were considered concerning. It was generally acknowledged that unusual but legal sexual behaviours were not evidence of sexual deviance. However, several participants highlighted that these could inform a fuller picture of an offender's sexual profile.

Frankie: *"It's been kind of drummed into me from early on "pay attention to unusual sexual interests" that might be legal but that might suggest that somebody just has a very broad church of interests that might also include children for example or other illegal activities."*

Participants described behaviour as an important starting point for the risk assessment of sexual deviance. This initial focus on behaviour seemed to be reinforced by the focus on the ARAI's. Almost all participants began interviews by referring to sexual deviance in

behavioural terms. However, even at the early stages of interviews, there was recognition of nuances as to what constituted a deviant sexual act.

Recognising sources of bias. The reference to abnormal or unusual sexual behaviours, interests and preferences implies the existence of a “normal” alternative. Participants highlighted the influence of social norms in shaping what is legally and colloquially viewed as sexually deviant behaviour and referenced this could change across cultures and time. Charlie emphasised cultural norms, *“I think there are things where you could argue that things have limited harm but they are just seen as completely unacceptable by our society, but other societies view them as acceptable”*. Eidann explored how norms can change with age and generation *“one thing that we don’t have a handle on, or I don’t have a handle on, is what is normal sexual behaviour for teenagers at the moment”*.

Within this discussion, participants recognised that their life experience and the norms they were exposed to, meant that they held their own view of what was normal or abnormal. It was felt that this had the potential to bring bias into their assessment. Alex highlighted that, *“you can’t help but think about your own sexual interests when you make these assessments and try to make some sense of it”*. Sam shared that *“you do make these unconscious biases.”*

Participants recognised that there were other sources of bias that could affect their job roles (including developing a professional relationship with the offender). However, bias was primarily explored in relation to how it affected participants understanding of sexual deviance and their assessment. Charlie emphasised the importance of this, *“Because really, you’re wanting to be thinking about, what suggests risk to other people or risk in future offending not ‘do we think it’s weird what people get up to in their private time’*.

These early discussions indicated that social norms were important in shaping legal decisions about what behaviours are considered socially acceptable/unacceptance. The implication is that the word deviance is referring to something that is socially “deviant”. However, there was recognition in discussion about the differences between “unusual” and “illegal” acts, that sexual deviance as a risk factor has to be more than just something that violates social norms. Participants recognised that social norms paired within individual differences (e.g., age and experience) can create differing views of what is “normal” sexual behaviour. Participants highlighted that this could create a bias which could mislead assessments. Therefore, they

referred to various methods for checking their bias, including drawing on the experience of their colleagues, grounding their decisions in research and using actuarial tools.

Using the Actuarial Risk Assessment Instruments. The use of actuarial tools (ARAI's) was the most common way of checking initial subjective thoughts/assumptions. This is likely because their use is an embedded part of the assessment process. Allowing participants to reflect on their biases was viewed as one of the primary benefits of using ARAI's.

Alex: *"I suppose that's one of the big reasons why risk assessment tools are helpful. They do a lot of things but one of the big things is that they help us to not make an assessment based on our own biases."*

Several other benefits were attributed to the use of these tools, including that they provide a starting point for conversation with offenders as mentioned by Charlie, *"I think it's good as it gives you broad areas of things to think about, things to talk about with offenders."* Participants also highlighted they create a shared language with other professionals.

Frankie: *"we wanted to have common language around risk assessment, intervention and needs... If some of that common language is a bit dodgy [shrugs]. It might not be 'sexual deviance' when we use the words sexual deviance as defined by SA07, but I still think it's really helpful to have [that shared language], however ill-defined it is."*

As demonstrated by Frankie, despite these benefits, there was a sense that these tools were not capable of fully capturing sexual deviance as a risk factor.

Billie: *"the risk assessments we have are mainly about who the victims were rather than about what was going on for the offender... it's really categorical, doesn't have any nuance, doesn't care about what they were doing really. So... I don't really feel like it captures things particularly well. So, I think RSVP has more scope for it, but it's not an assessment that we would be doing."*

Furthermore, participants provided examples where the tools had increased the complexity of a risk assessment by presenting conflicting outcomes.

Davi: *“we use the RM2000 and the Stable and they look at very different things. I think my biggest issue is that they often come out showing wildly different things. It’s not uncommon for somebody to be scoring low on the RM2000 but actually moderate-high on the Stable”*

Therefore, while participants viewed the use of these tools as helpful and necessary for grounding bias and creating a shared language across professions, they felt that they needed to be part of a more comprehensive risk assessment process, rather than sufficient in themselves. Sam said, *“I mean it’s how you interpret what’s going on really, at the end of the day. You can have all the assessments in the world but what does it actually mean.”*

From the point of being presented with a case the participants described making an initial interpretation of sexual deviance based on behaviour. They discussed the need to be aware of subjective judgements, which were instinctive and human but could bias assessment. They indicated that the use of ARAI’s were a helpful, objective tool which allowed them to check their biases and ensure they were grounding their assessment and decision-making in the evidence-base. However, these tools were not felt to be sufficient to capture sexual deviance as a risk factor. Although at this stage in the process, behaviour was a primary focus of discussion, it became clear that participants felt the need to look beyond behaviour to underlying processes.

Considering pathway to offence. Participants demonstrated that the same offending behaviour could be related to different underlying processes, as Billie clearly explains in the following example.

Billie: *“I have guys who use sex as a weapon and who in order to assert dominance will rape people... and it’s not really about sex for them but it still probably comes under sexual deviance... and other guys who, who it is a sexual thrill for them to be causing pain to people, either as part of sex or as sex adjacent action... I’ve got other guys who are going for a consensual sexual encounter, and then the other person withdraws consent and they just have rage and cause a huge amount of damage to them, em, and that’s not them maybe being interested in violence, it’s them having... anger issues that are completely separate from their sexual functioning but the end effect on the victim can be very similar for all three.”*

All participants highlighted sexual offences could be driven by a range of emotional and situational factors which had nothing to do with sexual deviance.

Frankie: *“I mean I see things as a continuum and for some people their offending is based on time and place. If you take away the time and place, then they will probably never offend again and would probably never have offended in the first place.”*

Eidann: *“[some] people sexually offend because there is a lack of something else going on or several deficits going on in their life and they’re making up for it or trying to respond to these more severe deficits through the offending behaviour. That does account for a lot of people.”*

Importantly, participants agreed that sexual deviance was not a pre-requisite for sexual offending to occur. Frankie highlighted, *“it could be a deviant act but that doesn’t mean that the person has sexual deviance.”*. Sam shared that, *“You might have done this horrendous offence but actually, you’re pretty boring when it comes to sexual thoughts and fantasies and behaviour.”*

They represented that understanding what has led to an offending behaviour can be difficult. In these discussions they demonstrated that it was important not to assume an underlying interest or attraction based on the offence alone. Participants emphasised the need to explore the underlying processes and drivers of an offence to understand the presence of sexual deviance and its role in risk of recidivism. Two participants clearly framed this process as trying to understand “why?” someone committed an offence. Billie highlighted that, *“I suppose what I mean, is we’re not thinking just about that behaviour anymore, we’re thinking about what is it they’re doing?... ‘Why do I think this person committed this offence?’, ‘What evidence do we have?’”*. Sam shared that, *“When you interview, for a report assessment, you don’t want to be like “this is deviant”, you want to be like “tell me more about what you do and why and do you do some stuff like this?”*

Therefore, although behaviour was an important element of this assessment process, participants felt that sexual deviance was more related to an underlying motivating process. To understand sexual deviance, they had to understand what motivated the act. Participants language when referring to underlying processes and internal drives provided insight into how they were understanding sexual deviance.

Understanding Sexual Deviance

Throughout interviews participants referred to sexual deviance when talking about behaviour, arousal, attraction, interest, and preference. The use of the term “sexual deviance”, “sexually deviant” and “deviant” were often used to refer to each of these components in the same interview. Therefore, attending to the context of each reference was important.

Thinking about Sexual Arousal. Half of the participants used the term arousal when considering an offender’s sexual deviance. This was referred to in the context of understanding the pathway to the offence. The point at which arousal occurred was described as an important indicator as to whether sexual deviance was present. Billie highlighted this, asking *“Which came first, was he feeling violent and used sex as a weapon or was he feeling aroused and within that became violent? You know? What’s the route into this offending?”*

Eidann: *[In a case of domestic sexual abuse] I have inferred that really, sexual assault was used as a tool in that relationship as another means of power and control and that’s deviant. That’s seriously [a deviant behaviour] and that can be really persistent behaviour. Now it becomes thorny to untangle, is it the coercion that’s sexually arousing or would they not use that if something else was present or is it just a tool?”*

Understanding the order and the pairing of sexual arousal and action was felt to provide insight into the presence of sexual deviance. The presence of arousal in the offence was not taken to indicate sexual deviance. The point at which it occurred and the role it played in the offence was thought to provide an indication of motivation for offending and that could indicate the presence of sexual deviance.

Sexual Interest. When considering sexual deviance as a factor that could motivate sexually deviant behaviour, the participants most consistently referred to the role of deviant sexual interests. Jo said, *“I suppose what I’d be looking for is a more enduring interest”*. Alex shared, *“I suppose you just generally think about [Sexual Deviance] in terms of what your own sexual interests are, what people’s sexual interests might be.”* While Eidann expressed the view that, *“The whole concept of deviance is that it’s a persistent interest.”*

It became clear from the examples that participants provided, that they understood sexual deviance as sexual interest in a being (e.g., children, animals) or act (e.g., coercion, pain, violence) that was unacceptable or illegal. Some participants explicitly referenced that there

had to be some degree of harm embedded in the interest. Ollie said that they “*would describe [a person with sexual deviance] as someone who receives some kind of sexual gratification through the harm of others or through causing harm to others.*” However, especially with internet offending the role of harm became difficult to understand. An offender may have sexual deviance but not have the intention or desire to directly perpetrate the harm. Additionally, the use of the internet may allow them to minimise or dismiss the harm, which would suggest it was not an important element of their interest. Participants also discussed that degree of intended versus actual harm in an offence could be difficult to interpret in assessment. Therefore, while harm was an important consideration, not all participants felt it was an essential component in a sexually deviant interest.

Participants reflected uncertainty that a deviant sexual interest could be sufficient to lead to sexual offending. Alex spoke about this in relation to sexual interest in children, “*Well somebody can have a paedophilic interest in children and, you know, there are people out there, ehh, who have a paedophilic interest in children who have never offended.*”

Participants shared the belief that sexually deviant interest are likely common in non-offending populations. Charlie shared, “*I guess my understanding is that a lot of people might have sexual deviance [referring to interests], but they don’t act on them.*” While Quinn wondered, “*how many more people might be thinking about it but not doing it.*”

Therefore, participants felt that there had to be something else moderating the link between having a deviant sexual interest and acting on that interest.

Sexual Preference. Some participants suggested that the strength or dominance of an interest could impact the risk associated with it. Billie stated, “*I would say that sexual deviance in terms of risk factor is either sexual functioning which in itself is an offence and that somebody’s core interest is in doing something which is an offence.*”

There was a suggestion that the increased strength of an interest or the more embedded it was in an individual’s sexual identity, the more this was perceived as a risk factor.

Davi: “*Whereas, somebody who has lots of interests, actually there’s a lot more stuff you can do there to just whittle out the stuff that’s more problematic and kind of develop distraction techniques, making sure that they are just indulging in the things that are, again, consensual. So, you’ve still got something that meets a need.*”

The participants' reference to stronger or core sexual functioning, resembled the definition in the literature of a sexual preference. Although participants used the terms interest and preference interchangeably, they indicated that they were considering the intensity of a sexual interest and the context of an individual's wider sexual functioning (e.g., other interests) when assessing sexual deviance as a risk factor. They viewed a stronger, more dominant deviant sexual interest was associated with higher risk.

Separating Sexual Deviance and Sexual Preoccupation. Participants highlighted that an offender's sexual functioning was more than their sexual deviance. Other aspects of their sexual functioning could represent distinct risk factors. Participants consistently indicated that sexual preoccupation was a separate risk factor from sexual deviance. As demonstrated by Eidann, "*The question that I'm asking in this case is whether it's a deviant sexual interest that's driving that or is it the sexual preoccupation*" and by Quinn, "*you're looking at kind of evidence of deviance, victim access, sexual preoccupation, how they're coping, how they use sex, if they use sex as coping and how their coping without it or coping with different sexual behaviours.*" While these were viewed as separate risk factors, there was recognition that the presence of other sexual risk factors (and other risk factors in general) influenced participants view of sexual deviance as a risk factor.

Defining through assessment. Thinking about the context around sexual deviance was a theme within interviews. When initially asked to define sexual deviance, most participants began with quite concrete terms and spoke about deviant behaviours. However, the more they discussed it, the more uncertain and nuanced the concept became. It was an interesting phenomenon that when trying to zoom in on "what is sexual deviance?" there was a tendency to zoom out and speak more about the process of trying to assess it. Therefore, the understanding of sexual deviance was constructed from the process of assessment. Several participants explicitly acknowledged the difficulties with defining this term. Frankie said, "*I think [Sexual Deviance] becomes a term that dissolves when you put your steely gaze on it... and I'm completely okay with that because I know it when I see it.*"

Frankie referenced knowing it "when I see it". This participant was one of three who expressed a degree of specialism in sexual offender assessments. This was associated with increased confidence in their ability to identify indicators of sexual deviance. While there were some differences in the depth to which these three participants spoke about certain

facets of sexual deviance, they described the same processes as those who referred to themselves as less confident or experienced.

When participants were asked how they “know it when they see it”, they referred to this process of starting with the behaviour and working backwards – considering what the tools told them about risk and what the offender’s pathway to the offence was, including thinking about underlying sexual motivators and their interplay with other risk factors.

Moving into actions. As has been demonstrated, when considering sexual deviance, participants were thinking most frequently about the presence of deviant sexual interests. However, the presence of these interests was felt to indicate different levels of risk depending on the context of the offender. Importantly, participants considered sexually deviant interests as more of a risk factor when they were associated with action. Charlie highlighted, *“it’s one thing to have a fetish about woman’s underwear, it’s another to steal them off of someone’s line... [deviant sexual interests] places you at a higher risk of sexual offending, assuming you’ve already got a sexual offence.”* This may explain why sexually deviant interests are considered more important and concerning in those who have demonstrated a capacity for offending behaviour than in those who have not.

Davi: *“Thinking about that everybody has their own sexual interests and sometimes those are safe and healthy and other times they become a bit more problematic and it’s when they become a problem that people might be more liable to offend”*

The research indicates that past violent offending is a risk factor for future sexual recidivism. However, unless someone has committed a sexual offence, risk assessment is unlikely to explore the presence of sexual deviance.

Gathering Evidence

The participants found it difficult to present a clear operationalised definition of sexual deviance. However, there was a consistent pattern in how they endeavored to understand it, in terms of deviant sexual interests, preferences and to some extent arousal. Participants indicated that within their risk assessment, understanding whether sexual deviance was a relevant risk factor was not a simple “yes”, “no” matter (as may be suggested by the actuarial tools). Instead, they described drawing upon a range of evidence sources that could speak to the presence and function of sexual deviance. Two primary sources of information discussed in interviews were self-report and past behaviour.

Using Self-Report. Participants referenced that self-report information could take different forms, from overt statements to more subtle inferences from behaviour and presentation. Eidann shared an example where, *“I was co-working him long-enough that he would admit when he was relapsing into deviant fantasies.”* It was felt that having a rapport with participants was helpful in accessing self-report. They also indicated that participants were more likely to engage in self-report at different points of their sentencing process.

Ollie: *“people at report stage will not tell you anything but when you go see them after they’ve been sentenced, they realise that there’s no point in pretending, once you’ve got a 10-year sentence or whatever. So, they tend to talk more.”*

Self-report was viewed as a useful source of information. Quinn highlighted the problems associated with a lack of self-report, *“this guy appears to be low risk but as he is unwilling or unable to talk about the alleged offence or any of the potential behaviours it’s kind of hard to identify – is there a risk there?”* However, all participants emphasised the challenges with relying on self-report.

Participants shared an expectation that offenders were likely to lie or conceal elements of their sexual interests. Frankie stated, *“People lie. Especially about this area.”* They proposed various reasons for this, including fear of consequences, feelings of shame, embarrassment and lack of insight.

Davi: *“people aren’t used to talking about their sexual desires or their sexual habits and you know, we have to ask them this and it can be quite a process for them, you know. Asking people, how often are you masturbating right now? It can be hugely disarming for them, and they don’t always necessarily want to talk about it.”*

Jo: *“I think one of the tricky things is that I think clients will often try to shut you down quite quickly because they feel uncomfortable, so it takes quite a bit of skill and patience to kind of probe that.”*

Therefore, while some value was placed on self-report as a source of information. It was also seen as limited, affected by fear of judgement or consequence and variable across offenders.

Considering Sexual Fantasies. Half the participants referred to sexual fantasies in their interviews as a source of information on sexual deviance, but it was rarely a key focus in

initial interviews. In follow-up respondent validation interviews participants were asked about this and they highlighted that although an understanding of someone's sexual fantasies could be informative to underlying sexual deviance, the only access to this was via self-report.

Jo: "I would say that probably raises the risk of them potentially engaging in sexually violent crime and obviously you'd need to explore that with them and if they're fantasising about this in real life... but again you need to have a really good relationship with that client to get to that point."

While it was felt that rapport could enhance access to self-reported information, there was still a limit to how much could be inferred from this, as self-report about underlying processes (such as fantasy and interest) could not be corroborated.

Looking for patterns in behaviour. The primary means that participants described for evidencing sexual deviance was through interpreting offender behaviours. Frankie stated, *"the best predictor of future behaviour is past behaviour. So, if somebody has offended 7 times, I pay more attention to that than to a person who has offended one time."* An analysis of past and present offending and sexual behaviours was thought to provide the most robust evidence of sexual deviance and its role in offending. This included thinking about the pathway of events into the offence itself to understand what motivated each aspect of the offender's behaviour. Davi explained this as, *"People have pathways, we look for that pathway and sometimes they make sense. Other times you're like "I've got no clue and we're just going to put in lots of risk management strategies and hope for the best"*.

Additionally, participants referenced considering the intention behind behaviours. This was especially the case for internet offenders or offenders who were potentially interrupted before an offence could take place. Charlie explored an example of this, *"So that there would make me more concerned about that man because he's clearly acted upon that chat online, and that chat online has been purposed to commit a contact offence against a child."*

Participants referenced that they felt most confident in their understanding of sexual deviance when there was evidence of longevity and/or chronicity. When they could see patterns of behaviour that occurred repeatedly over time.

Frankie: *“There has to be something about patterns of behaviour, there has to be something about chronicity over time, there has to be something... For me, what I look for in my work is something about patterns and chronicity regarding a sexual interest that’s got somebody into trouble.”*

Eidann: *“The whole concept of deviance is that it’s a persistent interest, it’s one that lasts over time and doesn’t go away... deviance that emerges at a certain point and then goes away again isn’t deviance I would argue. Deviance has to be persistent; it has to be something... the person’s sort of default mode... The multiple victims, the same activity, um over a long period of time, 5-6years.”*

Quinn: *“How much evidence do you need of a behaviour before you can say that I can define that behaviour as deviance? When somebodies only done something once, or a couple of times, it’s difficult to know how indicative that is.”*

Therefore, participants would explore the offence and previous patterns of behaviour to evidence (and understand) the possible presence and role of sexual deviance in the offending pathway. Although most of these examples referenced looking overtly for patterns of deviant sexual behaviour or functioning, some also indicated that inferences could be made about an offender’s sexual identity based on other relational risk factors. Frankie gave the following example, *“if somebody has never had an adult relationship and they’ve only offended against children. It doesn’t really matter what you call it, there’s just a problem there that we need to pay attention to.”*

In this way participants emphasised that information on other risk factors could inform their understanding of sexual deviance. They also clearly indicated that they viewed all risk factors in the context of how they interact with each other. The risk associated with sexual deviance was interpreted within the wider context of other risk factors.

Interpreting the evidence in the wider context. Throughout interviews, discussion of sexual deviance – what it is and how it is assessed – was intertwined with discussion about the risk assessment process as a whole and assessment of other factors. The interweaving nature of this discussion helped to emphasise the impact of the wider systems and influences on the process of understanding and assessing sexual deviance.

Assessing other risk factors. Participants are never assessing one risk factor at a time and highlighted that it did not feel meaningful to do so. The process of thinking about sexual deviance as a risk factor was intertwined with assessment of other risk factors. Eidann stated, *“It’s not that meaningful to think about [sexual deviance] on its own. It’s how it interacts with other parts of the personality and people’s history.”*

Participants emphasised that looking holistically at the wider context of an individual’s risk factors, helped to understand the link between sexual deviance and risk.

Charlie: *“I think the fantasy, the thoughts... They are a risk factor. But I think their relevance only comes in when you look at other risk factors around self-control and around a good life and around your own mental health and your own ability to control your behaviour.”*

Billie: *“I suppose in that sense it’s not really thinking about deviance, it’s thinking about the person and everything, trying to think of everything that’s going on for that individual. And if part of what is going on for that individual is deviance, then it should be coming out [during that process].”*

Quinn: *“you’re looking at [sexual interests] but you’re also looking at their, kind of, wider situation so are they are they emotionally stable or are things are falling apart for them...you know, how much support do they have, how isolated are they, various stressors and things like that which could potentially raise the risk of re-offending.”*

While participants referenced a variety of risk factors, including relationship status, employment, problem-solving, mental health and coping, one that came through as very important was a person’s impulse control or ability to control their behaviour. Frankie shared, *“that might just be that some sexual deviance, whatever it is, is easier to control than others. So, you might still have, whatever it is, but some people don’t struggle to control it.”* It was felt that having good impulse/behavioural control would facilitate someone’s ability to manage the drive to move from sexual interest into action.

Charlie: *“I guess, from my experience of managing risk it’s about someone’s general self-control, general self-management or ability to problem solve, their ability to recognise problems and take appropriate action... his ability to manage himself – this*

means it will stay a sexual fantasy because that's what it is, rather than be acted upon".

As can be seen by these quotes the ability of an offender to control their behaviour in relation to sexual deviance, was associated with the manageability of risk. The ability of the offender to self-manage risk behaviour or to be managed by supervision orders was an important theme throughout interviews.

Recognising the influence of Manageability. Reference to manageability occurred in quotes throughout every category of this model. Participants recognised that consideration of the manageability of risk was embedded into the risk assessment process as, ultimately, participants were risk assessing how manageable the offender would be in the community. This involved considering the person's ability to self-manage (linking into the role of impulse control) and a prediction of how effective the management tools available to social workers and other professionals were likely to be. As emphasised by Alex, *"well how much can this person manage their behaviour in the community without offending? What support can we give them to, to help them manage their behaviour in the community?"*

Billie: "I'll be thinking about how is this person going to meet their sexual needs? Or how is their sexual functioning going to affect their ability to stay in the community and what do we need to do to manage that?"

Frankie: "If it's situational, that's the only time that their sexual deviance manifests then I would work to see if we could control the situations that that person is in but if it seems to be internally driven then that might involve looking more at medication or some kind of behavioural intervention that actually cuts across situations."

The process of managing risk was not explored within the scope of this study, but likely overlaps with the processes described in this theoretical model. Consideration of manageability influenced participants assessment of sexual deviance. When thinking about sexual deviance in the context of other risk factors, they more commonly referred to risk factors which would influence manageability. Highlighting that this dual role influences their thinking.

Internet offending vs. Contact offending

Throughout interviews, participants referred to behaviours associated with contact sexual offences, in-person non-contact sexual offences (e.g., exposure and exhibitionism) and internet sexual offences (including accessing illegal pornography, soliciting underage sex, etc.). Participants indicated that there were differences between assessing internet sexual offenders and in-person offenders. They attributed this to the fact it is still a relatively new offence and therefore the evidence-base is still developing. Charlie shared that they, *“got quite interested in internet mediated sexual offending... cause it is obviously a relatively new area of sexual offending... researchers are trying to get up to speed with that and what that means”*

Davi: *“I think a lot of it has changed with the research because you know, at the moment, the way we view internet offenders tends to be lower risk in terms of harm ... sometimes [risk] can be more significant with them, but again in 2008 the research wasn't there for them.”*

Several participants indicated that they viewed sexual deviance as a less likely risk factor for internet offenders. Frankie thought that *“you can get the illusion of deviance with [internet offenders] sometimes”*. While Alex shared cases where they had wondered if, *“they've gone a bit down a rabbit hole and started to internet offend however, they're not particularly risky characters you know. They don't have a sexual interest in children.”*

This understanding may be because the research currently suggests that internet offenders are less likely to reoffend than contact offenders and therefore, they are viewed as inherently lower risk. It may also suggest that internet offenders are better able to mask specific interests or preferences within their viewing histories.

Jo: *“[Referring to internet offenders] Over the course of an hour, a couple of hours, somebody could view so many different paraphilia's for instance. What you need to explore then is, “is it curiosity? Were you aroused? Were there ones you were coming back to again and again?”*

Regardless, participants described applying the same processes for understanding sexual deviance and assessing risk of sexual recidivism to all types of sexual offending, including internet offending.

Respondent Validation

Those who participated in respondent validation reported that the model accurately represented their experience of understanding and assessing sexual deviance in practice. In discussion, two participants began referencing components of the model that had not yet been presented to them. This demonstrated the resonance and possible workability of the model. Although the term preference was not used distinctly in initial interviews, participants felt this was an appropriate conceptualisation in the mode. Interestingly, although 3 of these participants had referenced the value of self-report in their initial interviews, all showed hesitancy to endorse this as a reliable information source at validation.

Summary of the theoretical model

Participants viewed sexual deviance as an underlying factor that could act as a motivator for sexual offending when paired with other risk factors that facilitated action. Participants most commonly understood this motivating factor to be a deviant sexual interest, which presented higher risk when it was stronger or represented an offender's core sexual functioning. In their process of assessment, participants described gathering information from multiple sources, including self-report, actuarial tools and the offenders sexual and offending behaviour. Behaviour was viewed as the most trust-worthy source of information on sexual deviance as it was the least susceptible to intentional manipulation and reliance on behaviour to indicate sexual deviance is advocated by the use of ARAI's. However, participants highlighted that they were not just trying to ascertain whether a sexually deviant behaviour was present. Instead, they were trying to understand the processes underlying a behaviour – exploring the motivations and functions embedded in the pathway to offending. Sexual deviance was viewed as a stable and enduring factor, best evidenced through the presence of patterns of repeated behaviour. While sexual deviance was viewed as an important risk factor, participants didn't feel that any single risk factor could lead to sexual offending. They explained that for sexual deviance to lead to sexual offending it had to co-occur with other risk factors, such as behavioural disinhibition. This view of risk as multiply determined led participants to take a holistic approach to the assessment of sexual deviance, including other risk factors, contextual factors and sources of bias.

Discussion

This study used an inductive, qualitative methodology to explore the experience of Criminal Justice Social Workers (CJSW's) understanding and assessing sexual deviance as a risk factor for sexual recidivism. Using Social Constructivist Grounded Theory (Charmaz, 2006),

a theoretical model of two interlinked processes was developed (Figure 2). These findings are discussed in the context of the theories and evidence for the role of sexual deviance in sexual recidivism.

Understanding Sexual Deviance in Practice

Participants primarily understood sexual deviance in terms of sexually deviant interests in beings (e.g., children, animals) or acts (e.g., sadism, coercion) that are deemed inappropriate. By this definition, they saw sexually deviant interests as a risk factor that creates an underlying motivation to sexually offend. However, participants felt that the presence of sexually deviant interests alone were not sufficient to cause sexual offending. They felt that it was more likely to contribute to a sexual offence if it was of a certain intensity and if it was facilitated by other risk factors. This conceptualisation has been supported by the literature on sexually deviant interests in non-offending samples (Green, 2002; Joyal & Carpentier, 2017; Seto, 2008), phallometric studies on the relationship between interest and sexual recidivism (Freund & Watson, 1991; Harris et al., 1992; Knight & Thornton, 2007; Lalumière et al., 2003; Seto & Kuban, 1996) and theoretical models of sexual offending (Seto, 2008; 2018; Smid & Wever, 2019).

Participants consideration of the intensity of sexually deviant interests mirrored the literature on sexually deviant preferences (Moser, 2016; Seto, 2010; Quinsey, 2003). Although participants did not initially distinguish between the terms interest and preference, they resonated with this categorisation at respondent validation. Their interviews reflected the evidence-base that suggests deviant sexual preference is a stronger predictor of sexual recidivism than interest (Hanson & Bussiere, 1998; Hanson & Bourgon-Morton, 2005; Mann et al., 2010; Seto & Lalumière, 2001). However, participants assessment focused on interests more than preferences.

This may seem counter-intuitive but is likely reflective of the assessments methods available to them and may facilitate a fuller understanding of an offender's sexual risk. A sole focus on sexually deviant preference could lead to an underestimation of risk as it relates to multiple paraphilias (Mann et al., 2010) and in offender populations where the evidence of deviant sexual preference is weaker, such as in sexualised violence (Hanson & Bussiere, 1998; Mann et al., 2010; Stinson & Becker, 2008).

Participants did not clearly distinguish between sexually deviant interests and preferences in practice. As shown in the model, deviant sexual preferences were represented as an extension

of deviant sexual interests. The strength of an interest/lack of other interests led it to be viewed more as a preference. This was associated with increased risk, but participants reflected that a stronger sexual interest could relate to increased risk without being a dominant sexual preference. This conceptualisation reflects the way these are typically represented in the literature. However, if as proposed by Gannon (2021), these are distinct constructs with different developmental pathways, there will be a strong rationale for distinguishing between them in risk assessment. Particularly if the proposition that sexual interests are more malleable receives empirical support as this would have significant implication for manageability and intervention. However, accessible methods for distinguishing between them, would be required.

Assessing Sexual Deviance in Practice

The understanding of deviant sexual preferences as a risk factor for sexual recidivism has its foundations in phallometric studies (Lalumière & Quinsey, 1994; Mann et al., 2010; Seto, 2010). Researchers argue phallometry allows the assessment of whether someone exhibits greater arousal to one stimulus over another, therefore allowing them to infer a sexual preference. However, these measures are not accessible in routine forensic practice (Seto & Lalumière, 2001; Stinson & Becker, 2008). Although the evidence for deviant sexual interest has similar foundations, there has been more development of proxy measures (Akerman & Beech, 2012; Smit et al., 2011).

ARAI's and SPJ's aim to make the assessment of sexual deviance more accessible in practice and minimise bias. Some have shown concurrent validity with phallometric measures (Canales et al., 2009; Seto & Lalumière, 2001). However, they do not assess the same construct (e.g., sexual arousal). This has led authors to question their value in practice (Lalumière et al., 2003; Marshall & Fernandez, 2000; Seto, 2001).

Participants felt that ARAI's added value to their practice by providing a concrete procedure and definition. This helped them to recognise the influence of bias and minimise its impact on their assessments. However, there was a consensus that they offered limited insight into sexual deviance. They emphasised the importance of understanding sexual deviance and its function for the individual offender. A key criticism of ARAI's is that they offer little scope for consideration of the individual (Berlin et al., 2003; Darjee et al., 2016). They also highlighted an awareness of the risks of using routine tools with offenders not represented in

the tools normative sample, such as internet offenders. This impacted their confidence in this tool's contribution to these cases.

Participants reflected that information on sexual and offending behaviour felt the most useful and reliable for assessment. A person's history of offending is one of the strongest predictors of sexual recidivism, second only to phallometric measures (Hanson & Bussiere, 1998). This has informed the development of ARAI's, using behavioural indicators for rating different risk factors (e.g., number of offences). However, participants demonstrated that instead of just considering the presence or absence of a behaviour, they were engaging in an analytic process to understand its function and meaning. They described the process of reviewing the behavioural evidence of a risk factor and exploring hypotheses for underlying motivations and drivers before returning to check the validity of these hypotheses against the evidence. Participants highlighted that this was the most effective way to understand sexual deviance but recognised the difficulties of inferring sexually deviant interests and preferences from behaviour.

They reflected that repeated patterns of behaviour over time could indicate whether a sexual interest was present and give some indication of preference. However, they also highlighted that the same observable behaviour could be underpinned by different motivations.

Therefore, while behaviour was viewed as an important source of information, they felt that there were limits as to what could be conclusively inferred from behaviour regarding sexual deviance. Participants highlighted that confidence in their assessment could be impacted by the quantity and quality of information that they had access to.

Self-report was presented as a potentially valuable source of information. Several participants described cases where they felt having a good rapport with an offender facilitated enough self-report to significantly contribute to assessment and understanding of their sexual deviance. However, all participants emphasised the dangers of relying on self-report given societal pressures and the risk of negative consequences associated with disclosure (Abel et al., 1998; Seto, 2010). The literature has indicated that self-report could provide information on aspects of sexual functioning that are not accessible by other means. For example, self-reported sexual fantasies could inform assessment of sexual deviance and risk of recidivism (Bartels & Beech, 2016; Gee, Devilly & Ward, 2004). However, some participants indicated the inability to validate the information, made this less of a priority in assessments.

Self-report could provide a means of distinguishing between deviant sexual interests and preferences (Abel, et al., 1998). However, participants felt that even if offenders shared some information pertaining to their sexual interests, it was likely to be minimised or presented in a more socially desirable light. Participants reflected the view that a deviant sexual preference would be more stigmatising than a deviant sexual interest. Therefore, while self-report could theoretically be a good source of information regarding someone's deviant sexual interests and preferences, it could not be relied upon.

Although the evidence-base views arousal as an important element in sexual deviance (Hart & Kropp, 2008; Stinson & Becker, 2008), participants did not focus on this in their assessment. Although it could inform behavioural analysis, without a formal assessment of arousal separate to the offence, this was not felt to be a meaningful component in practice.

Some authors have advocated that the only way to effectively translate research to practice is to increase access to phallometric assessment (Kingston et al., 2010; Seto, Lalumière & Kuban, 1999). While these are not infallible (Renaud, 2009; Seto & Lalumière, 2001; Stinson & Becker, 2008), it would at least create a consistency between what the research says and what is being measured in practice. Perhaps a more realistic alternative is for the research base to continue to explore accessible ways of assessing sexual deviance. The ARAI's which participants routinely use (e.g., SA07), define sexual deviance in terms of sexual interest. This may contribute to their conceptualisation of sexual deviance in these terms. Therefore, developing tools which have clear and accurate definitions of sexual deviance could contribute to clarity of understanding as well as accuracy in assessment.

The primary perceived value of ARAI's was as a means of grounding participant understanding of sexual deviance in the literature by presenting a definition and a structure for assessment. They felt this allowed them to manage potential bias in their assessment, returning to the tool to ground subjective opinion. However, they identified that subjective interpretation of behavioural information was necessary for an accurate assessment of an individual. This allowed them to consider the relevance of the tools to the individual offender (e.g., internet offenders) as well as the way in which risk factors functioned to increase individual risk. In this way, SPJ tools may be more appropriate to this process as they are designed to facilitate and minimise bias in the subjective aspects of assessment. (Darjee et al., 2016; Harris & Hanson, 2010; Hart et al., 2003). ARAI's do not offer any indication of how

to navigate these subjective processes. Therefore, the responsibility falls more on individual CJSW's and there is scope for increased bias.

The Motivation-Facilitation Model

There was a consensus among participants that sexually deviant interests or preferences alone were unlikely to lead to sexual offending without interaction with other risk factors or the opportunity to offend. Participants specifically referenced the role of behavioural inhibition or self-control. This may in part be reflective of their dual role of risk assessment and risk management. Factors relating to behavioural inhibition were viewed as a significant indicator for the manageability of any offender's risk. However, the emphasis on behavioural inhibition was reflective of the key theories of sexual deviances role in sexual offending (Gannon, 2021; Seto, 2018; Smid & Wever, 2019).

The process which participants described using to assess sexual deviance and understand its role in overall risk, significantly overlapped with Seto's (2008; 2018) motivation-facilitation model. They saw sexually deviant interests or preferences as creating a drive to commit a sexual offence but felt that this would only come to fruition if their behavioural inhibition was in some way compromised and they were presented with the opportunity to offend. This parallel between theoretical conceptualisations based on the evidence and the processes described by participants is promising for the value of CJSW's practice. Despite the controversies around whether they have access to the most appropriate tools for the assessment of sexual deviance, the way in which they are approaching assessment reflects the evidence-base.

Difficulties in the literature

The core difficulties in the practice of assessing sexual deviance reflect the core difficulties in the literature. There needs to be a consistent way of defining sexual deviance and a gold-standard assessment determined from which to validate other more accessible tools. Some currently argue that phallometry is the gold-standard assessment (Kingston et al., 2010; Seto, Lalumière & Kuban, 1999) but that depends on the definition being used – creating a circular problem.

In the literature the term sexual deviance is used interchangeably and inconsistently (Laws & O'Donohue, 2008; Smid et al., 2011) and this was mirrored in participant interviews. Sexual deviance was used to refer to different aspects of sexual functioning throughout interviews and context was key to understanding the meaning. This tendency is inherent in the term

sexual deviance, as the relevant aspect of sexual functioning is not specifically referenced. It is therefore open to subjective interpretation and use. The word deviance introduces further scope for subjective bias as it is accompanied by the assumption that there is a norm or accepted standard from which an aspect of someone's sexual functioning can deviate.

Participants clearly demonstrated that the word deviance introduced consideration of different types of norms - drawn from their own experience, societal influence or that of the law. For the first two, participants demonstrated an understanding of the subjective and fluid nature of these norms (e.g., how they could vary across individual, time and culture). The implicit association to a changing normative standard in what is meant to be an objective assessment is problematic in itself. The law arguably has more relevance to sexual deviance as a risk factor for sexual offending, but it directs attention to sexually deviant behaviour, which is not inclusive.

The non-specific nature of the term “sexual deviance” creates difficulties in the assessment and communication of risk. As the underlying processes of risk factors are not currently understood, perhaps it would be more appropriate/effective to refer specifically to the aspect of sexual functioning that is being assessed and what is being inferred from this. Sexually deviant behaviour, sexually deviant arousal, sexually deviant interests and sexually deviant preferences all provide important and distinct information related to assessing risk of sexual recidivism (Hanson & Morton-Borugon, 2005; Mann et al., 2010; Smit et al., 2011). Recognising these as interlinked, yet separate constructs could be key to clarifying communication and understanding.

Sexual offenders are known for being a significantly heterogeneous population (Noteborn, 2022) and the research in this area has been consistently criticised for not representing this enough in their research designs and participant samples (Conroy, 2006; Gannon, 2021; Lalumière & Quinsey, 1994). This likely impacts the ability to develop an understanding of sexual deviance across different offence presentations. The research on sadism and non-consent has likely been grouped together because in offending populations they will both most commonly be represented by those charged with rape. However, these may be representative of distinct interests that require studied separately to fully understand. This could explain to some extent the weaker evidence of sexual preferences in sexualised violence (Kingston & Yates, 2008; Lalumière et al., 2017).

The strongest evidence of sexual deviance has been gathered in relation to paedophilia and it has been assumed that these findings are generalisable to other types of sexual interest and preference. However, several authors have suggested that the demonstration of a strong, persistent interest into a physical characteristic of a person, such as their age, may function and develop in a similar way to sexual orientation (Moser, 2016; Seto, 2012). There is hesitancy to use this term due to fear of legitimising attraction to children (which is not the intention of this author; Ganon, 2021). However, fear of looking at paedophilia through this lens may make it harder to understand the underlying process of it as a risk factor. Gannon (2021) has suggested that a sexual interest and sexual preference in the same stimuli (children) might be underpinned by vastly different processes. It is not outside the realms of possibility that the processes underlying a sexual interest in a physical characteristic may differ from that of a sexual interest in an activity or act.

Considerations for Internet Offenders

The focus of this study was not to consider sexual deviance across different types of sexual offending. Nonetheless, some interesting distinctions between in-person sexual offenders and internet sexual offenders were raised.

Participants shared the view that internet offenders were less likely to reoffend and less likely to be sexually deviant than in-person (especially contact) offenders. The first proposition is in line with the literature ((Babchishin, Hanson & Hermann, 2011; Wakeling, Howard & Barnett, 2011). However, in a meta-analysis comparing online and in-person sexual offenders, Babchishin, Hanson & Hermann (2011) found that online offenders demonstrated significantly more sexually deviant interests than in-person offenders in studies using phallometry, the Stable-2007 and self-reported fantasies. When exploring the heterogeneity within internet sexual offenders, Seto and colleagues (2012) found that child pornography users compared to online-solicitation offenders and contact offenders, were more likely to show deviant sexual arousal, scored higher on the sexual deviance items of the SA07 and were more likely to self-report a deviant sexual interest in children.

Online offenders also scored higher on items relating to sexual preoccupation on the SA07 which participants viewed as more relevant when considering the pathway to online sexual offending. There may be methodological reasons that online sexual offenders score higher on various ratings of sexual deviance than contact offenders. For example, the SA07 was not developed for use with internet offenders and so may have limited validity.

The participants view of online sexual offenders as less sexually deviant may be influenced by the tools. It could also suggest that the perception of harm caused by the offender influences participants perception of risk. When discussing harm some participants viewed sexual deviance as arousal towards the infliction of harm. However, the nature of the internet can create a sense of separation from the harm their actions are contributing to (Howitt & Sheldon, 2007; Quayle & Taylor, 2002). Their perpetration is indirect and therefore their interest may be perceived as less related to the causation of harm, rather than to other aspects of the experience. Regardless, there is an important difference between the way sexual deviance in internet sexual offenders is viewed in research and practice. This merits increased training in practice and consideration in research.

Strengths and Limitations

The strengths of this study are embedded in its adherence to Social Constructivist Grounded Theory and the adoption of quality standards into research design. The data from this study has been presented to allow the reader to make their own judgements of fit and credibility (I.e., how the theory fits the codes and the codes fit the categories). Respondent validation demonstrated the resonance of the model and to some extent its workability (I.e., its ability to predict behaviour). It presents a novel theoretical understanding of how risk assessment of sexual deviance is undertaken in practice and transparency of construction should allow for this model to be explored in other population and modified appropriately. Thus, enhancing its generalisability.

Some may consider the level of saturation a weakness of this project. Theoretically, a grounded theory is not complete until all its components are fully saturated (Glaser & Strauss, 1967). However, from a Social Constructivist standpoint, this researcher questions whether a theory can ever be fully saturated. Holding the belief that there is no one underlying truth, contradicts the concept of saturation as there is always another perspective to be found. Strauss and Corbin (1998) acknowledged that saturation is a matter of degree. This study has endeavoured to present the extent of saturation with transparency. This should facilitate critical appraisal of the model presented and direct future studies towards the areas which are less robust. The intertwined nature of the processes in this model indicates that they likely take place within the context of other processes associated with risk assessment which were beyond the scope of this study to explore.

To facilitate sufficient saturation a small, cohesive sample was recruited (Charmaz, 2006; Ritchie et al., 2003). Therefore, the generalisability of the findings has not yet been assessed. Future studies will be required to see if this model represents the experiences of other CJSW's and professionals assessing sexual deviance in practice.

In line with the principles of grounded theory, the researcher was relatively naïve to the concept of sexual deviance during interviews and analysis. Therefore, their understanding of sexual deviance was shaped by participants understanding and then honed by the literature in a retrospective review. This naivety reduced the level of bias within analysis and construction of the model. However, without a pre-existing concept of sexual deviance, the researcher may have missed opportunities for model development which someone more familiar with the literature may have recognised. Nonetheless, this approach was important in protecting the resonance of the theory.

Clinical implications and Future Directions

The findings of this study provide a clear indication into how assessors are understanding sexual deviance in practice. Previous research has highlighted what tools are used in practice (Harris, Boccaccini & Rice, 2017) but not how they are used to inform the wider assessment process. Importantly, it appears that the tools and methods available for assessment can shape the way assessors understand that which is being assessed. This emphasises the importance of a shared, consistent definition of sexual deviance and clear communication of the extent and limitations associated with each method. Assessment tools need to be clearer in what they are claiming to assess (“how”).

The results indicate that SPJ tools may be more appropriate for this assessment than ARAI's. Participants described a process of combining evidence-based risk factors (as scored in ARAI's) with subjective judgement to understand risk in the context of the individual. SPJ tools are better placed to facilitate this by providing an evidence-based, structured approach to the subjective processes. Although SPJ tools are more resource intensive, an incorrect assessment of risk is ultimately more so.

Despite advancements in understanding what risk factors are associated with sexual recidivism, there is more to learn (“what”). Future research should recognise the heterogeneity of sexual offenders and endeavour to recruit populations appropriate to the research question. This would allow results which can be more meaningfully interpreted. Exploring specific facets of sexual deviance separately (e.g., interest/preference in coercion

separate to paedophilia) would allow a deeper understanding of the comparability of these constructs and whether they can be captured under an overarching term such as sexual deviance, or whether they are better understood as distinct concepts. Furthermore, the research in sexual deviance and risk assessment as a whole, would benefit from a greater focus on the processes underlying risk factors (“why”). Seto (2008;2018), Smid & Wever (2019) and Gannon (2021) present theoretical models resonant to the experience of assessing sexual deviance in practice and offer testable hypotheses that could offer insight into how sexual deviance functions to influence risk of recidivism.

Conclusion

Participants understanding of sexual deviance in this way seemed to be shaped by what they were realistically able to assess in practice. They conceptualised sexual deviance primarily in terms of sexually deviant interests with recognition that this was not sufficient alone for offending and needed facilitated by other risk factors and opportunity to offender. They also felt that there was a relationship between the intensity of the sexually deviant interest and level of perceived risk. Participants risk assessment went beyond assessing the presence of risk factors. They emphasised the importance of understanding the function of risk factors in each individual offender. Their assessment involved a holistic approach, using objective tools to ground subjective processes. Participants understanding and approach to risk assessment, mirrored that of theoretical models of sexual deviance, which suggests their approach is grounded in the evidence-base. However, their practice also reflected the difficulties of the literature which highlighted the need for a clear, consistent, operational definition of sexual deviance.

References

- Abel, G.G., Mittelman, M., Becker, J.V., Rathner, J. & Rouleau, J.L. (1988). Predicting child molesters' response to treatment. *Annals of the New York Academy of Sciences*.
- American Psychiatric Association (APA) (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Arlington, VA: American Psychiatric Association.
- Akerman, G., & Beech, A. R. (2012). A systematic review of measures of deviant sexual interest and arousal. *Psychiatry, Psychology and Law*, 19(1), 118.
- Andrews, D.A., & Bonta, J. (1998). The psychology of criminal conduct (2nd ed.). Cincinnati, OH: Anderson.
- Babchishin, K. M., Karl Hanson, R., & Hermann, C. A. (2011). The characteristics of online sex offenders: A meta-analysis. *Sexual Abuse*, 23(1), 92.
- Bartels, R., & Beech, A. (2016). Theories of deviant sexual fantasy. In Boer, D.P. (Ed.), *The Wiley Blackwell Handbook on assessment, treatment and theories of sexual offending*. Wiley, Chichester.
- Bartels, R. M., & Gannon, T. A. (2011). Understanding the sexual fantasies of sex offenders and their correlates. *Aggression and Violent Behavior*, 16(6), 551.
- Bartosh, D.L., Garby, T., Lewis, D. & Gray, S. (2003). Differences in the Predictive Validity of Actuarial Risk Assessments in Relation to Sex Offender Type. *International Journal of Offender Therapy and Comparative Criminology*, 47(4), 422.
- Berlin, F. S., Galbreath, N. W., Geary, B., & McGlone, G. (2003). The use of actuarials at civil commitment hearings to predict the likelihood of future sexual violence. *Sexual abuse: a journal of research and treatment*, 15(4), 377.
- Boccaccini, M. T., Murrie, D. C., Caperton, J. D., & Hawes, S. W. (2009). Field validity of the STATIC-99 and MnSOST-R among sex offenders evaluated for civil commitment as sexually violent predators. *Psychology, Public Policy, and Law*, 15(4), 278.
- Boswell, J. F., Kraus, D. R., Miller, S. D., & Lambert, M. J. (2015). Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions. *Psychotherapy research*, 25(1), 6.

Breckenridge, J., & Jones, D. (2009). Demystifying theoretical sampling in grounded theory research. *Grounded Theory Review*, 8(2).

Canales, D.D., Olver, M.E. & Wong, S.C.P. (2009) Construct Validity of the Violence Risk Scale-Sexual Offender Version for Measuring Sexual Deviance. *Sexual Abuse: A Journal of Research and Treatment*, 21(4), 474.

Charmaz, K. (2006). *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. Sage Publications Ltd.

Conroy, M. A., & Murrie, D. C. (2007). *Forensic assessment of violence risk: A guide for risk assessment and risk management*. John Wiley & Sons.

Cooke, D.J. & Michie, C. (2013) Violence Risk Assessment from prediction to understanding – or from what? To why? In C.Logan & L. Johnston (Eds.), *Managing Clinical Risk: A Guide to Effective Practice*. London: Routledge.

Craig, L. A., & Beech, A. (2009). Best practice in conducting actuarial risk assessments with adult sexual offenders. *Journal of Sexual Aggression*, 15(2), 193.

Darjee, R., Russell, K., Forrest, L., Milton, E., Savoie, V., Baron, E., & 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 Southeast Scotland*. Report to Risk Management Authority and NHS Scotland Forensic Network. Retrieved from:

https://www.rma.scot/wpcontent/uploads/2018/04/Risk_for_Sexual_Violence_Protocol_-_RSVP.pdf.

Dempster, R.J. & Hart, S.D. (2002). The relative utility of fixed and variable risk factors in discriminating sexual recidivists and non-recidivists. *Sexual Abuse: A Journal of Research and Treatment*, 14(2), 121.

Doren, D., & Elwood, R. (2009). The diagnostic reliability of sexual sadism. *Sexual Abuse*, 21(3), 251.

Douglas, K. S., & Kropp, P. R. (2002). A prevention-based paradigm for violence risk assessment: Clinical and research applications. *Criminal Justice and Behavior*, 29(5), 617.

Dunsieth, N. W., Nelson, E. B., Brusman-Lovins, L. A., Holcomb, J. L., Beckman, D., Welge, J. A., Roby, D., Taylor, B., Soutullo, C. A., & McElroy, S. L. (2004). Psychiatric and

legal features of 113 men convicted of sexual offences. *Journal of Clinical Psychiatry*, 65, 293.

First, M.B. (2014). DSM-5 and Paraphilic Disorders. *The Journal of the American Academy of Psychiatry and the Law*, 42(2), 191.

Freund, K., & Watson, R. J. (1991). Assessment of the sensitivity and specificity of a phallometric test: An update of phallometric diagnosis of pedophilia. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 3(2), 254.

Gannon, T. A. (2021). A compositional explanatory theory of pedophilia. *Aggression and violent behavior*, 61.

Gee, D. G., Devilly, G. J., & Ward, T. (2004). The content of sexual fantasies for sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 16(4), 315.

Glaser, B.G. & Strauss, A.L. (1967) *The discovery of grounded theory*. Chicago: Aldine.

Green, R. (2002). Is pedophilia a mental disorder?. *Archives of sexual behavior*, 31(6), 467.

Green, B., Carroll, A., & Brett, A. (2010). Structured risk assessment in community forensic mental health practice. *Australasian Psychiatry*, 18(6), 538.

Gross, M. (2014). Paraphilia or perversion? *Current Biology*, 24(17), 777.

Hammersley, M. (2012). The issue of quality in qualitative research. *International Journal of Research & Method in Education*, 30(3), 287.

Hanson, K.R. (2000) Will they do it again? Predicting Sex-Offense Recidivism. *Current Directions in Psychological Science*, 9(3), 106.

Hanson, R. K., & Bussiere, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of consulting and clinical psychology*, 66(2), 348.

Hanson, R. K., Harris, A. J. R., Scott, T.-L., & Helmus, L. (2007). *Assessing the risk of sexual offenders on community supervision: The Dynamic Supervision Project* (User Report No. 2007-05). Ottawa, Ontario: Public Safety Canada. Retrieved from <http://www.publicsafety.gc.ca/cnt/rsrscs/pblctns/ssssng-rsk-sxl-ffndrs/index-eng.aspx>

Hanson, R. K., & Harris, A. J. (2001). A structured approach to evaluating change among sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 13(2), 105.

- Hanson, R.K. & Morton-Bourgon, K.E. (2005). The Characteristics of Persistent Sexual Offenders: A Meta-Analysis of Recidivism Studies. *Journal of Consulting and Clinical Psychology, 73*(6), 1154.
- 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.
- Harris, A., Phenix, A., Hanson, R. K., & Thornton, D. (2003). *Static 99: Coding rules revised 2003*. Ottawa, Ontario: Solicitor General Canada.
- Harris, P. B., Boccaccini, M. T., & Rice, A. K. (2017). Field measures of psychopathy and sexual deviance as predictors of recidivism among sexual offenders. *Psychological Assessment, 29*(6), 639.
- Harris, A.J. R. & Hason, K. (2010) Clinical, actuarial and dynamic risk assessment of sexual offenders: Why do things keep changing? *Journal of Sexual Aggression, 16*(3), 296.
- Harris, G. T., Rice, M. E. & Quinsey, V. L. (2007). Abandoning evidence-based risk appraisal in forensic practice: Comments on Hart et al. *British Journal of Psychiatry, 15* August, 2007. Electronic letters sent to the journal available from <http://bjp.rcpsych.org/cgi/eletters/190/49/s60#6574>
- Hart, S. D. (1998). The role of psychopathy in assessing risk for violence: Conceptual and methodological issues. *Legal and criminological psychology, 3*(1), 121.
- Hart, S. D., 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 for sexual violence*. Vancouver, Canada: BC Institute Against Family Violence.
- Hart, S.D & Kropp, P.R. (2008). Sexual Deviance and the Law. In D.R. Laws & W.T. O'Donohue (Eds.), *Sexual Deviance: Theory, Assessment and Treatment*. Guilford Press.
- 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. *The British Journal of Psychiatry, 190*(49), 60.
- Heilbrun, K., Douglas, K. S., & Yasuhara, K. (2009). Controversies in violence risk assessment. *Psychological science in the courtroom: Controversies and consensus, 333*.

- Howitt, D., & Sheldon, K. (2007). The role of cognitive distortions in paedophilic offending: Internet and contact offenders compared. *Psychology, Crime & Law*, 13(5), 469.
- Joyal, C.C. & Carpentier, J. (2016). The Prevalence of Paraphilic Interests and Behaviours in the General Population: A Provincial Survey. *The Journal of Sex Research*, 54(2), 161.
- Judge, J., Quayle, E., O'Rourke, S., Russell, K. & Darjee, R. (2014) Referrers' views of structured professional judgement risk assessment of sexual offenders: A qualitative study. *Journal of Sexual Aggression*, 20(1), 94.
- Kenny, M. & Fourie, R. (2015). Contrasting Classic, Straussian and Constructivist Grounded Theory: Methodological and Philosophical Conflicts. *The Qualitative Report*, 20(8), 1270.
- Khiroya, R., Weaver, T., & Maden, T. (2009). Use and perceived utility of structured violence risk assessments in English medium secure forensic units. *Psychiatric Bulletin*, 33(4), 129.
- Kingston, D.A., Firestone, P., Seto, M.C. & Bradford, J.M. (2010). Comparing Indicators of Sexual Sadism as Predictors of Recidivism Among Adult Male Sexual Offenders. *Journal of Consulting and Clinical Psychology*, 78(4), 574.
- Kingston, D. A., & Yates, P. M. (2008). Sexual sadism: Assessment and treatment. In D. R. Laws & W. T. O'Donohue (Eds.), *Sexual deviance: Theory, assessment, and treatment* (pp. 231–249). The Guilford Press.
- Knight, R. A., & Thornton, D. (2007). *Evaluating and improving risk assessment schemes for sexual recidivism: A long-term follow-up of convicted sexual offenders*. Rockville, MD: National Criminal Justice Reference Service.
- Lalumière, M. L., Quinsey, V. L., Harris, G. T., Rice, M. E., & Trautrimas, C. (2003). Are rapists differentially aroused by coercive sex in phallometric assessments?. *Annals of the New York Academy of Sciences*, 989(1), 211.
- Lalumière, M.L. Fairweather, A., Harris, G.T., Suschinsky, K.D. & Seto, M.C. (2017). Genital Responses to Rape Vignettes Among Young Men: The Influence of Mood and Directed Attention. *Archive of sexual Behaviour*, 46, 685.

- Lalumière, M. L., & Quinsey, V. L. (1994). The discriminability of rapists from non-sex offenders using phallometric measures: A meta-analysis. *Criminal Justice and Behavior*, 21(1), 150.
- Långström, N. (2004). Accuracy of actuarial procedures for assessment of sexual offender recidivism risk may vary across ethnicity. *Sexual Abuse: A Journal of Research and Treatment*, 16(2), 107.
- Lauridsen, E. I., & Higginbottom, G. (2014). The roots and development of constructivist grounded theory. *Nurse researcher*, 21(5).
- Laws, D.R. & O'Donohue, W.T. (Eds). (2008). *Sexual deviance: Theory, Assessment and Treatment*. Guilford Press.
- Levenson, J. S., & Morin, J. W. (2006). Factors predicting selection of sexually violent predators for civil commitment. *International Journal of Offender Therapy and Comparative Criminology*, 50(6), 609.
- Mann, R. E., Hanson, R. K., & Thornton, D. (2010). Assessing risk for sexual recidivism: Some proposals on the nature of psychologically meaningful risk factors. *Sexual Abuse*, 22(2), 191.
- Marshall, W. L. (2000). Revisiting the use of pornography by sexual offenders: Implications for theory and practice. *Journal of Sexual Aggression*, 6(1), 67.
- Marshall, W. L., & Fernandez, Y. M. (2000). Phallometric testing with sexual offenders: Limits to its value. *Clinical Psychology Review*, 20(7), 807.
- Mason, M. (2010). Sample Size and Saturation in PhD Studies Using Qualitative Interviews. *In Forum: Qualitative Social Research*, 11(3).
- Masters, K.B. & Kebbell, M.R. (2019). Police Officers' Perceptions of a Sex Offender Registration Scheme: Identifying and Responding to Risk. *Psychiatry, Psychology and Law*, 26(3), 396.
- McCann, K., & Lussier, P. (2008). Antisociality, sexual deviance, and sexual reoffending in juvenile sex offenders: A meta-analytical investigation. *Youth Violence and Juvenile Justice*, 6(4), 363.
- Menzies, R., Webster, C. D., McMain, S., Staley, S., & Scaglione, R. (1994). The dimensions of dangerousness revisited. *Law and Human Behavior*, 18(1), 1.

- Monahan, J. (2007). Clinical and actuarial predictions of violence. In D. Faigman, D. Kaye, M. Saks, J. Sanders, E. Cheng (Eds.), *Modern Scientific Evidence: The Law and Science of Expert Testimony* (pp. 125-147). St Paul, MN: West Publishing Company.
- Moser, C. (2016). Defining Sexual Orientation. *Archives of Sexual Behaviour*, 45(3), 505.
- Noteborn, M. G. (2022). Adult Male Contact Sexual Offenders: Challenges in Classification and Theoretical Perspectives. In C. Garofalo & Sijtsema, J.J. (Eds.), *Clinical Forensic Psychology* (pp. 415-443). Palgrave Macmillan, Cham.
- Waite, M. (Ed.). (2012). *Paperback Oxford English dictionary*. Oxford University Press.
- Prentky, R. A., Lee, A. F., Knight, R. A., & Cerce, D. (1997). Recidivism rates among child molesters and rapists: A methodological analysis. *Law and human behavior*, 21(6), 635.
- Quayle, E., & Taylor, M. (2002). Child pornography and the Internet: Perpetuating a cycle of abuse. *Deviant behavior*, 23(4), 331.
- Quinsey, V. L. (2003). The etiology of anomalous sexual preferences in men. *Annals of the New York Academy of Sciences*, 989(1), 105.
- Quinsey, V. L., Jones, G. B., Book, A. S., & Barr, K. N. (2006). The dynamic prediction of antisocial behavior among forensic psychiatric patients: A prospective field study. *Journal of Interpersonal Violence*, 21(12), 1539.
- Renaud, P., Chartier, S., Rouleau, J. L., Proulx, J., Trottier, D., Bradford, J. P., Fedoroff, P., Bouchard, S. & Décarie, J. (2009). Gaze behavior nonlinear dynamics assessed in virtual immersion as a diagnostic index of sexual deviancy: preliminary results. *JVRB-Journal of Virtual Reality and Broadcasting*, 6(3).
- Rettenberger, M., & Eher, R. (2007). Predicting reoffence in sexual offender subtypes: A prospective validation study of the German version of the Sexual Offender Risk Appraisal Guide (SORAG). *Sexual Offender Treatment*, 2(2), 1.
- Rice, M. E., Harris, G. T., Lang, C., & Cormier, C. (2006). Violent sex offenses: How are they best measured from official records?. *Law and human behavior*, 30(4), 525.
- Risk Management Authority Scotland. (2013). *The Evaluation of the Implementation of Stable and Acute 2007 in Scotland*. Paisley, UK.

Ritchie, J., Lewis, J. & Elam, G. (2003). *Designing and selecting samples*. In J. Ritchie & J. Lewis (Eds.), *Qualitative Research Practice. A Guide for social science students and researchers* (pp. 77-108) Thousand Oaks, CA: Sage.

Scottish Government (2022) *Recorded Crime in Scotland, 2021-22*. Retrieved from: <https://www.gov.scot/binaries/content/documents/govscot/publications/statistics/2022/06/recorded-crime-scotland-2021-2022/documents/recorded-crime-scotland-2021-22/recorded-crime-scotland-2021-22/govscot%3Adocument/recorded-crime-scotland-2021-22.pdf>

Seto, M.C. & Lalumiere, M.L. (2001). A Brief Screening Scale to Identify Pedophilic Interests Among Child Molesters. *Sexual abuse: A journal of Research and Treatment*, 13(1), 15.

Seto, M. C. (2005). Is more better? Combining actuarial risk scales to predict recidivism among adult sex offenders. *Psychological assessment*, 17(2), 156.

Seto, M.C. (2010). Child Pornography Use and Internet Solicitation in the Diagnosis of Pedophilia. *Archives of Sexual Behaviour*, 39(3), 591.

Seto, M. C. (2018). The motivation-facilitation model of sexual offending. *Sexual Abuse*, 31(1), 3.

Seto, M. C. (2012). Is pedophilia a sexual orientation?. *Archives of sexual behavior*, 41(1), 231.

Seto, M. C., & Kuban, M. (1996). Criterion-related validity of a phallometric test for paraphilic rape and sadism. *Behaviour research and therapy*, 34(2), 175.

Seto, M. C., & Lalumière, M. L. (2001). A brief screening scale to identify pedophilic interests among child molesters. *Sexual Abuse: A Journal of Research and Treatment*, 13(1), 15.

Smallbone, S., & Cale, J. (2016). Situational theories. *The Wiley Handbook on the Theories, assessment and treatment of sexual offending*, 289.

Smallbone, S. W., & Wortley, R. K. (2004). Criminal diversity and paraphilic interests among adult males convicted of sexual offenses against children. *International Journal of Offender Therapy and Comparative Criminology*, 48(2), 175.

Smid, W., Beek, D. & Troelstra, J. (2011). Proxy Measures of Sexual Deviance. In Boer, D.P., Eher, R., Craig, L.A., Miner, M.H. & Pfafflin, F. (Eds.), *International Perspectives on*

- the Assessment and Treatment of Sexual Offenders* (pp. 171-192). West Sussex: Wiley-Blackwell.
- Smid, W. J., & Wever, E. C. (2019). Mixed emotions: An incentive motivational model of sexual deviance. *Sexual Abuse, 31*(7), 731.
- Stinson, J. D., Sales, B. D., & Becker, J. V. (2008). *Sex offending: Causal theories to inform research, prevention, and treatment*. Washington, DC, USA: American Psychological Association.
- Strauss, A. & Corbin, J. (1998). *Basics of Qualitative Research Techniques*. Thousand Oaks, CA: Sage.
- Wakeling, H. C., Howard, P., & Barnett, G. (2011). Comparing the validity of the RM2000 scales and OGRS3 for predicting recidivism by Internet sexual offenders. *Sexual Abuse, 23*(1), 146.
- Ward, T. (2014). The explanation of sexual offending: From single factor theories to integrative pluralism. *Journal of sexual aggression, 20*(2), 130.
- Ward, T., & Fortune, C. A. (2016). The role of dynamic risk factors in the explanation of offending. *Aggression and violent behavior, 29*, 79.
- Ward, T., & Siegert, R. J. (2002). Toward a comprehensive theory of child sexual abuse: A theory knitting perspective. *Psychology, crime and law, 8*(4), 319.
- Zillmann, D. (1996). The psychology of suspense in dramatic exposition. *Suspense: Conceptualizations, theoretical analyses, and empirical explorations*, 199.
- Zimring, F. E., Piquero, A. R., & Jennings, W. G. (2007). Sexual delinquency in Racine: Does early sex offending predict later sex offending in youth and young adulthood? *Criminology & Public Policy, 6*(3), 507.
- Zimring, F. E., Jennings, W. G., Piquero, A. R., & Hays, S. (2009). Investigating the continuity of sex offending: Evidence from the second Philadelphia birth cohort. *Justice Quarterly, 26*(1), 58.

Appendices

Appendix 1: Sage Journal of Assessment - Submission Guidelines

Manuscript Submission Guidelines:

The editor invites high quality manuscripts covering a broad range of topics and techniques in the area of psychological assessment. These may include empirical studies of assessment of personality, psychopathology, cognitive functions or behavior, articles dealing with general methodological or psychometric topics relevant to assessment, or comprehensive literature reviews in any of these areas. This journal encourages submissions evaluating a) new assessment methodologies and techniques for both researchers and practitioners, b) how assessment methods and research informs understanding of major issues in clinical psychology such as the structure, classification, and mechanisms of psychopathology, and c) multi-method assessment research and the integration of assessment methods in research and practice. Additionally, the journal encourages submissions introducing useful, novel, and non-redundant instruments or demonstrating how existing instruments have applicability in new research or applied contexts. All submissions should provide strong rationales for their efforts and articulate important implications for assessment science and/or practice.

Research participants may represent both clinical and nonclinical populations. Manuscripts should include how sample size has been determined, all data exclusions, all manipulations, and all measures in the study.

In general, regular articles should not exceed 30 pages of text, excluding Title Page, Abstract, Tables, Figures, Footnotes and Reference list.

Authors submitting manuscripts to the journal should not simultaneously submit them to another journal, nor should manuscripts have been published elsewhere, including the World Wide Web, in substantially similar form or with substantially similar content.

Manuscript Submission:

Manuscripts must be submitted in Microsoft Word or Rich Text Format (rtf) electronically at <https://mc.manuscriptcentral.com/asmnt>. Figures may be submitted using any of the formats listed below. If requesting a masked anonymize review, please ensure that both a manuscript file with no identifying author information and a separate title page with author details are included in your submission. Questions should be directed to the ASSESSMENT Editorial Office by email: assessment.editorial@gmail.com.

If you or your funder wish your article to be freely available online to nonsubscribers immediately upon publication (gold open access), you can opt for it to be included in SAGE Choice, subject to the payment of a publication fee. The manuscript submission and peer review procedure is unchanged. On acceptance of your article, you will be asked to let SAGE know directly if you are choosing SAGE Choice. To check journal eligibility and the publication fee, please visit [SAGE Choice](#). For more information on open access options and compliance at SAGE, including self/author archiving deposits (green open access) visit [SAGE Publishing Policies](#) on our Journal Author Gateway.

Preparation of Manuscripts:

Authors should carefully prepare their manuscripts in accordance with the following instructions.

Authors should use the [Publication Manual of the American Psychological Association](#) as a guide for preparing manuscripts for submission. All manuscript pages, including reference lists and tables, must be typed double-spaced.

The first page of the paper (the title page) should contain the article title, the names and affiliations of all authors, authors' notes or acknowledgments, and the names and complete mailing addresses of the corresponding author. If requesting a masked anonymize review, the first page should contain only the article title and the title page should be uploaded as a separate document.

The second page should contain an abstract of no more than 150 words and five to seven keywords that will be published following the abstract.

The following sections should be prepared as indicated:

Tables. Each table should be fully titled, double-spaced on a separate page, and placed at the end of the manuscript. Tables should be numbered consecutively with Arabic numerals.

Footnotes to tables should be identified with superscript lowercase letters and placed at the bottom of the table. All tables should be referred to in the text.

Figures. Electronic copies of figures can be submitted in one of the following file formats: TIFF, EPS, JPEG, or PDF. All figures should be referred to in text. Each figure should appear on a separate page at the end of the manuscript but before the tables, and all titles should appear on a single, separate page.

Endnotes. Notes should appear on a separate page before the References section. Notes should be numbered consecutively, and each endnote should be referred to in text with a corresponding superscript number.

References. Text citations and references should follow the style of the [Publication Manual of the American Psychological Association](#).

Appendix 2: Prediction model Risk of Bias Assessment Tool (PROBAST)

| | | |
|--|--|------------------|
| Domain 1: Participants | | |
| A. Risk of Bias | | |
| <i>Describe the sources of data and criteria for participant selection:</i> | | |
| | | Validation Study |
| 1.1 Were appropriate data sources used? | | |
| 1.2 Were all inclusions and exclusions of participants appropriate? | | |
| Risk of bias introduced by selection participants | RISK: <i>(low/high/unclear)</i> | |
| <i>Rationale of bias rating:</i> | | |
| B. Applicability | | |
| <i>Describe included participants, setting and dates:</i> | | |
| Concern that the included participants and setting do not match the review question | CONCERN: <i>(low/high/unclear)</i> | Low |
| <i>Rationale of applicability rating:</i> | | |

| | | |
|---|--|------------------|
| Domain 2: Predictors | | |
| A. Risk of Bias | | |
| <i>List and describe predictors included in the final model, e.g., definition and timing of assessment:</i> | | |
| | | Validation Study |

| | | |
|--|--|------------|
| 2.1 Were predictors defined and assessed in a similar way for all participants? | | |
| 2.2. Were predictor assessments made without knowledge of outcome data? | | |
| 2.3 Are all predictors available at the time the model is intended to be used? | | |
| Risk of bias introduced by predictors of their assessment | RISK: <i>(low/high/unclear)</i> | |
| <i>Rational of bias rating:</i> | | |
| B. Applicability | | |
| <i>Describe included participants, setting and dates:</i> | | |
| Concern that the definition, assessment or timing of predictors in the model do not match the review question | CONCERN: <i>(low/high/unclear)</i> | Low |
| <i>Rationale of applicability rating:</i> | | |

| | |
|---|------------------|
| Domain 3: Outcome | |
| A. Risk of Bias | |
| <i>Describe the outcome, how it was defined and determined, and the time interval between predictor assessment and outcome determination:</i> | |
| | Validation Study |
| 3.1 Was the outcome determined appropriately? | |
| 3.2 Was a pre-specified or standard outcome definition used? | |

| | | |
|--|--|--|
| | | |
| 3.3 Were predictors excluded from the outcome definition? | | |
| 3.4 Was the outcome defined and determined in a similar way for all participants? | | |
| 3.5 Was the outcome determined without knowledge or predictor information? | | |
| 3.6 Was the time interval between predictor assessment and outcome determination appropriate? | | |
| Risk of bias introduced by the outcome or its determination | RISK: <i>(low/high/unclear)</i> | |
| <i>Rationale of bias rating:</i> | | |
| B. Applicability | | |
| <i>At what time point was the outcome determined:</i> | | |
| <i>If a composite outcome was used, describe the relative frequency/distribution of each contributing outcome:</i> | | |
| Concern that the outcome, its definition, timing or determination do not match the review question | CONCERN: <i>(low/high/unclear)</i> | |
| <i>Rationale of applicability rating:</i> | | |

| |
|---|
| Domain 4: Analysis |
| A. Risk of Bias |
| <i>Describe numbers of participants, number of candidate predictors, outcome events and events per candidate predictor:</i> |

| | |
|--|------------------------------------|
| <i>Describe how the model was developed (for example, in regard to modelling technique (e.g. survival or logistic modelling), predictor selection and risk group definition):</i> | |
| <i>Describe whether and how the model was validated, either internally (e.g., bootstrapping, cross validation, random split sample) or externally (e.g. temporal validation, geographical validation, different settings, different type of participants):</i> | |
| <i>Describe the performance measures of the model, e.g. (re)calibration, discrimination, (re)classification, net benefit and whether they were adjusted for optimism:</i> | |
| <i>Describe any participants who were excluded from the analysis:</i> | |
| <i>Describe missing data on predictors and outcomes as well as methods used for missing data:</i> | |
| | Validation Study |
| 4.1 Were there a reasonable number of participants with the outcome? | |
| 4.2 Were continuous and categorical predictors handled appropriately? | |
| 4.3 Were all enrolled participants included in the analysis? | |
| 4.4 Were participants with missing data handled appropriately? | |
| 4.6 Were complexities in the data accounted for appropriately? | |
| 4.7 Were relevant model performance measures evaluated appropriately? | |
| Risk of bias introduced by analysis | RISK: (low/high/unclear) |
| <i>Rational of bias rating:</i> | |

Appendix 3: Ethical Approval

Re: Ethics Application: Assessing Sexual Deviance - A Criminal Justice Social Work Perspective

CLINICAL PSYCHOLOGY Research Ethics <submitting.ethics@ed.ac.uk>

Thu 04/03/2021 14:14

To: FORBES Claire

 1 attachments (129 KB)

Assessing Sexual Deviance - Ethics Form - 11.12.2020_Forbes_final.docx

Dear Claire,

Thank you for your revised application and related documents; and apologies for the delay in responding. The ethics mailbox has been very busy following the holidays, so we've been working through a backlog. Even so your reminder was very timely and we appreciate it, as for some reason your application was sorted into amendments. I have now retrieved it and checked through your changes.

Based on your responses the project meets requirements for favourable opinion from Clinical Psychology Ethics. If you require a formal letter of ethics approval (this is only required if you are approaching third parties, NGOs etc) then please contact the ethics mailbox requesting this and a formal letter of approval will follow in due course.

However, in the interim please use this email and signed application as proof of approval.

Good luck with your project.

Best wishes,
Ingrid

Ingrid Obsuth, PhD
Lecturer in Clinical Psychology
Ethics & Integrity Lead

Appendix 4: Participant Information Handout

Participant Information Sheet

Project Title: Assessing Sexual Deviance: A Criminal Justice Social Work Perspective

Name of Researcher: Claire Forbes, Trainee Clinical Psychologist

Email:

Tel:

My name is Claire Forbes, and I am undertaking this study as part of my Doctorate in Clinical Psychology. I'd like to thank you for considering participating. Before you decided to do so, I want to be sure that you understand 1) Why I am doing this study? and 2) What it will involve for you?

Please read through the following information and be sure to ask any questions you might have.

What is the purpose of this study?

The research around risk assessments for sexual recidivism suggests that risk is predicted by an interplay of different factors. The evidence suggests that sexual deviance is a strong predictor for sexual recidivism, however different research studies use different definitions and ways to assess this. There seems to be a lack of clarity around practice in this area. This study seeks to gather the views of the people who actually make these risk decisions every day, and how they consider sexual deviance and related constructs when assessing risk.

Who is doing this study?

The study is run by Claire Forbes, a psychology postgraduate student at the University of Edinburgh and a Trainee Clinical Psychologist in NHS Lothian. The research is funded by NHS Education for Scotland.

Claire is being supervised by Dr David Gillanders (Head of Clinical & Health Psychology, The University of Edinburgh) and Dr Louise Tansey (Consultant Forensic Clinical Psychologist, Serious Offenders Liaison Service & Forensic Clinical Psychologist, The Orchard Clinic).

Why have I been asked to take part?

All Criminal Justice Social Workers who have been qualified for at least a year are invited to take part in this study.

Do I have to take part?

Taking part in this study is completely voluntary. It is up to you to decide whether or not to take part. If you decide to participate you can change your mind before and during the study and withdraw. Any information you have given up until that point will be deleted.

Once you have completed the interview, the information you have given will be pseudonymised and incorporated into larger themes where it will no longer be identifiable. This means that after the interview has been given you will not be able to withdraw.

Before taking part, you will be asked to read and sign a consent form which explains what the study involves and what your rights are as a participant. If you have any questions, don't hesitate to ask at this time.

What does taking part involve?

If you opt-in to the study, you will meet with Claire Forbes either in person, via Microsoft Teams or by phone.

Before the interview starts, Claire will recap the information in this sheet with you and remind you what is involved with the interview process, including how your data is recorded and stored securely. You will have time to ask questions and discuss any concerns you may have. You will then be asked to sign a consent form either using paper and pen or via email.

In the interview you will be asked about your experience working in risk assessment for sexual recidivism and your thoughts around what sexual deviance is. It is important that you know there are no wrong answers. The aim is just to better understand your perspective and experience.

The interview will be recorded so that the researcher can later transcribe the interview. The transcription will be pseudonymised. Please see the sections on “Confidentiality and Data Protection” for more information.

You can ask for the interview to be stopped at any time and you do not have to give a reason. You can also ask to talk to the researcher again at a later date if you would like to talk about the interview process.

After the interview you will have some time to talk to the researcher about how it felt and to debrief together. You will also be asked if you would be willing to opt-in to be interviewed again later down the line and if you would be willing to look at the conclusions the researcher comes to towards the end of data collection. This helps to check the credibility of the research to your lived experience. You do not have to participate in these additional segments if you do not want to and again you will be able to withdraw from these if you change your mind.

Face-to-Face Interviews

If you would prefer a face-to-face interview this will only occur if acceptable under the government restrictions and local authority guidance in place at the time. Additionally, during you interview, the risk of exposure to COVID-19 for yourself and Claire will be mitigated in line with the most up to date Scottish Government guidance at the time (i.e., physical distancing arrangements, arrangements for hand washing/ sanitisation and drying, procedures for cleaning of surfaces and communal areas, use of face coverings and/or maintenance of 2meter physical distancing).

Further, you will only interact with researchers who have experienced no COVID-19 symptoms nor had any known contact with COVID-19 positive individuals for the 14 days prior to the research interaction.

If you feel unwell, experience COVID-19 related symptoms, or have been in contact with a COVID-19 positive individual in the past 14 days, then please contact the Claire to rearrange or cancel the interview. If you experience COVID-19 related symptoms, and/or have a positive COVID-19 test following the research interaction, please follow the Scottish Government guidance (or local equivalent).

What are the possible benefits of taking part?

There are unlikely to be direct or immediate benefits to you, though I hope you would find it interesting. The aim of this study is to benefit our understanding of best risk assessment practice, leading to improvements in training, and risk management research.

Confidentiality and Data Protection

The information that you provide will be kept securely and confidentially. The interviews will be recorded on a device that will be kept in a locked case when there are recordings on it. The researcher aims to transcribe each recording within a week of the interview at which point the recording will be deleted.

During transcription your information will be pseudonymised and stored under a unique participant number. This means that identifiable information shall be removed from the transcript, including names and locations. Additionally, any anecdotes or references that may make you identifiable will be redacted. During the study, the pseudonymised transcripts and working documents for the research will be stored in password protected documents on the Researcher's University OneDrive – a private document store.

The demographic information which you provide (for example, preferred gender, age, years of experience) will be stored separately in a document with your participant number. This information is kept as it allows the researcher to describe the participant sample at the end and to observe potential biases or patterns which arise during the research process. To prevent identification, your demographic information will be stored in broader categories, for example, instead of “age 42” this would be stored as “40-50 years old”. This document, along with copies of consent forms and any other documents with participant information will be stored separately, in password protected files on the Researchers secure NHS Drive.

After the study's completion the pseudonymised data will be stored for 10 years. It will be compressed into a zip file and stored in a password protected NHS drive, with restricted access.

The researcher will only break confidentiality if something that you disclose requires them to act under their duty of care to the public, for example if you disclosed an unreported crime or abuse of power.

If you have any questions or concerns about confidentiality or want to know more about how your data will be stored, please do not hesitate to ask the researcher at any time.

What will happen to the results of the study?

The results of the study will be available by May 2022. The results will talk about the experiences of groups of people who have participated and not identify any one person. The researcher will offer to present the findings to your teams and a written summary shall also be circulated via email to the teams who participated.

Who has reviewed the study?

This study has received ethical approval from the University of Edinburgh Ethics Committee and the Ethics Committee associated with your Local Authority Area.

Where can I lodge feedback or a complaint about this study?

If you would like to raise a complaint about the research, please contact Matthias Schwannauer, Head of School:

Appendix 5: Example of Line-by-Line Coding

| | |
|---|---|
| <p>So, there's something else within that fantasy or interest that makes it go from being a "normal" interest or fantasy to having something deviant about it?</p> <p>Yeah, so it may be something about the underlying thought processes and underlying emotional cues, rather than the act that they're wanting to do. I mean sometimes it's the act that they're wanting to do as well, obviously, but eh... but I think that's where there is more nuance than there used to be. Cause when I started, somebody who was into, you know, people in latex or kinda balloon stuff, you'd be like "That counts as deviant" probably. But, em, yeah, I think it's a lot more nuanced now.</p> | <p>Exploring underlying processes</p> <p>Looking beyond content</p> <p>Considering more than action</p> <p>Recognising nuance</p> <p>Perceiving shift in practice</p> |
| <p>Yeah, and can you explain a bit more about the "emotional bits" you mentioned?</p> <p>I suppose, I think that some people are... some people want to engage in rough sexual behaviour for example, because it feels elicit to them, or because it feels dangerous or something and some people want to cause pain to somebody else and it's really about the pain and control that is... that's what they're looking for rather than the kind of sexual aspect to it? But it's the same actions. And similarly, em, some people if you think about the em... the pickup artist type chaps who are really about control and... often it's about control and demeaning the person that they are with but what they're wanting to do is go out to a bar, pick up a em, person of the opposite gender and have fairly normal sex with them, probably. And so that isn't a particularly deviant sexual interest, but what they are looking for is this power dynamic and this control and probably humiliation or that feeling of you know, manipulation or something. I think that actually really brings it into that sort of deviant sexual world a bit more. Em, yeah, so I suppose what I mean, is we're not thinking just about that behaviour anymore, we're thinking about what is it they're doing? How would we target it, if we were looking at it on a porn site what category would you put it under? Yeah, em, yeah.</p> | <p>Considering motivations for action</p> <p>Looking beyond fantasy and behaviour</p> <p>Recognising different pathways to offending</p> <p>Deviant motives vs. Deviant acts</p> <p>Thinking about deviant sexual interests</p> <p>Differing interests from motives</p> <p>Hypothesising drives to understand in assessment</p> <p>Looking beyond behaviour</p> <p>Understanding behaviour in societal categories</p> |

Appendix 6: Samples from Researchers Memo's

13.07.2021 – Just coded participant 2

Note to self: while coding, keep this memo set open too so that I can jot down thoughts to capture revelations and understanding as they occur.

It's hard not to compare but following this interview, I was struck by the difference in his approach to participant 1's. I remember that at the time of the interview as well. Even just in terms of the amount of speech he produced. He felt more reflective and excited by the topic.

My overall view from him was recognising that sexual deviance is part of a wider whole in understanding offending. Actually, a lot of his work was based in behavioural practice – it was about looking beyond the offending behaviour and understanding the underlying purpose/reward to this – sometimes it was about sexual deviance and sometimes it was about something else entirely, which in his definition was a secondary type of sexual deviance.

He thought of sexual deviance as a core sexual functioning that is an offense in itself (does that relate to interest or just behaviours?) and of sexual deviance as those who use sexual behaviours to meet other needs/achieve other ends, such as pain and control. This is the first time I've thought about this distinction as so far, the focus has been on sexual offending for sexual gratification. I found the discussion around the same act with different underlying motives and progressions having totally different risk profiles very interesting.

- 22/11/21 – Interesting as participant 10, also felt very confident working in this area but saw SD as very much rooted in the presence/absence of gratification.

29.07.21 - Thesis Supervision – Clinical Input

Presenting early models that are coming through in the initial coding that we can then aim to saturate over further interviews.

Liz and Louise were keen to consider more about:

- Personality
How much the social workers are influenced by the personality of the offender?
I'm not as sure how this fits in with my research questions – going to see how it comes through in the interviews I've collected so far
- Management
Are they considering how they can manage sexual behaviour vs. how they can manage sexual deviance? If the first one how much is consideration of deviance is actually happening?

- Formulation

How much are they formulating the risk beyond the risk factors?

So far this seems to be mediated by interest and confidence within the team

Currently there are two sections to the codes:

- What – how are they defining/referring to sexual deviance?
- How – how are they incorporating that into their risk assessments?

Some initial thoughts of models coming through are here.

Liz highlighted that the public perception model is also a loop as the social work decisions feed back into public perception and court sentencing.

They were quite interested in the lilac model of what other factors help them to understand the risk associated with sexual deviance.

- 22.11.2021 - Looking back, personality traits have not come out of participant data, though management and formulation have. It's helpful that the way in which my clinical supervisors thought about these things helped me to notice when these concepts came up but have to be careful after every supervision not to be pulled into what they observe in their clinical practice and away from my participants experiences.