Risk for Sexual Violence Protocol (RSVP):

A real world study of the reliability, validity and utility of a structured professional judgement instrument in the assessment and management of sexual offenders in South East Scotland

Originally submitted: June 2015 Published online: January 2016

Authors:

Rajan Darjee, Katharine Russell, Lauren Forrest, Erica Milton, Valerie Savoie, Emily Baron, Jamie Kirkland & Stewart Stobie

NHS Lothian Sex Offender Liaison Service Orchard Clinic, Royal Edinburgh Hospital Morningside Terrace, Edinburgh, EH10 5HF Tel: 0131 537 5866 Email: rajan.darjee@nhs.net katharine.russell@nhs.net

Research partly funded by Risk Management Authority and the NHS Scotland Forensic Network

Abstract/Executive Summary

The Risk for Sexual Violence Protocol (RSVP; Hart, Kropp, & Laws; Klaver, Logan, & Watt, 2003) is a Structured Professional Judgment (SPJ) instrument for the assessment and management of individuals considered to pose a risk of sexual violence. It is widely used in Scotland in criminal justice and forensic mental health settings, particularly for more complex or concerning cases. However there have been very few studies on the reliability, predictive validity and utility of the RSVP. In the Risk Management Authority's (RMA) Risk Assessment Tools Evaluation Directory (RATED; Risk Management Authority Scotland, 2015) the RSVP is assessed as 'awaiting validation'.

The current study was of 109 individuals considered to pose a risk of sexual violence who were assessed using the RSVP by the Sex Offender Liaison Service (SOLS), based in Edinburgh, between 2006 and 2013, and then prospectively followed up for an average of 3 ¹/₄ years. Multi-agency records were used to ascertain whether there were further allegations, charges or convictions for sexual and non-sexual offending during follow-up, as well as the level of risk management delivered to cases during follow up.

The cases referred to SOLS were an unusual group of sexual offenders, in that they appeared higher risk and more complex than other sex offenders. During follow-up, 11.9% received a further sexual conviction, but when considering unconvicted allegations and charges, the actual rate of further sexual offending during follow-up was 23.9%. Eleven cases were independently rated by two assessors to ascertain reliability. Individual item reliability was good to excellent, and the reliability of summary risk judgements was excellent. There was a strong correlation between RSVP ratings and other measures, such as Risk Matrix 2000 (RM2000; Thornton, 2010) and the Psychopathy Check List-Revised (PCL-R; Hare, 1991).

RSVP Total Scores and Summary Judgments did not predict sexual offending during followup using ROC analysis, but using survival analysis Case Prioritization ratings did predict time to further sexual offending. RSVP Total Scores and Summary Judgements did predict serious sexual offending, violent offending and serious offending (whether sexual or non-sexual).

When taking into account level of risk management during follow-up, high risk cases who did not receive a high level of risk management reoffended more frequently and more quickly. Low risk offenders had low rates of reoffending when they received low levels of risk management. Moderate risk cases had relatively high rates of reoffending when they received medium levels of risk management.

The risk scenarios generated by the RSVP assessments were a good match to the actual subsequent sexual offences committed by recidivists with respect to victim age, gender, relationship and severity of offence committed.

This is the first prospective validation study of the RSVP and the first study of the RSVP not just based on case records. The cohort studied is unusual and at the more extreme end of the spectrum of sexual offenders. Our findings support the use of the RSVP to assess and manage risk of serious harm in sexual offenders. The RSVP may have a specific role to play in the management of more concerning cases, such as individuals being considered for an Order for Lifelong Restriction (OLR) or those managed under Multi-Agency Public Protection Arrangements (MAPPA) who are considered to pose a high or very high risk. This role fits with the principles of the Framework for Risk Assessment Management and Evaluation (FRAME; Risk Management Authority Scotland, 2011), where the RSVP would be used with sexual offenders who require a 'scrutinise' assessment and more intensive risk management. Given the limited research on the RSVP, further studies are required.

Content

Introduction	1
-Risk for Sexual Violence Protocol	3
-Researching SPJ tools	4
-Predictive validity of the RSVP	5
-Reliability of the RSVP	6
-Utility of the RSVP	6
-Study aims and objectives	6
Method	8
-Setting and service	8
-Assessment process	8
-Ethical and management approval	9
-Sample	9
-Measures used at initial assessment	9
<u>-Risk Matrix 2000 (RM2000)</u>	9
-Psychopathy Checklist Revised (PCL-R)	10
<u>-International Personality Disorder Examination (IPDE)</u>	10
-Structural Behavioural Assessment of Paraphilias	10
-Risk for Sexual Violence Protocol (RSVP)	11
-Other baseline data	12
-Follow-up data	12
-Follow-up offending, incidents, recall	12
<u>-Risk management level</u>	13
-Analysis and statistical methods	13
-Description of baseline and follow-up data	13
- <u>Reliability</u>	13
-Relation to other measures	14
<u>-Predictive validity of RSVP</u>	14
-Relationship between RSVP total score or	1.4
summary juagments and outcomes	14
-Relationship between RSVP summary	
judgements and outcomes taking into account	14
Predictive Value of Pisk Scenarios	14
-rieucuve value of Kisk Scenarios	15
Results	16
-Baseline characteristics of the sample	16
-Description of offending during follow-up	20
-RSVP ratings	22
-Reliability of RSVP Items, Totals and Summary Judgements	23
-Convergent validity with other instruments	26
-Predictive validity of RSVP	29
<u>-ROC Analysis</u>	29
-Offending during full follow-up	29
-Any sexual offending	29
-Any violent offending	29
-Any serious offending	29
-Any serious sexual offending	29

-Any offending at all	30
-Conviction during full follow-up	30
-Any sexual conviction	30
-Any violent conviction	31
-Any serious conviction	31
-Any serious sexual conviction	31
-Any conviction	31
-Breach of conditions	32
-Survival analysis	32
-Comparison of RSVP Summary Judgements and Risk	
Management Level	44
-Predictive Validity of Scenario Planning	54
Discussion	55
-Main Findings	58
-How reliable are ratings of items, Sections,	
total scores and summary judgments?	58
-How do RSVP ratings correspond with ratings using	
other instruments, such as Risk Matrix 2000 and the	
Psychopathy Check List-Revised?	59
-How do RSVP ratings (scores and summary judgments)	
predict further offending (including convicted and unconvicted	
offending; and including both sexual and non-sexual offending)?	60
-How are RSVP summary judgements related to further	
offending after taking into account the level of management	
cases are subject to?	62
-Do the risk scenarios generated by clinicians using the	
RSVP include the types of sexual violence that recidivist	
offenders go on to perpetrate?	64
-Summary of findings	65
-Research Implications	65
-Practice Implications	66
References	68
Appendix	75

Tables

Table 1. Categories of cases ascertained by taking into account both	
assessed level of risk and risk management received.	15
Table 2. Baseline characteristics of sample of 109 cases	17
Table 3. Details of index offence	18
Table 4. Offending history	19
Table 5. Offending behaviour (alleged, charged or convicted) during	
whole follow-up period ($N = 109$)	21
Table 6. Offending during whole follow-up period	21
Table 7. Follow up offending for specific durations of follow-up	22
Table 8. RSVP summary judgements	22
Table 9. Inter-rater Reliability of Past RSVP Items $(N = 11)$	23
Table 10. Inter-rater Reliability of Recent RSVP Items $(N = 11)$	24
Table 11. Inter-rater Reliability of Relevance RSVP Items ($N = 11$)	25
Table 12. Reliability of RSVP Summary Judgements, Total Scores, (N = 11)	26
Table 13. Correlations (RSVP Total Scores, Section totals and other Instruments)	27
Table 14. ANOVA exploring the association between Case Prioritisation	
ratings and ratings on other assessment tools.	28
Table 15. ANOVA exploring the association between Risk of Serious	
Physical Harm ratings and ratings on other assessment tools.	28
Table 16. ANOVA exploring the association between Immediate Action	
Required ratings and ratings on other assessment tools.	29
Table 17. ROC analysis: relationship between total scores and summary	
judgments, and offending (including convictions, charges and	
allegations) over the whole follow-up period.	30
Table 18. ROC analysis: relationship between total scores and summary	
judgments, and conviction over the whole follow-up period.	31
Table 19. ROC Analysis: relationship between total scores and summary	
judgments, and breaches of conditions at 6 moths, 12 months	
and 5 years.	32
Table 20. Assessed Case Prioritisation Level by Modal Managed Level,	
Frequency (%)	45
Table 21. Assessed Risk of Serious Physical Harm Level by Modal	
Managed Level, Frequency (%)	45
Table 22. Assessed Case Prioritisation by Managed Mode at 1 Year	
for Any Sexual Offending (%)	46
Table 23. Assessed Case Prioritisation by Managed Mode for Any	
Sexual Offending in Follow up (%)	46
Table 24. Assessed Risk of Serious Physical Harm by Managed Mode	
for Any Serious Sexual Offending (%)	47
Table 25. Scenario Planning Consensus Frequency (%)	54
Table i. RSVP individual items: means (medians)	75
Table II. RSVP individual items: number and percentage of cases who	
received each of the three ratings levels (no, query or yes) for	
each of the four types of rating (Past presence, Recent presence,	
Ever present, future Relevance).	76

Figures

Figure	1. Kaplan-Meier survival curves for time to any sexual offence	
	in the three RSVP case prioritisation groups	33
Figure	2. Kaplan-Meier survival curves for time to any sexual offence	
	in the three RSVP risk of serious physical harm groups	34
Figure	3. Kaplan-Meier survival curves for time to any sexual offence	
	in the three RSVP immediate action required groups	35
Figure	4. Kaplan-Meier survival curves for time to any violent offence	
	in the three RSVP case prioritisation groups	36
Figure	5. Kaplan-Meier survival curves for time to any violent offence	
	in the three RSVP risk of serious physical harm groups	37
Figure	6. Kaplan-Meier survival curves for time to any violent offence	
	in the three RSVP immediate action required groups	38
Figure	7. Kaplan-Meier survival curves for time to any breach in the	
	three RSVP case prioritisation groups	39
Figure	8. Kaplan-Meier survival curves for time to any breach in the	
	three RSVP risk of serious physical harm groups	40
Figure	9. Kaplan-Meier survival curves for time to any breach in the	
	three RSVP immediate action required groups	41
Figure	10. Kaplan-Meier survival curves for time to any sexual offence	
	for RSVP Ever Present Total scores divided into three groups	42
Figure	11. Kaplan-Meier survival curves for time to any violent offence	
	for RSVP Ever Present Total scores divided into three groups	43
Figure	12. Kaplan-Meier survival curves for time to breach for RSVP	
	Ever Present Total scores divided into three groups	44
Figure	13. Kaplan-Meier survival curves for time to any sexual offence	
	in the discrepancy between assessed level and managed level	
	for those who committed any sexual offence during follow up	48
Figure	14. Kaplan-Meier survival curves for time to any sexual offence in	
	the discrepancy between assessed level and managed level for those	
	who committed any serious sexual offence during follow up	49
Figure	15. Kaplan-Meier survival curves for time to any violent offence in	
	the discrepancy between assessed level and managed level for those	
	who committed any sexual offence during follow up	50
Figure	16. Kaplan-Meier survival curves for time to any violent offence in	
	the discrepancy between assessed level and managed level for those	
	who committed any serious sexual offence during follow up	51
Figure	17. Kaplan-Meier survival curves for time to breach in the discrepancy	
	between assessed level and managed level for those who committed	
	any sexual offence during follow up	52
Figure	18. Kaplan-Meier survival curves for time to breach in the discrepancy	
	between assessed level and managed level for those who committed	
	any serious sexual offence during follow up	53

Introduction

There is longstanding debate in the literature about the best approach to risk assessment. Research has identified a range of factors that are associated with risk of offending. However, individual offending behaviour remains an intrinsically unpredictable phenomenon. Therefore, it is vital that professionals use valid and reliable procedures that enable the effective identification, understanding and management of risks posed by clients who offend. Three distinct generations of risk assessment have been described by Doyle and Dolan (2002): (1) The unstructured clinical judgement approach; (2) the actuarial approach and (3) the structured clinical judgement approach.

Unstructured clinical judgement is based on expert opinions of experienced professionals. This approach was once accepted as a sufficient basis for legal and clinical decision making. However, it has little reliability, validity, transparency or evidence base and is highly susceptible to individual subjectivity. Therefore a structured, consistent, transparent and evidence based approach is required.

Actuarial approaches use group projections based on empirical evidence to make predictions about risk of future violence. Evaluators are required to rate risk factors of assigned weight, which are then combined in an algorithm to estimate level of risk. The algorithm is based on data collected from following up large groups of offenders over specified time periods, producing a model that identifies the factors, that when combined, predict those who reoffend. Actuarial measures may appear to have the following advantages: evidence-based; transparent; systematic; objective; standardised and cost-effective (de Vogel, 2005). A potential additional advantage is that they can help to allocate resources when dealing with a large caseload. The main disadvantage is that the assessment provides very little information about the actual individual being assessed. It merely allows the assessor to say that amongst a group of offenders with similar characteristics and risk factors a particular number would be expected to reoffend. The specific proportion is unlikely to be directly applicable to a case in a different jurisdiction, context or time from the original sample. It does not allow the assessor to consider the nature, imminence, severity of harm, context or frequency of potential future reoffending. Hart, Michie, and Cooke (2007) examined the margins of error for risk estimates made using actuarial methods and found that these margins of error were unacceptably large. They concluded that mathematical models based on group data are

difficult, if not impossible, to apply with any precision or utility to an individual case. Furthermore, actuarial methods are of limited practical value as they do not inform case formulation or identify changes in risk level or guide risk management (Sutherland et al., 2012). If actuarial methods are used, they should be interpreted with caution and should not form the sole basis of risk judgements.

The Structured Professional Judgement (SPJ) approach is used to provide comprehensive risk assessments that are based on the scientific and professional literature. Like actuarial instruments SPJ instruments have a specific set of defined risk factors that an assessor must rate objectively. However they allow freedom of clinician decision making whilst maintaining consistency, transparency and a degree of objectivity. SPJ tools may be particularly useful when used with more complex cases, for example offenders with underlying personality disorders (Hart & Logan, 2011). The inclusion of a formulation of risk is a critical component in the process that enhances the ability of the SPJ tool to produce an individualised understanding of risk. Evaluators are guided through the process of assessment, formulation and risk management planning. There is an emphasis on understanding and managing risk as opposed to predicting future offending. SPJ tools have been rated highly with respect to utility. Green, Carroll, and Brett (2010) and Khiroya, Weaver, and Madden (2009) found that SPJ risk instruments were used widely in forensic mental health settings. They were considered to inform risk management to a greater extent when compared with actuarial tools.

The Risk Management Authority (RMA) recommends the use of SPJ methods for more complex and challenging cases. The RMA specifically stipulates that SPJ tools should be used to aid legal decision-making for offenders being considered for an Order for Lifelong Restriction (OLR) - a sentence of lifelong supervision for high risk violent and sexual offenders (Risk Management Authority Scotland, 2015; Criminal Justice (Scotland) Act, 1995). The RSVP is the most commonly used SPJ tool for sexual violence when sexual offenders are assessed where courts are considering imposing an OLR (Darjee & Russell, 2011), and is in widespread use in forensic mental health and prison in Scotland and elsewhere. However, due to the limited number of empirical studies, the RSVP is categorised as awaiting validation by the RMA (Risk Management Authority Scotland, 2013).

Risk for Sexual Violence Protocol

(RSVP; Hart, Kropp, & Laws; Klaver, Logan, & Watt, 2003)

The RSVP is an SPJ risk assessment tool, developed following a systematic review of the sexual recidivism literature. The RSVP defines sexual violence as "actual, attempted or threatened sexual contact with another person that is non-consensual" (Hart et al., 2003). It evolved from earlier SPJ tools such as the precursor of the RSVP, the Sexual Violence Risk-20 (SVR-20) and the widely used violence risk assessment tool, the Historical Clinical Risk-20 (HCR-20; Webster, Douglas, Eaves & Hart, 1997). The RSVP can be used with men aged 18 and older who have a known or suspected history of sexual violence. The RSVP is intended to help evaluators conduct a comprehensive assessment of risk of sexual violence in clinical and forensic settings. The evaluator must gather comprehensive case information from multiple sources and assess the offender in relation to twenty-two individual risk factors as well as any additional case-specific risk factors. The twenty-two factors are divided into five Sections: Sexual Violence History, Psychological Adjustment, Mental Disorder, Social Adjustment, and Manageability. Each item is coded three times: for presence in the Past, Recent presence and future Relevance. Each of these ratings is on a three point scale: no evidence, partial evidence, or definite evidence. The evaluator must determine the relevance of the individual risk factors with respect to potential future sexual violence and the development of risk management plans, describe the most plausible scenarios of future sexual violence, and recommend strategies for managing sexual violence risk in light of the relevant factors and scenarios.

The RSVP manual stipulates that those using the tool should have an appropriate level of experience, competence and knowledge. Important features of the RSVP manual are that it provides an evidence-based rationale for each item, clear assessment guidelines and detailed operationalisation of terms and ratings. Specialist training workshops are provided to practitioners but are not mandatory to use the instrument. Formal training in the use of the RSVP is recommended (Hart et al., 2003) and there is evidence that such user training programmes enhance inter-rater reliability of assessment measures (Reichelt, James & Blackburn, 2003; Muller & Wetzel, 1998; Sutherland et al, 2012). Similarly, according to Darjee and Russell (2012), it is important that those who use these assessment instruments know their strengths and limitations, and have received appropriate training in their use and interpretation. It is important that they know how to interpret the output of any tool in order to

reach appropriate conclusions and plan risk management appropriately. So an important feature of SPJ tools like the RSVP, as opposed to actuarial tools, is that they depend not only on the manual and scoring of items but also on the practitioner who is using the instrument. They structure the practitioner in their task, they do not replace them.

Researching SPJ tools

In the research into the validity of SPJ tools, the use of total scores rather than SPJ summary judgement ratings (Heilbrun, Douglas, & Yasuhara, 2009) has been criticised, particularly as this is evaluating the tool in a different way to that which it is intended for use. Much of the research in the risk assessment literature has focused on the accuracy of risk prediction rather than the evaluation of risk management planning and the goal of violence prevention (Guy, 2009; Hart 1998; Douglas & Kropp, 2002). Douglas and Kropp (2002) proposed a research paradigm for SPJ tools highlighting that the use of outcome data fails to take into account the level of management or intervention that the offender received during the follow-up period. There has been very little written on the effect of matching intervention level to risk level through the use of SPJ tools. Vincent, Guy, Gershenson, & McCabe (2012) found that the matching of resource and intervention to level of risk in a group of juveniles, using the Structured Assessment of Violence Risk in Youth (SAVRY; Borum, Bartel, & Forth, 2006), resulted in a reduction in the use of resource without an increase in reoffending. Belfrage, Strand, Storey, Gibas, Kropp & Hart (2012) found that risk management level mediated the association between risk assessment and recidivism, i.e. high levels of intervention were associated with decreased recidivism in high risk Intimate Partner Violence (IPV) cases using the Spousal Assault Risk Assessment (SARA) Guide (Kropp, Hart, Webster & Eaves, 1995). Interestingly they found a high level of intervention in low risk cases was associated with increased recidivism

Retrospective assessments by researchers using case records constitute a weaker form of design because the outcome is already known prior to the assessment taking place (Guy, 2009). Furthermore, the use of reconviction as the only way of measuring reoffending gives a conservative reflection of further violent or criminal behaviour (Pedersen, Rasmussen, & Elsass, 2010). Other outcomes such as further breaches, charges or recall could give a more realistic picture of further problematic behaviour. There is little research on SPJ instruments as interventions to reduce risk (Douglas & Kropp, 2002) and little or no published research on formulations, scenarios and risk management strategies.

Predictive validity of the RSVP

There is a significant body of research exploring the validity and reliability of SPJ tools, but limited research specifically on the RSVP. The predictive validity of a tool is considered to be a useful method of assessing its efficacy. Structured professional judgement tools have been reported to perform better than unstructured methods but less well than (or sometimes as well as) actuarial tools using this criterion (Hanson & Morton-Bourgon, 2009). It should be noted that the predictive paradigm in research on risk assessment instruments should favour actuarial tools as they are developed mathematically using this paradigm. However, it has been argued by some that risk assessment tools are required to do more than simply predict recidivism. They should also inform treatment and risk management (Hart & Logan, 2011). Furthermore, because the emphasis of SPJ tools is on the development of risk management strategies that reduce risk, the risk level that an offender poses may not always be reflected in recidivism data. Appropriate risk management should reduce recidivism rates.

We are not aware of any published peer reviewed studies of the predictive validity of the RSVP and according to the Risk Management Authority's (RMA) Risk Assessment Tools Evaluation Directory (RATED) (Risk Management Authority Scotland, 2015) the RSVP awaits validation. Hart and Boer (2010) summarised the research on the predictive validity of the SVR-20 and RSVP. They quoted two unpublished studies of the RSVP by the developers of the instrument that were presented at conferences: Kropp (2001; as cited in Hart & Boer, 2010) and Hart and Jackson (2008; as cited in Hart & Boer, 2010). In a sample of 53 sexual offenders from the USA (of whom 15 recidivated sexually), Kropp (2001; as cited in Hart & Boer, 2010) found Case Prioritization was significantly associated with sexual recidivism but the RSVP total score was not. In a sample of 90 sexual offenders who had completed a community treatment programme in Canada (of whom 18% sexually recidivated over an average of 4 years follow-up), Hart and Jackson (2008; as cited in Hart & Boer, 2010) found Case Prioritization groups had significantly different recidivism rates, that Case Prioritization correlated with recidivism as well as ratings on other risk instruments, and that Case Prioritization ratings had some unique predictive power for recidivism even after controlling for numerical risk scores on the RSVP.

Vojt (2013) studied a small sample of mentally disordered sexual offenders in secure hospital care in Scotland and found no association between RSVP total scores and recidivism. However in that study patients were not released to the community, summary judgements were not ascertained and RSVP ratings produced by clinical teams were used without ascertaining reliability.

Reliability of the RSVP

Hart and Boer (2010) provided an overview of the literature on the inter-rater reliability of the SVR-20 and the RSVP. They pointed to three unpublished studies that examined the interrater reliability of the RSVP (Hart, 2003; Watt, Hart, Wilson, Guy, & Douglas, 2006; Watt & Jackson, 2008). All studies found that inter-rater reliability of ratings for individual presence and Relevance factors was good (ICC1 .5 - ICC1 .74) to excellent (ICC1 > .75), with the majority excellent (Hart & Boer, 2010). Sutherland et al. (2012) investigated the inter-rater reliability of the RSVP with a sample of 28 forensic mental health professionals in Scotland. The participants used the RSVP to assess six case vignettes. Inter-rater reliability was fair to good, and agreement was highest when the participants were highly trained in forensic risk assessment. These studies indicate the RSVP can be used to make reliable judgments.

Utility of the RSVP

A study by Judge, Quayle, O'Rourke, Russell, and Darjee (2013) aimed to explore the realworld clinical practice of the SPJ risk assessment approach through qualitative investigation of the accounts of referrers to the Sex Offender Liaison Service (SOLS). It did this by exploring whether the risk management recommendations made using the RSVP were perceived as useful and changed the way in which offenders were managed by criminal justice agencies. Five themes emerged from this analysis. The RSVP was useful in terms of informing risk management, confirming what was known and giving weight, understanding personality, treatment, and the usefulness and limitations of risk assessment. The participants reported that the assessments were influential with respect to risk management.

Study aims and objectives

There is limited research on the validity and reliability of the RSVP, although it is used extensively in Scotland to inform sentencing, institutional management, release decision making and community management in both forensic mental health and correctional settings.

The purpose of the study is to ascertain the inter-rater reliability, prognostic validity and usefulness (in informing management) of the assessments undertaken in an applied clinical setting using the RSVP.

Specifically the following research questions were addressed:

1. How reliable are ratings of Items, Sections, Total Scores and Summary Judgments?

2. How do RSVP ratings correspond with ratings using other instruments, such as Risk Matrix 2000 and the Psychopathy Checklist-Revised?

3. How do RSVP ratings (scores and summary judgments) predict further offending (including convicted and unconvicted offending; and including both sexual and non-sexual offending)?

4. How are RSVP summary judgements related to further offending after taking into account the level of management cases are subject to?

5. Do the risk scenarios generated by clinicians using the RSVP include the types of sexual violence that recidivist offenders go on to perpetrate?

Method

Setting and service

The Sex Offender Liaison Service (SOLS) provides clinical consultation, assessment and management advice to help criminal justice agencies manage complex and/or high risk sexual offenders in the community. The service was established in 2007 with the introduction of Multi-Agency Public Protection Arrangements (MAPPA) in Scotland. The service is a joint clinical psychology/psychiatry led service and promotes a psychological formulation and SPJ approach to risk assessment and risk management. Most cases are referred by criminal justice social work or police offender management units, with a small number of referrals from mental health services, courts and child protection services. The service covers the Lothian and Borders Community Justice Authority (CJA) Area. The population was 939,020 in 2010 (Scottish Government, 2010). In 2010, 599 sexual offenders were registered and at liberty in the CJA area, corresponding to 64 registered sexual offenders per 100,000 of the population (Scottish Government, 2010).

Assessment process

The SOLS assessment involves gathering comprehensive case information from interviews with offenders (usually at least two interviews, total duration at least 4 hours), case records (from court, police, criminal justice social work, mental health and prison as appropriate), interviews with staff and sometimes interviews with relatives of offenders. Structured instruments (see below), including the RSVP, are applied to cases. Full details of the assessment process are described elsewhere (Russell & Darjee, 2013). Assessments are undertaken by two members of the team trained in using the RSVP, at least one of whom is a qualified clinical psychologist or psychiatrist. Assessments are presented to the wider clinical team for discussion and supervision before reports are finalised. Reports are produced which include a case history, risk factors, a case formulation, future risk scenarios and management recommendations. These reports are provided to referrers from other agencies along with verbal feedback on cases. Assessment reports are then used to guide the management of cases.

SOLS uses a number of assessment tools for the purpose of assessing individuals, including: Risk for Sexual Violence Protocol (RSVP), Risk Matrix 2000 (RM2000; Thornton, 2010); Psychopathy Checklist-Revised (PCL-R; Hare, 1991); International Personality Disorders Examination (IPDE; Loranger, 1999); Sexual Sadism Scale (SeSaS; Nitschke, Osterheider, & Mokros, 2009); and Screening Scale for Pedophilic Interests (SSPI; Seto & Lalumiere, 2001). In addition, other SPJ instruments are used where in addition to sexual violence the individual poses a risk of other offending behaviours, e.g. Historical Clinical Risk 20 (HCR-20;Webster et al., 1997), the Spousal Assault Risk Assessment (SARA; Kropp, Hart, Webster, & Eaves, 1995), and Stalking Assessment & Management (SAM; Kropp, Hart & Lyon, 2008). Data from SPJ instruments other than the RSVP is not reported in this study.

Ethical and management approval

The service obtained a letter from the NHS Ethics Committee confirming that the use of this data for evaluation and research, as it is gathered and used by the service, does not require formal ethical approval. However, permission was gained from the police and local authorities regarding collecting follow-up data. The research was also given management approval within NHS Lothian.

Sample

The sample was 109 cases. The first 110 individuals who were assessed by the service using the RSVP were selected for the study however one case was removed because of his young age. Therefore the total number of individuals analysed in this study is 109. Most individuals had a conviction for sexual offending. Some had other convictions (e.g. murder or assault) where there was considered to be a significant sexual component to the offence. Some had charges or allegations of sexual offending, but had not been convicted. A small number had no charges or convictions but were assessed using the RSVP as they expressed fantasies of or urges towards sexual violence.

Measures used at initial assessment

The following measures were routinely used by SOLS to assess cases.

Risk Matrix 2000 (RM2000)

Risk Matrix 2000 is an actuarial risk assessment instrument for convicted sexual offenders. It has a sexual (S) scale used to estimate the risk of sexual offending and a violence (V) scale used to estimate the risk of violent offending. For each scale a small number of variables are coded and a case is placed into one of four risk categories: low; medium; high; or very high. A validation study of the RM2000 has been undertaken in Scotland (Grubin, 2011).

Psychopathy Checklist Revised (PCL-R)

The PCL-R has 20 items, each scored on a 3 point scale from 0 to 2, giving a total score ranging from 0 to 40. It was originally developed as a measure of the extent to which an individual matched Cleckley's (1976) description of the prototypical psychopath, and has been found to be a good predictor of violent recidivism (Dolan & Doyle, 2000). It has been suggested that the cut-off to make a diagnosis of psychopathy is culturally mediated (Hare, 2003; Cooke, Michie, Hart, & Clark, 2005). For the UK, Cooke and Michie (1999) suggested that a score of 25 or above was diagnostic of psychopathy, and a score of 15-24 indicated a moderate degree of psychopathy. However, for the current study we used the cut-offs for the PCL-R set out in the RSVP manual for the item 'psychopathic personality disorder' (definite psychopathy indicated by a score of 30 or above and partial psychopathy indicated by a score of 21-29).

International Personality Disorder Examination (IPDE)

The International Personality Disorder Examination (IPDE) is a semi-structured clinical interview developed to assess the personality disorders in the International Statistical Classification of Diseases and Health Related Problems, 10th Revision (ICD-10; WHO, 1992) and the Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV; American Psychiatric Association, 1994) classification systems. Data on the reliability of the instrument can be found in the IPDE manual (Loranger, 1999). The DSM-IV version was used in the current study.

Structural Behavioural Assessment of Paraphilias

Two assessment instruments which aim to look objectively at behaviours that might be indicative of paedophilia and sexual sadism are the Screening Scale for Pedophilic Interests (SSPI; Seto & Lalumiere, 2001) and the Sexual Sadism Scale (SeSaS; Nitschke, Osterheider, & Mokros, 2009). Both are relatively new measures which are primarily based on patterns of behaviour and offending that do not require self report by the individual being assessed. Scores on the SSPI have been shown to correlate highly with sexual arousal to children as assessed by penile plethysmography (Seto, Harris, Rice, & Barbaree, 2004). There have been some initial studies indicating the validity of the SeSaS in identifying sexual sadists (Nitschke et al., 2009).

Risk for Sexual Violence Protocol (RSVP)

RSVP assessments undertaken by SOLS were used for the current study.

For individual items there were three ratings: presence in the past (*Past*), presence recently (*Recent*) and relevance to future management (*Relevance*). So three ratings were available for each item; each rating was either 'Yes', 'Partial' or 'No'. For coding purposes and to generate 'scores' (see below) each of these ratings was scored 2, 1 or 0 respectively. These ratings were available for the 22 specific items.

For each of the five Sections of the RSVP (A. Sexual Violence History, B. Psychological Adjustment, C. Mental Disorder, D. Social Adjustment, and E. Manageability), three scores were generated based on adding the item ratings within that Section. The three scores generated for each Section were based on adding Past presence, Recent presence, and Future Relevance ratings.

For the RSVP as a whole, Total Scores were generated using the 22 specific items. Total scores were generated based on adding the three types of item ratings described above. So there were three total scores altogether: Past presence, Recent presence, and Relevance. For some analyses a further total score, labelled the Ever present total, was created by adding the highest of the Past presence and Recent presence item scores for each of the ratings.

The Summary Judgments generated by the RSVP assessment were also coded. This involved ratings of three areas of risk: Case Prioritization, Risk of Serious Physical Harm and Immediate Action Required. Each was rated on a three point scale in accordance with the RSVP manual. Case Prioritization was determined based on the "level of effort or intervention required to prevent the person from committing acts of sexual violence" as stipulated in the manual. The coding is broken down into High/Urgent, Moderate/Elevated and Low/Routine. Risk of Serious Physical Harm is assessed using a three-point measure (High, Moderate, Low) to identify "the risk that any future sexual violence will involve serious or life-threatening physical harm".

Other baseline data

In addition to these measures data were available covering the following: referral source, MAPPA level (where applicable), demographic details, offending history (including allegations, charges and convictions) and mental health history. At the end of the assessments risk ratings were also made in accordance with the four MAPPA risk levels (low, medium, high and very high; Scottish Government, 2014) and the three RMA risk levels used for potential OLR cases (high, medium, low; Risk Management Authority, 2013).

Follow-up data

Follow-up offending, incidents and recall

Police, criminal justice social work and the NHS granted permission to access outcome data from their records and systems. All offenders were followed up from the date of assessment, (2007 - 2012) to 2013, except for 4 individuals who died during follow-up. Outcomes included: (1) Further conviction (sexual, violent, general and breach); (2) Charges, allegations and incidents (which have not led to conviction); (3) Breach of legal orders/recall to custody (whether or not the breach led to a charge or conviction). The follow-up period ranged from 6 months to 5 years. We report some outcomes for one year follow-up to allow our sample to be compared with other samples reported in the literature.

In reporting these outcomes we will use the term 'conviction' to refer to further convictions by a criminal court during follow-up, and we will use the terms 'offending' or 'behaviour' to refer to any incidents whether convicted or unconvicted. *Any* follow-up offending refers to any convictions, charges, allegations and breaches during the follow-up period unless otherwise specified. Offences referred to as *serious sexual* incorporate any contact sexual offence including rape and sexual homicide. A *serious* offence is defined as serious sexual behaviour and any non-sexual serious violence including homicide cases. Reference to *anything at all* refers to any convictions, charges, allegations and breaches and breaches and any type of offence (sexual, violent, general and breach) during follow-up.

It should be noted that the service has assessed some individuals being considered for release from custody who have not subsequently been released, but this is an outcome in itself and may be related to the risk posed by the individual. Some unreleased individuals committed offences or had allegations/charges in custody.

Risk management level

The level of risk management and changes in risk management level during follow-up were also collected and coded. For every day of follow-up the level of risk management was recorded. Risk management levels were coded in seven categories:

- 1. No statutory intervention
- 2. Registered sex offender (RSO) only
- 3. Criminal justice social work statutory supervision (probation or parole) OR Sexual Offences Prevention Order (SOPO)
- 4. Criminal justice social work statutory supervision (probation or parole) AND Sexual Offences Prevention Order (SOPO)
- 5. On any statutory supervision and seen by staff daily
- 6. 24 hour supervision in the community
- 7. In prison or secure hospital with no unescorted access to the community

For analysis the risk management levels were collapsed as follows: 1 or 2 = low (i.e. no to minimal monitoring and supervision); 3, 4 or 5 = *medium* (intermediate monitoring/supervision); 6 or 7 = high (i.e. constant monitoring and supervision).

Analysis and statistical methods

Description of baseline and follow-up data

Descriptive statistics used were numbers and proportions (percentages) for categorical variables; and means, medians and ranges for continuous variables.

Reliability

Consensus decision making is a method of enhancing assessment reliability and validity (de Vogel & de Ruiter, 2006). All the RSVP assessments were completed by two clinicians, in consultation with multi-agency colleagues and the wider SOLS team. To assess the reliability of the assessments, 11 recent cases were rated prospectively with the two clinicians who assessed the cases rating the items and developing risk scenarios independently, before developing the consensus ratings and scenarios. This allowed an evaluation of the reliability of item ratings, total scores, summary judgments and scenarios, and allowed an evaluation of the merits of a consensus approach. Intra-class correlation coefficients (ICC) were used to

assess reliability. This is the same method used in other research (i.e., Douglas & Belfrage, 2014; Sutherland et al., 2012).

Relation to other measures

Associations between the RSVP items, total scores and summary judgments, and other measures (i.e. Risk Matrix 2000, PCL-R, SSPI, SeSaS) were assessed. Pearson's correlations were used for relationships between continuous variables (e.g. correlation between RSVP *Past presence* total score and PCL-R score). When comparing three summary judgment groups (i.e. high, medium, low) on a continuous variable (e.g. RM2000 score) one-way analysis of variance (ANOVA) was used.

Predictive validity of RSVP

Relationship between RSVP total score or summary judgments and outcomes

To analyse the relationship between RSVP Total Scores and outcomes, and between Summary Judgments and outcomes (e.g. sexual violence during follow-up) the area under the curve (AUC) of the receiver operating characteristic (ROC) was used. In addition, the rates of the occurrence of an outcome (e.g. sexual violence) in the summary judgement groups (e.g. low, medium or high priority) was calculated and compared using the chi-square test.

Kaplan-Meier survival analysis was used to analyze time to outcome (e.g. sexual violence) in low, medium and high risk groups (according to RSVP summary judgements).

Relationship between RSVP summary judgements and outcomes taking into account risk management level

The effect of risk management level on outcome was examined by comparing the rate of outcomes in the different risk management level groups. Groups were compared using chi-square tests. The association between risk management level and assessed level of risk was also ascertained using the chi-squared test.

There were two ways in which risk management level was determined: if an individual committed an offence during follow-up then the level at which they were managed at the time of the offence was chosen, otherwise if they did not commit an offence during follow-up, it

was determined by the management level at which the offender spent most of their time throughout the follow-up period.

To take into account risk management levels when considering recidivism, offenders were grouped into nine "management by assessment" categories as set out in Table 1. Rates of offending outcomes were ascertained in each of the nine groups. Due to the low numbers in individual cells no analytic statistic was used for significance testing.

Table 1. Categories of cases ascertained by taking into account both assessed level of risk and risk management received.

		Assessed level of risk		
		Low	Medium	High
Risk	Low	0	-1	-2
management	Medium	+1	0	-1
received	High	+2	+1	0

In addition Kaplan-Meier survival analysis was used to analyze time to outcome (e.g. sexual violence) in five groups. These five groups were created from the nine groups above by taking the level of risk management away from the assessed level of risk. There were offenders overmanaged by 2 levels (i.e. low risk cases managed at a high level; +2), over-managed by 1 level (i.e. low risk cases managed at a medium level or medium risk cases managed at a high level; +1), managed at the "correct level" (low-low, medium-medium or high-high; 0), undermanaged by 1 level (i.e. medium risk cases managed at a low level or high risk cases managed at a medium level; -1), and under-managed by 2 levels (i.e. high risk cases managed at a low level; -2).

Predictive Value of Risk Scenarios

In cases where there were further sexual allegations, charges or convictions the extent to which the actual offending matched any plausible scenario articulated in the RSVP report was described using percentage agreement. The variables used for this were victim age, victim gender, victim relationship and severity of sexual offence. Each item was independently rated by two assistant psychologists in the SOLS service and compared against the risk scenarios described in the reports. Any discrepancies in the ratings were discussed between the raters and a consensus rating was agreed upon.

Results

Baseline characteristics of the sample

Over three quarters of the sample were referred to SOLS by police and social work (Table 2). 18% were on no current criminal justice order or sentence (6 RSO; 3 SOPO; 3 under mental health legislation). 28.4% were in prison about to be released or being considered for release into the community, and 7.3% were on bail or remand awaiting sentencing. The rest were on criminal justice supervision in the community. Many offenders were on a legal order, some on more than one, e.g. an RSO and subject to a SOPO. However there were also offenders who were on no order. A quarter of the referrals were for people not subject to MAPPA. For the rest, almost 20% had been at MAPPA Level 3 at one point. In terms of describing the level of risk these men were considered to pose, almost a third were described as being High or Very High Risk MAPPA cases at the point of assessment and 28.4% were described as being a High RMA Risk Level case with respect to the criteria for potential OLR cases.

In terms of index offences (i.e. the last offence committed prior to the referral to SOLS), over three quarters had committed a contact offence and 11.9% had committed an internet offence (Table 3). In 95.4% of cases the offence had resulted in a conviction with the rest resulting in either a charge or else there was an allegation that had not led to a formal criminal charge being made. In one case the person had come forward to report concerning thoughts and fantasies. In over a quarter of cases (28.4%) the overall index offence constituted both sexual and non-sexual offences. In the majority of cases (67.9%) the index offence was a sexual offence only. In 2.8% of cases the index offence was nonsexual, however there had been previous concerning sexual behaviour or offending, or else there was a sexual aspect to the offence that had not been recognised in the conviction. In 59.6% of cases the index offence was deemed to be seriously harmful. 11.9% of referrals were given a life sentence and 28.4% were given a sentence of greater than four years duration. Only 1.8% had received a mental health disposal. In 52.3% of cases the victim of the index offence was a child and in a further 6.4% there was a child victim and an adult victim. In 65.1% of cases, the victim of the index offence was female and in a further 1.1% there were male and female victims. Around half of the victims were strangers.

Mean age (range)	38.6 years (17 - 68)
Highest MAPPA Level	
N/A	28 (25.7)
Level 1	9 (8.3)
Level 2	51 (46.8)
Level 3	21 (19.3)
Referral Agency	
CJSW	58 (53.2)
Police	29 (26.6)
NHS	17 (15.6)
Court	5 (4.6)
Current Criminal Justice Sanction	
Nothing	20 (18.3)
Prison	31 (28.4)
Parole	23 (21.1)
Probation	27 (24.8)
Bail	6 (5.5)
Prison Remand	2 (1.8)
Current Legal Order	
RSO	70 (64.2)
SOPO	22 (20.2)
Mental health legislation	4 (3.7)
Life sentence	11 (10.1)
MAPPA Risk Level	
N/A	4 (3.7)
Low	34 (31.2)
Medium	36 (33)
High	25 (22.9)
Very High	10 (9.2)
RMA Risk Level	
Low	38 (34.9)
Medium	40 (36.7)
High	31 (28.4)

Table 2. Baseline characteristics of sample of 109 cases

Index Offence Category	
No Index Offence	1 (9)
Internet Offending	13(11.9)
Non-Contact	11(101)
Contact	84 (77 1)
Type of Index Offence	0+(11.1)
Conviction	104 (95 4)
Charge	2(18)
Allegation	2(1.0) 2(1.8)
Concerning Thoughts/Eantasies	2(1.0) 1(9)
Mean Number of Index Offences (range)	226(0-48)
Index Offence Sex/Non Sex	2.20 (0 - 48)
No Index Offence	1 (0)
No index Offence	1(.3) 74(67.0)
Non sexual only	(07.9)
Non-sexual only Sexual and Non-sexual	3(2.0) 31(284)
Seriously Hermful Index Offence	51(20.4)
Sentence	05 (59.0)
No sontonco	6 (5 5)
INO Semence	(5.3)
> 1 years imprisonment	(11.7) 31 (28 4)
4 years imprisonment	25(22.4)
< 4 years imprisonment	20(183)
Community Service	20(18.5)
Mental health disposal	2(1.0) 2(1.8)
Awaited	10(92)
Victim Age	10 (9.2)
No Victim	1 (9)
Child	(.)
Adult	$A_{1}(32.3)$
Child and Adult	7(64)
Victim Gender	7 (0.4)
No Victim	1 (9)
Male	26 (23 9)
Female	71 (65 1)
Male and Female	11(101)
Victim Relationship	
No victim	1 (9)
Biological Relative	15 (13.8)
Sten Relative	3 (2.8)
Shouse/Partner	3 (2.8)
Well Known	11 (10.1)
Acquaintance	17 (15.6)
Professional/Staff	2(1.8)
Stranger	57 (52.3)

Table 3. Details of index offence

When looking at the whole offending history (index offence, previous offences and allegations) of the cases, they ranged from downloading child abuse images to non-contact to contact offences, some including homicide (Table 4). There were high rates of contact offending and almost half had a history of penetrative offending. 7.3% had committed homicide and almost half had committed a violent offence, with 62.4% having committed a non-sexual, non-violent offence. Just over two thirds had sexually offended against a child and half had sexually offended against an adult. Many more offenders had committed sexual offences against females than males. Most offenders had committed two sexual offences and most had not committed a violent offence, although some had committed many, making the mean number of violent offences almost 2. 64.2% had been subject to sexual allegations which had not led to conviction, indicating the potential for there to have been more past offences than the conviction data indicated.

Internet Child Pornography Offending	18 (16.5)
Type of Internet Offending	
None	91 (83.5)
Downloading	16 (14.7)
Distribution	2 (1.8)
Internet Grooming	0 (0)
Other Non Contact Offending	34 (31.2)
Contact Offending	88 (80.7)
Penetrative Offending	51 (46.8)
Homicide	8 (7.3)
Non Sexual, Violent Offending	52 (47.7)
Non Sexual and Non Violent Offending	68 (62.4)
Seriously Harmful Offending	70 (64.2)
Sexual Offending against Children	74 (67.9)
Sexual Offending against Adults	55 (50.5)
Sexual Offending against Males	37 (33.9)
Sexual Offending against Females	84 (77.1)
Mean Number of Sexual Convictions (median)	2.79 (2)
Mean Number of Violent Convictions (median)	1.91 (0)
Mean Number of Other Convictions (median)	8.26 (2)
Non Convicted Sexual Allegations	70 (64.2)

Table 4. Offending history

A small number of those referred were not eligible for certain assessment tools. This was due to young age or lack of conviction. For those who were eligible for assessment using RM2000 (n = 105), all were assessed as falling within the medium (39.4%) or high (56.8%) risk category on the sexual scale. For the RM2000 violence scale, most fell within the low

(32.1%) and medium (56%) category with a small number being in the high (8.3%) category. For those who had a PCL-R or a PCL-SV administered (n = 108), 11% met the threshold for psychopathy (PCL-R score \geq 30, PCL-SV score \geq 18) and 20.2% had partial psychopathy (PCL-R score = 21-29, PCL-SV score = 13-17). An HCR-20 was completed in a quarter of cases due to concern about violence risk as well as sexual violence risk. 5.5% met criteria for sexual sadism using the SeSaS criteria.

A high percentage of cases (81.7%) had previous contact with mental health services. 11% had been detained at some point using mental health legislation and 7.3% had spent time in a secure hospital. Just over a third had been in contact with a forensic mental health service in the past and 40.4% had been in contact with a general adult service. A third had been in contact with NHS Psychology and 38.5% had been seen as a child or adolescent by mental health services. 21.1% had their first contact with mental health services as a child, 18.3% as an adolescent, and 42.2% as an adult. In terms of mental disorder, 11.9% had a diagnosis of major mental illness, 10.1% had a diagnosis of learning disability, and 3.7% had a diagnosis of Autistic Spectrum Disorder.

Using DSM-IV criteria, 63.3% cases met criteria for a definite diagnosis of any personality disorder with a further 22% meeting criteria for a probable diagnosis of personality disorder. Some cases met criteria for more than one personality disorder. Antisocial Personality Disorder (40.4%) was the most common specific personality disorder diagnosis, followed by Narcissistic (22.9%) then Schizoid (16.5%). Almost half of the cases (49.5%) met criteria for any paraphilia as diagnosed by DSM-IV. Again offenders could meet criteria for more than one paraphilia. The most common paraphilia was Paedophilia (39.4%) followed by Sexual Sadism (6.4%) followed by Voyeurism (4.6%)

Description of offending during follow-up

There were 105 out of 109 offenders who were followed up for a minimum of 6 months. 96 offenders were followed up for 12 months or more and 23 offenders were followed up for over 5 years. The mean follow-up period was 3.25 years (SD = 1.77 years), the median was also 3.25 years indicating that our data is not skewed. The range of follow-up was from 0.08 years to 7.15 years.

Of the 109 offenders, 47 (43.1%) went on to have *any* type of further offending behaviour (allegation, charge or conviction) during the follow-up period (Table 5). This was broken down into *any* sexual offending (23.9%), *any* violent (7.3%), *any serious* offending (12.8%), *any serious sexual* (11%) and *any* breaches (26.6%). In terms of convictions, 28.4% were convicted for *any* kind of offence, 11.9% received a sexual conviction and 6.4% were convicted for *any serious* offence (Table 6).

Table 5. Offending behaviour (alleged, charged or convicted) during whole follow-up period (N = 109)

Any	47 (43.1)
Sexual	26 (23.9)
Violent	8 (7.3)
Serious	14 (12.8)
Serious Sexual	12 (11)
Breaches	29 (26.6)

Table 6. Offending during whole follow-up period.

	Convictions	Charges	Allegations
Any	31 (28.4)	18 (16.5)	13 (11.9)
Sexual	13 (11.9)	7 (7.3)	7 (6.4)
Internet	4 (3.7)	3 (2.8)	1 (.9)
Noncontact	8 (7.3)	5 (4.6)	2 (1.8)
Contact	5 (4.6)	3 (2.8)	5 (4.6)
Rape	1 (.9)	2 (1.8)	2 (1.8)
Sexual Homicide	1 (.9)	0 (0)	0 (0)
Against Adult Male	0 (0)	0 (0)	2 (1.8)
Against Adult Female	5 (4.6)	4 (3.7)	3 (2.8)
Against Male Child	3 (2.8)	4 (3.7)	0 (0)
Against Female Child	8 (7.3)	2 (1.8)	2 (1.8)
Violence	5 (4.6)	3 (2.8)	4 (3.7)
Serious Violence	3 (2.8)	1 (.9)	1 (.9)
Homicide	1 (.9)	0 (0)	0 (0)
General	4 (3.7)	5 (4.6)	2 (1.8)
Breach of order	18 (16.5)	10 (9.2)	2 (1.8)
Serious Sexual Offending	5 (4.6)	3 (2.8)	5 (4.6)
Any Serious Offending	7 (6.4)	3 (2.8)	6 (5.5)

Table 7 shows the re-conviction rates for 6 month, 12 month and 5 year follow-up periods. Breaches were also recorded in this way, with 14/105 (13.3%) having a breach within 6 months, 16/96 (16.7%) by 12 months and 5/23 (21.7%) by 5 years.

	Convictions	Offending (conviction, charge or
		allegation)
Sexual		
6 months $(n = 105)$	13 (12.4)	26 (24.8)
12 months (n = 96)	13 (13.5)	25 (26.0)
5 years $(n = 23)$	4 (17.4)	8 (34.8)
Violent		
6 months (n=105)	5 (4.8)	10 (9.5)
12 months (n = 96)	5 (5.2)	10 (10.4)
5 years $(n = 23)$	0 (0.0)	1 (4.3)

Table 7. Follow-up offending for specific durations of follow-up

RSVP ratings

In terms of the case summary judgements made using the RSVP, around a quarter (24.8%) of the sample were assessed to need High Case Prioritization and about a quarter (25.7%) were identified as posing a High Risk of Serious Physical Harm (Table 8). For both Case Prioritization and Risk of Serious Physical Harm the majority of the cases were in the Moderate rating (43.1% and 39.4% respectively). The majority of cases (59.6%) did not require immediate action while 16.5% did. In 45% of cases other risks were highlighted requiring further assessment, indicating that many of these offenders are not 'specialist' sex offenders.

Case Prioritization	
Low	35 (32.1)
Moderate	47 (43.1)
High	27 (24.8)
Risk of Serious Physical Harm	
Low	38 (34.9)
Moderate	43 (39.4)
High	28 (25.7)
Immediate Action required	
No	65 (59.6)
Possibly	26 (23.9)
Yes	18 (16.5)
Other Risks Indicated	
No	60 (55)
Yes	49 (45)

 Table 8. RSVP summary judgements

Tables i and ii (Appendix) present data on individual item ratings (means, medians, and frequencies).

Reliability of RSVP Items, Totals and Summary Judgements

Both ICC1 and ICC2 are reported, ICC1 being indicative of the reliability of one rater and ICC2 the measure of the average reliability from several raters (Douglas & Belfrage, 2014). Landis and Koch (1977) provided guidelines to address varying levels of agreement using a range of values. 'Almost Perfect' reliability was indicated within the range .81 - 1.00; 'Substantial' .61 - .80; 'Moderate' .41 - .60; 'Fair' .21 - .40; 'Slight' .00 - .20. ICC1 and ICC2 values for each item are reported in Tables 9 – 11. The majority of the Past, Recent, and Relevance items were considered 'Almost Perfect'. The risk items with the highest inter-rater reliability amongst Past, Recent, and Relevance scoring were *Minimization/Denial*, *Psychopathic PD*, *Problems with Planning* and *Problems with Supervision* (ICC \ge .81).

Criterion	ICC1	ICC2
RSVP 1 Past	.762 (.3593)	.865 (.5296)
RSVP 2 Past	1.00 (.)	1.00 (.)
RSVP 3 Past	.845 (.5396)	.916 (.6998)
RSVP 4 Past	.941 (.8198)	.970 (.8999)
RSVP 5 Past	.878 (.6197)	.935 (.7598)
RSVP 6 Past	.805 (.4194)	.892 (.5897)
RSVP 7 Past	.554 (0686)	.713 (1292)
RSVP 8 Past	.808 (.4594)	.894 (.6397)
RSVP 9 Past	.231 (4772)	.375 (-1.8084)
RSVP 10 Past	.835 (.5095)	.910 (.6798)
RSVP 11 Past	.754 (.3293)	.860 (.4896)
RSVP 12 Past	1.00 (.)	1.00 (.)
RSVP 13 Past	.800 (.4494)	.889 (.6197)
RSVP 14 Past	1.00 (.)	1.00 (.)
RSVP 15 Past	.706 (.2191)	.828 (.3595)
RSVP 16 Past	.565 (05586)	.722 (1293)
RSVP 17 Past	.091 (5464)	.167 (-2.3178)
RSVP 18 Past	1.00 (.)	1.00 (.)
RSVP 19 Past	1.00 (.)	1.00 (.)
RSVP 20 Past	1.00 (.)	1.00 (.)
RSVP 21 Past	.833 (.5095)	.909 (.679 8)
RSVP 22 Past	.833 (.5195)	.909 (.689 8)
Other Considerations	.752 (.3193)	.859 (.4796)
Past		

Table 9. Inter-rater Reliability of Past RSVP Items (N = 11)

Note: Almost Perfect .81 – 1.00 Substantial .61 - .80 Moderate .41 - .60 Fair .21 - .40 and Slight .00 .20 (Landis & Koch, 1977)

Criterion	ICC1	ICC2
RSVP 1 Recent	1.00 (.)	1.00 (.)
RSVP 2 Recent	Zero variance	Zero variance
RSVP 3 Recent	Zero variance	Zero variance
RSVP 4 Recent	.000 (5858)	.000 (-2.7273)
RSVP 5 Recent	Zero variance	Zero variance
RSVP 6 Recent	.861 (.5896)	.925 (.7398)
RSVP 7 Recent	.440 (2680)	.571 (6989)
RSVP 8 Recent	.848 (.5596)	.918 (.7198)
RSVP 9 Recent	.796 (.3994)	.886 (.5697)
RSVP 10 Recent	.719 (.2492)	.837 (.3896)
RSVP 11 Recent	.921 (.7598)	.959 (.8599)
RSVP 12 Recent	1.00 (.)	1.00 (.)
RSVP 13 Recent	Zero variance	Zero variance
RSVP 14 Recent	.762 (.3593)	.865 (.5296)
RSVP 15 Recent	Zero variance	Zero variance
RSVP 16 Recent	.932 (.7898)	.965 (.8899)
RSVP 17 Recent	.737 (.2692)	.848 (.4196)
RSVP 18 Recent	.932 (.7898)	.965 (.8899)
RSVP 19 Recent	.783 (.40 – .94)	.878 (.5797)
RSVP 20 Recent	1.00 (.)	1.00 (.)
RSVP 21 Recent	.825 (.4995)	.904 (.6697)
RSVP 22 Recent	1.00 (.)	1.00 (.)
Other Considerations	.755 (.3193)	.860 (.4796)
Recent		

Table 10. Inter-rater Reliability of Recent RSVP Items (N = 11)

Note: Almost Perfect .81 – 1.00 Substantial .61 - .80 Moderate .41 - .60 Fair .21 - .40 and Slight .00 -20 (Landis & Koch, 1977)

Criterion	ICC1	ICC2
RSVP 1 Relevance	.885 (.6397)	.939 (.7798)
RSVP 2 Relevance	.603 (.0987)	.752 (.1693)
RSVP 3 Relevance	.630 (.0988)	.773 (.1794)
RSVP 4 Relevance	.720 (.2692)	.837 (.4196)
RSVP 5 Relevance	.681 (.1790)	.810 (.2995)
RSVP 6 Relevance	.859 (.5596)	.924 (.7198)
RSVP 7 Relevance	.464 (1882)	.634 (4390)
RSVP 8 Relevance	.701 (.2491)	.825 (.3895)
RSVP 9 Relevance	.700 (.2191)	.824 (.3495)
RSVP 10 Relevance	.900 (.6797)	.947 (.8099)
RSVP 11 Relevance	.921 (.7598)	.959 (.8599)
RSVP 12 Relevance	1.00 (.)	1.00 (.)
RSVP 13 Relevance	1.00 (.)	1.00 (.)
RSVP 14 Relevance	.945 (.8299)	.972 (.9099)
RSVP 15 Relevance	.561 (04686)	.719 (09593)
RSVP 16 Relevance	.167 (3265)	.286 (9479)
RSVP 17 Relevance	.082 (5163)	.151 (-2.1177)
RSVP 18 Relevance	.717 (.2592)	.835 (.4096)
RSVP 19 Relevance	1.00 (.)	1.00 (.)
RSVP 20 Relevance	.904 (.7097)	.949 (.8299)
RSVP 21 Relevance	.687 (.1791)	.815 (.3095)
RSVP 22 Relevance	1.00 (.)	1.00 (.)
Other Considerations	.719 (.2492)	.837 (.3896)
Relevance		

Table 11. Inter-rater Reliability of Relevance RSVP Items (N = 11)

Note: Almost Perfect .81 – 1.00 Substantial .61 - .80 Moderate .41 - .60 Fair .21 - .40 and Slight .00 .20 (Landis & Koch, 1977)

All the Summary Judgements were classified by Landis and Koch's (1977) criteria as 'Almost Perfect' with the lowest value from Risk of Serious Physical Harm (ICC1 = .95) (Table 12). The Mean ICC values for the RSVP Past items were: ICC1 = .80 and ICC2 = .85; for the RSVP Recent items: ICC1 = .84 and ICC2 = .89; and for the RSVP Relevance items: ICC1 = .73 and ICC2 = .82.

Criterion	ICC1	ICC2
RSVP Case Prioritization	1.00 (1.00)	1.00 (1.00)
RSVP Risk of Serious Physical	.951 (.8499)	.975 (.9199)
Harm		
RSVP Immediate	1.00 (.)	1.00 (.)
Action Required		
RSVP Other	1.00 (.)	1.00 (.)
Risks Indicated		
RSVP Past	.811 (.4395)	.896 (.6097)
Total Items 1-22		
RSVP Recent	.914 (.7198)	.955 (.8399)
Total Items 1-22		
RSVP Relevance	.828 (.4995)	.906 (.6697)
Total Items 1-22		

Table 12. Reliability of RSVP Summary Judgements, Total Scores, (N = 11)

Note: Almost Perfect .81 – 1.00 Substantial .61 - .80 Moderate .41 - .60 Fair .21 - .40 and Slight .00 - .20 (Landis & Koch, 1977)

Convergent validity with other instruments (see Tables 13-16)

The RM2000 Sexual scale was significantly correlated with all of the RSVP total scores, although less so for Recent total. It was significantly correlated with the RSVP Section A (Sexual Violence History) Past and Relevance totals; with all three Section C (Mental Disorder) totals; with Section D (Social Adjustment) Past and Relevance totals; and with all four Section E (Manageability) totals. So the only Section uncorrelated with RM2000 Sexual was Section B (Psychological Adjustment). The RM2000 Violence scale was significantly correlated with all three RSVP Total Scores, but also less significant for the Recent total. It was significantly correlated with all three Section E totals; and Section E Past and Relevance totals. So the only section uncorrelated with RM2000 Violence was Section A (Sexual Violence History). There was a significant association between Case Prioritization and RM2000 Sexual and Violence scales. There was a significant association between RM2000 Violence, but not RM2000 Sexual, and Risk of Serious Physical Harm ratings. There was no association between the RM2000 scales and Immediate Action Required ratings.

	RM2000S	RM2000V	PCL-R	SeSaS	SSPI
			Total	Part 1	Total
			Score	Score	Score
RSVP Past Total Items 1-22	.412**	.472**	.765**	.278**	050
RSVP Recent Total Items 1-22	.308**	.244*	.454**	085	.061
RSVP Relevance Total Items 1-22	.434**	.484**	.751**	.267**	021
RSVP Section A Past Total	.314**	.099	.389**	.233**	.215*
RSVP Section A Recent Total	.029	030	.127	167	.224*
RSVP Section A Relevance Total	.349**	.124	.350**	.173	.269**
RSVP Section B Past Total	.142	.305**	.533**	.160	088
RSVP Section B Recent Total	.159	.248*	.449**	.019	020
RSVP Section B Relevance Total	.102	.317**	.519**	.136	107
RSVP Section C Past Total	.293**	.426**	.583**	.124	027
RSVP Section C Recent Total	.358**	.308**	.371**	.040	.100
RSVP Section C Relevance Total	.352**	.437**	.571**	.175	015
RSVP Section D Past Total	.236*	.519**	.622**	.180	272 **
RSVP Section D Recent Total	.182	.088	.135	223*	075
RSVP Section D Relevance Total	.211*	.457**	.536**	.195*	272 **
RSVP Section E Past Total	.472**	.408**	.632**	.268**	123
RSVP Section E Recent Total	.276**	.174	.358**	.050	044
RSVP Section E Relevance Total	.418**	.404**	.643**	.226*	118

 Table 13. Correlations (RSVP Total Scores, Section totals and other Instruments)

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

PCL-R total score was significantly correlated with the RSVP totals and with all the section totals. PCL-R total score was not significantly correlated with the Section D (Social Adjustment) Recent total. The correlations with the PCL-R were stronger than for RM2000. All PCL-R scores were significantly associated with Case Prioritization, Risk of Serious Physical Harm ratings and Immediate Action Required ratings.

The SeSaS showed weaker but significant correlations with the RSVP totals, except for the Recent total. There was a significant correlation with the Past Section A total; a significant negative correlation with the Section D Recent total but a significant weak positive correlation with the Section D Relevance total; and a significant positive correlation with the Section E totals, except for the Recent total. The SeSaS was significantly associated with

Case Prioritization, Risk of Serious Physical Harm ratings and Immediate Action Required ratings.

The SSPI was not significantly correlated with any of the overall total scores. There was a weak significant positive correlation with all the Section A totals; and a weak significant negative correlation with the Section D totals except the Recent total. The SSPI was not associated with Case Prioritization, Risk of Serious Physical Harm or Immediate Action Required.

Case Prioritization ratings were significantly associated with the RM2000 Sexual scale, the RM2000 Violence scale, PCL-R score and SeSaS score (Table 14). Risk of Serious Physical Harm ratings were significantly associated with RM2000 Violence scale, PCL-R score and SeSaS score (Table 15). Immediate Action Required ratings were significantly associated with PCL-R and SeSas scores (Table 16).

Table 14. ANOVA exploring the association between Case Prioritization ratings and ratings on other assessment tools.

	Case Prioritization				
	Low	Moderate	High	F	Р
RM2000 Sexual Scale	2.4	2.6	3.1	4.97	.009
RM2000 Violence Scale	1.7	2.0	2.5	.511	.008
PCL-R Total Score	7.9	16.4	25.0	38.78	.000
SeSaS Part 1 Score	1.3	1.3	2.6	9.47	.000
SSPI Total Score	2.0	1.9	1.7	.15	.857

Bolded figures are used to show the results of post-hoc significance testing using the Tukey test. Where one figure in a row is bold, that indicates that that group had a score on the particular scale which was significantly different from the other two groups. Where two figures are bold in any row, that indicates that the scores on a particular scale for those two groups were significantly different from each other.

Table 15. ANOVA exploring the association between Risk of Serious Physical Harm ratings and ratings on other assessment tools.

	Risk of Serious Physical Harm				
	Low	Moderate	High	F	Р
RM2000 Sexual Scale	2.6	2.7	2.7	.18	.836
RM2000 Violence Scale	1.8	2.0	2.5	5.75	.004
PCL-R Total Score	9.5	16.9	22.1	18.02	.000
SeSaS Part 1 Score	1.1	1.4	2.8	14.72	.000
SSPI Total Score	2.2	1.8	1.6	.78	.461

Bolded figures are used to show the results of post-hoc significance testing using the Tukey test. Where one figure in a row is bold, that indicates that that group had a score on the particular scale which was significantly different from the other two groups. Where two figures are bold in any row, that indicates that the scores on a particular scale for those two groups were significantly different from each other.

0					
	Immediate Action Required				
	No	Possibly	Yes	F	Р
RM2000 Sexual Scale	2.56	2.68	2.94	1.18	.311
RM2000 Violence Scale	1.98	1.92	2.50	2.47	.089
PCL-R Total Score	13.30	16.25	24.84	11.54	.000
SeSaS Part 1 Score	1.51	1.31	2.61	5.29	.006
SSPI Total Score	1.62	2.42	2.06	1.71	.186

Table 16. ANOVA exploring the association between Immediate Action Required ratings and ratings on other assessment tools.

Bolded figures are used to show the results of post-hoc significance testing using the Tukey test. Where one figure in a row is bold, that indicates that that group had a score on the particular scale which was significantly different from the other two groups. Where two figures are bold in any row, that indicates that the scores on a particular scale for those two groups were significantly different from each other.

Predictive validity of RSVP

ROC Analysis

Offending during full follow-up (Table 17)

Any sexual offending The only measure with a significant ROC for sexual offending during the full follow-up was the SeSaS, which had a negative association with sexual offending. None of the RSVP total or Section scores or Summary Judgements yielded a significant ROC. This was also the case for the RM2000 and PCL-R.

Any violent offending The RSVP Past and Relevance totals were significantly associated with *any* violent offending, as were the PCL-R total score and the summary judgement of Risk of Physical Harm on the RSVP.

Any serious offending There was a similar finding when using *any serious* offending as an outcome however the Risk of Serious Physical Harm judgment was not associated with this outcome. In addition, there was an association with the RM2000 Violence scale score.

Any serious sexual offending When using any serious sexual offending as an outcome, the Immediate Action Required Summary Judgement had a good association as did the RSVP Past and Relevance totals, PCL-R total score, and the RM2000 Violence scale total.
Any offending at all When taking anything at all as an outcome a smaller range of predictors were found to have an association. These were the summed totals of the RSVP Recent items and the Relevance items.

Predictor Variables	Any Off	Any SexualAny ViolentAny SeriousOffendingOffendingOffending		Any Serious Sexual Offending		Anyth	ing at all			
	AU	95%	AU	95%	AU	95%	AUC	95%	AUC	95%
	C	C.I.	C	C.I.	C	C.I.		C.I.		C.I.
RSVP Case Prioritization	.59	.4871	.61	.4279	.62	.4777	.63	.4680	.58	.4768
RSVP Risk of Serious Physical Harm	.53	.4065	.69*	.5782	.60	.4674	.60	.4876	.50	.3961
RSVP Immediate Action Required	.55	.4268	.51	.3170	.63	.4679	.68*	.5185	.54	.4365
RSVP Past Total Items 1-22	.58	.4671	.69*	.5485	.69*	.5285	.71*	.5389	.59	.4869
RSVP Recent Total Items 1-22	.61	.4774	.55	.3377	.62	.4281	.66	.4685	.70**	.6080
RSVP Relevance Total Items 1-22	.60	.4872	.70*	.5485	.70*	.5684	.72*	.5787	.63*	.5374
PCL-R Total Score	.55	.4169	.73*	.5691	.76* *	.6291	.79**	.6295	.59	.4870
RM2000 Sexual Scale	.52	.4064	.54	.3574	.52	.3767	.54	.3770	.54	.4266
RM2000 Violence Scale	.54	.4166	.70	.4892	.72* *	.5687	.70*	.5388	.57	.4568
Sexual Sadism Scale (SeSaS) Part 1 Score	.37*	.2648	.42	.2461	.42	.2857	.44	.2859	.39	.2750
SSPI Total Score	.60	.4774	.33	.1452	.41	.2457	.44	.2662	.54	.4265

Table 17. ROC analysis: relationship between total scores and summary judgments, and offending (including convictions, charges and allegations) over the whole follow-up period.

Note. *p < .05; **p < .01; AUC = Area under the curve; C.I. = Confidence interval

Conviction during full follow-up (Table 18)

Any sexual conviction Sexual conviction was not found to be significantly associated with any RSVP or other measures.

Any violent conviction A number of predictors were associated with *any violent* conviction during follow-up. These include a RSVP Past and Relevance totals, the PCL-R Total and RM2000 Violence scale.

Any serious conviction No measures were significantly associated with serious convictions during follow-up. The largest AUCs were for RSVP Recent totals and RM2000V.

Any serious sexual conviction No measures were significantly associated with serious sexual convictions during follow-up. The largest AUCs were for the RSVP Immediate Action Required judgment and the RSVP Recent total.

Any conviction RSVP Recent and Relevance summed totals were associated with *any* convictions at follow-up as well as the summary judgement Case Prioritization.

conviction over the whole ronow-up period.										
Predictor Variables	Any	Sexual	Any	Violent	Any	Serious	Any	Serious	Con	Any
	Col	IVICTION	Col	IVICTION	Con	victions	Convictions		Convictions	
	AUC	95% C.I.	AUC	95% C.I.	AUC	95% C.I.	AUC	95% C.I.	AUC	95% C.I.
RSVP Case Prioritization	.60	.4673	.60	.3586	.59	.3978	.60	.3586	.62*	.5173
RSVP Risk of Serious Physical Harm	.48	.3164	.69	.5186	.49	.2968	.46	.2172	.50	.3862
RSVP Immediate Action Required	.58	.4175	.55	.2585	.60	.3883	.72	.5094	.60	.4872
RSVP Past Total Items 1- 22	.46	.3062	.76*	.5899	.52	.2875	.51	.2083	.55	.4466
RSVP Recent Total Items 1- 22	.66	.4784	.60	.2596	.64	.3593	.73	.44-1.0	.73* *	.6184
RSVP Relevance Total Items 1-22	.53	.3768	.80*	.6099	.57	.3776	.57	.3283	.63*	.5274
PCL-R Total Score	.45	.2862	.78*	.6293	.58	.3582	.58	.2690	.61	.4973
RM2000 Sexual Scale	.43	.2858	.61	.4182	.51	.3270	.54	.3177	.50	.3862
RM2000 Violence Scale	.49	.3564	.82*	.6798	.65	.4386	.60	.3485	.60	.4872
Sexual Sadism Scale (SeSaS) Part 1 Score	.35	.2248	.36	.1756	.36	.1953	.36	.1756	.43	.3155
SSPI Total Score	.51	.3368	.35	.0862	.38	.1660	.45	.1773	.47	.3460

Table 18. ROC analysis: relationship between total scores and summary judgments, and conviction over the whole follow-up period.

Note. *p < .05; **p < .01; AUC = Area under the curve; C.I. = Confidence interval

Breach of conditions (Table 19)

A number of predictors were associated with a Breach at 6-months follow-up including RSVP summed Recent items and Relevance Items, PCL-R Total score and RM2000 Violence score. Fewer items were associated with a Breach at 12 months (RSVP summed Recent items and RM2000 Violence scale). The PCL-R Total score was associated with a Breach at 5 years.

Predictor Variables	Breach,		Breach,		Breach,		
	6 m	6 months		12 months		5 years	
	AUC	95% C.I.	AUC	95% C.I.	AUC	95% C.I.	
RSVP Case Prioritization	.64	.4979	.63	.4977	.79	.6197	
RSVP Risk of Serious Physical Harm	.57	.4073	.52	.3668	.66	.3695	
RSVP Immediate Action Required	.61	.4578	.56	.4072	.51	.2379	
RSVP Past Total Items 1-22	.63	.4878	.58	.4473	.63	.4185	
RSVP Recent Total Items 1-22	.76**	.6588	.74**	.6385	.79	.571.0	
RSVP Relevance Total Items 1-22	.68*	.5582	.63	.5077	.79	.6197	
PCL-R Total Score	.71*	.5785	.64	.4979	.79*	.6198	
RM2000 Sexual Scale	.54	.3870	.53	.3868	.61	.3687	
RM2000 Violence Scale	.68*	.5285	.66*	.5182	.62	.3489	
Sexual Sadism Scale (SeSaS) Part 1 Score	.54	.3871	.53	.3869	.51	.2180	
SSPI Total Score	.45	.2961	.41	.2656	.75	.45-1.0	

Table 19. ROC Analysis: relationship between total scores and summary judgments, and breaches of conditions at 6 moths, 12 months and 5 years.

Note. *p < .05; **p < .01; AUC = Area under the curve; C.I. = Confidence interval

Survival analysis

Survival curves for time to sexual offending, time to violent offending and time to breach in three groups defined by each of the RSVP summary judgements (Case Prioritization, Risk of Serious Physical Harm, Immediate Action Required) are shown in Figures 1 to 9. These need to be interpreted with caution as the survival curves for different groups overlap at points. Case Prioritization was significantly related to time to sexual offending (Log Rank [Mantel-Cox] test: chi-square=7.850, p=0.020) and to time to breach (Log Rank [Mantel-Cox] test: chi-square=6.424, p=0.040), but not to time to violent offence. The curves indicate that there was a particular differentiation between the low prioritisation group and the other two groups.

Risk of Serious Physical Harm was close to being significantly related to time to violent offending (Log Rank [Mantel-Cox] test: chi-square=5.955, p=0.051), but was not related to time to sexual offending or to breach. Again there was a particular differentiation between the low risk group and the other two groups. Immediate Action Required ratings were not significantly related to time to any of these three outcomes.





Figure 2. Kaplan-Meier survival curves for time to any sexual offence in the three RSVP Risk of Serious Physical Harm groups



Figure 3. Kaplan-Meier survival curves for time to any sexual offence in the three RSVP Immediate Action Required groups







Figure 5. Kaplan-Meier survival curves for time to any violent offence in the three RSVP Risk of Serious Physical Harm groups



Survival Functions

Figure 6. Kaplan-Meier survival curves for time to any violent offence in the three RSVP Immediate Action Required groups











Figure 9. Kaplan-Meier survival curves for time to any breach in the three RSVP Immediate Action Required groups



Survival curves for the same three outcomes (i.e. time to sexual offending, violent offending and breach) for three groups defined by the RSVP Ever Present total score, rather than by summary judgments, are shown in Figures 10, 11, and 12. The three low, medium and high risk groups were created by taking the highest scoring third and lowest scoring third of the sample, leaving a medium group. Although for all three outcomes the low risk group faired better than the other groups, this difference was not significant, except for time to breach (Log Rank [Mantel-Cox] test: chi-square=7.384, p=0.025). With time to breach the low risk group did best, followed by the high risk group, with the medium risk group fairing worst.

Figure 10. Kaplan-Meier survival curves for time to any sexual offence for RSVP *Ever Present* Total scores divided into three groups



Figure 11. Kaplan-Meier survival curves for time to any violent offence for RSVP *Ever Present* Total scores divided into three groups



Figure 12. Kaplan-Meier survival curves for time to breach for RSVP *Ever Present* Total scores divided into three groups



Comparison of RSVP Summary Judgements and Risk Management Level

Two RSVP Summary Judgements (Case Prioritization; Risk of Serious Physical Harm) determined at the time of assessment were compared with the overall management level for each offender during follow-up. These management levels are based on the nine "management by assessment" categories presented in Table 1 (see Method section, p.15).

The number and percentages of offenders for each of the nine categories of assessment by management levels are reported in Tables 20 and 21. Overall the majority of offenders were managed at the level that corresponded to their Case Prioritization and Risk of Serious Physical Harm rating. A minority of offenders were over or under managed by two levels. A significant percentage of offenders were managed at either one level higher or lower than their assessed level.

	Low Assessed	Moderate Assessed	High Assessed
Low Managed	24 (22)	19 (17.4)	0 (0)
Medium Managed	9 (8.3)	21 (19.3)	7 (6.4)
High Managed	2 (1.8)	7 (6.4)	20 (18.3)

Table 20. Assessed Case Prioritization Level by Modal Managed Level, Frequency (%)

Table 21. Assessed Risk of Serious Physical Harm Level by Modal Managed Level, Frequency (%)

	Low Assessed	Moderate Assessed	High Assessed
Low Managed	21 (19.3)	19 (17.4)	3 (2.8)
Medium Managed	13 (11.9)	14 (12.8)	10 (9.2)
High Managed	4 (3.7)	11 (10.1)	14 (12.8)

Case Prioritization and Risk of Serious Physical Harm were then compared with the overall management level for each offender who committed a sexual offence in follow-up.

Case Prioritization was first compared with management level at one year follow-up for individuals who had committed a sexual offence (N = 13). Due to the limited number of seriously harmful offences recorded during the one year follow-up, it was not possible to run the analysis for Risk of Serious Physical Harm. As seen in Table 22, 5.3% of offenders assessed and managed at a Low level had committed a sexual offence. 7.1% of offenders assessed as Low Case Prioritization but managed at a Medium level committed a sexual offence during one year follow-up. Of those assessed as Moderate and managed as Low, 15% had committed a sexual offence after one year. 26.1% of individuals assessed and managed at a Medium level had sexually offended at one year follow-up. Individuals managed at a High level did not commit a sexual offence within one year of follow-up, regardless of their assessment level. At one year follow-up, 28.6% of those offenders assessed as High case prioritisation and managed as Medium had committed a sexual offence. On the whole, offenders assessed as 'Medium' tended to reoffend more often even if they were managed accordingly. Overall those offenders who were managed one level lower than assessed had a higher rate of sexual offending during one year follow-up (e.g. High Assessed/Medium Management). Those offenders managed at levels higher than their assessed Case Prioritization revealed low percentages of sexual offending in one year follow-up (e.g. Moderate Case Prioritization/High Management).

U × /						
	Low Assessed	Moderate Assessed	High Assessed			
Low Managed	1 (5.3)	3 (15)	0 (0)			
Medium Managed	1 (7.1)	6 (26.1)	2 (28.6)			
High Managed	0 (0)	0 (0)	0 (0)			

Table 22. Assessed Case Prioritization by Managed Mode at 1 Year for Any Sexual Offending (%)

Note: N = 13; Chi-Square value 11.046; Note. **p* < .05

Case Prioritization was compared with management level for individuals who had committed a sexual offence at any point during follow-up (N = 26) (see Table 23). Throughout followup, 8.3% of those offenders assessed and managed at a Low level had committed a sexual offence. 22.2% of offenders assessed as Low Case Prioritization but managed at a Medium level committed a sexual offence during follow-up. Of those assessed as Moderate and managed as Low, 27.3% had committed a sexual offence. 45% of individuals assessed as Moderate and managed at a Medium level sexually offended. Offenders who were assessed as Low or Moderate and managed at a High level did not commit a sexual offence during follow-up. During follow-up, 50% of those offenders assessed as High and managed as Medium had committed a sexual offence. 15.8% of individuals assessed as High and managed as High did sexually offend. The results indicate that individuals managed at one level below their assessed Case Prioritization were more likely to commit a sexual offence during followup than those managed at level commensurate with assessed risk (e.g. High Assessed/Medium Managed). However, a high percentage of moderate Case Prioritization/Medium management offenders sexually offended within follow-up. Again, those offenders managed one or two levels above their assessed Case Prioritization showed relatively low percentages of sexual offending in follow-up (e.g. Moderate Case Prioritization/High Management).

Table 23. Assessed Case Prioritization by Managed Mode for Any Sexual Offending in						
Follow-up (%)						
	Low Assessed	Moderate Assessed	High Assessed			

	Low Assessed	Moderate Assessed	High Assessed
Low Managed	2 (8.3)	6 (27.3)	0 (0)
Medium Managed	2 (22.2)	9 (45.0)	4 (50.0)
High Managed	0 (0)	0 (0)	3 (15.8)

Note: N = 26; Chi-Square value 14.146*; Note. *p < .05

Risk of Serious Physical Harm was compared with management level for individuals who had committed a serious sexual offence during total follow-up (N = 12) (*see Table 24*). During follow-up, 4.3% of those offenders assessed and managed at a Low level had committed a serious sexual offence. 8.3% of offenders assessed as Low Risk of Serious Physical Harm but managed at a Medium level committed a serious offence during follow-up. Individuals who

were assessed as Low and managed as High did not commit a serious sexual offence. Of those assessed as Moderate and managed as Low, 10% had committed a serious sexual offence. 23.1% of individuals assessed and managed at a Medium level seriously sexually offended. One offender (accounting for 9.1% of the sample) who was assessed as Medium and managed as High committed a serious sexual offence. 66.7% of those individuals assessed as High and managed at a Low level committed a serious sexual offence in follow-up. 10% of those offenders assessed as High and managed as Medium had committed a sexual offence. One offender, accounting for 7.1% of the sample, assessed and managed as High did seriously sexually offend. Again, offenders assessed and managed as Moderate had the highest frequency of committing a serious sexual offence. A large percentage of those undermanaged (e.g. High Assessed/Low Managed) seriously sexually offended..

Table 24. Assessed Risk of Serious Physical Harm by Managed Mode for Any Serious Sexual Offending (%)

	Low Assessed	Moderate Assessed	High Assessed
Low Managed	1 (4.3)	2 (10.0)	2 (66.7)
Medium Managed	1 (8.3)	3 (23.1)	1 (9.1)
High Managed	0 (0)	1 (10.0)	1 (7.1)

Note: N = 12; Chi-Square value13.205; Note. *p < .05

Survival analysis was undertaken for the three outcomes; time to sexual offence; time to violent offence; and time to breach. Cases were grouped by the degree to which a case was over-managed or undermanaged with respect to assessed Case Prioritization or Risk of Serious Physical Harm. Again, these need to be interpreted with caution as the curves overlap at points. For sexual and violent offending, under-management was associated with quicker offending for both sexual and violent offences when both approaches to grouping cases are applied (Figures 13 – 18). For Risk of Serious Physical Harm/sexual offending (Log Rank [Mantel-Cox] test: chi-square=20.983, p=0.000), Case Prioritization/violent offending (Log Rank [Mantel-Cox] test: chi-square=14.952, p=0.002) and Risk of Serious Physical Harm/violent offending (Log Rank [Mantel-Cox] test: chi-square=14.952, p=0.002) and Risk of Serious Physical Harm/violent offending (Log Rank [Mantel-Cox] test: chi-square=14.952, p=0.002) and Risk of Serious Physical Harm/violent offending (Log Rank [Mantel-Cox] test: chi-square=14.952, p=0.002) and Risk of Serious Physical Harm/violent offending (Log Rank [Mantel-Cox] test: chi-square=16.102, p=0.003) differences between the undermanaged group and the other groups were significant, i.e. the undermanaged group offended much quicker.

For breaches the opposite effect was found. Over management was significantly related to shorter time to breach. Test statistics were: Case Prioritization/breach (Log Rank [Mantel-

Cox] test: chi-square=19.766, p=0.000) and Serious Physical Harm/breach (Log Rank [Mantel-Cox] test: chi-square=17.468, p=0.002).

Figure 13. Kaplan-Meier survival curves for time to any sexual offence in the discrepancy between assessed level and managed level for those who committed any sexual offence during follow-up



Figure 14. Kaplan-Meier survival curves for time to any sexual offence in the discrepancy between assessed level and managed level for those who committed any serious sexual offence during follow-up



Figure 15. Kaplan-Meier survival curves for time to any violent offence in the discrepancy between assessed level and managed level for those who committed any sexual offence during follow-up



Figure 16. Kaplan-Meier survival curves for time to any violent offence in the discrepancy between assessed level and managed level for those who committed any serious sexual offence during follow-up



Figure 17. Kaplan-Meier survival curves for time to breach in the discrepancy between assessed level and managed level for those who committed any sexual offence during follow-up



Figure 18. Kaplan-Meier survival curves for time to breach in the discrepancy between assessed level and managed level for those who committed any serious sexual offence during follow-up



Predictive Validity of Scenario Planning

There were 26 cases in the sample where sexual allegations, charges or convictions were reported during follow-up. Results show that 65.4% of victims were female, with the majority being children (53.8%). 57.7% of the offenders were strangers to the victim and 53.8% of the behaviour was noncontact. Nearly half of the noncontact offences (8 of 14) were internet offences. It is worth noting that 19.2% of the severity of the sexual behaviour during follow-up was contact with serious harm. This includes violence to the victim and/or penetration.

In terms of scenarios matching actual offences, the majority of the scenarios matched the behaviour for each of the 4 categories when compared to further instances of sexual offending in follow-up. Gender was correctly identified for 96.2% of offences, victim age matched 76.9%, victim relationship 69.2%, and level of severity of the behaviour was correctly identified 61.5% of the time (see Table 25).

	U I		0 1	
	Victim Gender	Victim Age	Victim	Level of
			Relationship	Severity
Not Correct	1 (3.8)	5 (19.2)	2 (7.7)	2 (7.7)
One off	N/A	1 (3.8)	6 (23.1)	8 (30.8)
Correct	25 (96.2)	20 (76.9)	18 (69.2)	16 (61.5)

Table 25. Scenario Planning matching actual offences during follow-up. Frequency (%)

Discussion

The purpose of this study was to evaluate the RSVP's validity, reliability and utility in a Scottish offending population. This has been done within a service based in South East Scotland that takes referrals from criminal justice agencies who are currently managing or will soon be managing sex offenders in the community. It is therefore a real-world study where assessments inform ongoing clinical practice and ongoing management. It should be noted that the cases referred are usually ones perceived by referrers to be complex and/or high risk.

The service was able to access information about the level at which each subject was managed throughout the follow-up period following the RSVP assessment. This allowed management level to be factored into the analysis. This aspect differentiates this study from many others studies, most of which are retrospective studies that follow up offenders who are not being managed or are studies where the level of management that each offender receives is either not available to the researchers or not factored into the analysis.

One of the strengths of this research is that the RSVP was not applied purely for the purposes of research. The RSVP has been applied prospectively rather than retrospectively. The RSVP assessments in the current study were available to agencies who were managing the cases. This study therefore reflects how clinicians are using the RSVP in their clinical practice to inform the development of case formulation and risk management strategies. However, a consequence of this is that the follow-up does not allow for the analyses of what happens to the offenders, who are assessed as posing different levels of risk, when they are left to be in the community without supervision. Therefore, the follow-up is complicated by the fact that the cases in the sample present a range of risks and are receiving a varying range of risk management strategies. Another important point is that the RSVP has not been completed at the point of sentencing in these cases. It has been applied once the subject is in the community, either having left prison or else once a community order or legal order has been imposed. The result of this is that the service has made risk management recommendations as part of the RSVP assessment but the agencies managing the case may or may not have been able to put these recommendations into practice. As a result some cases were being managed with less or more restrictions than were recommended by the RSVP assessment.

Overall, the fact that we have conducted the study on a group of managed offenders is important when considering how the RSVP and other instruments are related to outcome in this study. Given that this is a real world study it would not have been appropriate or ethical to assess these cases, leave them unmanaged and then monitor outcome, purely for the purposes of research. Our analyses demonstrate that there are high rates of further offending in the groups that are under-managed compared to those that are managed at a level consistent with their assessed risk or over-managed.

A small number of cases did not have a sexual conviction at the point that the RSVP was applied. The service is aimed at individuals who pose a risk of sexual violence whether or not they have a specific sexual conviction. Some subjects had convictions for other types of offending, e.g. violence, non-sexual offending. In many of these cases an offence that was not legally determined as sexual clearly had a sexual element, e.g. a number of homicide cases. In all cases there was deemed to be a risk of sexual violence that had to be managed. Unlike other studies our study was concerned with individuals who appear to pose a risk of sexual violence rather than individuals with a conviction for a legally defined sexual offence.

As mentioned above, an aspect of the population from which this sample is drawn that might skew the findings is that the service tends to be referred more complex and often more high risk sex offenders, as these are the cases that are felt to require additional advice and consultation. This is not therefore a representative sample of sex offenders. They are more likely to be assessed as being at a higher level of risk and therefore requiring higher levels of management. In particular they are likely to pose a higher risk of serious harm. Part of the reason for referral may be that agencies feel standard assessment and management plans are insufficient.

When considering our sample, it should be noted that most were referred by criminal justice agencies, with 30% being in prison at the point of assessment. Cases tended to be at MAPPA Levels 2 and 3 with a third considered to pose a significant ongoing risk of serious harm. Three quarters were contact offenders, half had raped, 1 in 10 were subject to life sentences, 7% had killed, and half had stranger victims. One in ten were PCL-R psychopaths, the majority had personality disorders, half had paraphilias, and 80% had previous contact with

mental health services. These baseline characteristics emphasise that this is an unusual, complex, and high risk sample.

It is well recognised that sexual offending is under-reported and under-convicted. Unlike some other studies, we ascertained whether there had been allegations or charges during follow-up that did not lead to conviction. This enabled us to have a broader picture of sexual offending during follow-up and meant that for the whole sample there was a relatively high rate of follow-up sexual offending (23.9%) and conviction (11.9%). For comparison with other studies we undertook some analyses using convictions only, alongside analyses using any offending.

The relatively small sample size made some analyses difficult. For example, there was a need to factor management level into the analyses of further offending during follow-up. Dividing the sample into the different management levels led to small numbers in each group. In addition, some offenders had a relatively short follow-up depending on when they were assessed, compared to others. This increases the chances of Type II errors.

Main Findings

How reliable are ratings of items, Sections, total scores and summary judgments?

Overall the RSVP was reliably rated. The ratings of the Past and Recent items fell within the 'Almost Perfect' category. The majority of the ratings of the Section and Total Scores of the RSVP were rated as 'Almost Perfect'. The Relevance items were less reliably rated than the Presence Items although they were still mostly within the Substantial range. As far as we are aware this is the first study of the RSVP to examine reliability based on a full assessment including interview. All other studies have been based on records or vignettes only. We had a relatively small number of reliability cases. The number of cases was sufficient for us to conclude that our ratings were reliable so as to then progress to analyses of validity.

Our reliability findings for presence and relevance of items are similar to findings by Douglas & Belfrage (2014) with the HCR-20 V3 but different to Hart (2003) and Watt et al (2006) who found that relevance and presence ratings were equally reliably rated for the RSVP. The high inter-rater reliability of items found in our study is similar to findings for other SPJ tools such as HCR-20 V2 and SVR-20 (Douglas & Reeves, 2010; Hart & Boer, 2010).

Psychopathic PD was one of the items with the highest reliability. This is in keeping with studies on the reliability of the PCL-R (Hare, 2003). However Problems with Minimisation/Denial, Problems with Supervision and Problems with Planning, where the rater has to use their judgement when rating the item, were also very reliably rated.

Sutherland et al (2012) assessed the reliability of the RSVP by asking 28 Scottish mental health professionals with varying levels of experience and training to rate 6 case vignettes. They found a lower level of inter-rater reliability than we did. Average inter-rater reliability for items was fair, with individual item reliabilities ranging from poor to excellent. For Summary Judgments inter-rater reliability was good except for Immediate Action Required, which was fair. The lower level of reliability found by Sutherland et al. (2012) compared to our findings, and the findings of 3 other studies of the RSVP (Hart, 2003; Watt et al, 2006; Watt & Jackson, 2008), may be due to their methodology where they had a small number of vignettes rated by many assessors and the fact that their assessors had a range of experience and training.

The Summary Judgement ratings were found to be very reliable. This concurs with findings from other RSVP studies which have found good to excellent reliability for Summary Judgments (Hart, 2003; Watt et al, 2006; Watt & Jackson, 2008; Sutherland et al, 2012). This is particularly interesting given the lack of clear definitions around how to apply the potential ratings of each judgement. Summary Judgements ratings have been investigated less often than Total Scores. However, overall they tend to show at least moderate reliability in the HCR-20 studies (Douglas & Reeves, 2010). The Summary Judgement rating may be a part of the process that clinicians under-value or omit as it seems a less important component in the SPJ process. This should give clinicians confidence that these ratings can be made reliably.

Given the relatively small number of reliability cases in our study, what we can say is that the reliability levels we found allowed us to proceed to consider validity, and our findings are very similar to larger reliability samples (albeit those samples were retrospective case file studies) which found that Items and Summary Judgements overall had a good to excellent level of reliability when applied by well-trained staff.

How do RSVP ratings correspond with ratings using other instruments, such as Risk Matrix 2000 and the Psychopathy Check List-Revised?

The RSVP was found, overall, to be significantly correlated with the PCL-R and the RM2000. Different sub-scores on each measure correlated differentially with sub-scores of other measures. Given that the underlying factors being assessed by these measures are similar then it would be expected that there would be a correlation. For example, the RM2000 Violence score and the Serious Physical Harm Summary Judgement rating on the RSVP had a significant association. The lack of association between RM2000 scores and the Immediate Action Required rating in RSVP can be understood in the context that RM2000 is looking at static factors that do not change over time and therefore looking at risk in the medium to long term whereas the Immediate Action Required rating is related to factors that are acute or dynamic and therefore more proximal to offending. Other studies have found that the RSVP and SVR-20 Total Scores correlated strongly with actuarial tool scores such as VRAG, Static-99 and SORAG (Hart & Boer, 2010).

How do RSVP ratings (scores and summary judgments) predict further offending (including convicted and unconvicted offending; and including both sexual and non-sexual offending)? In order to answer this question we undertook two types of analysis: ROC analysis and Survival Analysis. ROC analysis allowed us to look at the association between a measure and an outcome. Survival Analysis allowed us to look at a measure and time to outcome. There are some discrepancies between the ROC findings and the Survival Analysis findings.

Interestingly, using ROC analysis, none of the RSVP Judgements or Totals were associated with *any sexual offending* or *sexual conviction* during follow-up. However the RM2000 scores and the PCL-R scores were not found to be associated with this outcome either. There may be a number of explanations for this. This is not a representative sample of sex offenders, as explained above. Alternatively, it may be because this is a group of managed offenders, where management plans may have been influenced by assessments including RM2000, PCL-R and RSVP. Our findings in this regard differ from the findings of Kropp (2001; as cited in Hart & Boer, 2010) and Hart & Jackson (2008; as cited in Hart & Boer, 2010). Using Survival Analysis, Case Prioritization was associated with time to *any sexual offending*. This suggests that Case Prioritization is not tapping into whether someone sexually offends but how quickly they will sexually offend.

With ROC analysis, we found the RSVP Past and Relevance Total Scores and Immediate Action Required Summary Judgement were associated with *any serious sexual offending*. This suggests the RSVP may be a better tool for assessing risk of serious sexual offending than more minor sexual offending. It is interesting to note that the RSVP performed almost as well as the PCL-R in this regard. With ROC analysis, RSVP Past and Relevance totals and Risk of Serious Harm Judgement were associated with *violent offending*. As expected, the PCL-R was also associated with *violent offending*. Survival Analysis showed that Risk of Serious Harm Judgment was associated with time to violent offending. It is also of note that the RSVP was associated with violent offending when it was not associated with sexual offending. This may relate to the non-specificity of types of interpersonal violence and the association of RSVP ratings with more serious sexual offending. More serious sexual offending. Perhaps the RSVP is a useful tool for assessing risk of serious harm (whether sexual or non-sexual) in sexual offenders, rather than being useful for assessing risk of sexual offending.

This fits with the ROC findings for serious offending of any type and would support the practice of using the RSVP for higher risk MAPPA cases and sexual offenders being considered for OLRs. Assessing risk of serious harm is an issue that criminal justice practitioners struggle with and is not addressed by the instruments used in Scotland for sexual offenders (i.e., Risk Matrix 2000 and Stable and Acute 2007). Our findings may indicate that the RSVP has a specific role or may add to the assessment of cases where serious harm is under consideration. The LSCMI (Level of Service/Case Management Inventory; Andrews, Bonta & Wormith, 2004) which is used to assess and plan the management of all offenders in Scotland guides practitioners to use appropriate specialist tools when an offender is screened as posing a potential risk of serious harm. Our findings would support the RSVP playing this role for sexual offenders. However when considering these findings it is important to bear in mind that our cases were a particularly complex and high risk cohort, so our findings may not generalise to less concerning cases.

In the ROC analysis, we did not find that Summary Judgements were better predictors than RSVP Total scores although some of the Summary Judgements had a reasonable level of association with some of the outcomes. This is contrary to research with the HCR-20 which has found that Summary Judgements have a similar or better association with re-offending (Douglas & Reeves, 2010) and to RSVP validity studies (Kropp, 2001; Hart & Jackson, 2008; as cited in Hart & Boer, 2010). Different judgements were associated with different outcomes, i.e., Immediate Action Required was associated with any serious sexual offending; Serious Physical Harm was associated with any violent offending and any serious offending. Conversely, when using Survival Analysis, Summary Judgements performed better than Total Scores and again the analysis showed that different Summary Judgements were associated with different outcomes. Case Prioritization was significantly associated with time to sexual offending and time to breach but not time to violent offending, whereas Risk of Serious Harm was associated with time to violent offending. The fact that the different Summary Judgements are pulling out different aspects of risk is an important argument supporting their use in planning risk management rather than using Scores or even one rating of High, Medium or Low risk. This is in keeping with FRAME (Risk Management Authority Scotland, 2011) which emphasises the multi-faceted nature of risk and the importance of recognising that different factors may be related to different facets of risk.

When comparing our results to other studies looking at ROC curves with other risk assessment tools, it is common to find one ROC curve statistic quoted for a tool. It is important to note with the RSVP that there is not just one ROC curve but potentially six, if all the Total Scores and Summary Judgements are taken into account. It is not straightforward to compare our findings with the findings of other instruments. In general the significant ROCs that we found for some of the RSVP ratings and offending outcomes were of a similar magnitude to those reported in the literature for other risk assessment instruments (Hanson & Morton-Bourgon, 2009; Otto & Douglas, 2010).

Interestingly, in the Survival Analysis, when using violent offending, sexual offending and breach as outcomes, the Low risk group were clearly differentiated from the other groups. In keeping with other findings, Low Risk offenders do not tend to re-offend. Fazel, Singh, Doll & Grann (2012), in a systematic review, found that risk assessment tools identify low risk offenders with high levels of accuracy but overall they have low to moderate positive predictive values. It could be argued that tools such as the RSVP are good for identifying low risk individuals who do not require risk management.

How are RSVP summary judgements related to further offending after taking into account the level of management cases are subject to?

We are only aware of one previous study that has attempted to look at the mediating effect of Risk Management Level on recidivism (Belfrage, Strand, Storey, Gibas, Kropp, & Hart, 2012). This study was looking at a larger sample than ours but one solely being assessed and managed by the police where the risk management plans were directly derived from the Spousal Assault Risk Assessment (SARA). In their study they found that the SARA score was a better predictor of recidivism than Summary Judgments. Their risk management level was a score derived from adding up the number of risk management strategies recommended. To analyse the interaction between assessment and management level and recidivism they undertook two analyses. The first analysis used logistic regression to see if a model using assessment and management level. They found that to be the case. In the second analysis, they constructed a 2x2 contingency table by dichotomizing numerical total scores and the number of management strategies recommended. High risk management was associated with decreased recidivism in high risk perpetrators but increased recidivism in low risk

perpetrators. Therefore they found risk management mediated the association between risk assessment and recidivism.

We attempted a similar type of analysis but with different data, i.e. three categories of risk level and three categories of risk management level. Given that there has only been one study (Belfrage et al, 2012) that has attempted to research this issue there is not a standard approach or methodology for us to adopt. We acknowledge that others may have attempted to use a different way of measuring assessed risk or risk management level. Our analysis would seem to suggest that low risk offenders require only low risk management and high risk offenders require high risk management.

The findings about those in-between High and Low were interesting. Offenders assessed as a Moderate risk and managed with a medium level of intervention had the highest frequency of further serious sexual offending. Does this indicate there is a 'messy middle'? Are clinicians confident about those that are Low Risk and those that are High Risk but there are a group of offenders where it is less clear? Or is it that the level of intervention that we rated as being a Medium level of intervention in this study was not sufficient to manage Moderate risk offenders? Having a SOPO or CPO in addition to being an RSO was defined as a Medium management strategy but from this data, it would appear that it may not be sufficient to manage risk in those that pose a Moderate risk of re-offending. Given that over-managing offenders in this sample did not lead to a greater frequency of re-offending, should those who are assessed as moderate risk be managed as High Risk? Alternatively, are these offenders actually High Risk offenders and due to the fact that this is a non-representative sample those who are clearly assessed as High risk are actually in an additional category of 'Very High' risk? Given that these are Judgements and not clearly defined terms within the tool they do rely on professional opinion and will doubtless depend on the clinicians experience and the types of cases they routinely assess. It is possible that the opinions of clinicians in this service have been skewed by only seeing more extreme cases and therefore are under-rating the risk in high risk offenders. Another point of note is that by labelling someone 'Moderate risk', the assessor is effectively saying they are not clearly High or Low risk. However the medium range of risk may cover quite a wide range of risks. Medium management as defined in the study ranged from relatively little intervention (on an order being seen once a month) to quite intensive intervention (Being seen every day). Two individuals receiving medium risk management could be managed in very different ways. Low and High management had less variability.

For outcomes related to breach of conditions, we found a different relationship with risk management level. Those who were over-managed were more likely to breach than those who were under-managed. It is likely that this was due to the high level of monitoring they received so breaches of conditions were more likely to be detected and acted on, potentially reducing the likelihood of an actual offence occurring. The more closely someone is monitored and supervised the more likely that a breach of conditions will be detected.

Do the risk scenarios generated by clinicians using the RSVP include the types of sexual violence that recidivist offenders go on to perpetrate?

This is an area of research that has been highlighted as lacking in the field of SPJ research. We therefore attempted an analysis to look at the "predictive accuracy" of the scenarios. Where subjects had gone on to re-offend the potential scenarios produced with the RSVP were compared to the actual event to see if there were similarities. Specific categories of the scenarios were chosen for the comparison. Age of victim appeared to be a category that showed a good match between scenario and outcome. The offending behaviour also showed a reasonably good match while the victim relationship and severity of harm were less well matched although still showed a match in around two thirds of the cases. Overall, most of the actual offences were in keeping with the scenarios set out during the assessment process on the four variables that we looked at. A lack of a match between a scenario and actual offence is not necessarily a bad outcome in itself as if a serious harm scenario has not occurred and the actual offence is a less serious offence that was not apparent in scenario planning then this may be as a result of the good risk management plan

Howard, Barnett, and Mann (2014) found that sex offenders who re-offend mostly go on to commit similar offences to previous offences and on that basis it would appear that scenario planning may be a straight forward process in many cases. From our own experience of training professionals in SPJ tools, when different groups are asked to generate scenarios based on the same case study they generate very similar scenarios. Other research has found that scenarios are reliably rated by different raters looking at the same case material (Wilson, 2013)

Summary of findings

This study provides further evidence that the RSVP is a reliable tool. This was true for both individual Item ratings, Total Scores and Summary Judgements. Given that clinicians do not use the RSVP to sum totals, this study should give clinicians confidence that Summary Judgements are a reliable method of summarising the risk an offender poses. There was evidence of convergent validity with the RM2000 and PCL-R. Predictive validity is complicated with a tool like the RSVP because practitioners do not use Total Scores; they should use the three Summary Judgements. Assessing predictive validity is also made complicated by the level of management that cases receive. Therefore answering a straight forward question about whether the RSVP has predictive validity for sexual offending or other offending is unrealistic. Rather, the answer to this question has to take into account the complexity of the tool and the potential confounding influence of management level. In our sample we also found differences when we used different approaches to analysing the outcome data, i.e. ROC analysis and survival analysis. Using ROC analysis, the RSVP Total scores and some of the Summary Judgements predicted violence, any serious offending, and serious sexual offending but did not predict any sexual offending. Using survival analysis, Case Prioritization predicted time to any sexual offending and breach. Unlike other studies we then attempted to factor in the risk management level. Of particular note was the finding that those that were identified as High Risk using the RSVP, who were not risk managed commensurate with that risk, very quickly re-offended. The importance of risk management level means that caution is required when interpreting predictive validity data that does not take into account the level at which cases are managed. When considering our findings, it has to be borne in mind that ours is an unusual, complex and high risk sample of sexual offenders

Research Implications

Our study, alongside a handful of other studies, provides some evidence that the RSVP can be reliably rated, has predictive validity for some types of further offending and that management of cases should be commensurate with the risk posed. However given that the RSVP was published in 2003 it is disappointing that there have been so few studies of it compared to other risk assessment instruments. We require further prospective research looking at the reliability and predictive validity that also takes into account risk management. A standardised methodology needs to be developed to research SPJ instruments such as the RSVP taking into account the level of risk management applied to cases. We also need
research on some of the relatively new developments in SPJ risk practice: formulations, scenario planning and risk management. It is important that research on such tools is applicable to how they are used in practice. The ultimate aim of any SPJ approach is to prevent violence. Therefore the ultimate test of the RSVP would be to have a randomised controlled trial where a group of sexual offenders receives standard management and another group gets management informed by the RSVP. Such research would have ethical and resource challenges. However given that the evidence relating to the assessment and management of sexual offenders is far from robust; that the impact of sexual abuse can be devastating; and that there is a high level of resource dedicated to the assessment and management of sexual offenders, it is incumbent on us to have high quality evidence on which to base practice. We should never ignore the possibility that accepted or apparently sensible approaches to assessment and management may be at best ineffective or at worst harmful. High quality research is required. Our study should be considered when RATED is next updated by the RMA. Whether our study, alongside the limited other research, provides sufficient evidence to indicate that the RSVP is reliable, valid and useful for risk practice in Scotland will be for others to consider.

Another aspect of SPJ risk assessment that requires further research is the characteristics of the professionals who use these tools. As mentioned above, SPJ instruments structure a professional in their task so the characteristics of the professional carrying out the assessment may be as important as the instrument. By characteristics, we refer to factors such as: experience, knowledge, discipline, training, supervision, adherence to the SPJ guidelines. Are there particular characteristics of assessors that make SPJ assessments more reliable and valid?

Practice Implications

Despite the limited prior research on the RSVP and the modest nature of the current study, we would tentatively suggest the following practice implications. The RSVP has a useful role to play in the assessment and management of more complex/higher risk sexual offenders, particularly when managing the risk of serious harm. Our findings support the use of the RSVP for the minority of sexual offenders who pose a risk of serious harm, such as those managed at MAPPA levels 2 and 3, and those being considered for an OLR. The use of the tool in this way is in keeping with the RMA's Framework for Risk Assessment Management

and Evaluation (FRAME); providing a more comprehensive assessment and a more individualised management plan for the more complex/severe cases. The RSVP assessments in this study were undertaken jointly by two staff and discussed with the wider team to develop consensus ratings. Practitioners should ensure that their RSVP item ratings and judgments are reliable by ensuring that staff who use the tool are appropriately experienced, knowledgeable, trained, skilled and supervised. Practitioners should use Summary Judgements in practice to guide conclusions about the different facets of risk. It is important that offenders are managed at a level commensurate with the risk posed. Low risk offenders should not be over managed. This is a waste of resources. High risk offenders require a high level of risk management; otherwise others are placed at risk. In some cases, in our study, high risk offenders could not receive a high level of intervention due to the limitations of the legal order they were subject to. For example there is limited scope for high intensity risk management where an offender is only subject to sex offender notification requirements in the community. There should perhaps be more emphasis on the use of the RSVP at sentencing where there is an apparent risk of serious harm, so that a sentence can be imposed which will support longer-term risk management. The use of the RSVP in potential OLR cases contributes towards this, but perhaps the RSVP should also be used more consistently where cases are being considered for extended sentences. The RSVP would appear to be a useful tool for assessing risk of serious harm in sexual offenders, and therefore potentially has a role in certain cases beyond the current mandatory instruments for sexual offenders (i.e. Risk matrix 2000, Stable and Acute 2007 and LSCMI). For a sub-sample of the cases in the current study we undertook a qualitative evaluation of the utility of the RSVP assessments from the perspective of front line criminal justice staff supervising cases (This is reported elsewhere; Judge et al. 2013). This indicated that front line staff felt the RSVP based assessments brought additional value to the assessment and management of their cases. This further supports the contention that the RSVP may have a role in the management of the minority of sexual abusers who pose a risk of serious harm.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders: Diagnostic criteria from DSM-IV*. American Psychiatric Association.

Andrews, D.A., Bonta, J., & Wormith, J.S. (2004). *The Level of Service/Case Management Inventory (LS/CMI)*. Toronto, Ontario: Multi-Health Systems.

Belfrage, H., Strand, S., Storey, J.E., Gibas, A.L., Kropp, P.R., & Hart, S.D. (2012). Assessment and management of risk for intimate partner violence by police officers using the Spousal Assault Risk Assessment guide. *Law and Human Behavior*, DOI 10.1007/s10979-011-9278-0.

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

Cleckley, H. (1976). The mask of sanity. St. Louis, MO: C.V. Mosby.

Cooke, D. J. & Michie, C. (1999) Psychopathy across cultures: North America and Scotland compared. *Journal of Abnormal Psychology*, *108*, 55-68.

Cooke, D. J., Michie, C., Hart, S. D., & Clark, D. (2005). Assessing psychopathy in the UK: concerns about cross-cultural generalisability. *The British Journal of Psychiatry*, *186*(4), 335-341.

Darjee, R. & Russell, K. (2011). The assessment and sentencing of high-risk offenders in Scotland: A forensic clinical perspective. In B. McSherry & P. Keyzer (Eds.), *Dangerous People: Policy, Prediction, and Practice* (pp. 217-232). Sussex: Routledge.

Darjee, R. & Russell, K. (2012). What clinicians need to know before assessing risk in sexual offenders. *Advances in Psychiatric Treatment*, *18*(6), 467-478.

de Vogel, V. (2005). Structured risk assessment of (sexual) violence in forensic clinical practice: the HCR-20 and SVR-20 in Dutch forensic psychiatric patients: academisch proefschrift. Retrieved from http://hdl.handle.net/11245/2.35006

de Vogel, V. & de Ruiter, C. (2006). Structured professional judgement of violence risk in forensic clinical practice: A prospective study into the predictive validity of the Dutch HCR-20. *Psychology, Crime & Law, 12*(3), 321-336.

Dolan, M. & Doyle, M. (2000). Violence risk prediction Clinical and actuarial measures and the role of the Psychopathy Checklist. *The British Journal of Psychiatry*, *177*(4), 303-311.

Douglas, K.S. & Belfrage, H. (2014). Interrater reliability and concurrent validity of the HCR-20 version 3. *International Journal of Forensic Mental Health*, *13*, 130-139.

Douglas, K. S. & Kropp, P. R. (2002). A prevention-based paradigm for violence risk assessment: Clinical and research applications. *Criminal Justice and Behavior*, 29, 617-658.

Douglas, K. S. & Reeves, K.A. (2010). Historical-Clinical-Risk Management-20 (HCR-20) Violence Risk Assessment Scheme. In R.K. Otto & K.S. Douglas (Eds.). *Handbook of violence risk assessment tools* (pp. 147-185). Milton Park, UK: Routledge.

Doyle, M. & Dolan, M. (2002). Violence risk assessment: combining actuarial and clinical information to structure clinical judgements for the formulation and management of risk. *Journal of Psychiatric and Mental Health Nursing*. *9*(6), 649 – 657.

Fazel, S., Singh, J.P., Doll, H., & Grann, M. (2012). Use of risk assessment instruments to predict violence and antisocial behaviour in 73 samples involving 24 827 people: systematic review and meta-analysis. *British Medical Journal*, *345*, e4692. doi: 10.1136/bmj.e4692.

Green, B., Carroll, A., & Brett, A. (2010). Structured risk assessment in community forensic mental health practice. *Australasian Psychiatry*, *18*(6), 538-541. doi:10.3109/10398562.2010.498513.

Grubin, D. (2011). A Large-Scale Evaluation of Risk Matrix 2000 in Scotland. *Sexual Abuse*, 23(4), 419-433.

Guy, L. (2009) *Performance Indicators of the SPJ Approach for Violence Risk Assessment: A Meta-Analytic Survey* (Unpublished thesis). Paper presented at the annual meeting of the American Psychology - Law Society, TBA, San Antonio.

Hanson, R.K. & Bussière, N.T. (1998). Predicting Relapse: A meta-analysis of Sexual Offender Recidivism Studies. *Journal of Consulting and Clinical Psychology*, *66*, 348-362.

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

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

Hare, R.D. (2003). *The Hare Psychopathy Checklist-Revised, 2nd Edition*. Toronto, ON: Multi-Health Systems.

Hart, S. D. (1998). The role of psychopathy in assessing risk for violence: Conceptual and methodological issues. *Legal and Criminological Psychology*, *3*, 121–137.

Hart, S. D. (2003). *Assessing risk for sexual violence: The Risk for Sexual Violence Protocol* (*RSVP*). Paper presented at the annual meeting of the International Association of Forensic Mental Health Services, Miami, Florida.

Hart, S. & Boer, D. (2010). Structured professional judgement guidelines for sexual violence risk assessment: The sexual violence risk-20 (SVR-20) and risk for sexual violence protocol. In R.K. Otto & K.S. Douglas (Eds.). *Handbook of violence risk assessment tools* (pp. 269-294). Milton Park, UK: Routledge.

Hart, S.D. & Jackson, K. (2008). *The predictive validity of the Risk for Sexual Violence Protocol (RSVP)*. Paper presented at the annual meeting for the International Association of Forensic Mental Health Services, Vienna, Austria.

Hart, S., Kropp, P.R., & Laws, D.R.; with Klaver, J., Logan, C, & Watt, K.A. (2003). *The Risk for Sexual Violence Protocol (RSVP): Structured professional guidelines for assessing risk of sexual violence*. Vancouver, BC: The Institute Against Family Violence.

Hart, S. & Logan, C. (2011). Formulation of violence risk using evidence based assessments: The structured professional judgement approach. In P. Sturmey & M. McMurran (Eds.), *Forensic case formulations* (pp.83-106). Chichester, UK: Wiley-Blackwell.

Hart, S.D., Michie, C., & Cooke, D.J. (2007). Precision of Actuarial Risk Assessment Instruments: Evaluating the "margins of error" of group v. individual predictions of violence. *British Journal of Psychiatry*. *190*(49), 60-65.

Heilbrun, K., Douglas, K. S., & Yasuhara, K. (2009). Controversies in violence risk assessment. In J. L. Skeem, K. S. Douglas, & S. O. Lilienfeld (Eds), *Psychological science in the courtroom: Controversies and consensus* (pp. 333-357). New York, NY: Guilford.
Howard, P.D., Barnett, G.D., & Mann, R.E. (2014). Specialization in and within sexual offending in England and Wales. *Sexual Abuse*, *26*(*3*), 225-251.

Judge, J., Quayle, E., O'Rourke, S., Russell, K., & Darjee, R. (2013). Referrers' views of structured professional judgement risk assessment of sexual offenders: A qualitative study. *Journal of Sexual Aggression: An international, interdisciplinary forum for research, theory and practice*, 20, 94-109.

Khiroya, R., Weaver, T., & Madden, T. (2009). Use and perceived utility of structured violence risk assessments in English medium secure units. *Psychiatric Bulletin*, *33*, 129-132.

Kropp, P. R. (2001, April). The Risk for Sexual Violence Protocol (RSVP). Paper

presented at the annual conference of the International Association of Forensic Mental Health Services, Vancouver, BC.

Kropp, P.R., Hart, S.D., & Lyon, D.R. (2008). *Guidelines for Stalking Assessment and Management (SAM) – Manual.* Toronto, Canada: Multi-Health Systems.

Kropp, P., Hart, S., Webster, C., & Eaves, D. (1995). *The Spousal Assault Risk Assessment* (*SARA*). Vancouver, British Columbia: British Columbia Institute Against Family Violence.

Landis, J. & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159-174.

Loranger, A.W. (1999). International Personality Disorder Examination: DSM-IV and ICD-10 Interviews. Odessa, FL: Psychological Assessment Resources, Inc.

Muller, M.J. & Wetzel, H. (1998). Improvement of inter-rater reliability of PANSS items and subscales by a standardized rater training. *Acta Psychiatrica Scandinavia*. *98*, 135-139.

Nitschke, J., Osterheider, M., & Mokros, A. (2009) A cumulative scale of severe sexual sadism. *Sexual Abuse: A Journal of Research and Treatment, 3*, 262–78.

Otto, R.K. & Douglas, K.S. (Eds.). *Handbook of violence risk assessment tools* (pp. 269-294). Milton Park, UK: Routledge.

Pedersen, L., Rasmussen, K., & Elsass, P. (2010). Risk Assessment: The Value of Structured Professional Judgments. *International Journal of Forensic Mental Health*, *9*, 74-8.

Reichelt, F.K, James, I.A., & Blackburn, I. (2003). Impact of training on rating competence in cognitive therapy. *Journal of Behaviour Therapy and Experimental Psychiatry*. *34*(2), 87-99.

Risk Management Authority Scotland. (2011) Framework for Risk Assessment Management and Evaluation (FRAME). Paisley, UK.

Risk Management Authority Scotland. (2013). *Standards and Guidelines for Risk Assessment*. Retrieved from http://www.rmascotland.gov.uk/standardsandguidelines/

Risk Management Authority Scotland. (2015). *Risk Assessment Tools Evaluation Directory* (*RATED*). Retrieved from http://rated.rmascotland.gov.uk

Russell, K. & Darjee, R. (2013). Practical assessment and management of risk in sexual offenders. *Advances in Psychiatric Treatment*, *19*(1), 56-66.

Scottish Government (2010). *MAPPA Edinburgh, Lothian and the borders multi agency protection arrangements*. Annual Report 2009/10 Retrieved 29 April, 2013, from http://www.scotland.gov.uk/Publications/2010/11/10112315/0

Scottish Government. (2014). Multi Agency Public Protection (MAPPA) National Guidance. Retrieved from http://www.gov.scot/Publications/2014/06/6359/7

Seto, M.C., Harris, G.T., Rice, M.E., & Barbaree, H.E. (2004). The screening scale for pedophilic interests predicts recidivism among adult sex offenders with child victims. *Archives of Sexual Behavior*, *33*(5), 455-466.

Seto, M.C. & Lalumiere, M.L. (2001). A brief screening scale to identify paedophilic interests among child molesters. *Sex Abuse*, *13*(1):15-25.

Sutherland, A.A., Johnstone, L., Davidson., K.M., Hart, S., Cooke, D.J., Kropp, P.R., Logan, C., Michie, C., & Stocks, R. (2012). Sexual violence risk assessment: An investigation of the interrater reliability of professional judgements made using the risk of sexual violence protocol. *International Journal of Forensic Mental Health*, *11*, 119-133.

Thornton, D. (2010). Scoring guide for the Risk Matrix 2000. 10/SVC. Unpublished manuscript.

Vincent, G.M., Guy, L.S., Gershenson, B.G., & McCabe, P. (2012). Does Risk Assessment Make a Difference? Results of Implementing the SAVRY in Juvenile Probation. *Behavioral Sciences and the Law*, *30*, 384–405.

Vojt, G. (2013, November). A validation study of the Risk of Sexual Violence Protocol following clinical implementation. Paper presented at the Forensic Network Research Conference: Moving Forward Together, Polmont, Scotland.

Watt, K.A., Hart, S., Wilson, C., Guy, L., & Douglas, K.S. (2006). *An evaluation of the risk of sexual violence protocol.* Paper presented at the annual Meeting of the International Association of Forensic Mental Health Services, Miami Florida.

Watt, K. A. & Jackson, K. (2008). *Inter-rater and structural reliabilities of the risk for sexual violence protocol.* Paper presented at the Annual Meeting of the International Association of Forensic Mental Health Services, Vienna, Austria.

Webster, C.D., Douglas, K.S., Eaves, D., & Hart, S.D. (1997). *HCR-20 assessing risk for violence (version 2)*. Burnaby, Canada: Mental Health, Law and Policy Institute, Simon Fraser University.

Wilson, C. (2013). *Reliability and Consistency of Risk Formulations in Assessments of Sexual Violence Risk*. Unpublished DPhil (Psychology) thesis. Simon Fraser University. Vancouver. Retrieved from http://summit.sfu.ca/item/13599

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines*. Geneva: World Health Organization.

Appendix

RSVP Items	Presence: Past	Presence:	Relevance:		
		Recent	Future		
1	1.31 (2)	.31 (0)	1.28 (2)		
2	.67 (0)	.15 (0)	.66 (0)		
3	.74 (0)	.19 (0)	.8 (1)		
4	1.06 (1)	.11 (0)	1.06 (1)		
5	.93 (1)	.35 (0)	.89 (1)		
6	1.06 (1)	.94 (1)	.79 (1)		
7	.9 (1)	.7 (1)	.83 (1)		
8	1.52 (2)	1.35 (1)	1.5 (2)		
9	1.65 (2)	1.35 (2)	1.6 (2)		
10	1.04 (1)	.96 (1)	1.02 (1)		
11	.94 (1)	.83 (1)	.94 (1)		
12	.44 (0)	.44 (0)	.47 (0)		
13	.21 (0)	.16 (0)	.19 (0)		
14	1.06 (1)	.46 (0)	1.03 (1)		
15	.83 (1)	.38 (0)	.72 (0)		
16	1.83 (2)	1.67 (2)	1.87 (2)		
17	1.75 (2)	1.57 (2)	1.64 (2)		
18	1.50 (2)	1.50 (2)	1.27 (1)		
19	1.07 (1)	.49 (0)	1.00(1)		
20	1.68 (2)	1.15 (1)	1.59 (2)		
21	1.47 (2)	1.21 (1)	1.42 (2)		
22	1.30 (2)	.83 (0)	1.35 (2)		
Other	1.76 (2)	1.72 (2)	1.74 (2)		
Considerations					
Weapon use	.39 (0)	.08 (0)	.35 (0)		
Age	.81 (0)	.42 (0)	.72 (0)		
Sexual	.95 (1)	.54 (0)	.91 (1)		
Preoccupation					
Personality	1.62(2)	1.62 (2)	1.62(2)		
disorder					

Table i. RSVP individual items: means (medians)

Note: Each item could be rated 0, 1 or 2 in a particular case.

RSVP Items	Presence: Past		Presence: Recent			Ever Present			Relevance: Future			
	No	Query	Yes	No	Query	Yes	No	Query	Yes	No	Query	Yes
1	29	17	63	87 (79.8)	10 (9.2)	12 (11)	27	17	65	27	24 (22)	58
	(26.6)	(15.6)	(57.8)				(24.8)	(15.6)	(59.6)	(24.8)		(53.2)
2	59	27	23	96 (88.1)	10 (9.2)	3 (2.8)	57	29	23	60 (55)	26	23
	(54.1)	(24.8)	(21.1)				(52.3)	(26.6)	(21.1)		(23.9)	(21.1)
3	60 (55)	17	32	93 (85.3)	11 (10.1)	5 (4.6)	52	22	35	53	25	31
		(15.6)	(29.4)				(47.7)	(20.2)	(32.1)	(48.6)	(22.9)	(28.4)
4	42	18	49 (45)	101 (92.7)	4 (3.7)	4 (3.7)	41	18	50	41	20	48
	(38.5)	(16.5)					(37.6)	(16.5)	(45.9)	(37.6)	(18.3)	(44)
5	50	17	42	87 (79.8)	6 (5.5)	16	48 (44)	19	42	51	19	39
	(45.9)	(15.6)	(38.5)			(14.7)		(17.4)	(38.5)	(46.8)	(17.4)	(35.8)
6	38	27	44	43 (39.4)	29 (26.6)	37	38	27	44	46	40	23
	(34.9)	(24.8)	(40.4)			(33.9)	(34.9)	(24.8)	(40.4)	(42.2)	(36.7)	(21.1)
7	42	36 (33)	31	50 (45.9)	42 (38.5)	17	42	36 (33)	31	46	36 (33)	27
	(38.5)		(28.4)			(15.6)	(38.5)		(28.4)	(42.2)		(24.8)
8	10 (9.2)	32	67	12 (11)	47 (43.1)	50	10 (9.2)	31	68	11	32	66
		(29.4)	(61.5)			(45.9)		(28.4)	(62.4)	(10.1)	(29.4)	(60.6)
9	9 (8.3)	20	80	18 (16.5)	35 (32.1)	56	9 (8.3)	19	81	11	22	76
		(18.3)	(73.4)			(51.4)		(17.4)	(74.3)	(10.1)	(20.2)	(69.7)
10	42	21	46	45 (41.3)	23 (21.1)	41	42	21	46	43	21	45
	(38.5)	(19.3)	(42.2)			(37.6)	(38.5)	(19.3)	(42.2)	(39.4)	(19.3)	(41.3)
11	48 (44)	20	41	52 (47.7)	23 (21.1)	34	48 (44)	20	41	48 (44)	19	42
		(18.3)	(37.6)			(31.2)		(18.3)	(37.6)		(17.4)	(38.5)
12	73 (67)	24 (22)	12 (11)	73 (67)	24 (22)	12 (11)	73 (67)	24 (22)	12 (11)	72	23	14
										(66.1)	(21.1)	(12.8)
13	91	13	5 (4.6)	97 (89)	7 (6.4)	5 (4.6)	89	15	5 (4.6)	94	9 (8.3)	6 (5.5)
	(83.5)	(11.9)					(81.7)	(13.8)		(86.2)		
14	46	10 (9.2)	53	77 (70.6)	14 (12.8)	18	45	10 (9.2)	54	49 (45)	8 (7.3)	52
	(42.2)		(48.6)			(16.5)	(41.3)		(49.5)			(47.7)

Table ii. RSVP individual items: number and percentage of cases who received each of the three ratings levels (no, query or yes) for each of the four types of rating (Past presence, Recent presence, *Ever present*, future Relevance).

15	52	24 (22)	33	80 (73.4)	17 (15.6)	12 (11)	51	25	33	58	24 (22)	27
	(47.7)		(30.3)				(46.8)	(22.9)	(30.3)	(53.2)		(24.8)
16	3 (2.8)	13	93	9 (8.3)	18 (16.5)	82	2 (1.8)	9 (8.3)	98	1 (.9)	12 (11)	96
		(11.9)	(85.3)			(75.2)			(89.9)			(88.1)
17	6 (5.5)	15	88	7 (6.4)	33 (30.3)	69	6 (5.5)	15	88	8 (7.3)	23	78
		(13.8)	(80.7)			(63.3)		(13.8)	(80.7)		(21.1)	(71.6)
18	18	18	73 (67)	15 (13.8)	24 (22)	70	8 (7.3)	13	88	15	50	44
	(16.5)	(16.5)				(64.2)		(11.9)	(80.7)	(13.8)	(45.9)	(40.4)
19	41	19	49 (45)	76 (69.7)	13 (11.9)	20	41	18	50	44	21	44
	(37.6)	(17.4)				(18.3)	(37.6)	(16.5)	(45.9)	(40.4)	(19.3)	(40.4)
20	9 (7.3)	19	82	28 (25.7)	37 (33.9)	44	8 (7.3)	19	82	9 (8.3)	27	73
		(17.4)	(75.2)			(40.4)		(17.4)	(75.2)		(24.8)	(67)
21	15	28	66	24 (22)	38 (34.9)	47	13	28	68	15	33	61
	(13.8)	(25.7)	(60.6)			(43.1)	(11.9)	(25.7)	(62.4)	(13.8)	(30.3)	(56)
22	33	10 (9.2)	66	55 (50.5)	18 (16.5)	36 (33)	27	14	68	25	21	63
	(30.3)		(60.6)				(24.8)	(12.8)	(62.4)	(22.9)	(19.3)	(57.8)
23 Other	9 (8.3)	8 (7.3)	92	9 (8.3)	13 (11.9)	87	9 (8.3)	8 (7.3)	92	9 (8.3)	10 (9.2)	90
Considerations			(84.4)			(79.8)			(84.4)			(82.6)
Weapons use	85 (78)	5 (4.6)	19	101 (92.7)	7 (6.4)	1 (.9)	85 (78)	6 (5.5)	18	84	12 (11)	13
_			(17.4)						(16.5)	(77.1)		(11.9)
Age	59	12 (11)	38	81 (74.3)	10 (9.2)	18	59	12 (11)	38	61 (56)	18	30
-	(54.1)		(34.9)			(16.5)	(54.1)		(34.9)		(16.5)	(27.5)
Sexual	45	24 (22)	40	66 (60.6)	27 (24.8)	16	45	24 (22)	40	47	25	37
Preoccupation	(41.3)		(36.7)			(14.7)	(41.3)		(36.7)	(43.1)	(22.9)	(33.9)
Personality	15	11	83	15 (13.8)	11 (10.1)	83	15	11	83	15	11	83
Disorder	(13.8)	(10.1)	(76.1)			(76.1)	(13.8)	(10.1)	(76.1)	(13.8)	(10.1)	(76.1)

Note: N = No; ? = Query; Y = Yes