# Reliability and Consistency of Risk Formulations in Assessments of Sexual Violence Risk

**by**

**Catherine M. Wilson**

M.A. (Psychology), Simon Fraser University, 2007
B.Sc., Carleton University, 2003

Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
Department of Psychology
Faculty of Arts and Social Sciences



**SIMON FRASER UNIVERSITY**

**Summer 2013**

# Approval

**Name:** **Catherine M. Wilson**

**Degree:** **Doctor of Philosophy (Psychology)**

**Title of Thesis:** ***Reliability and Consistency of Risk Formulations in Assessments of Sexual Violence Risk***

**Examining Committee:** **Chair:** Shannon Zaitsoff
Assistant Professor

**Stephen Hart**
Senior Supervisor
Professor

**Kevin Douglas**
Supervisor
Associate Professor

**Tonia Nicholls**
Supervisor
Adjunct Faculty

**Carmen Gress**
Internal Examiner
Adjunct Professor
School of Criminology

**Derek Perkins**
External Examiner
Head of Psychological Services
West London Mental Health NHS Trust
Forensic Services
London, England

**Date Defended/Approved:** 1 August 2013

## Partial Copyright Licence

## Ethics Statement


The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics,

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University;

or has conducted the research

c. as a co-investigator, collaborator or research assistant in a research project approved in advance,

or

d. as a member of a course approved in advance for minimal risk human research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Burnaby, British Columbia, Canada

update Spring 2010

# Abstract

Sexual violence represents an intrusion of personal boundaries that can be physically and psychologically traumatic for the victim. Assessing risk for sexual violence is an important process that can have serious consequences for the public (e.g., risk to public safety) and the individual assessed (e.g., indefinite commitment). Given the serious potential consequences, it is vital that assessments are conducted using empirically supported risk assessment measures. The Risk for Sexual Violence Protocol (RSVP; Hart, et al., 2003) is a measure to guide sexual violence risk assessments.

The RSVP provides a framework for case formulation, a process that gathers diverse case-specific information to guide decision-making. Formulation is an essential element of risk assessment, but has been neglected in research. The current study added to the literature base supporting the RSVP and addressed the gap in the literature concerning formulation. First, reliability of presence, relevance and summary risk judgments was examined. Second, the similarity of formulations made by different raters for the same cases was compared to that of formulations made by different raters for different cases.

Seventeen professionals completed an online risk assessment course on the administration of the RSVP and completed file-based RSVP assessments for six of ten cases. Rater agreement for presence and relevance ratings and summary judgments was poor to fair, whereas agreement for domain and total scores was fair to good. Similarity ratings (made by independent judges) for randomly selected pairs of formulations made by different raters for the same cases were significantly higher than those made by different raters for different cases. This was true both for global ratings of formulation similarity (i.e., causes of past sexual violence, scenarios of future sexual violence, recommended management strategies), as well as specific facets of formulations similarity (e.g., identification of motivating, disinhibiting, and destabilizing risk factors in past sexual violence; nature, severity of future sexual violence; monitoring, supervision, treatment, and victim safety planning tactics). The findings provide evidence that formulations of violence risk are consistent or similar across raters. Findings are discussed in the context of risk assessment practice, directions for risk assessment training, and future research.



**Keywords**: Risk for Sexual Violence Protocol, risk assessment, case formulation, interrater reliability, sexual violence

## Acknowledgements

I am extremely grateful to a number of individuals who provided assistance and support as I worked towards, and achieved, this goal.

First, I would like to thank my Senior Supervisor Dr. Stephen Hart. His mentorship and guidance throughout this project, as well as through my graduate program, was invaluable. He continuously challenged me to view issues from different perspectives, which is something I will take forward as I pursue a career in forensic psychology. On a personal level, his never-ending optimism was an important source of support as I tackled this project and other responsibilities. My committee members, Dr. Kevin Douglas and Dr. Tonia Nicholls provided ongoing support and helpful feedback. In addition, both have provided numerous professional opportunities throughout my graduate career. I am very thankful that I had the opportunity to work with them and look forward to new collaborations in the future.

There are a number of people who assisted with various aspects of this study. I am appreciative of the organizations that provided funding: Social Sciences and Humanities Research Council Canadian Graduate Doctoral Fellowship, Michael Smith Foundation for Health Research Trainee Award, American Academy of Forensic Psychology Dissertation Grant, and American Psychology-Law Society Grant-in-Aid. This project would not have been possible without the help and cooperation of Dr. Patricia Zapf. Patty provided the forum for the online training and enabled the program to proceed smoothly. I simply cannot thank her enough for everything she did to get this project up and running. Drs. Kelly Watt and Randy Kropp spent numerous hours reviewing the case files and completing assessments. Despite their incredibly busy schedules, both of them took the time to assist me and for that I am very grateful. I could not have asked for better research assistants than Yan Lim and Kate Bousfield. Without hesitation or complaint, they took on every request I made. I greatly appreciate their hard work to help meet my very tight deadlines.

I was fortunate to have support from many friends and family along the way. In particular, I would like to acknowledge Laura Guy, Kaitlyn McLachlan and Michelle Collins. These ladies have been amazing colleagues and even better friends. My

parents, Margaret and John, and sister, Heather, were a continuous source of support and love. They never stopped believing in me and always had words of encouragement when needed.

Finally, this work is dedicated to my children, Maya and Samuel, and husband, Patrick. Maya, my ray of sunshine, always provided a welcome distraction and her contagious laugh was uplifting during stressful times. Samuel, my little hero, kindly allowed me to submit my dissertation before arriving into this world. And last but not least, Patrick. There are simply no words to convey my appreciation for the support, love, encouragement and assistance he provided throughout this journey. I truly could not have done this without him. You three were, and continue to be, my inspiration.

# Table of Contents

# List of Tables

## List of Figures

# 1. Introduction

Violence risk assessment is an important aspect of many professionals' day-to-day responsibilities and is used within a number of correctional, mental health and forensic contexts. Within the correctional and mental health systems, violence risk assessments may be conducted to assist security classification, treatment, management, and release decisions. There are very serious potential consequences of conducting an improper risk assessment, such as an increased risk to public safety, or alternatively, an individual's loss of their right to liberty. In light of the potential consequences of a risk assessment, it is imperative that professionals engage in evidence-based assessments of violence risk and employ validated, reliable and empirically supported risk assessment measures. An evidence-based assessment of violence risk has been defined as the:

> [P]rocess of gathering information about people in a way that is consistent with and guided by the best available scientific and professional knowledge to (a) understand their potential for engaging in violence against others in the future and (b) determine what should be done to prevent this violence from occurring. (Hart, 2009, p. 148)

Fortunately, over the past 30 years, a wealth of knowledge has been accrued and important advances have been made with respect to risk assessment. Research has identified key risk factors, and a multitude of risk assessment guidelines and procedures have been developed and validated (Haggard-Grann, 2007; Hanson, 2000; for a review of risk assessment measures see Otto & Douglas, 2010). As a result, professionals now have a number of empirically based risk assessment tools available to assist the risk assessment process across multiple contexts and for specific types of risk or violence, including sexual violence risk.

Sexual violence is one specific type of violence that benefits from specialized risk assessment tools. Although general risk assessment measures may be appropriate to assess for risk of violent or non-violent recidivism by sex offenders, sexual recidivism is

best predicted by measures developed for this purpose (Hanson & Bussière, 1998; Hanson & Morton-Bourgon, 2009). The *Risk for Sexual Violence Protocol* (Hart et al., 2003) is a structured professional judgment (SPJ) guide developed specifically for assessing sexual violence risk. Briefly, the RSVP is similar to other SPJ tools such as the *Historical-Clinical-Risk Management-20* (HCR-20; Douglas, Hart, Webster & Belfrage, 2013; Webster, Douglas, Eaves, & Hart, 1997) and *Sexual Violence Risk-20* (SVR-20; Boer, Hart, Kropp, & Webster, 1997), in that it comprises clinically relevant and empirically supported risk factors that are coded based on their presence within an individual case and the evaluator makes a summary risk estimate. However, the RSVP contains unique elements, such as the inclusion of steps to develop case formulation, risk scenarios, and case management strategies. Case formulation is considered the integral link between assessment and management; it informs the management strategies and interventions while summarizing and communicating an individual's risk status (Doyle & Dolan, 2002; Hart et al., 2003; Heilbrun et al., 2000). The RSVP is a widely used SPJ measure to assess risk for future sexual violence (Hart & Boer, 2010); however, to date there is limited published empirical support available. Further, despite the emphasis on the importance of case formulation, it has remained relatively neglected in research.

The current study addressed some of the gaps in the literature with respect to the RSVP. First, the study examined the interrater reliability of item ratings and summary risk estimates. Second, the study investigated the quality of the risk assessments by comparison with a gold standard assessment. Finally, research has not yet examined whether case formulations can be developed consistently and reliably across evaluators for the same case. As this issue has not yet been investigated, the current study examined the reliability of case formulation by comparing formulations developed by different raters for the same case with formulations developed by different raters for different cases.

In the remainder of this chapter, the definition and problem of sexual violence will be discussed, followed by an overview of the different approaches to violence risk assessment, specifically focusing on the actuarial and structured professional judgment approach. The RSVP, the focus of the current study, will then be introduced, and an overview of case formulation will be provided. Next, research supporting the RSVP will

be detailed. Finally, the current study will be presented, including the research questions and corresponding hypotheses for each.

## 1.1. Sexual Violence

Sexual violence can be defined as “the actual, attempted or threatened sexual contact with another person that is non-consensual” (Hart et al., 2003, p. 2). In Canada, more than 21,800 sexual assaults were reported to police in 2011 (Statistics Canada, 2013); however, this statistic is likely an underestimate of the true rate of sexual offences. Self-report victimization data from the General Social Survey (GSS) indicate that in 2009, there were 677,000 self-reported sexual assaults, corresponding to a rate of 24 per 1000 population aged 15 and above (Perreault & Brennan, 2010). Of these self-reported offences, most (88%) were not reported to police (Perreault & Brennan, 2010). With respect to the severity of the sexual offenses, according to both police reported and self-reported statistics, the majority are categorized as ‘Level 1’, the least severe category (e.g., sexual touching, unwanted grabbing, kissing or fondling). However, according to GSS statistics, approximately 1 in 5 sexual assault incidents were sexual attacks that involved the use of threats or physical violence. Regardless of the specific severity level assessed by public safety authorities, sexual offenses are an intrusion of personal boundaries and can be traumatic for the victim.

### *1.1.1. Sexual violence legislation*

Given the nature of this type of offense, it often attracts significant media attention. As a result of horrific, well-publicized sexual violence cases and the public reaction to such offenses, regulatory controls have been enacted to deal with sex offenders. Examples of such controls include sex offender registries, community notification, and residence restrictions (Janus, 2003). Further, a number of countries or jurisdictions have enacted legislation that allows for the civil commitment or indeterminate institutionalization of individuals convicted of sexually violent offenses and deemed high risk (Larcombe, 2012). In certain Australian jurisdictions *preventative detention* legislation has been introduced, allowing for the continuing detention or community supervision of serious sexual offenders who are considered dangerous

(Doyle & Ogloff, 2009). Similarly, within the United Kingdom, offenders who meet the legal criteria for dangerousness can receive an extended or indefinite sentence (Bickle, 2008). Certain states within the United States have enacted Sexually Violent Predator (SVP) legislation. Under SVP laws, offenders that present a risk of future sexual violent offending can be civilly committed upon completing their criminal sentence (First & Hanlon, 2008; Janus, 2003; Levenson, 2004). Although the statutory language varies between states, SVP legislation contains four general criteria: (1) history of criminal sexual behaviour, (2) presence of a mental abnormality, (3) presence of volitional impairment, and (4) a high likelihood of future sexually violent behaviour as a result of the mental abnormality (Miller, Amenta, & Conroy, 2005).

In Canada, Dangerous Offender (DO) and Long-term Offender (LTO) legislation (*Criminal Code*, 1985) is available for dealing with offenders who commit serious violent offences, including sexual offences. Under this legislation, the Crown may petition the court to remand an offender for a DO or LTO evaluation after he (or she) has been convicted of an offence but prior to sentencing. To be declared a DO, the index offense must be a "serious personal injury offense" and the offender must constitute a "threat to life, safety, or physical or mental well-being of other persons on the basis of evidence establishing":

> (i) a pattern of repetitive behaviour by the offender, of which the offense for which he or she has been convicted forms a part, showing a failure to restrain his or her behaviour and a likelihood of causing death or injury to other persons, or inflicting severe psychological damage on other persons, through failure in the future to restrain his or her behaviour,
>
> (ii) a pattern of persistent aggressive behaviour, by the offender of which the offense for which he or she has been convicted forms a part, showing a substantial degree of indifference on the part of the offender respecting the reasonably foreseeable consequences to other persons of his or her behaviour, or
>
> (iii) any behaviour by the offender, associated with the offence for which he or she has been convicted, that is of such a brutal nature as to compel the conclusion that the offender's behaviour in the future is unlikely to be inhibited by normal standards of behavioural restraint.
>
> (*Criminal Code*, Section 753(1)(a) (i) – (iii))

Further, the act states that an offender can be declared a DO if he or she has committed a sexual offense and "has shown a failure to control his or her sexual

impulses and a likelihood of causing injury, pain or other evil to other persons through failure in the future to control his or her sexual impulses" (*Criminal Code*, 1985, Section 753(1)(b)). Under DO legislation, an offender can be sentenced to incarceration for an indeterminate period of time, with National Parole Board reviews being conducted seven years following the date the offender was taken into custody for the index offense and every two years thereafter, for offenders sentenced after August 1, 1997 (Solicitor General of Canada, 2001).

Long-term Offender legislation was enacted to ensure that sexual offenders are provided with long-term supervision when released to the community following incarceration. The criteria for declaring an individual as a LTO state it must be "appropriate to impose a sentence of imprisonment for two years or more" for the index offense, there must be "a substantial risk that the offender will reoffend" and there must be a "reasonable possibility of eventual control of the risk in the community" (*Criminal Code*, Section 753.1(1)(a) – (c)). If an offender is determined to be a LTO, the court will impose a sentence for the index offense, which must be a minimum imprisonment for two years, and order that the offender is subject to long-term supervision for a period that does not exceed 10 years (Section 753.1(3)).

In light of the seriousness of these decisions across numerous jurisdictions, accurate assessment of sexual violence risk is essential. Further, it is important that risk assessment informs the criteria for the specific legal issue at stake. For example, as discussed, the criteria for SVP legislation include a history of sexual violence, presence of mental abnormality, high risk for future sexual offending and a causal nexus between mental abnormality and risk. Not only should the assessment establish that the offender has a history of sexual violence and make a decision about the level of risk the person poses, but in this case it must address the presence and influence of mental abnormality (Hart, 2003; Vincent, Maney, & Hart, 2009). In Canada, an important distinction in determining the applicability of DO compared to LTO designation is whether an individual can be effectively managed in the community. To answer this question, a risk assessment can provide estimates of the risk posed by the person under a number of conditions, such as complete confinement compared to living in the community under various levels of restrictions (Hart, 2003). Risk assessment might also address issues such as the type(s) of violence an individual is at risk of perpetrating, likely or potential

victims, the degree of harm likely to be inflicted and imminence of violent acts, all important for responding to various legal questions.

### *1.1.2. Sexually Violent Recidivism*

Risk assessment also plays an integral role when an offender is being released to the community. Given the primary goal of preventing or minimizing the likelihood of reoffending, risk assessments are undertaken and individualized risk management strategies are implemented to reduce the risk of recidivism. "True" recidivism rates for sexual offending are difficult to determine. Issues such as the operational definition of recidivism (e.g., criminal charges, arrests, convictions), length of the follow-up period, the possibility of plea-bargaining (i.e., sexual offences being recorded as non-sexual convictions), type of sexual offender (e.g., different recidivism rates for rapists vs. child molesters, etc.) and decisions of victims not to report can lead to varying recidivism rates (Vess & Skelton, 2010). In an early review, Furby and colleagues (1989) reported that detected recidivism rates varied from 0 to 88% across 42 studies, with the majority being under 30%. Hanson and Bussière's (1998) meta-analysis on sexual offender recidivism studies found a sexual offending recidivism rate of 13% (*N* = 28,972, 61 studies) over a 5-year period. Hanson and Morton-Bourgon (2004) provided an update to Hanson and Bussière's meta-analysis and found a similar observed sexual recidivism rate of 14% (*N* = 20,444, 84 studies) over an average follow-up time of 5 to 6 years. This prevalence rate is also consistent with Hanson and Morton-Bourgon's (2009) most recent meta-analysis, which found an observed sexual recidivism rate of 11.5% (*N* = 28,757, 100 samples) and sexual or violent recidivism rates of 19.5% (*N* = 17,421, 50 samples) over an average follow-up period of 70 months.

Research indicates that sexual recidivism rates vary as a function of offender subtype. In their meta-analysis, Hanson and Bussière (1998) examined sex offense recidivism rates by offender type and found an average rate of 19% for rapists (*N* = 1,839) and 13% for child molesters (*N* = 9,603) over an average follow-up period of 4 to 5 years. Bartosh and colleagues (2003) used more narrow categories of offenders: rapists, incest offenders, extrafamilial child molesters, and hands-off offenders. Over a follow-up period of 60 to 66 months, the mean sexual offence recidivism rate was 12%. Base rates for each type of offender were as follows: 5% for rapists, 11% for incest

offenders, 14% for extrafamilial child molesters, and 35% for hands-off offenders. Vess and Skelton (2010) used yet another classification scheme. In their study, offenders were categorized as 'Only Adult Victims,' 'Only Child Victims,' and 'Mixed' (both adult and child victims), based on offense histories. Results demonstrated that 11.2% of Child Only Victim offenders sexually reoffended, 13.5% of Adult Only Victim offenders sexually reoffended and 17.7% of Mixed offenders sexually reoffended. Important to remember for all recidivism rates for sexual offences discussed here, these should be considered an underestimate of true recidivism rates as not all offenses are detected or reported.

Taken together, these findings indicate that sexual violence is a serious problem. Whether someone who has committed sexual offences is being considered for indefinite commitment, or will be released and supervised in the community, risk assessment plays a major role in protecting the rights of the individual as well as public safety. Prior to a focused discussion on sexual violence risk assessment, the primary approaches to violence risk assessment will be reviewed.

## 1.2. Approaches to Risk Assessment

As a SPJ guide, the RSVP represents one approach to risk assessment, but there are several approaches available to an evaluator. There continues to be an ongoing debate regarding the most appropriate method for conducting a risk assessment, including sexual violence risk assessments. Currently, the debate mainly focuses on two approaches, which have been referred to as discretionary and non-discretionary (Hart & Logan, 2011). Briefly, the discretionary approach involves an evaluator making decisions regarding the type of information to gather and consider, deciding how best to weigh that information and then combining the information to make a final risk estimate (Grove & Meehl, 1996). The non-discretionary approach involves the evaluator coming to a final decision based on a set algorithm (Meehl, 1996). Although there are a number of approaches that fall within these broad categories, this section will focus on the actuarial (non-discretionary) and clinical judgment (unstructured and structured; discretionary) approaches.

The actuarial approach generally involves the use of fixed and explicit rules to estimate an individual's level of risk. Measures developed using this approach are composed mostly of static risk factors, such as historical or demographic factors (Dvoskin & Heilbrun, 2001; Vincent, Maney, & Hart, 2009) and are designed to predict future behaviour (Hart & Logan, 2011; Vincent et al., 2009). Under this approach, weighted items are combined according to a specific set of rules, and summed to give a final total risk score that is used to estimate the risk of future behaviour; the human judge is eliminated (Hanson, 1998; Litwack, 2001).

Research has demonstrated that several actuarial instruments have good empirical support (Hanson & Morton-Bourgon, 2009; see Otto & Douglas, 2010). Also, actuarial tools have the advantage of requiring minimal resources for completion, often requiring only file information. However, a number of limitations to this approach have been cited. First, actuarial measures have limited ability to assist in planning treatment and identifying management strategies as they rely mostly on fixed (or static) factors that do not reflect changeable circumstances (Janus, 2003; Vincent, et al., 2009). Second, actuarial instruments require substantial time to develop and validate and some have argued that certain instruments may not be relevant because they were developed with samples from different countries or with a different population than that of the individual being assessed (Janus, 2003). Finally, most actuarial measures yield ostensibly precise likelihood estimates but do not provide the information necessary to understand the error inherent in these estimates. Consequently, it is easy to accord too much weight to information concerning estimated likelihood of recidivism provided by these instruments (Hart, Michie, & Cooke, 2007). Despite these limitations, actuarial instruments remain extremely popular and are widely used in risk assessments for sexual violence (Archer, Buffington-Vollum, Stredny, & Handel, 2006). For example, the Static-99 (Hanson & Thornton, 1999) is reported to be the most commonly used sexual violence risk assessment measure in Canada and the United States (Archer et al., 2006) and use of the Static-99 is mandated in specific legal proceedings in certain states in the U.S. (e.g., Code of Virginia; Florida Administrative Register & Florida Administrative Code). Other commonly used sexual violence risk assessment tools that use the actuarial approach include the Minnesota Sex Offender Screening Tool-Revised (MnSost-R; Epperson,

Kaul, Huot, Hesselton, & Alexander, 2000), Rapid Risk Assessment of Sexual Offence (RRASOR; Hanson, 1997) and Risk Matrix 2000-Sex (RMS; Thornton et al., 2003).

In contrast to the actuarial approach is the clinical or professional judgment approach. Unstructured professional judgment or unstructured clinical judgment relies on an evaluator's clinical experience, with no constraints or guidelines for the evaluator when making a decision (Douglas & Kropp, 2002). This approach does not impose any type of structure on the assessment, clinical judgment or decision. The assessment is based exclusively on the knowledge, experience and intuition of the assessor (Hanson, 1998) and decisions are justified based on the qualifications and experience of the assessor (Douglas & Kropp, 2002). Unstructured judgment has the advantages of being flexible, imposing minimal cost and being person-focused (Hart, 2008). However, a number of disadvantages have been identified that have led to this approach falling out of favour. Generally, there is a lack of scientific evidence supporting unstructured judgment (Hart & Logan, 2011). This approach has been criticized for lacking validity, in particular predictive validity (Lidz, Mulvey, & Gardner, 1993; McNiel & Binder, 1995; Monahan, 1984), reliability, and accountability (Harris, Rice, & Cormier, 2002; Quinsey, Harris, Rice, & Cormier, 1998). Unstructured assessments are broad in scope and the intuitive decisions are unimpeachable; that is, the assessor has difficulty explaining how the decision was made (Hart, 2008). As Hart and Logan (2011) state, unstructured professional judgment cannot be considered an evidence-based approach.

In contrast with the unstructured approach, the SPJ approach provides evaluators with a list of empirically supported and clinically relevant risk factors to assess. In addition, the evaluator may take into account any unusual or compelling evidence that may affect the individual's risk level (Hanson, 1998). Key to the SPJ approach is that structure is imposed on the risk factors to consider and the definition of these risk factors, but the combining or weighting of risk factors is left to the evaluator. This same approach and structure is applied to the final risk estimate as well. Rather than the evaluator summing or calculating the risk items according to a predetermined formula, as is done in the actuarial approach, the evaluator is encouraged to use his or her own judgment to come to a conclusion regarding the individual's level of risk (Heilbrun, Yasuhara, & Shah, 2010). Under the SPJ approach, it is not assumed that risk level is directly related to the number of risk factors present (although, it is generally true

that an increase in the number risk factors corresponds with an increase in risk level). Instead, it is recognized that it is possible that one or two risk factors can impact an individual's risk level. The final risk estimate reflects the best available research for the behaviour being assessed, as well as the evaluator's discretion to provide an opinion regarding the severity of risk and degree of effort or intervention required to manage this risk (Hart, 2008).

An important limitation of the SPJ approach is with respect to the resources required for completion. It requires considerable time and relies on interviews, collateral interviews, records and other sources of information (Heilbrun et al., 2010). Another limitation is the extensive training and expertise specific to the area of assessment (e.g., sexual deviance, sexual violence) that is required by professionals using SPJ instruments. However, the approach has many important advantages: it aims to improve the consistency and usefulness of decisions, and to improve transparency in the decision making process (Hart, 2008). In addition, SPJ measures go beyond defining risk in terms of probability of arrest for future sexual violence; they are useful for treatment and management planning as well. The SVR-20 is the most commonly used SPJ tool for assessing sexual violence risk assessment (Archer et al., 2006). The RSVP is similar in content to the SVR-20 and can be considered a parallel form of the guide; however, the RSVP has a greater emphasis on psychological risk factors and the development of case formulation and case management plans (Hart & Boer, 2010).

## 1.3. The Risk for Sexual Violence Protocol

The RSVP was developed to assist users to evaluate the risk individuals pose for perpetrating sexual violence and to determine the management strategies that would be most effective to reduce the risk and prevent sexual violence (Hart & Boer, 2010). The RSVP was developed based on literature focused on men aged 18 and older and this is the primary group for which the measure can be used; however, the RSVP may also be used with older male adolescents (16 or 17 years old) and adult women (Hart et al., 2003). For all target populations, the individual should have a known or suspected history of sexual violence. The RSVP can be used within a wide number of contexts, including pretrial and presentence assessments, correctional intake and discharge,

indeterminate commitment, child protection and family law and duty to warn or protect (Hart et al., 2003). The RSVP was based on the authors' experiences developing other forms of SPJ violence risk assessment measures (e.g., HCR-20, Webster et al., 1997; Spousal Assault Risk Assessment [SARA], Kropp, Hart, Webster, & Eaves, 1994), and in particular the SVR-20 (Boer, et al., 1997). That is, consistent with the measures developed prior to the RSVP, the RSVP was developed, in part, on a systematic review of the literature while taking existing standards of practice into consideration and can be considered a management-focused risk assessment instrument (Hart & Boer, 2010).

The RSVP manual states that clinicians should meet three qualifications to use the tool. First, users should have knowledge of sexual violence. That is, users should be familiar with the "professional and scientific literatures on the nature, causes, and management of sexual violence" (Hart et al., 2003, p. 21). Users should also have experience and training in individual assessment, including conducting clinical interviews with offender populations, completing collateral interviews and reviewing case histories. Finally, the user should have expertise in the assessment and diagnosis of mental disorder. If the evaluator does not meet this third requirement, the RSVP can still be used if the evaluator assesses the risk factors related to mental disorder: (1) under the supervision of/in consultation with someone with the requisite expertise, (2) by referring to the results of psychodiagnostic evaluations completed by those with the required expertise, (3) provisionally, with appropriate documentation, and discussing the need to have the provisional assessment confirmed by an expert. Alternatively, the evaluator can complete the RSVP but omit the items related to mental disorder; in this scenario, the evaluator should provide detailed documentation regarding the way(s) in which the incomplete assessment limited their final opinions (Hart et al., 2003).

In terms of the evaluation process, the RSVP guides evaluators through six stages. In the first step, evaluators gather case information. Hart and colleagues (2003) provide eight principles to assist evaluators in this step of collecting data: (1) gather case information regarding the individual's history of sexual violence, (2) gather information using multiple methods (e.g., interviews, observations, case files, physiological assessments), (3) gather case information from multiple sources (e.g., perpetrator, law enforcement, mental health professionals, family/friends/co-workers), (4) gather information that covers multiple domains of the individual's functioning (e.g., sexual,

interpersonal, intrapersonal, social), (5) gather information regarding static and dynamic aspects of risk, (6) update information concerning risk factors, (7) document case information, and (8) evaluate the adequacy of case information.

In the second step, evaluators assess the presence of risk factors (Table 1.1). Of note with the RSVP, the presence of risk factors is assessed for two time periods: more than one year prior to the evaluation (Past) and within the year prior to the evaluation (Recent). Presence is coded for the two time periods, for the 22 individual risk factors, as well as any case specific risk factors, and is coded on a 3-point scale (No, Possibly/Partially or Yes). The third step of the evaluation involves the assessment of the relevance of risk factors. For each of the 22 risk factors and case specific factors coded in the previous step, evaluators determine its relevance with respect to future risk and management strategies. According to the RSVP manual, one criterion of relevance is causality and a risk factor should be considered relevant if "in the evaluator's opinion, is likely to be functionally (i.e., causally) related to the individual's perpetration of sexual violence in the future" (Hart et al., 2003, p. 26). A risk factor should also be considered relevant, even if not causally related to sexual violence, if it is likely to impair the effectiveness of risk management strategies implemented to prevent future sexual violence. Similar to presence ratings, relevance is coded on a 3-point scale (No, Possibly/Partially or Yes).

The fourth and fifth steps of the evaluation process involve developing case formulation, plausible scenarios and management strategies and will be discussed in more detail below. Skipping ahead to the final step in the RSVP assessment procedure, consistent with other SPJ measures, evaluators communicate the summary risk judgments. The RSVP was the first manual to include summary judgments beyond an overall *Low*/*Moderate*/*High* decision.[1] When completing the RSVP, the evaluator is asked to communicate not only the overall level of risk for the case, including the level of effort that will be required to manage the case (Case Prioritization; High/Urgent, Moderate/Elevated, Low/Routine), but also the risk of serious physical harm (i.e., the

[1] The RSVP was the first to include this summary judgment approach; however, this approach has been incorporated into the Stalking Assessment and Management risk assessment guide (Kropp, Hart, & Lyon, 2008) and the HCR-20 $V^3$ (Douglas, et al., 2013) violence risk assessment measures.

severity of sexual violence in which the person might engage; High, Moderate, Low), need for immediate action (i.e., imminence of sexual violence; Yes, Possibly, No), and other risks indicated (i.e., whether the individual poses a risk for additional violent or general criminal behaviour; Yes, Possibly, No). RSVP evaluators also offer a case review timeline (i.e., date for case review and warning signs that would trigger an earlier assessment) (Table 1.2).

***Table 1.1 Risk Items, Definitions and Domains of the Risk for Sexual Violence Protocol***

| Domain and Risk Item | | Definition |
|---|---|---|
| **A. Sexual Violence History** | | |
| 1. | Chronicity of Sexual Violence | Persistence and frequency of sexual violence |
| 2. | Diversity of Sexual Violence | Diversity in the nature of offending (e.g., offense type and victim characteristics) |
| 3. | Escalation of Sexual Violence | Pattern of escalation in offending becoming progressively more frequent, serious or diverse over time. |
| 4. | Physical Coercion in Sexual Violence | Actual, attempted, or threatened physical harm that arises in the course of sexual violence or that is intended to further the commission of sexual violence. |
| 5. | Psychological Coercion in Sexual Violence | Acts committed in the course of sexual violence or to further sexual violence that involve either threatened loss or promised gain of status, privilege, favor, or affection. |
| **B. Psychological Adjustment** | | |
| 6. | Extreme Minimization or Denial of Sexual Violence | Denial of personal responsibility for past sexual violence or denial of serious consequences of past sexual violence. |
| 7. | Attitudes that Support or Condone Sexual Violence | Socio-political, religious, cultural or sub-cultural and personal beliefs and values that directly or indirectly encourage or excuse sexual violence. |
| 8. | Problems with Self-Awareness | Lack of self-appraisal of one's own general mental processes and reactions, particularly as they relate to one's history and risk of sexual violence. |
| 9. | Problems with Stress or Coping | Unstable psychosocial adjustment and susceptibility to external stressors. |
| 10. | Problems Resulting from Childhood | Serious problems in psychosocial adjustment that are the result of abuse experiences in childhood or adolescence and are associated with increased risk of sexual violence. |

| Domain and Risk Item | Definition |
|---|---|
| **C. Mental Disorder** | |
| 11. Sexual Deviance | Stable pattern of deviant sexual arousal. |
| 12. Psychopathic Personality Disorder | As defined and assessed by the Psychopathy Checklist-Revised (Hare, 1991, 2003), Screening Version of the PCL-R (Hart et al., 1995) or ICD-10 (Dissocial Personality Disorder) |
| 13. Major Mental Illness | Substantial impairment of cognition, affect or behaviour. |
| 14. Problems with Substance Use | Use of legal and illegal substances that impair psychosocial functioning. |
| 15. Violent or Suicidal Ideation | Thoughts, impulses, and fantasies about causing, as well as intent or attempts to cause, physical harm to self or others. |
| **D. Social Adjustment** | |
| 16. Problems with Intimate Relationships | Failure to establish or maintain stable romantic or sexual relationships with age-appropriate partners. |
| 17. Problems with Non-Intimate Relationships | Failure to establish or maintain a positive social network. |
| 18. Problems with Employment | Failure to establish and maintain stable employment. |
| 19. Non-Sexual Criminality | Serious criminal conduct that is not sexual in nature. |
| **E. Manageability** | |
| 20. Problems with Planning | Problems making and implementing prosocial life plans. |
| 21. Problems with Treatment | Failure to benefit from rehabilitative services designed to address deficits in psychosocial adjustment. |
| 22. Problems with Supervision | Failure to benefit from services designed to make it (more) difficult for the person to engage in further sexual violence. |

*Note.* Items and definitions were obtained from the Risk for Sexual Violence Protocol (RSVP) user's manual (Hart et al., 2003).

As can be seen from the description above, the RSVP has a number of similarities with other SPJ tools (e.g., HCR-20, SVR-20). However, there are a few unique aspects to this tool. In particular, the authors have integrated and emphasized case formulation and the development of case management plans as part of the RSVP assessment procedure. Case formulation within the RSVP occurs after assessing the relevance of each individual risk factor but prior to making the overall risk estimates. Before discussing the case formulation approach specific to the RSVP, it is important to understand the concept of case formulation and available approaches. As such, an overview and background of case formulation will be provided next, after which we will return to discuss case formulation as it specifically relates to completing an RSVP

assessment. First, a general overview and background of case formulation will be provided. Next, case formulation in forensic psychology will be discussed, including four major approaches used in forensic psychology. Finally, we will return to discuss case formulation as it specifically relates to completing an RSVP assessment.

***Table 1.2 Summary Judgments and Definitions of the Risk for Sexual Violence Protocol***

| | Summary Judgment | Definition |
|---|---|---|
| 1. | Case Prioritization | The level of effort or intervention that will be required to prevent the person from committing acts of sexual violence. |
| 2. | Serious Physical Harm | The risk that any future sexual violence will involve serious or life-threatening physical harm. |
| 3. | Immediate Action Required | Does the person pose an imminent risk and is immediate action required? |
| 4. | Other Risks Indicated | Evidence that the person poses a substantive risk of general (i.e., non-sexual) violence or criminality. |
| 5. | Case Review | The timeframe for scheduling a routine review (re-assessment). |

*Note.* Summary judgments categories and definitions were obtained from the Risk for Sexual Violence Protocol (RSVP) user's manual (Hart et al., 2003).

## 1.4. Case Formulation

In general, case formulation refers to a clinical skill that uses the primary case information to determine the most effective treatment strategies for each individual (Haynes & O'Brien, 2000). The skills required to conduct case formulation are considered among the most important skills acquired and refined throughout one's career (Fernando, Cohen, & Henskens, 2012) and a number of mental health organizations have included case formulation as a core competency for professionals (e.g., American Psychological Association Task Force on the Assessment of Competence in Professional Psychology, 2006; British Psychological Society Professional Practice Board, 2008). The aim of formulation is to explain an individual's problems and symptoms (Eels & Lombart, 2011) and has been applied to a wide variety of mental health problems (Sturmey, 2010). Such clinical case formulation involves the integration of relevant information and judgments about an individual's problems, the causal variables that influence these problems, and additional variables that can impact the treatment results (Haynes & Williams, 2003).

In the mental health field, case formulation serves multiple functions and is an important process in many ways. Case formulation can assist in organizing the vast amount of information about a person and his or her problems (Eells, 2007). Case formulation helps to move beyond describing symptoms, to developing an explanation that fills the gap between diagnosis and treatment (Eells & Lombart, 2011). Given that on its own, diagnosis does not provide sufficient information to guide treatment decisions, formulation guides treatment planning (Eells, 2007; Eells & Lombart, 2011) and can assist in tailoring treatment to meet the needs of a specific individual (Eells & Lombart, 2011). Overall, case formulation can assist the evaluator in deciding the most important problem behaviour(s) that should be targeted for treatment, which variables have the greatest impact for those problems and treatment, what additional information would be useful to make a treatment decision and which treatments most closely match the characteristics of the client (Haynes & O'Brien, 2000).

With respect to the field of forensic psychology, Sturmey (2010) suggests that case formulation serves two specific functions: to summarize the development of offending and other related behaviours (Gresswell & Hollin, 1992) and (potentially) to assist in determining the most effective treatment strategies for a particular client (Haynes & O'Brien, 2000). Together, formulation of assessment information can assist in understanding the complex interactions between different factors and in identifying possible causal mechanisms and can be used to select treatment and rehabilitation programs expected to reduce the likelihood of offending (Hart & Logan, 2011; Sturmey, 2010).

Formulation should be theory-driven (Hart, Sturmey, Logan & McMurran, 2011). Theory provides guidance for determining which factors, behaviours, and explanations are legitimate (Hart et al., 2011). As such, theories of violence are integral for violence risk assessment formulation. Reflecting the multiple theories of violence available there may be several formulation approaches possible; however, four approaches that have been discussed by others (see Hart & Logan, 2011) will be introduced here: Offence Paralleling Behaviour (OPB), Good Lives Model (GLM), Risk-Needs-Responsivity (RNR), and the SPJ approach, the approach adopted by the RSVP.

### *1.4.1. Offence Paralleling Behavior (OPB)*

Daffern and colleagues (2007) have defined Offence Paralleling Behaviour (OPB) as

> A behavioural sequence incorporating overt behaviours (that may be muted by environmental factors), appraisals, expectations, beliefs, affects, goals, and behavioural scripts, all of which may be influenced by the patient's mental disorder, that is functionally similar to behavioural sequences involved in previous criminal acts. (p. 267)

That is, OPB parallels a sequence of behaviour that resulted in, or may result in, similar criminal behaviour. According to Jones (1997), OPB essentially acts as a surrogate for offending behaviour. As Daffern and colleagues (2007) explain, the OPB needs to have functional equivalence to the chain of offending behaviour. The OPB approach relies on a systematic analysis of past violence. However, this approach has been refined as it forces evaluators to consider the psychological functions of violence (i.e., in what ways is violence potentially rewarding for the individual?) and the evaluator looks for evidence of behaviour that parallels past violent offences.

Hart and Logan (2011) state the positives of this approach include its clinical utility and provision of guidance for identifying risk factors and developing case management plans. The OBP approach is limited in that it does not provide a list of theory-derived principles or empirically based risk factors on which to focus, nor does it facilitate thinking about the causal roles of risk factors and it also does not guide the evaluator to speculate about future violence in light of the risk factors identified (Hart & Logan, 2011).

### *1.4.2. Good Lives Model (GLM)*

Briefly, Ward and Brown (2004) state that the GLM

> (1) Is a positive strength-based approach to treating offenders, (2) conceptualizes dynamic risk factors as distortions in the internal and external conditions required for the acquisition of human goods, (3) outlines the basic skills and capacities necessary to engage in treatment (i.e., treatment readiness), and (4) explicitly addresses the question of clinicians' attitudes toward offenders and the relationship between these attitudes and factors such as forgiveness, evil, and the therapeutic alliance. (p. 244)

The framework of the GLM views violence as a problematic means of trying to obtain primary goods (i.e., actions, states of affairs, characteristics, experiences, and states of mind that are intrinsically beneficial to human beings) and problems occur when the strategies cannot obtain the primary goods. In the case of criminal behaviour, problems arise from the means used to secure goods, and a lack of scope within a good lives plan, or the presence of conflict among goals (Ward & Stewart, 2003). For example, an offender who is seeking intimacy and mastery (i.e., primary goods) engages in a sexual relationship with a child (Ward & Stewart, 2003). The overall goal of GLM is not only to reduce the risk an offender presents, but also to assist the offender to live a better kind of life, thereby reducing the likelihood of committing future offenses (Ward & Stewart, 2003). Treatment planning using this model considers the offender's estimated level of risk, identifies dynamic risk factors that are directly linked to antisocial behaviour (i.e., criminogenic needs), and identifies the primary goods related to the offense process (Ward & Stewart, 2003).

Hart and Logan (2011) suggest the strengths of GLM are the focus on strengths and resources in the development of case management plans, and that it is well articulated. However, the limitations of this approach include the failure to provide guidance concerning risk factors most related to violence, the lack of structure for considering the causal roles of risk factors, and the lack of procedure for speculating about future violence or developing management plans in light of the risk factors identified (Hart & Logan, 2011).

### *1.4.3. Risk-Needs-Responsivity (RNR)*

The RNR (Andrews, Bonta, & Hoge, 1990) approach is based on three core principles (risk, need, responsivity) for the purposes of effective treatment.[2] The risk principle states that the level of services delivered to offenders should match the level of risk an offender poses to re-offend. That is, individuals assessed as high risk for reoffending should receive more intensive management and extensive services, whereas low risk offenders are best served with minimal or no monitoring or intervention. The needs principle focuses on criminogenic needs, or dynamic risk factors that are directly linked to criminal behaviour that, when changed, are associated with changes in risk of recidivism. According to the needs principle, effective treatment must focus on criminogenic needs in order to reduce risk of offending. Finally, the responsivity principle suggests that management strategies and treatment programs should be delivered in a style or mode that is consistent with the ability and learning style of the offender. For example, the delivery of services should consider issues such as learning styles, motivation and mental health, which may impact treatment response (Andrews & Bonta, 2003).

Andrews and Bonta have developed various tools to assist in the assessment and management of risk for general reoffending following the RNR approach (Level of Service Inventory-Revised, Andrews & Bonta, 1995; Level of Service/Case Management Inventory (LS/CMI), Andrews, Bonta, & Wormith, 2004). Broadly, an evaluator uses the measure to conduct a structured assessment. Based on the findings, relevant case management strategies can be recommended and implemented. The RNR approach has been extensively researched, with results supporting the principles of the model (Andrews & Bonta, 2006). In addition, the RNR approach provides a clear structure for identifying risk factors and thinking about the causal roles (Hart & Logan, 2011). However, Hart and Logan (2011) have noted certain limitations of the RNR model when used within the narrower context of violence risk assessment. First, this approach may be limited in that it is based on a theory of general criminality and not violence, and as a result some risk factors associated with violence may be undervalued. Second, the RNR

[2] A fourth principle focuses on the role of the professional to exercise discretion in making decisions (Andrews & Bonta, 2003).

approach has relied heavily on research from adult offenders, which might reduce its generalizability to other, more diverse populations. This approach tends to rely on the assumption that violence will resemble past violence and does not guide evaluators to speculate about other future violent scenarios (Hart & Logan, 2011). Finally, Ward and Brown (2004) criticize the RNR model for its lack of focus on a positive, strength-based approach to treatment (however, it should be noted that recent measures developed by Andrews and colleagues include a consideration of positive or strength factors, e.g., LS/CMI, Andrews, et al., 2004; Youth Level of Service/Case Management Inventory, Hoge & Andrews, 2006).

### *1.4.4. Structured Professional Judgment*

This approach to case formulation developed from the application of SPJ guidelines (Hart & Logan, 2011) and is incorporated into the RSVP risk assessment. As such, a more thorough description of this case formulation approach is provided. Under this approach, formulation is structured in two ways: (1) analyzing past violence using a decision theory framework, and (2) speculating about future violence by developing case scenarios and related management plans (Hart, 2010; Hart & Logan, 2011; for a full review of this approach and a case example, see Hart & Logan, 2011).

The decision theory framework can be viewed as the Psychology of Criminal Conduct theory adapted specifically to violence (Hart & Logan, 2011). Using this theory, violence is viewed as a choice or purposeful behaviour intended to achieve one or more goals. Specifically, a person who engages in violence has made the choice to commit a violent act, the victim of the act, the type of violence to perpetrate and the timing of the violence. As such, the task of risk assessment is to determine the causal roles of risk factors regarding the decision to engage in violence (Hart, 2010). Drawing from this, the key question becomes why the person decided to commit this violence at this time and with this victim? Case formulation reflects the decision-making process in determining the relevance of risk factors, the types of violence an individual will commit, and how to prevent such an act from occurring (Hart, 2012).

The decision theory framework suggests that prior to engaging in violence, an individual has completed a four-step process that concludes with engaging in the violent

behaviour (Hart, 2012; Hart & Logan, 2011). First, the individual had an urge, impulse or idea to engage in violence. That is, the individual entertained the idea of acting violently in a given situation, instead of rejecting or dismissing the idea. Next, the individual evaluated the potential positive consequences of the violent act and made the decision that engaging in violence might result in personal reward or benefit. Third, the individual considered the negative consequences of engaging in violence and decided the costs were acceptable. The are numerous potential costs that one could consider such as feelings of guilt or remorse for committing such an act, the time and energy to engage in an act, possible punishment one could face, and the possibility of losing respect and love from others. Yet, an individual who decides to engage in a violent act has considered these costs and decided they are acceptable. In the final step of the process, the individual has evaluated the options for committing violence and determined it is feasible. Regardless of the various barriers that may be in place with respect to committing violence, the individual has evaluated these and decided they are able to overcome these barriers (Hart, 2012; Hart & Logan, 2011).

Under the decision theory framework, risk assessment is intended to assist the evaluator to understand how and why a person decided to engage in violence and how factors impacted the decision making process (Hart & Logan, 2011). Specifically, risk factors influence decision-making regarding whether or not to engage in violence in at least three ways (Hart & Logan, 2011). First, as a motivator, a risk factor increases the perceived rewards or benefits of violence. With respect to sexual violence specifically, some types of motivation include sexual gratification, to control or gain compliance from someone, or to enhance an individual's sense of status or self-esteem. Second, as a disinhibitor, a risk factor decreases the perceived costs or negative consequences of violence. Potential disinhibitors include negative attitudes (e.g., devaluing another person), negative self-concept (e.g., regarding oneself as a bad or evil person), or lack of insight. Finally, as a destabilizer, a risk factor disturbs or disrupts an individual's ability to monitor and control his or her decision-making. Destabilizers include disturbed attention, disturbed perception, impaired memory, or impulsive thinking (Hart, 2012; Hart & Logan, 2011).

Identifying the primary risk factors and the causal role they play is the essential aspect of formulation; it provides a theory of sexual violence for a specific person. Using

this decision theory approach and developing an explanation of an individual's history of violence has two main implications (Hart, 2012). First, in attempting to understand how and why an individual decided to commit violence in the past, the evaluator can try to determine how and why the particular individual might decide to commit a violent act in the future. Second, based on an understanding of how the person has gone through the decision-making process in the past, the evaluator can consider methods for encouraging the person to decide *not* to commit violence in the future (Hart, 2012). That is, risk management is intended to identify effective strategies to encourage prosocial decisions and behaviour while discouraging decisions to act violently (Hart & Logan, 2011). These two implications take us to the next components of the SPJ approach.

The second component of the SPJ approach to case formulation involves speculating about future violence risk, considering plausible future violence scenarios and developing management strategies. Scenario planning is:

> A process of positing several informed, plausible and imagined alternative future environments in which decisions about the future may be played out, for the purpose of changing current thinking, improving decision making, enhancing human and organization learning and improving performance. (Chermack & Lynham, 2002, p. 376)

It has been argued that scenario planning is appropriate for situations in which decisions must be based on incomplete knowledge, which is applicable to violence risk assessments (Hart & Logan, 2011). Scenario planning in violence risk assessment considers and speculates about what might occur, but it does not attempt to *predict* what will happen. As Chermack and Lynham (2002) explain, scenarios are not necessarily concerned with getting the future "right." The goal is to challenge current paradigms of thinking and direct attention to issues that might have been otherwise overlooked.

Scenario planning involves the development of multiple possible futures, and there may be a seemingly endless number of possibilities; however, there are only a few distinct scenarios that are evaluated as reasonable and plausible relative to what is known about a case (Hart, 2012). With that said, it is important to take into account a wide range of possible futures as this may assist an evaluator in avoiding tunnel vision

(i.e., gambling on one future), help to strive for desired outcomes, and avoid undesired outcomes (Hart, 2012).

Chermack and van der Merwe (2003) explain that good scenarios are those that are relevant to the concerns of the decision makers, challenge the existing assumptions and take them beyond what is currently believed to be plausible, can withstand scrutiny and are based on analysis and understanding of the factors that drive the future and the range of behaviour these factors may display. The consideration of four basic scenario types provides a good starting point when developing case scenarios (Hart, 2012; Hart & Logan, 2011). First, a *repeat scenario* is one that considers the person's history of violent acts and judges which act(s) are of most concern that the individual will commit again in the future (Hart, 2012). Second, constructing a *worst case scenario* involves considering the possibility of escalation and whether this individual could commit a more serious, possibly life-threatening, act of (sexual) violence. In a *twist* or *sideways trajectory* scenario, the individual could be imagined to engage in an act of (sexual) violence that changes in nature or evolves. There are a number of changes an evaluator might consider, including change in motivation, change in victim type or how a victim is selected, and change in behaviour or the way the offender perpetrates an offense (e.g., different type of coercion). Finally, the *better case scenario* is one in which the individual commits a less serious act or desists (Hart, 2012; Hart et al., 2003; Hart & Logan, 2011). Each scenario developed should be reviewed and evaluated, considering whether the scenario is plausible (e.g., does it fit the facts of the case, does it appear to be consistent with what is known about this particular offender), whether the scenario is useful (e.g., does the scenario provide ideas of how to prevent violence, does it lead to the development of specific risk management plans), and the degree of reliability or consistency (e.g., do others agree the scenario makes sense; do others agree about the primary scenario(s)) (Hart, 2012). Based on the experience of authorities that have used this approach, it has been suggested that 3 to 5 scenarios are generally sufficient to capture the plausible future events for any given case (Hart, 2012; Hart et al., 2003; Hart & Logan, 2011). Further, each scenario that is retained should include a detailed description. It is not sufficient for an evaluator to state, "I believe this individual will commit a rape." Rather, the scenario should include a detailed description of the nature (e.g., kind of [sexual] violence, likely victims), severity (e.g., psychological and physical

harm to the victims, likelihood of escalating to life-threatening violence), imminence (e.g., how soon the person might engage in [sexual] violence, any warning signs that risk is increasing), frequency/duration (e.g., how often the [sexual] violence will occur, chronic or acute risk), and likelihood of violence (e.g., overall, how common is this type of violence?, based on individual history, how likely is it to occur?) (Hart et al., 2003; Hart & Logan, 2011). Following this step of developing detailed scenarios and reviewing and evaluating each scenario, an evaluator is left with plausible scenarios that can be used to develop case management strategies (Hart, 2012; Hart & Logan, 2011).

Based on the fully developed risk scenarios, the evaluator develops case management plans to address the specific elements for each scenario. As discussed above, one of the primary functions of case formulation is to guide treatment or management decisions. Hart and colleagues (2003) suggest evaluators should consider four approaches to risk management: *monitoring*, *treatment*, *supervision* and *victim safety planning*.

Monitoring focuses on surveillance of the individual, rather than on control or restriction, which are the focus of supervision. Strategies that fall under monitoring include contact with the client and other relevant people (e.g., victims, family, supervisors), home visits, electronic surveillance, polygraphs, and drug testing. In addition, under monitoring, the evaluator should discuss the frequency and type of contacts (e.g., weekly, face-to-face) and any triggers that might indicate an increase in risk.

Treatment is intended to improve deficits in the individual's psychosocial adjustment through the provision of rehabilitative services. Treatment services may address multiple problem areas (e.g., treatment for mental disorder, anger management, vocational skills training, coping-focused treatment) but the evaluator should determine which deficits are highest priorities and focus on those.

As briefly mentioned above, the goal of supervision is to control or restrict the individual in order to make it difficult to engage in sexual violence. The evaluator must decide whether institutionalizing the individual in a correctional or health care facility is warranted, or if the person can be managed in the community, and in the latter case, the

types of restrictions that should be implemented (e.g., controlling the individual's movement, associations, or activities).

Finally, victim safety planning is implemented to minimize the likelihood or negative consequences of potential future sexual violence through the improvement of the victim's resources. Of course, in some scenarios the potential victims are unknown, in which case there are limited options for planning; victim scenario planning is most effective in cases where violence is targeted and potential victims are known. Victim safety strategies might include recommending that the physical environment of the victim be improved, providing strategies to assist the victim in responding to situations that pose risk, counselling a victim to increase their awareness, providing treatment to address psychosocial deficits that impair the victim's ability to protect him or herself, recommending completion of self-defence classes, and identifying strategies to be employed if the individual attempts to contact or approach the victim (Hart, 2012; Hart et al., 2003).

Hart (2012) stresses that it is not sufficient to consider only management strategies and tactics; it is important to also think about logistics and how a management plan will be implemented. For example, if an evaluator decides a sex offender treatment program is required, the evaluator also should consider the type of sex offender treatment program, whether the kinds of activities that would help this particular individual meet his goals, and program accessibility and availability.

The main strengths of this approach are its theoretical foundation, the structure provided for identifying risk factors and conceptualizing their causal roles, speculating about future violence and developing case management plans (Hart & Logan, 2011). In addition, the SPJ approach focuses on risk factors that are related to violence, rather than general criminal behaviour as per the RNR approach, and focuses on both past incidents of violence and future potential scenarios (Hart & Logan, 2011).

Case formulation is important not only for clinical decisions, such as identifying important treatment targets, but is essential for responding to legal questions. As discussed above, sex offender legislation enacted in various jurisdictions include specific considerations beyond the level of risk an individual poses. For example, SVP legislation

includes a provision regarding the presence of a mental abnormality and the necessary association between the mental abnormality and future sexual violence risk. A sexual violence risk assessment that provides only an estimate of risk (e.g., “high risk”) does not sufficiently respond to this issue. Case formulation is a critical step in evaluating whether a mental abnormality is present and whether it plays a causal role in the risk posed by the individual. If the evaluator determines there is no evidence that mental abnormality plays a causal role in the offender’s risk, he does not meet the criteria for designation as a SVP. In Canada, an important distinction between Dangerous Offender and Long-Term Offender legislation focuses on whether the offender can be managed effectively in the community. Again, providing a risk estimate does not necessarily clarify this issue. Case formulation, and specifically the development of plausible management strategies, can assist in determining the level of restriction required to manage this individual (e.g., whether the offender can be managed in the community). A recent qualitative study provided preliminary evidence that clinicians view formulation favourably and clinicians use formulation in practical ways to assist with risk management (e.g., communication with the offender) (Judge, Quayle, O’Rourke, Russell, & Dargee, 2013). Finally, a good risk assessment is not only about giving a level of risk (e.g., high risk, Bin 4, etc.), it is important to know the likely nature of violence, potential victims, possible causes, and effective management strategies, and case formulation gives the structure and ability to provide this information.

### *1.4.5. Case Formulation in the RSVP*

As previously discussed, the SPJ approach to case formulation is integrated within the guidelines of the RSVP. The RSVP case formulation process follows the steps outlined above, starting with the identification of primary risk factors and their causal role(s) in the perpetration of sexual violence. Although not covered in the current version of the RSVP manual, the measure’s authors have included this process as part of training on the use of the RSVP (e.g., Hart, 2012). Next, the evaluator speculates about what might happen in the future and develops plausible risk scenarios, including a detailed description (e.g., nature of violence, likely victim, etc.) of each plausible scenario. Finally, the evaluator develops risk management strategies to prevent each risk scenario posited in the previous step.

To summarize the RSVP assessment process, an evaluator is required to gather and obtain case information relying on multiple sources of information, which is used to assess the presence (past and recent), as well as the relevance (future) for each risk factor. Next, the evaluator develops a case formulation, determining the primary risk factors and causal role(s) of each for the individual's past sexual violence and developing scenario and related management strategies for future sexual violence. Finally, the evaluator must come to a conclusion regarding the summary risk ratings, including overall risk, severity of harm, imminence and whether other types of violence or criminal behaviour are of concern. Next, a review of the empirical support for the RSVP is provided.

## 1.5. Empirical Evaluation of the RSVP

Although limited, there is empirical support for the RSVP as a tool for sexual violence risk assessment. The majority of available research has focused on reliability and validity, with little research examining case formulation. A review of available research is presented, focusing first on reliability, then validity, and finally case formulation.

### *1.5.1. Interrater Reliability*

The interrater reliability of the RSVP has been examined in a handful of studies and, in general, the measure has demonstrated good reliability. Most of these studies indexed interrater reliability using intraclass correlation coefficients or ICCs. There are actually a number of different methods used to calculate ICCs and, confusingly, several different conventions for denoting them. (I will return to this point in the Method section.) In the rest of the Introduction, I will use "ICC" to refer to any ICC calculated for single raters and for absolute agreement. Following Fleiss (1981), ICCs were interpreted as follows: < .39 = *poor*, .40 to .59 = *fair*, .60 to .74 = *good*, and > .75 = *excellent*.

Hart (2003) evaluated RSVP assessments completed by two raters based on file review in a sample of 50 high-risk male adult sex offenders under community supervision in Canada (Hart, 2003). The interrater reliability of the item ratings were

generally good to excellent, with ICCs ranging from .58 to .97 (*Mdn* = .91) for Presence-Past, .62 to 1.00 (*Mdn* = .87; 1 item could not be evaluated) for Presence-Recent, and .65 to .95 (*Mdn* = .88) for Relevance-Future. Total and domain scores were calculated from the item scores and interrater reliability was examined for each of these. In general, the summary scores demonstrated excellent reliability. The interrater reliability (ICC) of Total scores across the three time periods was excellent: Presence-Past = .99, Presence-Recent = .96, and Relevance-Future = .98. For domain scores, ICCs ranged from .85 for the Relevance-Future domain of Social Adjustment to .98 for the Presence-Past Manageability domain. Finally, the interrater reliability of the summary risk rating (i.e., Case Prioritization) was good, ICC = .68.

Watt and colleagues (2006) also examined the interrater reliability for the RSVP in a sample of high-risk male sex offenders under community supervision. Consistent with Hart's (2003) findings, interrater reliability of the item ratings was good to excellent: Presence-Past = .58 to .97 (*Mdn* = .91), Presence-Recent = .62 to 1.00 (*Mdn* = 87, 1 item could not be evaluated), and Relevance-Future = .65 to .95 (*Mdn* = .88). The Total RSVP scores and Domain scores also demonstrated excellent reliability, with Total scores ranging from .96 to .99 and Domain scores ranging from .85 (Relevance-Future Social Adjustment) to .98 (Presence-Past Manageability). In contrast with the findings from Hart's (2003) study, the interrater reliability of the summary risk rating (i.e., Case Prioritization) was excellent in Watt et al.'s study (ICC = .92).

Watt and Jackson (2008) examined interrater reliability in a sample of 90 male adult sex offenders who had completed a community sex offender treatment program. Consistent with previous research, interrater reliability of individual risk items was mostly good to excellent: Presence-Past = .58 to .95 (*Mdn* = .68), Presence-Recent = .32 to .83 (*Mdn* = .69, 3 items could not be evaluated), and Relevance-Future = .62 to .92 (*Mdn* = .84). Total scores had excellent interrater reliability, Presence-Past = .95, Presence-Recent = .85, Relevance-Future = .91. Interrater reliability of the Domain scores was good to excellent, with a low of .65 for Presence-Recent Sexual Violence to .92 for Presence-Past Manageability. Interrater reliability for Case Prioritization was good (.75). In addition, Watt and Jackson examined the interrater reliability of the Risk of Serious Physical Harm (.85) and Immediate Action Required (.81).

Finally, Sutherland and colleagues (2012) examined the interrater reliability of the RSVP. Twenty-eight qualified forensic mental health professionals completed RSVP assessments for six case vignettes. Interrater reliability for the individual risk items ranged from poor (Stress or Coping, Relevance-Future = .05) to excellent (Major Mental Illness, Presence-Recent = .78), with an average ICC of .51 (fair). With the exception of Immediate Action Required (.43, Fair), the summary risk estimates demonstrated good reliability: Case Prioritization = .62; Risk of Serious Physical Harm .69; and Other Risks, = .66.

### *1.5.2. Validity*

Evidence of concurrent validity was provided by research demonstrating the RSVP to be positively associated with other measures of sexual violence risk. For example, Jackson and Healey (2008) found that the RSVP was strongly correlated with the SVR-20, as expected as the RSVP was developed as a parallel form of the SVR-20. Total scores for lifetime presence ratings were highly correlated between the two measures ($r$ = .97). Similarly, the domain scores were positively correlated. The Psychological Adjustment domain demonstrated the smallest correlation, $r$ = .57, and the remaining domains demonstrated correlations greater than .77. Evidence of concurrent validity also exists for RSVP scores and risk estimates and scores on the MnSOST-R (Hart, 2003; Klaver et al., 2002; Kropp, 2001), SORAG (Kropp, 2001), Static-99 (Hart, 2003; Jackson & Healey, 2008; Klaver et al., 2002; Watt et al., 2006), VRAG (Watt et al., 2006) and Static-2002 (Jackson & Healey, 2008). Two studies have examined the predictive validity of the RSVP. Kropp (2001) found that RSVP case prioritization rates were significantly correlated with sexually violent recidivism ($r$ = .40, $p$ < .05) in a sample of 53 offenders; RSVP total scores did not significantly correlate with recidivism. Hart and Jackson (2008) demonstrated the RSVP has good predictive validity by examining the association between RSVP ratings and recidivism. Results demonstrated that approximately one-third of risk factors from each time period (Presence-Past, Presence-Recent, Relevance-Future) were positively and significantly associated with future sexual violence recidivism. In addition, Hart and Jackson examined the association between case prioritization and recidivism. Relative to the low-priority group, the moderate group had 1.90 times greater likelihood of recidivism and the high risk group

had a 9.50 times greater likelihood of recidivism, which was a statistically significant difference, $\chi^2$ (1, $N$ = 54) = 9.95, $p$ = .002.

### *1.5.3. Case Formulation*

There is only one known study that has investigated case formulation using the RSVP. In their study of qualified mental health professionals who completed RSVP assessments for six case vignettes, Sutherland and colleagues (2012) investigated the reliability of case scenarios and case management strategies. An important point about this study is that case scenarios and case management strategies were not completed as intended. Instead, participants were required to respond to forced-choice questions designed for the study in order to capture key items from the case scenario and case management steps. For the case scenario step, clinicians were asked to identify plausible repeat and escalation scenarios and to respond to a series of questions for each scenario (e.g., type of offense: sexual breach of peace, indecent exposure, indecent assault, rape (without serious violence), rape (with serious violence), and sexual homicide; likely victim of scenario: prepubescent male, prepubescent female, adolescent male, adolescent female, adult male, or adult female; etc.). For case management, clinicians were required to make forced choice recommendations of the most appropriate monitoring and supervision strategies (e.g., recommended level of supervision: community outpatient (no supervision in place), community outpatient (supervision in place), inpatient (non-forensic), inpatient (low-secure), inpatient (forensic medium secure), inpatient (forensic high secure; etc.).

Sutherland and colleagues conducted reliability analyses and results of the interrater reliability analyses for case scenarios and case management strategies showed that ICCs ranged from poor (Level of Psychological Harm Escalation scenario = .25) to excellent (Recommended Level of Supervision = .87). The items that achieved good to excellent interrater reliability included Nature of Scenario (repeat), Victim in Scenario (repeat and escalation), and Recommended Level of Supervision. Poor reliability items were Level of Psychological Harm (escalation scenario), Estimated Imminence of Scenario (escalation), and Estimated Frequency of Scenario (escalation). Overall, better agreement was obtained for repeat scenarios (mean ICC = .59) regarding details of the scenario compared to escalation scenarios (mean ICC = .46). Although the

scenarios and management strategies were not completed in the manner recommended by the measure authors, overall the case scenario and management strategy process demonstrated good reliability and this study provides a good first step in evaluating the reliability of the case formulation process of the RSVP.

Unfortunately, in general, little research has focused on case formulation, which has led to substantial criticism (Sturmey, 2010). The majority of research has focused on interrater (or interformulator) reliability (Mumma, 2011). Mumma (2011) suggests that interformulator reliability might be a more appropriate term, rather than interrater reliability, as case formulation involves much more than making numerical ratings. More specifically, components such as target behaviours, relevant causal variables and their interrelationships are involved.

The lack of empirical evaluation is surprising given that case formulation is identified as a core skill by different professional organizations (Kuyken, 2006). Overall, studies examining reliability of case formulation have shown mixed findings (Sturmey, 2010). Mumma (2011) reviewed research on interformulation reliability conducted across numerous mental health contexts and noted that reliability of psychodynamic formulation ranged from good to excellent (i.e., kappas > .75) in some studies, but showed poor reliability (as low as .41) in other studies. Studies of cognitive-behavioural case formulations have demonstrated that clinicians are able to identify most problems and issues, but showed poor reliability when identifying the underlying mechanisms of the problems (Persons & Bertagnolli, 1999; Persons, Mooney, & Padesky, 1995). A number of limitations regarding studies examining reliability of case formulation have been identified, such as using a limited range of clinical materials or raters, constraints on the clinician's behaviour, resulting in development of formulation that is not reflective of actual practice, and using predetermined categories for clinicians to use (Hart, Sturmey, Logan, & McMurran, 2011; Mumma, 2011). As Sturmey (2010) discusses, future research should investigate the reliability of elements of case formulation, such as target behaviours, identifying current and historical variables that affect target behaviours, and translating formulation into an idiographic treatment plan.

## 1.6. Current Study

The present study addressed several gaps in the empirical support underlying the RSVP. To achieve the study's objectives, specialized sexual violence risk assessment training on the RSVP was provided to 17 mental health professionals. Trainees were required to complete six file-based RSVP assessments (randomly selected from 10 possible RSVP case files), which provide the data for this study. First, because there is limited research on the reliability of the RSVP, a primary goal of this study was to contribute to the small literature base on the reliability of the tool. The second objective was to examine the quality of the assessments by comparing the risk factor codings and summary judgments completed by 17 trainees with gold standard assessments completed by RSVP experts (two of four experts completed consensus ratings for each case). Investigating this issue, Sutherland et al. (2012) reported the average level of agreement (percent agreement) between item ratings made by clinicians and experts was 65%. The average level of agreement for summary judgments (Case Prioritization, Risk of Serious Physical Harm, Immediate Action Required and Other Risks Indicated) was 51%. The current study aims to replicate these findings. Finally, although case formulation is an integral component of the risk assessment process, it has received limited research attention. In particular, research has not yet established whether formulations developed by evaluators are consistent or reliable. As such, the current study will investigate the consistency and reliability of RSVP case formulations.

The following reflects the specific research questions and corresponding hypotheses that are investigated in the current study.

*Research Question 1*: What is the interrater reliability of individual factors (Presence-Past/Recent combined and Relevance-Future), domain scores, and summary risk estimates of RSVP assessments completed by trained mental health professionals?

*Hypotheses*. Based on previous research, it is expected that individual items and summary judgments will demonstrate fair to good reliability (ICC > .40). Composite scores, such as domain and total scores are expected to be higher, within the good to excellent range (ICC > .60).

*Research Question 2*: What is the quality of the assessments completed by trained mental health professionals? That is, what is the level of agreement between assessments of risk factors, domain scores, and summary risk estimates completed by recently trained mental health professionals and gold standard assessments completed by RSVP experts?

*Hypotheses*: Based on past research (i.e., Sutherland et al., 2012), it is expected there will be moderate agreement (mean percent agreement > 60%) between the trained professionals' and experts' RSVP ratings.

*Research Question 3*: What is the reliability, or similarity, of RSVP based case formulations (including case scenarios and case management strategies) between evaluators rating the same case compared to evaluators rating different cases?

*Hypothesis*: Due to the lack of research in the area, exploratory analyses will be conducted to answer this research question. However, it is expected that the mean similarity rating for within case comparisons (i.e., comparing formulations for two evaluators assessing the same case) will be significantly higher than mean similarity ratings for across-case comparisons (i.e., comparing formulations for two evaluators assessing different cases).

# 2. Methods

## 2.1. Overview

The current study recruited professionals for a sexual violence risk assessment course. Professionals were trained on administering the RSVP, their knowledge of using the RSVP was tested (Part 1) and they were sent six case files to conduct RSVP evaluations (Part 2). The completed cases were used to evaluate the interrater reliability of item, domain, and summary judgment ratings. In addition, Research Assistants conducted similarity ratings of formulations developed by different raters for the same case (Within Case) and different raters for different cases (Across Case). These similarity ratings were compared (Within vs. Across) to evaluate the reliability of case formulation. Full details of the study procedure are provided below.

Prior to commencing the study, ethical approval was obtained from the Simon Fraser University Office of Research Ethics.

## 2.2. Participants

Information pertaining to a newly developed, online, sexual violence risk assessment course was disseminated through multiple forms of media to professionals working in mental health and/or corrections settings. First, members of two professional organizations with a focus on mental health and law/forensics (i.e., the American Psychology-Law Society and the International Association of Forensic Mental Health Services) were sent an email providing information on the course. The same information was posted to the PSYLAW listserv, an online forum for professionals to discuss forensic and mental health clinical and research related issues. Finally, information on the sexual violence risk assessment course was posted to Dr. Hart's professional website, published in a newsletter distributed to members of the Canadian Association of Threat

Assessment Professionals organization, and posted on the training course host website *Consolidated Continuing Education and Professional Training* (CONCEPT; www.secure.concept-ce.com, discussed below).

The information disseminated included a brief description of the training course, including the type of work involved and the expected number of hours for completing all course components (i.e., Part 1 questionnaire and all six practice assessments), and notification that upon successful completion they would be eligible for 25 Continuing Education credits and a certificate of completion from CONCEPT (Appendix A). Individuals also were informed that the registration fee would be reimbursed upon successful completion of the course if they allowed their training materials to be used for the current study.

Individuals who requested further details about the research opportunity were informed that they were required to complete both parts of the training, provide consent to share the training materials with the researcher, and complete the materials by November 1, 2012. The deadline date was later extended to April 1, 2013. The extension was provided due to length of time participants required to complete each of the training cases and the time I required to complete and return the individualized feedback. As of the deadline, 48 individuals had registered for the course; 39 individuals had completed Part 1, and 18 participants had completed Parts 1 and 2. One of the 18 individuals failed to submit completed worksheets (e.g., lack of evidence for item ratings, did not generate detailed scenarios) and, therefore, did not provide data for the current study. The remaining 17 individuals comprise the final sample.

The majority of participants were female (*n* = 13; 77%). Most participants were 26 to 30 years of age (*n* = 6, 35%) or 31 to 35 years of age (*n* = 5, 29%). The most common profession was psychology (*n* = 15; 88%). The participants were internationally based, with most residing and working in the United States (*n* = 7; 41%), followed by Australia and Canada (*n* = 4; 24% each). The professionals had been working an average of 5 to 10 years in their current profession (*n* = 8, 47%); four participants had greater than 10 years of experience in their current profession. Most participants had less than 10 years of forensic risk assessment experience (*n* = 13, 76%). All participants had previously completed some type of sexual violence risk assessment training (e.g., Static-99, SVR-

20). On average, the participants completed 28 (*SD* = 40.69; range = 1 to 150) general violence risk assessments per year and 14 (*SD* = 12.29; range = 2 to 40) sexual violence risk assessments per year (Table 2.1). [3]

***Table 2.1 Demographics and Risk Assessment Experience of Study Sample***

| Demographic | *n* (%)* |
|---|---|
| Gender | |
| Male | 3 (18) |
| Female | 13 (77) |
| Age Range | |
| 26 – 30 years | 6 (35) |
| 31 – 35 years | 5 (29) |
| 36 – 40 years | 2 (12) |
| 41 – 45 years | 1 (6) |
| 46 – 50 years | 1 (6) |
| Over 60 years | 1 (6) |
| Profession | |
| Psychology | 15 (88) |
| Social Work | 1 (6) |
| Education | |
| Bachelor's degree | 1 (6) |
| Master's degree | 7 (41) |
| Doctorate degree | 8 (47) |
| Country | |
| United States of America | 7 (41) |
| Canada | 4 (24) |
| Australia | 4 (24) |
| United Kingdom | 1 (6) |
| Years of Experience in Current Profession | |
| Less than 5 | 4 (24) |
| 5 to 10 years | 8 (47) |
| 11 to 15 years | 3 (18) |
| Over 25 years | 1 (6) |

[3] Demographic information was not available for one participant.

| Demographic | *n* (%)* |
|---|---|
| Years of Forensic Risk Assessment Experience | |
| Less than 5 | 6 (35) |
| 5 to 10 years | 7 (41) |
| 11 to 15 years | 1 (6) |
| Over 25 years | 1 (6) |
| Prior Sexual Violence Risk Assessment Training | |
| ERASOR[a] | 4 (24) |
| MnSOST[b] | 1 (6) |
| RRASOR[c] | 2 (12) |
| Risk Matrix 2000[d] | 1 (6) |
| SORAG[e] | 3 (18) |
| SVR-20[f] | 7 (41) |
| Stable/Acute[g] | 0 (0) |
| Static-99 or Static-2002[h] | 4 (82) |

| | *M* (*SD*) |
|---|---|
| Number of Risk Assessments Conducted per Year | |
| General Violence | 28.44 (40.70) |
| Sexual Violence | 14.13 (12.29) |

*Note*: [a] ERASOR = Estimate of Risk of Adolescent Sexual Offense Recidivism (Worling & Curwen, 2001); [b] MnSOST-R = Minnesota Sex Offender Screening Tool – Revised (Epperson, Kaul, & Hesselton, 1998); [c] RRASOR = Rapid Risk Assessment for Sex Offense Recidivism (Hanson, 1997); [d] Risk Matrix 2000 (Thornton et al, 2003); [e] SORAG = Sex Offender Risk Appraisal Guide (Quinsey, Harris, Rice, & Cormier, 2006); [f] SVR-20 = Sexual Violence Risk-20 (Boer et al., 1997); [g] Stable/Acute = STABLE-2007/ACUTE-2007 (Hanson, Harris, Scott, & Helmus, 2007); [h] Static-99 (Hanson & Thornton, 1999), Static-2002 (Hanson, & Thornton, 2003).
* Demographic information was not available for one participant.

## 2.3. Materials

A two-part specialized course on sexual violence risk assessment, focused on the use of the RSVP, was developed. The first component was didactic and involved a review of fundamental issues of the SPJ approach to risk assessment and the administration of the RSVP. The second component was practical and involved the administration of the RSVP for six cases, with individualized feedback provided for each case.

### *2.3.1. Part 1 Materials*

#### Video Presentation

The information for Part 1 of the training course was presented in a series of four PowerPoint video presentations (Appendix B). I developed the slide presentations together with Dr. Hart, who narrated the videos. The four videos covered the following topics, in order: (1) Sexual violence and risk assessment (including relevant definitions); (2) SPJ approach to risk assessment; (3) the RSVP; and a (4) Sample case (see Appendix B for presentation slides). The length of the training was approximately 4.5 hours, including 3 hours of video presentation and 1.5 hours for reviewing the RSVP user manual and sample case (within the manual). To assist the trainees as they proceeded through the videos, a handout of the slides was available for recording additional notes.

#### Evaluation of Competence

To assess trainees' understanding of the information provided, they were asked to complete a 15-item open-ended questionnaire after watching the four videos in Part 1 of the training. The questions focused on general risk assessment concepts, such as the SPJ approach, and administration of the RSVP (Appendix C).

#### Feedback Form

A form was developed to provide feedback to trainees regarding their responses on the Evaluation of Competence. Responses to each of the 15 questions were assessed as *Acceptable* or *Not Acceptable*. For those responses deemed *Not Acceptable*, additional comments were provided to participants to educate them about the rationale regarding the appropriate response. The location of the information (training slide and manual page) was provided for the trainee to further review the issue.

### *2.3.2. Parts 1 and 2 Materials*

#### Risk for Sexual Violence Protocol User Manual and Worksheet.

All participants were provided an electronic copy of the RSVP user manual and worksheet. The manual provides an introduction to the RSVP and instructions on its

administration. Further, the manual provides risk factors and their definitions (Table 1.1). Trainees were given the manual for use as a reference guide while completing both parts of the course.

The RSVP Worksheet (Appendix D) is completed when administering the measure. The worksheet comprises six sections. The first section corresponds with Step 1 of the evaluation: the evaluator gathers the relevant background information and records identifying information. The second section of the worksheet corresponds with Steps 2 & 3, rating the presence and relevance of the risk factors. In this section of the worksheet, the evaluator records the relevant evidence to support each risk factor and to rate (*Yes*, *Possibly*/*Partially*, *No*) the Presence (Past and Recent) and Relevance (Future) for the 22 RSVP items, and also identifies any case specific risk factors. The next section of the worksheet is for case formulation and asks the evaluator to provide an account of the person's sexual violence by identifying the primary risk factors and their causal roles (i.e., motivators, disinhibitors, and destabilizers). Next, is the development of risk scenarios. That is, the evaluator is asked to describe the most plausible scenarios of future sexual violence, including (a) the nature of the sexual violence (e.g., type of violence, likely victim and motivation for such violence); (b) the severity of psychological and physical harm that is likely and whether there is a chance the violence could escalate to serious or life-threatening violence; (c) the imminence of the sexual violence; (d) the frequency and duration of the violence and (e) the likelihood of this type of sexual violence. The fifth section prompts the evaluator to develop case management strategies that are tailored to the individual and scenarios described in the previous step. The evaluator should identify the monitoring, treatment, supervision and victim safety planning strategies that should be implemented to reduce the risk of future sexual violence. In the final section of the worksheet, the evaluator makes judgments regarding case prioritization, level of harm, imminence of violence, and other risks. The evaluator also provides a case review timeline.

### *2.3.3. Part 2 Materials*

#### Case Files

Ten case files of male offenders with a history of sexually violent offending were used in the practical component of the training (Part 2). The cases were developed from actual cases from Dr. Hart's professional clinical work. All cases had been adjudicated and were anonymized, with all personal identifiers (including names, dates of birth, and locations) changed. The written case materials varied in length from 16 to 67 pages and included multiple types and sources of information, such as police records, correctional records, psychological evaluations, social history reports, and psychosexual assessments. The type of information provided varied between cases, depending on the information available in the original case file. The cases were selected to represent a range of offenders (e.g., different mental health issues, type of violence, victim type; see Appendix E) and were meant to represent, as much as possible, the density of information that might be provided to an evaluator for an actual assessment.

#### Gold Standard Assessments

A *gold standard* RSVP assessment was developed for each of the ten case files. To complete the gold standard assessments, two evaluators reviewed each case: me and one of three experienced, expert evaluators, each of whom was one of the RSVP co-authors. Evaluators made independent RSVP ratings and then met with the second evaluator for the case to discuss item ratings, case formulations, risk scenarios, management strategies and conclusory opinions. The gold standard assessments reflected consensus ratings between the two evaluators.

#### Feedback Form

Individualized feedback was provided for each case assessment (Appendix F) and based on the gold standard assessment scoring. Feedback evaluations for the first 15 cases submitted were completed by me and reviewed by Dr. Hart; after 15 cases, Dr. Hart indicated that I was able to complete the feedback myself. I continued to consult with Dr. Hart to clarify issues for cases I found difficult to evaluate.

The first three sections of the feedback form evaluated the presence and relevance ratings (*Correct*, *1-Off*, or *2-Off*). A general summary was provided that

informed the trainee of item definitions, reasons for expert ratings, and evidence that was overlooked during his or her assessment.

Next, trainee's case formulation, scenarios and management strategies were evaluated. The trainee was provided feedback regarding the motivators, disinhibitors, destabilizers, and primary causal risk factors identified. The risk scenarios were evaluated (*Yes*, *Possibly*/*Partially*, *No*) using four questions:

1. Did the trainee generate multiple risk scenarios?
2. Were the scenarios described in a plausible, detailed manner?
3. Were the scenarios linked to management strategies?
4. Were the scenarios consistent with those generated by experts?

Judgments of plausibility were based on whether the scenario was realistic and deserving of attention based on the facts of the case and the scenarios developed by the experts. Next, scenarios were evaluated on whether sufficient information and details were provided to develop management plans and strategies. Finally, the degree of overlap between the participant's scenarios and those of the expert evaluators were compared. An overall summary of the risk scenarios was provided.

Management plans were assessed (*Yes*, *Possibly*/*Partially*, *No*) with three questions:

1. Did the trainee document detailed management plans?'
2. Did the trainee identify appropriate strategies?'
3. Did the trainee identify appropriate tactics?

The 2nd question (appropriate strategies) was based on whether the participant included management techniques that corresponded with the risk factors and concerns outlined in the scenarios (and with those suggested by experts). Tactics were evaluated based on whether the evaluator provided sufficient details to implement the strategies noted. The final section of the form evaluated the conclusory opinions. Again, an evaluation key of *Correct*, *1-off*, and *2-off* was used to evaluate Case prioritization,

Serious Physical Harm, Immediate Action Required, and Other risks Indicated. A *Yes*/*No* rating was used for Case Review to reflect whether the recommended timeframe was within an acceptable range. Acceptable range was defined as approximately 2 weeks to 1 month.

### Demographics Survey

A brief online survey was developed to gather information about participants' demographics and risk assessment experience. The survey consisted of questions regarding gender, age, country of residence/employment, number of years in current position, number of years of experience with violence risk assessments, risk assessment experience and sexual violence risk assessment experience/training.

### Case Formulation Similarity Coding Form

A similarity rating form was developed (Appendix G) that consisted of three sections. The first section focused on case formulations, including a general overall similarity rating and individual similarity ratings regarding the motivators, disinhibitors, and destabilizers. The second section focused on the case scenarios: an overall similarity rating and individual ratings regarding the nature, level of harm, imminence, frequency/duration and likelihood. Finally, the third section focused on case management, with an overall comparison for the case management strategies followed by ratings for the monitoring, treatment, supervision, and victim safety plans. All ratings were made on a scale of 1 to 10, with 1 = *Not Similar at All,* 4 = *Different in many important respects, but similar in a few*, 7 = *Similar in many important respects, but different in a few* and 10 = *Very Similar*.

## 2.4. Procedure

Consolidated Continuing Education and Professional Training (CONCEPT; www.secure.concept-ce.com) hosted the risk assessment course on its professional website. CONCEPT provides continuing education and professional training for psychologists and other mental health professionals and is approved by the American Psychological Association (APA), the Canadian Psychological Association (CPA), and Canadian Association of Threat Assessment Professionals (CATAP) to sponsor

continuing education for its members (it may also be recognized by other organizations). All trainees, regardless of whether they later agreed to participate in the research component were required to register and pay for the course through the CONCEPT website. Part 1 cost $150 USD and Part 2 cost $350 USD. Trainees could register for both parts at one time, or could register for Part 1 and return to register for Part 2 at a later time (registrants could not register for Part 2 without registering for Part 1). Upon registration, the trainee was provided access to the materials for Part 1 of the course. Part 2 materials were made available upon successful completion of Part 1.

### *2.4.1. Completion of Part 1*

The trainee was required to watch the first three instructional videos in sequence. The instructional portion of the training was divided into multiple videos for ease of moving between topics and to provide as much flexibility as possible to the trainee; each video could be completed individually, at a time convenient for the trainee. After viewing the third video, the trainee was instructed to review the RSVP user manual. In particular, the trainee was encouraged to carefully review the RSVP risk factors and definitions, as these were not thoroughly discussed in the training videos. Also, the trainee was encouraged to review the practice case in the RSVP user manual. After reading the manual, the trainee watched the fourth video.

After watching all four videos and reading the RSVP user manual, the trainee completed the Evaluation of Competency form and e-mailed it to the CONCEPT RSVP email address. Completed forms were forwarded to me; I reviewed them and provided individual feedback. If a response was deemed *Not Acceptable*, the trainee was directed to the specific training video/slide and corresponding section within the manual to review the material. At the time the feedback was returned, the trainee was provided a unique code to enter at the RSVP CONCEPT training website to obtain the certificate of completion, a record of the continuing education credits obtained for Part 1 (7 credits), and Part 2 materials (if applicable).

### 2.4.2. Completion of Part 2

Each trainee was randomly assigned six of 10 possible case files and was required to rate the RSVP. The trainee was provided one case file at a time and upon completion and submission of the RSVP worksheet, the next case was provided. Attempts were made to provide feedback for each case prior to the trainee proceeding with the next case assignment; however, this was not possible in all cases. Due to time constraints or concerns of the professional completing the training (e.g., wanting to complete the course within a certain timeframe as he or she was aware of upcoming professional commitments that would limit time available to complete the course), as well as the time required to complete the feedback form, many trainees requested and proceeded with their next case prior to receiving feedback for the previous case. All participants received some case feedback prior to completing all six cases. With respect to initial feedback (i.e., first feedback received by participant), 47% (*n* = 8) received their first feedback prior to submitting their second case, 29% (*n* = 5) received their first feedback after they had submitted their second case and 18% (*n* = 3) received their first feedback after submitting their third case; one person (*n* = 1, 6%) received their feedback the same day they submitted their second case.

Trainees who successfully completed the six cases were sent a new code to download their certificate of completion and receive continuing education credits via the CONCEPT website. At this time, the trainee was also e-mailed the informed consent (*n* = 17) document, which described the purpose of the current research and asked the trainee if he/she would be willing to share their training materials (i.e., Evaluation of Competency and six case assessments) for research purposes. Trainees who signed and returned the consent form (*n* = 17, 100%) were reimbursed the registration fee. Also at this time, the participants (i.e., trainees who consented to the research) were provided a link to the online demographic survey and asked to complete it.

### 2.4.3. Coding Case Formulation Similarity

To investigate the reliability and consistency of participants' case formulations, similarity ratings were completed for a randomly selected set of case pairings. Case pairings were categorized into two groups: *Within Case* pairings included pairings of the

same case completed by different raters and *Across Case* pairings were different cases and different raters. A power analysis was conducted to determine the sample size required to detect a moderate effect size ($d$ = .5) with 90% power and alpha set at .05. This analysis determined that 69 pairings within each group (i.e., Within Case and Across Case) were required. A minimum of six participants completing 10 cases each was required to provide sufficient data to complete the analyses. As such, 69 pairs of Within Case and 69 pairings of Across Case were randomly selected. Each case was assigned a new case ID, which was a four-digit randomly generated number, to prevent similarity raters from identifying cases completed by the same participant.

Pilot testing was conducted to assist in finalizing the case formulation similarity-rating coding form and develop coding guidelines. Once the form and the guidelines were established, two Research Assistants (RAs; 1 Master's level psychology student; 1 Honours level undergraduate student) were trained to complete similarity ratings using 10 practice cases (5 Within Case and 5 Across Case). The RAs were provided the coding form, guidelines for coding and examples of *High* and *Low* similarity formulations. The RAs were then assigned the first practice case to complete independently. Upon completion, ratings were compared, coding differences discussed, and consensus ratings were determined. The second and third practice cases were completed in the same manner, with discussions occurring immediately after completing each case. The RAs completed the next 3 cases independently and discrepancies in ratings (in particular, ratings that differed by 2 or more points) were discussed. Next, the RAs completed case seven and coding discrepancies were discussed upon completion. Finally, the research assistants completed the remaining practice cases and discrepancies were again discussed. Across all cases, approximately 75% of the ratings were within 2-points of each other and over the final three cases, 83% of ratings were within 2-points of each other. The RAs were provided the 138 case pairings to be used for data analysis. Similarity ratings were completed independently, but to prevent drift between RAs, consensus coding was completed for every fifth case (after completing independent ratings).

## 2.5. Statistical Analyses

All analyses were conducted using IBM SPSS Statistics, Version 20. Ratings were coded into numeric values for research purposes (*Yes* = 2, *Possibly*/*Partially* = 1, *No* = 0 for risk factors; *High* = 2, *Moderate* = 1, *Low* = 0 for Case Prioritization and Level of Harm ratings). For all analyses, Presence: Past and Recent ratings were combined into one variable *Presence*. That is, the timing of the risk factor was disregarded and a single variable was created to capture whether the risk factor was *ever* present for the particular individual (i.e., *highest* score of Past and Recent was retained; if coded Yes for either Past or Recent, Yes was retained; if Past and Recent coding was a combination of Possibly/Partially and No, Possibly/Partially was retained; if Past and Recent were both coded No, No was retained). Although analyzing the data using this approach does not reflect the intended administration of the RSPV, the primary reason for combining Past and Recent was due to the fact the offenders within the 10 case files were institutionalized for the year prior to the assessment. As a result, few items would demonstrate the requisite level of variability to conduct reliability analyses.

To provide context for the case files, case descriptives are provided. Case descriptives include the mean item, domain, and total scores for the Presence and Relevance ratings for all cases combined, as well as each case.

To evaluate interrater reliability (IRR; Research Question 1), intraclass correlations coefficents (ICCs) were calculated. Much has been written about various statistics for indexing the interrater reliability of psychological tests and measures (Bartko, 1976; McGraw & Wong, 1996; Shrout & Fleiss, 1979). There is some consensus regarding the criteria that make an index most useful. First, to assist interpretation, it should be unaffected by the metric of the raw scores, and so yield values between a fixed minimum and maximum (e.g., 0 to 1, or -1 to 1). Second, it should correct for chance agreement. Third, when used with interval- or ratio-type variables, the index should be able to account for between-rater differences with respect to both rank and anchor point.

Two families of statistics that meet these criteria are commonly recommended in the psychological literature. The first is kappa, used to index nominal-or ordinal-level

variables. The second is ICC, used for interval- or ratio-level variables. Both are chance-corrected indexes of agreement with values that range from -1 (perfect disagreement) to 1 (perfect agreement), with 0 indicating chance agreement. They can be referred to as "families" because they can be calculated in a number of different ways, depending on the method used to make observations. Also, the mathematics underlying the two families of statistics are closely related.

RSVP ratings may be considered ordinal- or interval-level variables. In this study, I was most interested in estimating the absolute agreement of ratings made by a single rater, corrected for chance. Three aspects of the study method were important. First, RSVP ratings were made by a group of raters who were assumed to be representative of the entire universe of possible raters, rather than representative of a select or special group. Second, a different set of raters made ratings for each case. Third, the assignment of raters to cases was random. Accordingly, I indexed agreement using Case 1 (also referred to as Type 1) ICCs for single raters and absolute agreement (Nichols, 1998; Shrout & Fleiss, 1979). This type of ICC treats raters as a random effect and ratings as a fixed effect. Due to the incomplete nesting of raters within cases, it is impossible to examine Rater x Case interactions (Nicholls, 1998; Shrout & Fleiss, 1979). Calculated this way, the ICC is mathematically equivalent to a simple weighted kappa coefficient. With respect to interpretation of ICCs, Fleiss (1981) recommended that values less than .40 indicated poor agreement; values between .40 and .59 indicated fair agreement; values between .60 and .74 indicated good agreement; and values of .75 and greater indicated excellent agreement.

As trainees were assigned a random selection of six of a possible 10 cases, the number of completed assessments differed across cases. Six participants completed Case 6, the lowest number of raters for a case. As such, the maximum number of assessments per case that could be included in the IRR analyses was six. As each of the remaining nine cases had more than six completed assessments, six assessments were randomly selected for the analyses. Cases to include in the analyses were selected in a random, yet systematic manner. The objective of this approach was to have randomly selected cases, but not have any evaluators over-represented or under-represented in the analyses. Based on the total number of assessments required to complete the analyses (10 cases x 6 assessments/case = 60 assessments), it was first

determined that no evaluator should have less than three cases or more than four completed cases selected (60 assessments ÷ 17 raters = 3.5 assessments/rater; 9 raters with 4 cases included, 8 raters with 3 cases included). As such, the evaluators were assigned to have either three or four cases included in alternate fashion. To do this, evaluators were listed in order of participant ID, the first evaluator was assigned to have 4 cases included, the next evaluator to have 3 cases included, and so on. From there, cases were randomly selected for each evaluator so that each case had six completed assessments. A random number generator was used to randomly select cases for each participant (according to assigned number of cases discussed above, i.e., three or four cases). A few adjustments were required in order to obtain six assessments for each case. For example, if a case reached its maximum of six assessments but was selected for a participant, a new case was randomly selected for the participant or one of the previously selected assessments was replaced (and that participant had a new case randomly selected). In the end, we had a dataset that included at least three and no more than four assessments from each evaluator.

Comparisons between professional RSVP assessments and expert RSVP assessments (Research Question 2) were completed using percent agreement for Presence and Relevance ratings and summary risk judgments for each case individually and across all cases as a group. To examine agreement for domain and total scores, *t*-tests were conducted and Cohen's *d*'s were calculated to examine differences between the mean scores of the professionals' and experts' assessments for each case.

Finally, reliability of case formulations was evaluated by comparing similarity ratings (as assessed by Research Assistants) between Within Case pairings (same case, different raters) with Across Case pairings (different case, different raters). RSVP formulations are not quantified and could not be examined using traditional statistical indexes of reliability. As such, to conduct these analyses, similarity between formulations were compared, with similarity reflecting the consistency of formulations between raters and, ultimately, this was used to represent reliability. Similarity ratings averaged across the two RAs were used for these analyses. Mean inter-item correlations were computed for four groups of ratings: (1) overall similarity of case formulation, case scenario, and case management strategies; (2) case formulation specific ratings (i.e., motivators, disinhibitors and destabilizers); (3) case scenario specific ratings (i.e., type of violence,

victims, level of harm, imminence); and (4) case management specific ratings (i.e., monitoring, treatment, supervision, victim safety planning). The mean inter-item correlations were moderate for all analyses. Therefore, four one-way multivariate analysis of variance analyses were conducted to investigate differences in mean similarity ratings between the groups for each set of ratings (Research Question 3). Univariate analysis of variance analyses were completed to determine where the differences between groups occurred.

# 3. Results

## 3.1. Competency Evaluation

The participants performed very well on the competency evaluation. The mean total score was 14.65 out of a possible score of 15 (*SD* = .79; range = 12 to 15).

To investigate whether participants with more risk assessment experience demonstrated a greater understanding of the training material, as assessed by the questionnaire, bivariate correlation analyses were conducted. Analyses examined the association between the questionnaire Total score and demographic variables, including number of years in current profession, number of years of forensic risk assessment experience, and number of sexual violence risk assessments completed per year. No significant associations were found.

The results from the questionnaire suggested that the participants had an adequate understanding of conducting SPJ risk assessments and, in particular, administration of the RSPV. Based on this, all participants were able to proceed to the next stage of the training and complete the six practice cases.

## 3.2. Case Descriptives

Analyses were conducted to examine the mean Total, Domain, and Item RSVP scores (Presence and Relevance, separately) for each case. Again, ratings were recoded into numeric values for research purposes (*Yes* = 2, *Possibly*/*Partially* = 1, *No* = 0 for risk factors; *High* = 2, *Moderate* = 1, *Low* = 0 for Case Prioritization and Level of Harm ratings). The average Total score across all cases was 33.53 (*SD* = 6.63). The majority of cases had high Total Presence scores; four of the 10 cases had mean Total scores greater than 35 out of a possible score of 44 (Table 3.1). Case 1 had the highest mean Total score, 38.56 (*SD* = 3.28). Only three cases had a mean Total score under

30, and only one of these cases had a mean Total score under 20 (Case 4: *M* = 29.73, *SD* = 5.50; Case 6: *M* = 29.50, *SD* = 6.60; Case 8: *M* = 18.43, *SD* = 8.16). Similarly, the mean Domain scores for presence ratings were mostly in the upper range of possible scores (Table 3.1). In particular, most cases had high scores within the Manageability and Social Adjustment domains. For the Manageability domain, all cases, with the exception of Cases 6 and 8, had an average domain score greater than 5, out of a possible maximum score of 6. For Social Adjustment, other than Case 2 and Case 8, the remaining cases had an average domain score greater than 7, of a possible maximum score of 8. Case 1 had an average Social Adjustment score of 8, indicating all evaluators gave the maximum rating for each item within this domain. Mean ratings of individual items were examined for extreme scores (i.e., < .5 or > 1.5; Appendix H). Many items demonstrated high average scores, with 15 of the 22 items having an average rating greater than 1.5 (no items had a mean score less than .5). Consistent with the findings for the average Domain scores discussed above, all items within the Social Adjustment and Manageability Domains had mean scores greater than 1.5 across all cases.

***Table 3.1 Mean Scores and Standard Deviations for Presence Total Score and Domain Scores per Case***

| Case (*n*) | Total Score | Sexual Violence History | Psychological Adjustment | Mental Disorder | Social Adjustment | Manageability |
|---|---|---|---|---|---|---|
| | Range 0 – 44 | Range 0 – 10 | Range 0 – 10 | Range 0 – 10 | Range 0 – 8 | Range 0 – 6 |
| | *M* (*SD*) | *M* (*SD*) | *M* (*SD*) | *M* (*SD*) | *M* (*SD*) | *M* (*SD*) |
| Case 1 (9) | 38.56 (3.28) | 7.11 (1.69) | 9.22 (.97) | 8.44 (.73) | 8.00 (.00) | 5.78 (.67) |
| Case 2 (11) | 32.55 (3.50) | 7.00 (2.05) | 8.45 (1.69) | 6.27 (1.62) | 5.72 (1.19) | 5.09 (1.30) |
| Case 3 (9) | 34.56 (2.07) | 6.22 (1.39) | 9.11 (.78) | 5.78 (1.20) | 7.89 (.33) | 5.56 (.53) |
| Case 4 (12) | 29.73 (5.50) | 5.73 (2.15) | 5.45 (2.16) | 5.18 (1.94) | 7.73 (.65) | 5.63 (.67) |
| Case 5 (9) | 31.44 (3.50) | 7.56 (2.19) | 6.44 (2.19) | 5.00 (1.11) | 7.22 (1.30) | 5.22 (.97) |
| Case 6 (6) | 29.50 (6.60) | 6.50 (2.07) | 6.67 (1.75) | 4.00 (1.90) | 7.50 (.84) | 4.83 (1.17) |
| Case 7 (14) | 36.57 (4.75) | 6.36 (2.31) | 9.21 (1.05) | 7.36 (1.65) | 7.71 (.61) | 5.93 (.27) |

| Case (*n*) | Total Score | Sexual Violence History | Psychological Adjustment | Mental Disorder | Social Adjustment | Manageability |
|---|---|---|---|---|---|---|
| | Range 0 – 44 | Range 0 – 10 | Range 0 – 10 | Range 0 – 10 | Range 0 – 8 | Range 0 – 6 |
| | *M* (*SD*) | *M* (*SD*) | *M* (*SD*) | *M* (*SD*) | *M* (*SD*) | *M* (*SD*) |
| Case 8 (7) | 18.43 (8.16) | 5.71 (1.80) | 5.86 (1.95) | 4.14 (2.27) | 1.71 (1.98) | 1.00 (1.29) |
| Case 9 (12) | 37.75 (3.28) | 7.67 (1.37) | 9.00 (1.41) | 7.83 (.72) | 7.50 (.80) | 5.75 (.62) |
| Case 10 (14) | 37.43 (2.47) | 8.14 (1.29) | 8.64 (.74) | 7.64 (1.01) | 7.07 (1.32) | 5.92 (.27) |
| Overall Mean | 33.53 (6.63) | 6.88 (1.95) | 7.99 (2.01) | 6.43 (1.98) | 6.96 (1.84) | 5.27 (1.43) |

Relevance scores demonstrated a similar pattern, although, in general, Relevance mean scores were lower than Presence mean scores (Table 3.2). The mean Relevance Total score across all cases was 28.91 (*SD* = 7.20). In contrast to the Presence Total scores, where Case 1 had the highest mean score, Case 7 had the highest mean Relevance Total score, *M* = 35.00 (*SD* = 5.48). Case 8 had the lowest average score, *M* = 17.57 (*SD* = 6.13). Six cases had mean scores under 30. Relevance Domain scores were lower compared to Presence scores, as well. Consistent with Presence findings, Social Adjustment (*M* = 5.81, *SD* = 2.09, out of possible score of 8) and Manageability (*M* = 4.39, *SD* = 1.84, out of possible scores of 6) appeared to be the most problematic domains, with mean scores in the upper range of possible scores. For individual items, compared to Presence ratings, fewer items had Relevance ratings that fell into the extreme range (< .5 or > 1.5). Six of the 22 items had an average score greater than 1.5; no items had an average score less than 0.5 (Appendix H).

***Table 3.2 Mean Scores and Standard Deviations for Relevance Total Score and Domain Scores Per Case***

| | Total Score | Sexual Violence History | Psychological Adjustment | Mental Disorder | Social Adjustment | Manageability |
|---|---|---|---|---|---|---|
| **Case (*n*)** | **Range 0 – 44** | **Range 0 – 10** | **Range 0 – 10** | **Range 0 – 10** | **Range 0 – 8** | **Range 0 - 6** |
| | ***M* (*SD*)** | ***M* (*SD*)** | ***M* (*SD*)** | ***M* (*SD*)** | ***M* (*SD*)** | ***M* (*SD*)** |
| Case 1 (9) | 34.22 (5.31) | 6.67 (1.93) | 8.11 (1.54) | 7.22 (1.20) | 7.22 (1.09) | 5.00 (1.80) |
| Case 2 (11) | 30.36 (4.43) | 6.82 (1.72) | 7.64 (1.43) | 5.91 (1.97) | 5.09 (1.45) | 4.91 (1.38) |
| Case 3 (9) | 28.89 (4.46) | 4.89 (2.52) | 7.33 (1.32) | 4.89 (.93) | 6.78 (1.20) | 5.00 (.87) |
| Case 4 (12) | 26.63 (6.14) | 5.45 (1.86) | 4.64 (1.91) | 4.82 (1.89) | 7.36 (.81) | 4.36 (1.36) |
| Case 5 (9) | 24.00 (3.87) | 6.56 (2.13) | 3.78 (1.56) | 4.33 (1.00) | 6.11 (1.62) | 3.22 (1.09) |
| Case 6 (6) | 25.83 (6.55) | 6.67 (2.50) | 5.83 (2.04) | 3.83 (1.94) | 4.83 (1.94) | 4.67 (1.75) |
| Case 7 (14) | 35.00 (5.48) | 6.43 (2.38) | 8.43 (1.79) | 7.29 (1.59) | 7.14 (.95) | 5.71 (.61) |
| Case 8 (7) | 17.57 (6.13) | 5.86 (.90) | 5.57 (1.51) | 4.00 (1.83) | .86 (1.21) | 1.29 (1.38) |
| Case 9 (12) | 23.75 (5.59) | 6.92 (1.68) | 5.58 (1.98) | 4.58 (1.16) | 4.33 (1.44) | 2.33 (1.72) |
| Case 10 (14) | 34.64 (3.73) | 7.43 (1.65) | 8.21 (1.19) | 6.86 (1.23) | 6.29 (1.20) | 5.86 (.53) |
| Overall Mean | 28.91 (7.20) | 6.43 (2.02) | 6.69 (2.23) | 5.59 (1.91) | 5.81 (2.09) | 4.39 (1.84) |

For summary judgments of Case Prioritization and Risk of Serious Physical Harm, the distribution of High, Moderate and Low ratings was examined (Table 3.3). For seven of the 10 cases, High was the modal Case Prioritization rating. Two of the three remaining cases received mostly Moderate Case Prioritization ratings and the final case was evenly split between High and Moderate ratings. Although five cases received Low Case Prioritization ratings from at least one assessor, no cases received a majority of Low ratings. Regarding the distribution of scores for the Risk of Serious Physical Harm ratings, three cases had High as the modal rating, Moderate was the modal rating for 5

cases, and two cases were mostly rated as Low, indicating the Risk of Serious Physical Harm expected from these cases was more evenly distributed and mostly at a moderate level. Examining the average score for the Case Prioritization ratings (Table 3.4), six cases had an average score greater than 1.50. The average Case Prioritization score across all cases was 1.50 (*SD* = .63).

***Table 3.3 Frequency of Summary Judgments (High, Moderate, Low) for Case Prioritization and Level of Harm***

| (*n*) | Case Prioritization | | | Risk of Serious Physical Harm | | |
|---|---|---|---|---|---|---|
| | High n (%) | Moderate n (%) | Low n (%) | High n (%) | Moderate n (%) | Low n (%) |
| Case 1 (9) | 6 (67) | 3 (33) | 0 (0) | 6 (67) | 2 (22) | 1 (11) |
| Case 2 (11) | 10 (91) | 1 (9) | 0 (0) | 8 (73) | 3 (27) | 0 (0) |
| Case 3 (9) | 5 (56) | 4 (44) | 0 (0) | 0 (0) | 5 (56) | 4 (44) |
| Case 4 (12) | 1 (18) | 8 (73) | 2 (9) | 0 (0) | 3 (27) | 8 (73) |
| Case 5 (9) | 1 (22) | 6 (67) | 2 (11) | 0 (0) | 5 (56) | 4 (44) |
| Case 6 (6) | 5 (83) | 1 (17) | 0 (0) | 2 (33) | 3 (50) | 1 (17) |
| Case 7 (14) | 10 (71) | 3 (43) | 1 (14) | 2 (14) | 11 (79) | 1 (7) |
| Case 8 (7) | 3 (43) | 3 (43) | 1 (14) | 6 (86) | 1 (14) | 0 (0) |
| Case 9 (12) | 6 (50) | 5 (42) | 1 (8) | 0 (0) | 4 (33) | 8 (67) |
| Case 10 (14) | 11 (78) | 3 (21) | 0 (0) | 1 (7) | 9 (64) | 4 (29) |
| **Overall** | 58 (57) | 37 (36) | 7 (7) | 25 (25) | 46 (45) | 31 (30) |

Taking the above findings together, the results indicated the cases were, in general, high risk cases and the offenders presented with multiple problems across risk factors and risk domains. Although there was not as much variability in the cases as one might hope, the case descriptives demonstrated the cases were not exclusively high risk and there was some variability in the total and domain scores, indicating the cases were appropriate for conducting further analyses, including reliability analyses.

***Table 3.4 Means and Standard Deviations for Case Prioritization, Risk of Serious Physical Harm, and Imminence scores for All Cases***

| | Case Prioritization | Risk of Serious Physical Harm | Imminence |
|---|---|---|---|
| (*n*) | *M* (*SD*) | *M* (*SD*) | *M* (*SD*) |
| Case 1 (9) | 1.67 (.50) | 1.56 (.73) | 1.33 (.71) |
| Case 2 (11) | 1.91 (.30) | 1.73 (.47) | 1.64 (.50) |
| Case 3 (9) | 1.56 (.53) | .56 (.53) | 1.44 (.53) |
| Case 4 (12) | 0.91 (.54) | .27 (.47) | .82 (.87) |
| Case 5 (9) | 0.89 (.60) | .56 (.53) | .89 (.78) |
| Case 6 (6) | 1.83 (.41) | 1.17 (.75) | 1.00 (.89) |
| Case 7 (14) | 1.64 (.63) | 1.07 (.47) | 1.07 (.73) |
| Case 8 (7) | 1.29 (.76) | 1.86 (.38) | 1.00 (.58) |
| Case 9 (12) | 1.42 (.69) | .33 (.49) | .67 (.79) |
| Case 10 (14) | 1.79 (.43) | .79 (.58) | 1.36 (.63) |
| **Overall Mean** | 1.50 (.63) | .94 (.74) | 1.13 (.74) |

## 3.3. Interrater Reliability: Ratings of Presence, Relevance and Summary Risk

### *3.3.1. Agreement among Raters*

First, analyses examined individual RSVP Presence (Table 3.5; Figure 3.1) and Relevance (Table 3.6; Figure 3.2) ratings. Intraclass correlation coefficients (ICCs) were interpreted following Fleiss' (1981) guidelines: $\leq$ .39 = *poor*, .40 to .59 = *fair*, .60 to .74 = *good*, and $\geq$ .75 = *excellent*. The average ICC across all 22 Presence items was fair (.46), whereas the mean ICC for the corresponding Relevance ratings was poor (.39). Across all Presence and Relevance risk factor ratings, the interrater reliability was mostly poor (19 of 44 items had ICCs below .40), or fair (17 of 44 items had ICCs between .40 to .60). Of the Presence ratings, 8 of 22 items were fair, 4 items had good interrater reliability, one item had excellent interrater reliability and the remaining 8 items had poor reliability (ICC for one item could not be calculated due to lack of variance). For Relevance ratings, nine items showed fair reliability, one item had good reliability, and the remaining 12 items had poor reliability.

### *Table 3.5 Mean, Standard Deviation, and Intraclass Correlation Coefficients (ICCs) for RSVP Presence Risk Factors for All Cases*

| RSVP Risk Item | Presence Ratings | | | |
|---|---|---|---|---|
| | *M* (*SD*) | ICC | p | 95% CI |
| 1. Chronicity of Sexual Violence | 1.63 (.44) | .27 | .004 | .056, .630 |
| 2. Diversity of Sexual Violence | 1.15 (.42) | .11 | .103 | -.046, .458 |
| 3. Escalation of Sexual Violence | 1.07 (.11) | .31 | .001 | .087, .666 |
| 4. Physical Coercion in Sexual Violence | 1.60 (.67) | .62 | <.001 | .372, .859 |
| 5. Psychological Coercion in Sexual Violence | 1.47 (.55) | .31 | .001 | .082, .661 |
| **Sexual Violence History Domain** | | .32 | | |
| 6. Extreme Minimization/Denial of Sexual Violence | 1.67 (.41) | .26 | .005 | .045, .616 |
| 7. Attitudes that Support or Condone Sexual Violence | 1.52 (.47) | .42 | <.001 | .168, .742 |
| 8. Problems with Self-Awareness * | 0.82 (.15) | | | |
| 9. Problems with Stress or Coping | 1.60 (.39) | .16 | .045 | -.019, .515 |
| 10. Problems Resulting from Child Abuse | 1.15 (.63) | .42 | <.001 | .174, .747 |
| **Psychological Adjustment Domain** | | .32 | | |
| 11. Sexual Deviance | 1.48 (.42) | .29 | .003 | .067, .643 |
| 12. Psychopathic Personality Disorder | 1.03 (.63) | .42 | <.001 | .175, .748 |
| 13. Major Mental Illness | 0.85 (.74) | .55 | <.001 | .295, .824 |
| 14. Problems with Substance Use | 1.45 (.71) | .63 | <.001 | .383, .864 |
| 15. Violent or Suicidal Ideation | 1.45 (.58) | .46 | <.001 | .210, .773 |
| **Mental Disorder Domain** | | .47 | | |
| 16. Problems with Intimate Relationships | 1.18 (.31) | .63 | <.001 | .342, .872 |
| 17. Problems with Non-Intimate Relationships | 1.72 (.51) | .59 | <.001 | .334, .843 |
| 18. Problems with Employment | 1.63 (.55) | .53 | <.001 | .273, .812 |
| 19. Non-sexual Criminality | 1.55 (.66) | .64 | <.001 | .397, .869 |
| **Social Adjustment Domain** | | .60 | | |
| 20. Problems with Planning | 1.70 (.44) | .39 | <.001 | .150, .728 |
| 21. Problems with Treatment | 1.72 (.50) | .49 | <.001 | .238, .791 |
| 22. Problems with Supervision | 1.80 (.58) | .91 | <.001 | .801, .972 |
| **Manageability Domain** | | .60 | | |
| **Overall Mean ICC** | | .46 | | |

*Note*. *Unable to calculate ICC due to lack of variance.

***Figure 3.1 Intraclass Correlation Coefficients (ICCs) and 95% Confidence Intervals for RSVP Presence Items***

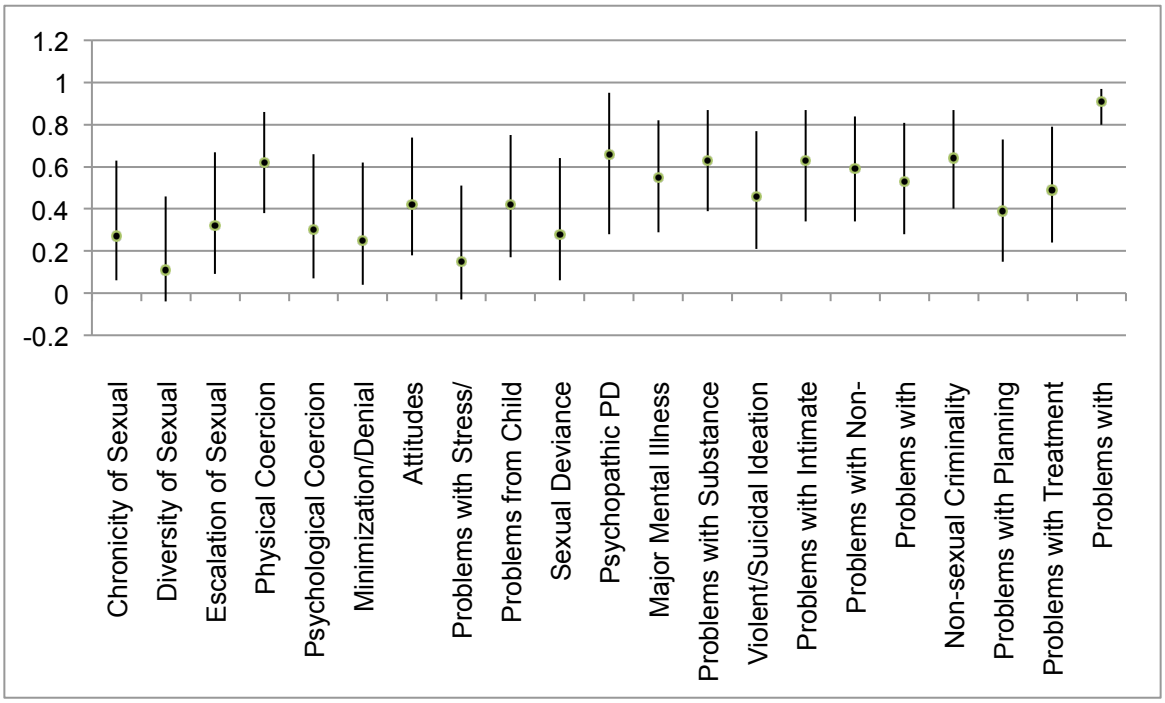
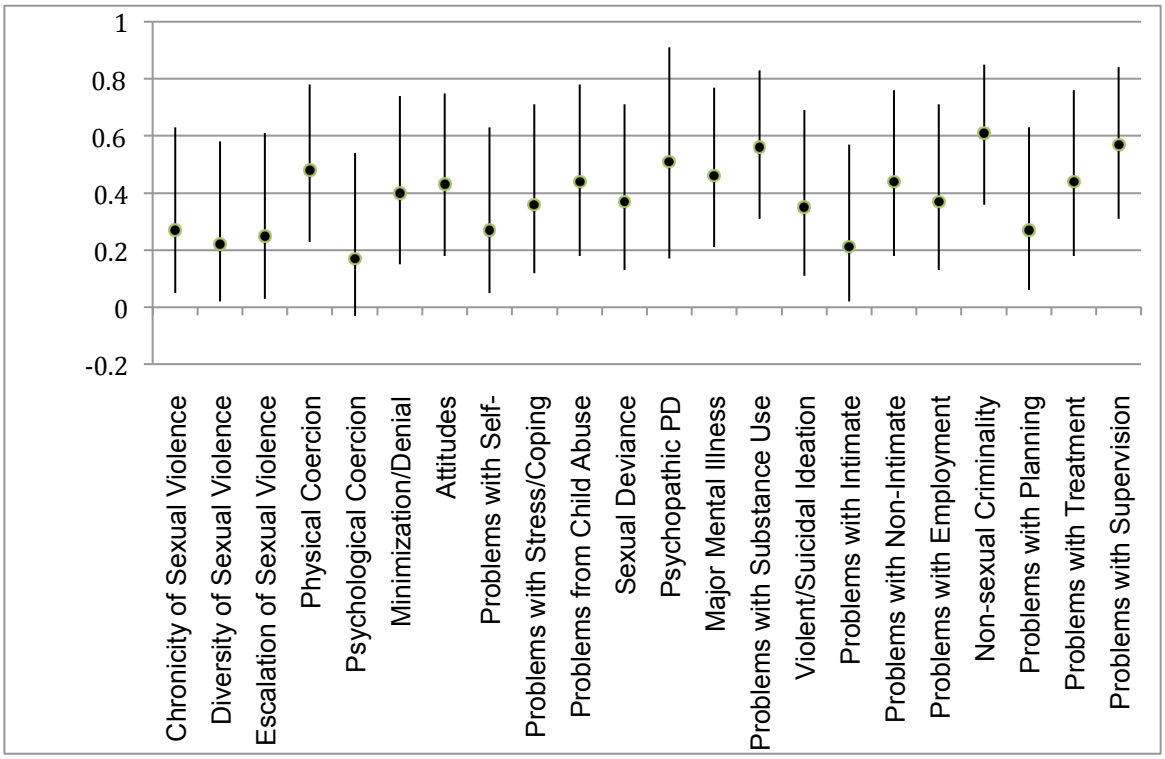



***Table 3.6 Mean, Standard Deviation, and Intraclass Correlation Coefficients (ICCs) for RSVP Relevance Risk Factors for All Cases***

| RSVP Risk Item | Relevance Ratings | | | |
|---|---|---|---|---|
| | *M* (*SD*) | ICC | p | 95% |
| 1. Chronicity of Sexual Violence | 1.50 (.43) | .27 | .005 | .058, .633 |
| 2. Diversity of Sexual Violence | 1.07 (.47) | .22 | .012 | .023, .584 |
| 3. Escalation of Sexual Violence | 1.08 (.47) | .25 | .006 | .040, .609 |
| 4. Physical Coercion in Sexual Violence | 1.47 (.62) | .48 | <.001 | .222, .781 |
| 5. Psychological Coercion in Sexual Violence | 1.23 (.45) | .18 | .030 | -.006, .537 |
| **Sexual Violence History Domain** | | .28 | | |
| 6. Extreme Minimization/Denial of Sexual Violence | 1.43 (.50) | .41 | <.001 | .159, .735 |
| 7. Attitudes that Support or Condone Sexual Violence | 1.23 (.50) | .43 | <.001 | .183, .754 |
| 8. Problems with Self-Awareness | 1.57 (.42) | .27 | .003 | .058, .632 |
| 9. Problems with Stress or Coping | 1.37 (.54) | .36 | <.001 | .125, .705 |
| 10. Problems Resulting from Child Abuse | 0.91 (.64) | .39 | <.001 | .148, .726 |
| **Psychological Adjustment Domain** | | .37 | | |

| RSVP Risk Item | Relevance Ratings | | | |
|---|---|---|---|---|
| | *M* (*SD*) | ICC | p | 95% |
| 11. Sexual Deviance | 1.38 (.48) | .37 | <.001 | .132, .711 |
| 12. Psychopathic Personality Disorder | 1.30 (.64) | .46 | <.001 | .206, .770 |
| 13. Major Mental Illness | 0.78 (.67) | .46 | <.001 | .211, .774 |
| 14. Problems with Substance Use | 1.22 (.65) | .55 | <.001 | .298, .825 |
| 15. Violent or Suicidal Ideation | 1.12 (.55) | .35 | <.001 | .111, .691 |
| **Mental Disorder Domain** | | .44 | | |
| 16. Problems with Intimate Relationships | 1.67 (.35) | .20 | .018 | .010, .564 |
| 17. Problems with Non-Intimate Relationships | 1.50 (.54) | .44 | <.001 | .187, .757 |
| 18. Problems with Employment | 1.32 (.52) | .37 | <.001 | .131, .711 |
| 19. Non-sexual Criminality | 1.03 (.68) | .61 | <.001 | .355, .852 |
| **Social Adjustment Domain** | | .41 | | |
| 20. Problems with Planning | 1.55 (.44) | .27 | .003 | .058, .633 |
| 21. Problems with Treatment | 1.37 (.51) | .44 | <.001 | .193, .761 |
| 22. Problems with Supervision | 1.45 (.56) | .58 | <.001 | .320, .836 |
| **Manageability Domain** | | .43 | | |
| **Overall Mean ICC** | | .39 | | |

***Figure 3.2 Intraclass Correlation Coefficients (ICCs) and 95% Confidence Intervals for RSVP Relevance Items***




With the exception of the Sexual Violence History domain (Presence ICC = .07; Relevance ICC = .12), the interrater reliability of the Domain scores was fair to excellent for Presence (Table 3.7) and Relevance ratings (Table 3.8) (Range ICC = .43 to .78). Total scores for Presence and Relevance ratings had fair interrater reliability (Presence ICC = .56, *p* = .000, 95% CI [.310, .829]; Relevance ICC = .55, *p* = .000, 95% CI [.292, .821]) (Figure 3.3).

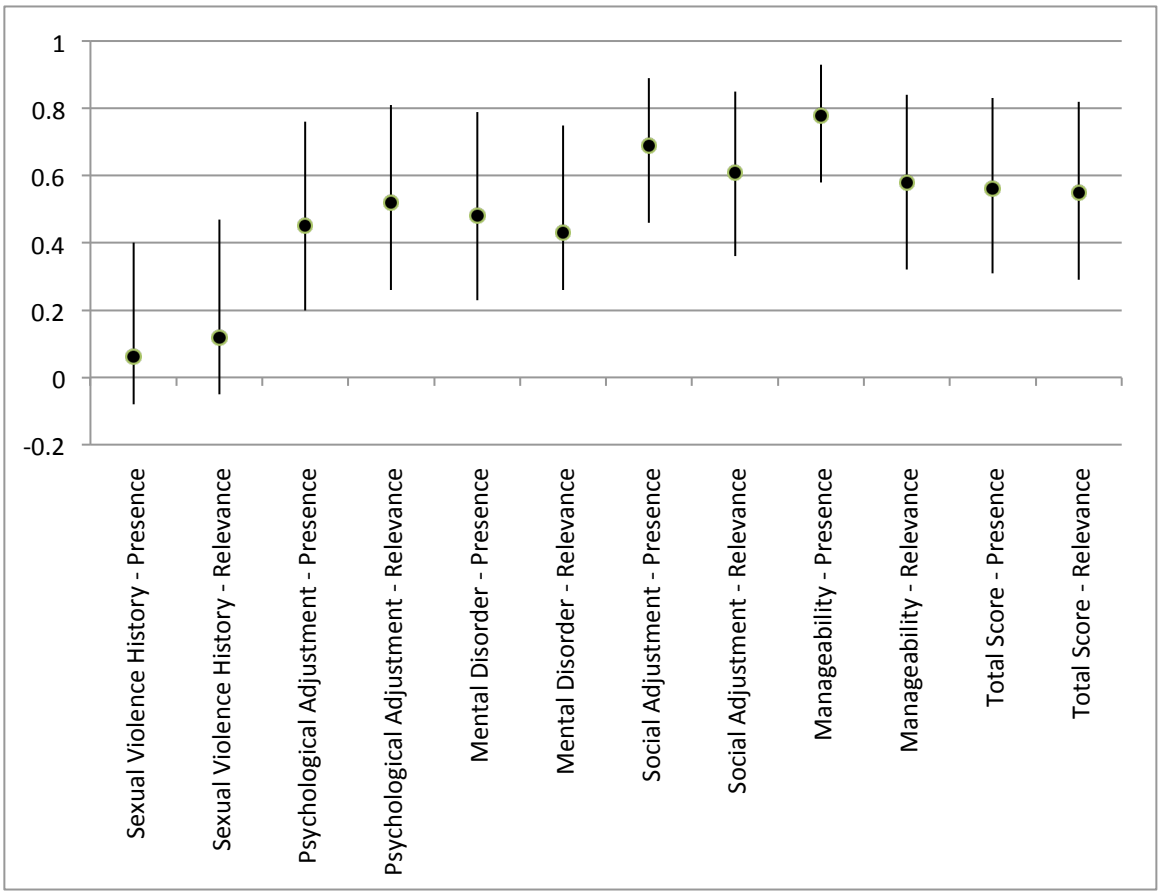
### *Table 3.7 Mean, Standard Deviation, and Intraclass Correlation Coefficients (ICCs) for Presence RSVP Domain and Total Scores*

| Domain | Presence Ratings | | | |
|---|---|---|---|---|
| | *M* (*SD*) | ICC | p | 95% CI |
| Sexual Violence History | 6.92 (0.94) | .07 | .204 | -.072, .396 |
| Psychological Adjustment | 7.73 (1.49) | .45 | <.001 | .196, .764 |
| Mental Disorder | 6.13 (1.61) | .48 | <.001 | .227, .784 |
| Social Adjustment | 6.75 (1.84) | .69 | <.001 | .457, .891 |
| Manageability | 5.22 (1.46) | .78 | <.001 | .585, .928 |
| **Total Score** | 32.75 (5.80) | .56 | <.001 | .305, .829 |

### *Table 3.8 Mean, Standard Deviation, and Intraclass Correlation Coefficients (ICCs) for Relevance RSVP Domain and Total Scores*

| Domain | Relevance Ratings | | | |
|---|---|---|---|---|
| | *M* (*SD*) | ICC | p | 95% CI |
| Sexual Violence History | 6.35 (1.11) | .12 | .87 | -.041, .471 |
| Psychological Adjustment | 6.48 (1.74) | .53 | <.001 | .269, .810 |
| Mental Disorder | 5.35 (1.42) | .43 | .<.001 | .179, .751 |
| Social Adjustment | 5.52 (1.88) | .60 | <.001 | .352, .851 |
| Manageability | 4.37 (1.46) | .58 | <.001 | .327, .839 |
| **Total Score** | 28.07 (5.84) | .55 | <.001 | .290, .821 |

***Figure 3.3 Intraclass Correlation Coefficients (ICCs) and 95% Confidence Intervals for RSVP Presence and Relevance Domains and Total Scores***

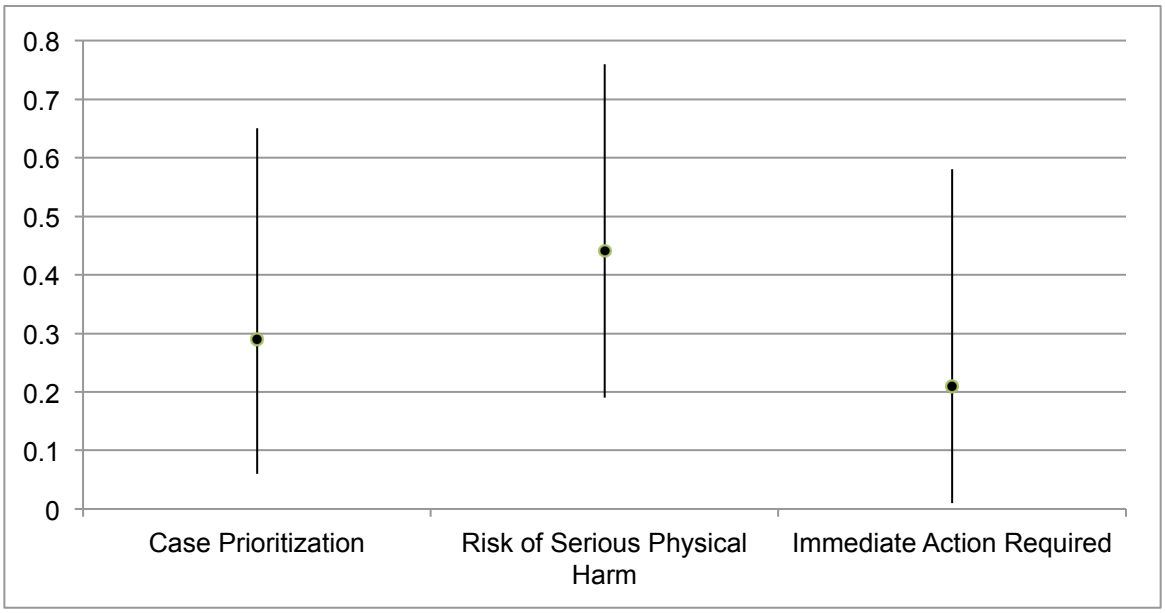



Finally, the interrater agreement for the summary risk ratings Case Prioritization, Serious Physical Harm and Imminence was examined (Table 3.9; Figure 3.4). Case Prioritization had poor reliability, ICC = .29, *p* = .002, 95% CI [.071, .648], as did Imminence, ICC = .21, *p* = .014, 95% CI [.017, .576]. Ratings for Level of Serious Physical Harm had fair agreement, ICC = .44, *p* = .000, 95% CI [.192, .761].

***Table 3.9 Mean, Standard Deviation and Intraclass Correlation Coefficients for RSVP Summary Judgments***

| Summary Judgment | *M* (*SD*) | ICC | p | 95% CI |
|---|---|---|---|---|
| Case Prioritization | 1.55 (.39) | .29 | .002 | .071, .648 |
| Risk of Serious Physical Harm | 1.02 (.56) | .44 | <.001 | .192, .761 |
| Immediate Action Required | 1.11 (.44) | .21 | .014 | .017, .576 |

***Figure 3.4 Intraclass Correlation Coefficients (ICCs) and 95% Confidence Intervals for RSVP Summary Judgments***




To investigate whether any case impacted the reliability findings, analyses were replicated by removing one case (i.e., analyses were repeated 10 times, each time removing a different case). Reliability remained stable across all analyses.

### *3.3.2. Agreement Between Raters and Gold Standard*

Using another approach to examine the reliability and quality of participants' assessments, analyses were conducted to examine the level of agreement between participants' RSVP assessments and experts' consensus gold standard ratings. To do so, percent agreement [4] was calculated for individual items and summary risk judgments for all cases combined, as well as by individual case. T-tests were conducted to compare mean participant and expert domain and total scores.

[4] Percent agreement was calculated by computing a difference score between Participant and Expert ratings for each item (i.e., Participant rating – Expert rating = Difference score) and calculating the number of '0' difference scores (i.e., number of 0 scores/number of 0 scores + number of other difference scores)

Agreement was better for presence ratings than for relevance ratings. The average item agreement across all cases for Presence ratings was 76% and for Relevance ratings was 64% (Table 3.10). Percent agreement for Presence item ratings ranged from 54% for Escalation (Item 3) to 93% for Problems with Supervision (Item 22). For Relevance items, percent agreement ranged from 49% for Attitudes that Support or Condone Sexual Violence (Item 3) to 79% for Problems with Intimate Relationships (Item 16). The item Psychopathic Personality Disorder was unique in that the agreement for the Relevance rating was greater than that of the Presence rating (65% vs. 60%).

***Table 3.10 Percent Agreement between Participant Ratings and Expert Ratings for RSVP Items Across All Cases***

| RSVP Item | % Agreement Across all Cases | |
|---|---|---|
| | Presence Rating | Relevance Rating |
| 1. Chronicity of Sexual Violence | 77 | 69 |
| 2. Diversity of Sexual Violence | 55 | 54 |
| 3. Escalation of Sexual Violence | 54 | 52 |
| 4. Physical Coercion in Sexual Violence | 82 | 78 |
| 5. Psychological Coercion in Sexual Violence | 71 | 60 |
| **Sexual Violence History Domain** | 68 | 63 |
| 6. Extreme Minimization/Denial of Sexual Violence | 79 | 64 |
| 7. Attitudes that Support or Condone Sexual Violence | 75 | 49 |
| 8. Problems with Self-Awareness | 75 | 72 |
| 9. Problems with Stress or Coping | 80 | 70 |
| 10. Problems Resulting from Child Abuse | 73 | 56 |
| **Psychological Adjustment Domain** | 76 | 62 |
| 11. Sexual Deviance | 62 | 57 |
| 12. Psychopathic Personality Disorder | 60 | 65 |
| 13. Major Mental Illness | 68 | 65 |
| 14. Problems with Substance Use | 85 | 57 |
| 15. Violent or Suicidal Ideation | 66 | 56 |
| **Mental Disorder Domain** | 68 | 60 |

| | % Agreement Across all Cases | |
|---|---|---|
| **RSVP Item** | **Presence Rating** | **Relevance Rating** |
| 16. Problems with Intimate Relationships | 91 | 79 |
| 17. Problems with Non-Intimate Relationships | 82 | 67 |
| 18. Problems with Employment | 86 | 59 |
| 19. Non-sexual Criminality | 82 | 64 |
| **Social Adjustment Domain** | 85 | 67 |
| 20. Problems with Planning | 85 | 71 |
| 21. Problems with Treatment | 88 | 74 |
| 22. Problems with Supervision | 93 | 63 |
| **Manageability Domain** | 89 | 69 |
| **Mean Percent Agreement** | 76 | 64 |

One sample *t*-tests were conducted to examine whether the participant mean Domain and Total scores differed significantly from mean Domain and Total scores of the experts (Table 3.11). Results showed that across all cases, experts had significantly higher scores for the Sexual Violence History (Presence and Relevance), Psychological Adjustment (Presence and Relevance), and Social Adjustment (Relevance) Domains. Experts also had significantly higher Total scores (Presence and Relevance). Effect sizes (*d*) were interpreted following Cohen's (1988) guidelines, where $d = 0.2$ is considered a small effect size, $d = 0.5$ is a medium effect size and $d = 0.8$ is a large effect size. The effect sizes for difference between mean participant and expert scores were mostly in the small to medium range.

***Table 3.11 Comparison of Participant and Expert Mean Domain and Total scores across all Cases***

| | | Participant Score | Expert Score | t | p | d |
|---|---|---|---|---|---|---|
| | | M (SD) | M (SD) | | | |
| **Sexual Violence History** | Presence | 6.90 (1.95) | 7.70 (1.24) | -4.13 | <.001 | .49 |
| | Relevance | 6.43 (2.02) | 7.08 (1.09) | -3.25 | .002 | .40 |
| **Psychological Adjustment** | Presence | 7.98 (2.00) | 8.59 (1.97) | -3.07 | .003 | .31 |
| | Relevance | 6.67 (2.21) | 7.81 (1.90) | -5.22 | <.001 | .55 |
| **Mental Disorder** | Presence | 6.43 (1.98) | 6.25 (1.59) | 0.92 | .358 | .10 |
| | Relevance | 5.60 (1.91) | 5.31 (1.62) | 1.52 | .132 | .16 |

| | | Participant Score | Expert Score | t | p | d |
|---|---|---|---|---|---|---|
| | | M (SD) | M (SD) | | | |
| **Social Adjustment** | Presence | 6.96 (1.84) | 7.12 (1.59) | -0.87 | .384 | .09 |
| | Relevance | 5.80 (2.09) | 6.43 (1.49) | -3.02 | .003 | .35 |
| **Manageability** | Presence | 5.27 (1.43) | 5.53 (1.53) | -0.18 | .074 | .18 |
| | Relevance | 4.44 (1.82) | 4.78 (1.61) | -1.88 | .062 | .20 |
| **Total Score** | Presence | 33.55 (6.64) | 35.18 (5.78) | -2.48 | .015 | .26 |
| | Relevance | 28.94 (7.21) | 31.42 (5.52) | -3.47 | .001 | .39 |

The percent agreement for summary risk ratings of Case Prioritization, Level of Harm and Imminence were examined (Table 3.12). Results showed similar agreement between participants and experts for Case Prioritization (59%) and Level of Harm (59%). There was poorer agreement for Imminence, with only 38% agreement between participants and experts across all cases.

***Table 3.12 Percent Agreement between Participant and Expert Ratings for Summary Judgments for All Cases***

| Summary Judgment | % Agreement |
|---|---|
| Case Prioritization | 59 |
| Risk of Serious Physical Harm | 59 |
| Immediate Action Required | 38 |

For the individual cases (Appendix I), Case 1 had the highest agreement for Presence ratings, with an average of 91% agreement across all items. Agreement for Presence ratings for Case 1 ranged from 56% (Psychological Coercion, Problems Resulting from Child Abuse) to 100% for 13 items (Physical Coercion in Sexual Violence, Extreme Minimization or Denial of Sexual Violence, Attitudes that Support or Condone Sexual Violence, Problems with Self-Awareness, Sexual Deviance, Psychopathic Personality Disorder, Problems with Substance Use, Violent or Suicidal Ideation, Problems with Intimate Relationships, Problems with Non-intimate Relationships, Problems with Employment, Non-Sexual Criminality, Problems with Supervision). Case 5 had the lowest average agreement for Presence ratings (65%), with item agreement ranging from 11% (Psychopathic Personality Disorder) to 100%

(Physical Coercion, Non-sexual Criminality, Problems with Supervision). For Case 5, six of the 22 items had rates of agreement below 50%. For Relevance ratings, Case 10 had the greatest agreement (73%). For Case 10, percent agreement ranged from 29% (Attitudes that Support or Condone Sexual Violence) to 100% (Problems with Planning). Again, Case 5 had the lowest agreement (53%). Agreement for items for Case 5 ranged from 22% (Problems Resulting from Child Abuse, Sexual Deviance, Psychopathic Personality Disorder) to 89% (Diversity of Sexual Violence, Problems with Supervision), and 12 items had an agreement level below 50%.

One sample *t*-tests were conducted to examine whether Domain and Total scores differed significantly between participant and expert ratings for individual cases. Of note, Case 5 had a number of significant differences between participant and expert scores. With the exception of Social Adjustment (Presence and Relevance) and Manageability (Relevance) Domain scores, participants' Domain and Total scores differed significantly from experts' scores. In contrast, Cases 1, 4, and 6 each had only one significant difference in mean Domain scores between participant and expert scores.

## Supplementary Analyses: Individual Rater Interrater Reliability

Additional analyses were conducted to examine the performance of individual participants to investigate whether any participant performed poorly relative to the gold standard assessments. To do this, ICCs were calculated for each rater by comparing the agreement between participant and expert Presence Total scores, Relevance Total scores and Case Prioritization ratings (Table 3.13). Across all raters and ratings, 24% (*n* = 12) demonstrated excellent reliability, 35% (*n* = 18) demonstrated good reliability, 16% (*n* = 8) demonstrated fair reliability, and 25% (*n* = 13) demonstrated poor reliability. Visual inspection of Table 3.13 demonstrates the variability in the reliability across Presence and Relevance scores and summary risk ratings. Whereas some raters had good to excellent reliability for the Presence and Relevance scores and poor to fair reliability for Case Prioritization summary risk rating, other participants had the opposite pattern.

***Table 3.13 Intraclass Correlation Coefficients (ICC) for Individual Raters***

| Rater ID | Total Presence Score | | | Total Relevance Score | | | Case Prioritization | | |
|---|---|---|---|---|---|---|---|---|---|
| | ICC | p | 95% CI | ICC | p | 95% CI | ICC | p | 95% CI |
| 1 | .70 | .046 | -.106, .952 | .67 | .060 | -.210, .947 | .71 | .038 | -.019, .952 |
| 2 | .68 | .059 | -.247, .950 | .64 | .074 | -.263, .940 | .29 | .294 | -.852, .867 |
| 3 | .55 | .118 | -.445, .924 | .65 | .067 | -.232, .943 | .71 | .038 | -.019, .952 |
| 4 | .22 | .343 | -.901, .849 | .64 | .079 | -.353, .942 | .75 | .017 | .101, .959 |
| 5 | .90 | .001 | .401, .985 | .49 | .004 | -.059, .903 | -.25 | .739 | -.740, .590 |
| 6 | .39 | .191 | .401, .985 | .56 | .102 | -.303, .922 | .83 | .011 | .277, .973 |
| 7 | .70 | .013 | -.058, .951 | .34 | .155 | -.231, .846 | 1.00 | | |
| 8 | .64 | .011 | -.117, .941 | .81 | .002 | -.050, .973 | .47 | .141 | -.372, .903 |
| 9 | .96 | .000 | .779, .994 | .96 | .000 | .786, .995 | .29 | .294 | -.852, .867 |
| 10 | .70 | .035 | .000, .950 | .62 | .084 | -.320, .936 | 1.00 | | |
| 11 | .51 | .012 | -.102, .907 | .61 | .000 | -.025, .938 | 1.00 | | |
| 12 | .35 | .124 | -.193, .846 | .52 | .078 | -.171, .908 | .00 | .500 | -.951, .777 |
| 13 | .43 | .184 | -.548, .897 | .67 | .064 | -.273, .948 | .21 | .251 | -.263, .786 |
| 14 | .86 | .009 | .280, .979 | .60 | .065 | -.135, 928 | .21 | .251 | -.263, .786 |
| 15 | .95 | .000 | .641, .993 | .88 | .001 | .066, .984 | .29 | .294 | -.852, .867 |
| 16 | .77 | .006 | .002, .966 | .73 | .001 | -.083, .961 | .71 | .038 | -.019, .952 |
| 17 | .52 | .116 | -.311, .913 | .32 | .237 | -.493, .862 | -.43 | .767 | -1.263, .607 |

An average ICC was calculated for each rater. One rater (6%) achieved excellent reliability, eight raters (47%) demonstrated good reliability, five raters (29%) demonstrated fair reliability, and three raters (18%) demonstrated poor reliability. Two of the three poor raters demonstrated poor to fair reliability across all three analyses: Rater 12 had poor reliability for Presence (ICC = .35) and Case Prioritization (ICC = .00) ratings and fair reliability for Relevance ratings (ICC = .52) and Rater 17 had poor reliability for Relevance (ICC = .32) and Case Prioritization (ICC = -.43) ratings and fair reliability for Presence ratings (ICC = .52). The third rater who demonstrated poor reliability, on average, had excellent reliability for Total Presence ratings, fair reliability for Relevance ratings and poor reliability for Case Prioritization. Interrater reliability analyses were conducted after removing or replacing raters 12 and 17. Due to the case distribution, analyses were conducted with five assessments per case. The removal of

these two raters did not have a significant impact on reliability, the only notable change was Case Prioritization increased from ICC = .29 to .42.

Correlation analyses were conducted to examine the association between participant ICCs and years of forensic risk assessment experience, the estimated number of sexual violence risk assessments completed per year, and Part 1 questionnaire score. No significant associations were found (Table 3.14).

***Table 3.14 Correlation between Individual Intraclass Correlations Coefficients (ICCs) and Participant Experience***

| Experience | Presence Rating *r* (*p*) | Relevance Rating *r* (*p*) | Case Prioritization *r* (*p*) |
|---|---|---|---|
| Years of Experience in Current Profession | .11 (.698) | -.31 (.960) | .26 (.339) |
| Years of Forensic Risk Assessment Experience | .01 (.960) | -.39 (.148) | .30 (.283) |
| Number of Sexual Violence Risk Assessments per year | .26 (.336) | -.21 (.444) | .05 (.857) |
| Part 1 Questionnaire Score | -.02 (.936) | .05 (.840) | -.47 (.059) |

Although a few participants did not achieve a moderate level of agreement with expert ratings, the removal of these raters did not have any substantial impact on rater agreement across cases and therefore, all raters were retained for all analyses including case formulation analyses (below).

## Summary of Interrater Reliability Findings

To summarize the above findings, overall the RSVP demonstrated moderate reliability. The ICCs found in this study were smaller than those found in previous studies; however, the lower levels of agreement in this study were not unexpected. The current study involved novice evaluators recently trained to use the RSVP, compared to previous research that has used two highly trained evaluators. Despite the limitations of the approach used in the current study, there was a moderate level of agreement across participants, as well as between participants and experts, for RSVP ratings and summary judgments. Based on these findings, it was possible to proceed to the next stage of the study: examining the reliability of case formulations. Had the results from the interrater reliability analyses been overwhelmingly poor, it would have been difficult

to expect participants would be able to develop reliable formulations. Fortunately, the findings indicated there was a moderate level of reliability and it was expected participants could develop reliable formulations as well.

## 3.4. Interrater Reliability: Case Formulations

The reliability of case formulation was investigated by comparing the similarity ratings (evaluated by Research Assistants) for pairs of formulations developed for the same case by different raters (Within Case) with pairs of formulations developed for different cases by different raters (Across Case). Analyses were conducted separately for the overall similarity ratings for formulation, scenarios and management strategies, as well as for the specific ratings within each section (i.e., within formulation, scenarios and management). For each set of analyses, mean inter-item correlations were calculated, followed by a multivariate analysis of variance (MANOVA) to determine whether group differences existed, and a univariate analysis of variance (ANOVA) to determine where the differences occurred. Effect sizes were indexed by multivariate $\eta^2$ and were interpreted by the following guidelines: multivariate $\eta^2$ = .01 is a small effect size, multivariate $\eta^2$ = .06 is a medium effect size and multivariate $\eta^2$ = .14 is a large effect size (Cohen, 1988).

### *3.4.1. Similarity of Overall Judgments of Case Formulation, Scenarios and Management Strategies*

First, the similarity of overall judgments of case formulation, scenarios and management was examined. The mean inter-item correlation for the overall formulation, scenario, and management strategies was .25. As the dependent variables were moderately correlated with each other, a MANOVA was calculated to determine whether there were significant differences between the Within Case and Across Case groups for case formulation, case scenario, and case management similarity ratings. There was a statistically significant difference in similarity ratings based on the pairing group, Wilk's $\Lambda$ = .770, $F(3, 134) = 13.32$, $p < .001$, multivariate $\eta^2$ = .230.

Analysis of variance (ANOVA) was conducted as a follow-up test and showed there was a statistically significant difference between the groups for case scenario similarity ratings, $F(1, 136) = 34.34$, $p < .001$, multivariate $\eta^2 = .202$, and for case management similarity ratings, $F(1, 136) = 16.56$, $p < .001$, multivariate $\eta^2 = .109$ (Table 3.15). After applying a Bonferroni correction (accepting statistical significant at $p < .017$), differences in formulation did not reach statistical significance, $F(1, 136) = 4.10$, $p = .045$, multivariate $\eta^2 = .029$. The Within Case similarity ratings were greater than the Across Case group ratings and multivariate $\eta^2$ indicated there was a large effect for case scenario group differences and a medium effect size for case management group differences.

***Table 3.15 Comparisons between Within Case and Across Case Groups for Case Formulation, Scenarios and Management Similarity Ratings***

| | Within Case *M* (*SD*) | Across Case *M* (*SD*) | *F* | *p* | multivariate $\eta^2$ |
|---|---|---|---|---|---|
| Case Formulation | 5.96 (1.73) | 5.37 (1.68) | 4.10 | .045 | .029 |
| Case Scenarios | 6.51 (1.54) | 5.01 (1.47) | 34.34 | <.001 | .202 |
| Case Management | 5.46 (1.50) | 4.46 (1.41) | 16.56 | <.001 | .109 |

The results showed that, overall, participants developed scenarios and management plans that showed some similarity when developed for the same case, and differed appropriately when different cases were involved. In general, the similarity ratings for the Within Case group were slightly lower than expected. In particular, the Within Group similarity ratings for case management were only slightly higher than the scale middle score, indicating that there were often important differences between the management strategies developed even when developed for the same case. Nevertheless, the results showed that participants showed modest reliability for this step of the evaluation process. The next step was to examine the specific ratings for each section to investigate whether certain aspects of formulation were more reliable than others.

### 3.4.2. Similarity of Case Formulation: Motivators, Disinhibitors and Destabilizers

The mean inter-item correlation for motivators, disinhibitors and destabilizers was .28. The results of the MANOVA showed there was a statistically significant difference between the Within and Across Case groups, Wilk's $\Lambda$ = .898, $F(3, 134) = 5.05$, $p = .002$, multivariate $\eta^2 = .102$. The univariate ANOVA follow-up test showed a statistically significant difference for destabilizers, $F(1, 136) = 12.43$, $p = .001$, multivariate $\eta^2 = .084$. There was no significant difference for motivators, $F(1, 136) = 3.33$, $p = .070$, multivariate $\eta^2 = .024$, or disinhibitors, $F(1, 136) = 0.12$, $p = .732$, multivariate $\eta^2 = .001$ (Table 3.16).

***Table 3.16 Comparisons between Within Case and Across Case Groups for Motivator, Disinhibitor and Destabilizer Similarity Ratings***

| | Within Case *M* (*SD*) | Across Case *M* (*SD*) | *F* | *p* | multivariate $\eta^2$ |
|---|---|---|---|---|---|
| Motivators | 5.44 (2.28) | 4.73 (2.30) | 3.33 | .070 | .024 |
| Disinhibitors | 5.64 (2.35) | 5.51 (2.37) | 0.12 | .732 | .001 |
| Destabilizers | 6.44 (2.42) | 4.98 (2.45) | 12.43 | .001 | .084 |

Within this aspect of case formulation, only destabilizers statistically differed between Within and Across groups, with a medium effect size. The Within Case similarity ratings were greater than Across Case ratings. With respect to disinhibitors, there was little difference between the groups for the mean score and for both groups the mean score was near the scale midpoint, suggesting for this aspect of formulation there were key similarities and differences regardless of whether the formulations were for the same case or different cases. These findings may be a function of the cases involved or related to the training and understanding of the participants, but at this time it is not possible to determine why there were such similarities and a lack of distinction between groups.

### *3.4.3. Similarity of Case Scenarios: Type of Violence, Victims, Level of Harm and Imminence*

For the case scenario specific ratings, the mean inter-item correlation was .28. The MANOVA showed a statistically significant difference in similarity ratings between the Within Case and Across Case groups, Wilk's $\Lambda = .840$, $F(4, 133) = 6.34$, $p < .001$, multivariate $\eta^2 = .160$. The follow-up univariate ANOVAs showed there was a significant difference between groups for type of violence, $F(1, 136) = 16.65$, $p < .001$, multivariate $\eta^2 = .109$, victims, $F(1, 136) = 15.76$, $p < .001$, multivariate $\eta^2 = .104$, and level of harm, $F(1, 136) = 7.74$, $p < .001$, multivariate $\eta^2 = .054$ (Table 3.17). No significant differences between groups were found for imminence of violence. Again, as expected, the Within Case similarity ratings were greater than the Across Group ratings. Type of violence and victim had medium to large effect sizes, and level of harm had a small effect size.

***Table 3.17 Comparisons between Within Case and Across Case Groups for Type of Violence, Victim, Level of Harm and Imminence Similarity Ratings***

| | Within Case *M* (*SD*) | Across Case *M* (*SD*) | *F* | *p* | multivariate $\eta^2$ |
|---|---|---|---|---|---|
| Type of Violence | 6.22 (2.13) | 4.78 (2.05) | 16.65 | < .001 | .109 |
| Victim | 6.17 (2.39) | 4.51 (2.52) | 15.76 | < .001 | .104 |
| Harm | 6.71 (1.73) | 5.81 (2.05) | 7.74 | < .001 | .054 |
| Imminence | 4.83 (2.67) | 4.87 (2.63) | 0.01 | .923 | .000 |

The results indicate evaluators are capable of agreeing on the likely type of violence, the likely victims, and the level of harm when assessing the same case, but a decision such as when the individual is likely to commit an offense (i.e., imminence) was more difficult. In fact, there was little difference between the groups for imminence ratings and for both groups the similarity ratings were below 5, indicating there were more differences in the assessment of imminence than similarities for both groups. Again, the lack of difference for imminence could reflect issues with the cases, training, or some other issue currently unaccounted for.

### *3.4.4. Similarity of Case Management: Monitoring, Treatment, Supervision, and Victim Safety Planning*

Finally, for the case management section, the mean inter-item correlation was .37. There was a statistically significant difference between the Within Case and Across Case groups for case management similarity ratings, Wilks' Λ = .856, $F(4, 133) = 5.58$, $p < .001$, multivariate $\eta^2 = .144$.

Univariate ANOVAs (Table 3.18) showed the groups differed significantly for treatment, $F(1, 136) = 8.80$, $p < .004$, multivariate $\eta^2 = .061$, and victim safety planning, $F(1, 136) = 15.45$, $p = .001$, multivariate $\eta^2 = .029$. Applying a Bonferroni correction (accepting statistical significant at $p < .013$), differences in monitoring, $F(1, 136) = 5.20$, $p = .024$, multivariate $\eta^2 = .037$ and supervision did not reach statistical significance, $F(1, 136) = 4.06$, $p = .046$, multivariate $\eta^2 = .029$. The Within Case group had significantly higher similarity ratings than the Across group, and the effect size for treatment and victim safety planning was medium.

***Table 3.18 Comparisons between Within Case and Across Case Groups for Monitoring, Treatment Supervision and Victim Safety Planning Similarity Ratings***

| | Within Case *M* (*SD*) | Across Case *M* (*SD*) | *F* | *p* | multivariate $\eta^2$ |
|---|---|---|---|---|---|
| Monitoring | 4.71 (2.13) | 3.91 (2.02) | 5.20 | .024 | .037 |
| Treatment | 5.61 (2.05) | 4.58 (2.03) | 8.80 | .004 | .061 |
| Supervision | 4.41 (2.12) | 3.71 (1.98) | 4.06 | .046 | .029 |
| Victim Safety Planning | 5.86 (2.32) | 4.33 (2.23) | 15.45 | < .001 | .102 |

Consistent with above findings, differences between the groups were not found consistently across all case management ratings. The Within Case group similarity ratings for monitoring and supervision were below 5 and therefore lower than expected. Treatment and victim safety planning Within Case group mean scores were slightly above the scale midpoint, again suggesting there were important differences in those formulations between raters, despite being developed from the same case material.

## Summary of Case Formulation Results

The above findings provide some evidence that case formulation (including scenarios and management strategies) can be similar, or consistent, when developed by two different evaluators who are assessing the same case (based on the same information). Significant differences between the groups were not found consistently across all ratings, so it appears that some aspects of formulation might be easier to understand or develop compared to others. The Within Case similarity ratings were, perhaps, lower than one might hope when two evaluators are assessing the same case (average ratings were all approximately 6.5 and below) indicating there are some important differences being noted between two formulations developed from the same case. However, the mostly low similarity ratings for the Across Case group, particularly for scenarios and management strategies, suggests there are clear differences in the formulations developed for different cases and that evaluators are considering some case-specific details when developing their formulations. Of note, Wilk's $\Lambda$ was quite high in the above analyses, indicating that there was a great amount of variance not explained by the group factor. The results suggest that it is not simply the Within Case/Across Case group that is influencing the ratings and there are other sources of variance, possibly including cases, participants (i.e., evaluators), the Research Assistants or any combination of these. This study cannot clarify the sources of variance. Overall, the findings are promising and suggest formulations, which are based on subjective decisions, can be reliably developed. Further issues and the implications of these findings will be discussed in the next section.

# 4. Discussion

The RSVP is a potentially useful and important tool available to professionals tasked with the responsibility of assessing and managing sexual violence risk. Although the RSVP is widely used internationally (Hart & Boer, 2010), there is limited direct empirical support for the measure. The purpose of this study was to add to the small empirical base supporting the RSVP and address an important gap in the literature. First, the study examined the interrater reliability of RSVP assessments completed by professionals trained on the measure. Second, the study examined the quality of the assessments completed by participants, relative to the gold standard assessments completed by expert RSVP raters. The results from these analyses provided the opportunity to proceed with an examination of case formulations, the first known study to investigate formulations completed in the manner intended. The rest of this section will discuss the primary research findings and related implications, as well as the limitations of the current study.

## 4.1. Competency Evaluation

Participants performed very well on the Part 1 questionnaire. The purpose of the questionnaire was to confirm that trainees had completed the training videos and had gained a basic level of understanding of the material presented in the training. The results suggested the participants had an acceptable understanding of the SPJ risk assessment approach and the process for administering the RSVP. Further, the results suggested the participants were prepared to continue with the practical component of the training.

Unfortunately, there was a limited range of scores on the questionnaire and the questionnaire might not have been sufficiently difficult to distinguish between participants who understood the material and training from those who, in fact, had difficulty with the

material. Future research should ensure the questionnaire used to assess an evaluator's level of understanding is appropriately difficult to make this distinction. The evaluation procedure should also focus on measure-specific issues, such as item definitions. The questionnaire developed for this study included broader, conceptual-type questions. A participant with experience with the SPJ approach to risk assessment might have been able to answer the questions adequately, but might not have had a good understanding of RSVP-specific issues.

Not only are these types of issues relevant to future research, they are applicable to practice as well. In practice, when providing training to a group of evaluators it is important to be able to identify individuals who are not truly understanding the concepts or process being discussed. Before a professional can proceed to conducting *true* risk assessments, it is important they are able to demonstrate a certain level of knowledge and expertise in the area of risk assessment and the specific approach being used. An appropriate evaluation of competency would be useful to ensure evaluators possess the needed level of understanding and for identifying those who require further training.

## 4.2. Case Descriptives

The results from this section showed that all cases were at elevated risk; that is either moderate or high risk. The goal was to include a diverse range of cases, but in terms of case risk level, this was not necessarily achieved. The lack of diversity in risk level was likely a reflection of how these cases were obtained. All cases were obtained from Dr. Hart's clinical practice and his expertise is requested for cases that are more difficult and high risk; he is not asked to consult for low risk, routine cases. Further, there was no attempt to alter the cases to adjust risk level. Fortunately, although all cases were assessed as moderate or high risk in the summary judgments, there was some variability in the total scores across cases and the cases were acceptable for examining reliability issues; however, this lack of variability might have had an impact on the reliability findings (this will be discussed further below).

Sutherland et al. (2012) found that case factors influenced interrater reliability. For example, Sutherland found that cases that were at either extreme of the

risk/complexity spectrum (i.e., low risk/low complexity or high risk/high complexity) obtained the highest average agreement for all summary judgments. Although it appears that the cases in the current study mostly were at the high end of the spectrum, the complexity of the cases was not assessed and as such, case complexity is unknown. Future research should consider case complexity and risk level prior to including a case in the study. Taking this step would ensure that participants are provided with a diverse range of cases and would help to minimize the impact of low variability on reliability analyses.

The inclusion of diverse cases is important for training from a practical perspective as well. Comprehensive training will expose professionals to as many different types of cases as possible. When a professional proceeds to conducting assessments in a real world setting, it is likely that he or she will be faced with assessing a broad range of individuals and cases. It is important that the evaluator not only feels comfortable and confident when faced with a unique case, but that the evaluator has demonstrated an acceptable level of competence at handling different types of cases. From this perspective, it is important that those who develop risk assessment training courses take steps to confirm the trainees will be provided a diverse range of cases that will appropriately challenge the trainees and confirm they are prepared to conduct assessments as part of their professional responsibility.

## 4.3. Interrater Reliability

This study used a couple of approaches to examine the reliability of the RSVP: investigating agreement among raters and agreement between raters and experts (i.e., gold standard assessments). The following sections will discuss the findings separately, comparing them with previous research and discussing possible explanations for differences (and some similarities) between findings. Finally, the implications of these results taken together will be discussed.

### *4.3.1. Agreement among Raters*

With respect to agreement among raters, the ICCs in the current study were lower than those found in most previous studies (i.e., Hart, 2003; Jackson & Hart, 2008; Watt et al., 2006), but showed similarities with those found by Sutherland and colleagues (2012).[5] Whereas Hart (2003), Jackson and Hart (2008) and Watt et al. (2006) found good to excellent agreement for items, domains, and total scores; this study and Sutherland's study found fair levels of agreement for these same ratings. There was a notable difference in the agreement between this study and previous research for summary judgments. Previous research has found summary judgments to be in the good to excellent range for agreement, but the results of this study indicated poor agreement between raters.

There are a couple of possible explanations for why the levels of interrater agreement did not reach the expected level in the current study. First, this study employed a new training approach used by the course trainer for the first time. The training was conducted entirely online, with little personal contact: participants were required to watch training videos and conduct assessments independently, with no interaction with the trainer or opportunity for feedback. It is possible that the training was not sufficient to ensure the participants completely understood the RSVP. Although participants demonstrated appropriate levels of understanding according to the results from the Part 1 questionnaire, as discussed, the questionnaire might not have been adequately difficult to differentiate between participants who truly understood the training and measure and those who did not. Further, the questionnaire focused on broad conceptual issues related to conducting sexual violence risk assessments and using the RSVP; the questionnaire did not assess the participants' understanding of the individual RSVP items. In fact, the training course did not spend substantial time discussing the items or their definitions. Participants were required to read the manual and item definitions independently and to refer to the definitions when completing an assessment.

[5] Of note, previous research has calculated ICCs using a two-way mixed effects model (i.e., Hart, 2003; Jackson & Hart, 2008; Watt et al., 2006) or a two-way random effects model (i.e., Sutherland et al., 2012). The analytical approach in the previous studies is conceptually different from the one-way random effects model used in the current study, but using the different models does not substantially alter the value of the coefficients and findings can be compared across studies.

It is possible that participants did not fully understand the meaning of some items or what they were intending to address or capture, accounting for some differences in item ratings.

Examining the demographics of the participants in this study, the majority were relatively young professionals with limited forensic risk assessment experience: most had less than 10 years of experience and, on average, completed less than 15 sexual violence assessments per year. The limited experience of the participants might explain the low rater agreement for the summary judgments. Further, reliability was evaluated based on training cases and participants were in the process of learning how to use the RSVP with these cases. It is possible that participants demonstrated improvement in their assessments with practice; however, the design used in this study (i.e., cases were randomly assigned so the cases and the order in which cases were assigned differed across participants) does not provide the opportunity to compare reliability from earlier practice cases to later practice cases.

Another possible explanation for the low rater agreement compared to previous research is with respect to the study methodology employed. Reliability, in general, tends to be better under certain circumstances, such as when interviews are conducted, with good case files, and fewer raters. None of the research to date has involved interviews, but the studies conducted by Hart (2003), Jackson and Hart (2008), and Watt and colleagues (2006) included good case files, with fewer raters (i.e., two highly trained evaluators) assessing a large number of cases (50 to 90 cases). As discussed, the findings from this study are relatively consistent with those from Sutherland's study. This might not be that surprising as there were important similarities in the methodology used across these two studies, which differed from the above noted studies. Both studies involved training a greater number of professionals (i.e., 22 raters and 17 raters) on the use of the RSVP and included fewer cases (i.e., 6 cases and 10 cases). This approach might have limited our ability to obtain the high levels of agreement found in previous research, but might be a more accurate representation of agreement with respect to its use in practice.

Finally, sample size and low variability limitations of the current study might have had an impact on the interrater agreement. A small number of cases (10 cases with six

raters each) were included in the analyses and a few differences in scores could have a substantial impact on overall agreement levels. Characteristics of the cases themselves might have also had an impact. Although an effort was made to select cases that represented a wide range of issues and risk levels, most cases were assessed as high risk and lacked variability in item ratings. The interaction of low variability and a small sample size can result in intraclass correlation coefficients that shift drastically for small disagreements. It is likely that this impacted the findings for the current study.

### *4.3.2. Agreement Between Raters and Gold Standard*

Sutherland et al.'s (2012) study is the only one that has examined agreement between rater assessments and gold standard assessments for the RSVP. The average level of agreement across item ratings was similar across studies (65% for Past, Recent and Future combined for Sutherland and 76% for Presence and Relevance here), but the specific risk factors that achieved highest and lowest reliability differed. In Sutherland et al.'s study, Extreme Minimization/Denial, Problems resulting from Child Abuse, Major Mental Illness, Non-sexual Criminality and Problems with Treatment obtained agreement greater than 70% across all three ratings. Consistent with Sutherland's findings, Problems with Treatment also achieved greater than 70% agreement across Presence and Future ratings but the remaining items differed. The current study found high reliability for Physical Coercion in Sexual Violence, Problems with Self-Awareness, Problems with Stress or Coping, Problems with Intimate Relationships, and Problems with Planning. Items that achieved the least reliability (under 55% across all ratings) were Psychological Coercion, Violent or Suicidal Ideation, and Problems with Planning in Sutherland et al.'s study, but in the current study the items were Diversity of Sexual Violence and Escalation of Sexual Violence.

It is unclear why different 'high' and 'low' reliability items were found in the two studies. Case characteristics within each study might account for some of the differences. The vignettes developed for Sutherland et al.'s study were two to four pages in length and followed a structured organization, with the types of information remaining consistent across cases. Participants in this study were provided case files that ranged from 17 to 65 pages, with the structure and content of information differing from case to case depending on the information available in the original file. The cases included in the

current study were presented in this manner to reflect the types of case files one might expect for a *true* assessment. Sutherland's vignettes were designed to represent the type of case that might be encountered in a forensic mental health setting. In contrast, the cases selected for the current study represented individuals within correctional settings. Further, the cases in Sutherland's study reflected a broad range of clinical complexity and risk of sexual violence, whereas the current study included mostly high-risk cases, limiting the variability of the cases. It is possible that differences in the case files provided to participants resulted in certain items being more difficult, or straightforward, to rate.

With respect to summary judgments, the current study found low to moderate agreement for Case Prioritization (59%), Risk of Serious Physical Harm (59%) and Imminence (38%). The current study also included *t*-tests to investigate whether the mean Domain and Total scores of participants differed significantly from the mean scores of experts, across all cases. The results showed that mean Sexual Violence History (Presence and Relevance), Psychological Adjustment (Presence and Relevance), Social Adjustment (Relevance only) and Total Scores (Presence and Relevance) significantly differed between participants and experts, with experts having significantly higher scores than the participants. It is unclear as to why these differences and this pattern occurred and the current study does not provide the opportunity to clarify this issue. It is possible that the explanations discussed previously regarding training issues (e.g., lack of emphasis on item definitions) and the sample involved (i.e., less experienced) apply to these findings as well.

## Supplementary Analyses: Individual Rater Interrater Agreement

Supplementary analyses were conducted to examine whether any of the participants performed poorly on the assessments compared to the gold standard assessments. Intraclass correlations coefficients were calculated for each rater, comparing their case assessments with the ratings provided by the experts. Two raters demonstrated concerning levels of rater agreement, with ICCs for Total Presence and Relevance scores and Case Prioritization ratings under .60. There was no evidence that performance on the assessments was associated with years of forensic risk assessment experience or experience conducting sexual violence risk assessments, indicating that

there was no evidence that greater experience within the field was associated with better quality assessments (as related to expert assessments). Removal of these participants did not have a significant impact on group level reliability analyses. However, this is an area where an improved evaluation following Part 1 might have been useful. If the evaluation had been able to identify evaluators who did not fully understand the measure, additional training could have been provided and it is possible the reliability results would have improved.

### *4.3.3. Implications of Interrater Reliability Findings*

Overall, the results from the interrater analyses demonstrated a moderate level of agreement. The ability to demonstrate reliability was an important step within this study. If a moderate level of agreement had not been established, it would have been difficult to continue with an evaluation of the formulations. For example, if evaluators could not demonstrate a basic level of agreement for rating risk factors, which involves more objective decision-making, it would be difficult to expect that evaluators would be able to develop similar formulations, which tends to involve more subjective decision-making.

These findings have important implications for future research and practice as well. Compared to previous research, the methodology employed in this study enhances the ecological validity and it is possible that the approach used in previous research (i.e., two highly trained evaluators) resulted in an overestimate of reliability. Future research should use an approach similar to that used in this study and continue to move to more field studies, rather than lab studies, to gain an understanding of the reliability of RSVP assessments as conducted in the field. It is difficult to tease apart whether the level of agreement achieved in this study is related to the sample, the training, the cases, or an interaction between any of these issues. This sample was a relatively young and inexperienced group of evaluators. Future research could attempt to replicate this study using expert evaluators, but it would be difficult to determine whether the level of reliability is attributable to the measure being used or to the fact that established experts completed the assessments. There are adjustments to the training that could improve reliability. For example, if researchers use an online system, which certainly has advantages in terms of recruiting participants, efforts should be made to make the trainer available to all trainees to respond to questions and provide clarification on training

related issues. Finally, case selection should be completed carefully, ensuring that a diverse range of cases is included. It might also be useful to include a greater number of cases and let participants train and practice with the first set of cases, and providing a second set of cases that will be used for research purposes.

Future research should conduct power analyses to confirm an adequate sample size is obtained to detect moderate effect sizes. In this study, pre-investigation power analyses were completed but the focus was on interpreting formulation similarity findings. To obtain adequate power for the interrater analyses, a greater number of raters or cases might have been required. In this study, the decision for participants to complete six practice cases was practical. Past experience by the measure's authors suggested six cases would provide the participants sufficient opportunity to practice the assessments and achieve an appropriate level of competence, but would not overburden the trainees (and thereby prevent them from registering or dropping out). Future research would need to take the burden on participants into consideration as well, so recruiting a greater number of evaluators would be important.

Many of the above points regarding future research can be applied to practice as well. Training is an essential component for professionals who will be undertaking risk assessments in the field and establishing a high-quality training program is critical. The training in this study required approximately a half-day of instruction followed by six practice cases and a moderate level of reliability was obtained. In practice, training might need to be extended in order to achieve greater reliability. A more thorough and in-depth instructional component could help participants to understand the important characteristics and unique features of the measure. Experience with the training in this study demonstrated that certain item definitions were a concern for some participants. For example, some participants had difficulty with Item 6 *Extreme Minimization or Denial of Sexual Violence* as they had difficulty defining the term *Extreme*. Therefore, spending more time discussing the risk factors and definitions is an area that should be emphasized in training. In addition, a greater number of cases should improve evaluator ability. Further, the cases should be delivered in various formats. That is, instead of providing only file-based cases, it would be useful to provide practice cases that include an interview with the individual being assessed. Diversifying the cases, providing information from several sources, and distributing the case information through multiple

means (e.g., audio or videotapes) reflects the experience an evaluator will have when conducting an assessment in the field.

Another important consideration for practice and training is providing feedback. Participants in this study received written feedback for each case. Unfortunately, my review of all assessments submitted indicated that not all evaluators appropriately integrated the feedback provided and it was not possible to confirm participants reviewed and understood the feedback. When conducting training, it might be useful to provide in-person feedback for each case (or select cases) to confirm the evaluator attends to the feedback and understands the errors he or she made in the assessment. If the training is provided in an online format, as was done here, options such as teleconferencing or videoconferencing are available that would allow for direct person-to-person communication. Moving beyond training, providing ongoing support and supervision is important for guaranteeing that evaluators continue to maintain a level of standard that is acceptable for professional practice.

## 4.4. Interrater Reliability: Case Formulations

The current study is the second study to attempt to examine the reliability of case formulations using the RSVP, but there are important differences in the approach used in these two studies. Sutherland and colleagues (2012) opted to use forced-choice questions to examine interrater agreement of case formulations, scenarios and management. That is, participants in Sutherland et al.'s study were required to respond to a series of forced choice items regarding case scenarios and case management (e.g., Nature (type of offense): sexual breach of peace, indecent exposure, indecent assault, rape (without serious violence), rape (with serious violence) and sexual homicide). To assess reliability, Sutherland calculated ICCs and found poor to good rater agreement for some formulation elements. For example, nature (type of offence) and victims within the scenarios had good reliability, whereas imminence, frequency and level of harm case scenario decisions had poor reliability. Although this provided a good first step to examine case formulation for sexual violence risk and provided some evidence of reliability for formulations, there were limitations to this approach. The primary limitation concerns the methodology for completing the case formulations and the implications for

ecological validity. Using a set of forced-choice questions does not reflect how evaluators would complete a case if conducting a true assessment.

Given the limitations to Sutherland's study, a different approach was used for this study. The current study attempted to investigate case formulation while maintaining the procedure advocated by the measure authors (Hart et al., 2003). Evaluators were asked to complete case formulations, scenarios and management strategies according to the procedure outlined in the user manual and reinforced within the training. To evaluate the reliability of these formulations, Research Assistants made similarity ratings from pairs of cases: Within Cases, which represented formulations developed for the same case by different raters and Across Cases, which represented formulations developed for different cases by different raters. As formulation is not quantified in the RSVP, similarity ratings were intended to reflect consistency of formulations developed by different raters for same or different cases, and this represented reliability. The similarity ratings for the Within Case group and Across Case group were compared and it was expected that the Within Case group would have higher mean similarity ratings compared to the Across Case group.

Although Within Case similarity ratings were not significantly higher than Across Case similarity ratings for all aspects of formulation there were promising findings. Despite lower than expected levels of agreement for item and case prioritization ratings, participants were able to develop similar formulations in some regards. For the overall global assessments, the Within Case group had significantly higher similarity ratings than the Across Case group for case scenarios and case management strategies. For specific case formulation, scenario and management ratings, destabilizers (case formulation), type of violence, victims, and level of harm (case scenario) and treatment, and victim safety planning (management) were significantly more similar in the Within Case group.

In terms of case formulation ratings, only destabilizers significantly differed between the two groups. The lack of difference between the groups for motivators and disinhibitors might be related to the level of agreement for item ratings. The ICCs from the reliability analyses indicated participants had fair agreement regarding the presence of risk factors and poor agreement regarding the relevance of risk factors. It is possible

that a good understanding of the items and good rater agreement with respect to scoring the items (particularly for relevance ratings) are prerequisites for developing reliable case formulations, including an understanding of the primary causal factors. If this is the case, it speaks to the importance of the training and required focus on the measure risk items. It would be interesting to see whether case formulation is, in fact, more reliable when there are greater levels of agreement for risk factor ratings. Another possible explanation for the lack of agreement for formulation is related to the use of mostly high-risk cases. Including offenders with several problem areas, and therefore many risk factors, may have resulted in participants including all (or several) risk factors into the formulation, without attempting to establish specific factors as potential root causes of sexual violence. As explained in the training, during this step of the formulation process, evaluators are encouraged to consider all of the major risk factors in the case, then review the risk factors and attempt to identify the simplest or smallest set of risk factors that seem to be able to explain why the person committed sexual violence. It is possible that evaluators in this study did not complete this particular step. Again, this might reflect training issues and that a greater emphasis needs to be placed on the formulation process (i.e., require a greater focus on the process of simplifying the set of risk factors).

With respect to case scenarios, there were significant differences between groups for violence type, victims, and level of harm. These findings provide support for Sutherland et al.'s (2012) study, where type of offence and victims obtained good levels of reliability. In fact, these findings expand on the results from the previous study as participants here were required to develop their own scenarios and did not respond to forced-choice questions. Further, Sutherland et al. found that *repeat* scenarios had the highest reliability, which might reflect the fact that developing a scenario that copies the offender's previous history is straightforward compared to other types of scenarios. However, anecdotally, it does not seem likely that the differences between groups were a result of participants including only *repeat* scenarios in their evaluations. Based on my review of the evaluations, repeat scenarios were often included in this step of the evaluation; however, many participants did attempt to "think outside the box" and considered twist, escalation, and best-case scenarios.

Participants were also able to develop similar treatment and victim safety planning strategies within the case management step. Given that participants generally

agreed on the likely victims when evaluating the same case, it seems logical that they were also able to develop similar strategies to keep these victims safe. For example, in scenarios where a potential victim was identifiable, reasonable strategies available to enhance the protection of the victim might have been easy to develop. Similarly, in scenarios where there was no identifiable victim, or the victim was likely to be a stranger, the participants might have easily agreed there were no viable victim safety planning strategies available. For the Within Case group, the average similarity rating for monitoring and supervision were lower than expected (i.e., under the scale midpoint). Based on a review of all assessments submitted, there appeared to be some confusion regarding what would be considered a monitoring strategy and what would be considered a supervision strategy and it appeared, at times, that participants used these two sections interchangeably. This might help explain why there were lower levels of agreement for the Within Case group for monitoring and supervision. Increased attention in training, including clear examples, might help to remedy this confusion.

The findings from this study provide an important first step in assessing case formulation and attempting to establish its reliability. Demonstrating a modest level of agreement in formulations has potential implications across a number of contexts. In light of the role of risk assessment within the legal context and its use to assist in responding to specific legal questions, the ability to develop reliable scenarios and management plans is important. As previously discussed, SVP legislation requires an individual be at risk for future sexual violence. If evaluators are able to go beyond an overall estimate of risk level and agree on plausible types of violence an offender is likely to commit, the likely victims and the extent of harm the offender might perpetrate, this could provide the trier of fact with greater support when making the decision whether to apply the designation. The same argument can be made with DO legislation in Canada, where the offender must be at risk for *threat to life, safety, or physical or mental well-being of other persons.* The decision of whether this designation is applicable could be supported if evaluators demonstrate agreement regarding the type(s) of violence an offender is at risk of perpetrating. Agreement for case management strategies can also have implications in the legal context. The primary distinction between Canadian DO and LTO legislation concerns the ability to manage the offender in the community. If evaluators can agree on 1) whether an individual can be managed in the community and

2) the types of strategies that would need to be implemented in order to manage an individual in the community, again the trier of fact is provided with additional support in making the legal decision.

With respect to future research, researchers should continue to consider new approaches to evaluating formulation reliability. This study examined the agreement between participants, but future research could examine the level of agreement between trainee and expert evaluations. There may be other approaches for evaluating case formulation available as well. For example, instead of comparing formulations with other risk assessment experts, it might be useful to engage professionals who work directly with the individual. A case formulation can be developed, based on a thorough risk assessment, followed by a review by a treatment team or correctional staff member to confirm whether the formulation addresses the concerns and main issues presented by the individual. In addition, the validity of formulations should be considered and investigated. This study did not address validity and it is not known whether the formulations were accurate or would have predictive validity. Future research might be able to examine whether issues such as the type of violence committed upon release was consistent with the formulation scenarios developed during a risk assessment prior to release. Or, it might be interesting to examine whether the management strategies suggested mapped well with the actual management strategies implemented, and the final outcome for the individual (e.g. whether the offender reoffends).

This study failed to show reliability across all aspects of case formulation demonstrating that further improvements are required in this area. This is consistent with previous case formulation research that has shown that not all clinicians are competent with respect to this skill (Hart et al., 2011). The lack of ability in applying this particular skill might reflect the type of training that is provided for formulation. Perhaps the first step for future research is to develop a training approach for case formulation skills that is empirically supported (Hart et al., 2011). There is some evidence that providing at least a minimum level of training improves the quality of formulations (Kendjelic & Eells, 2007); however, to date, there is a lack of evidence regarding the most appropriate method for teaching case formulation and how to appropriately evaluate the quality of formulations. It is possible that we will need to look outside of psychology or clinical formulation in order to develop improved methods for teaching and developing case

formulations. For example, it may be useful to move towards a more visual approach to case formulation, using a technique such as concept mapping. Concept maps are pictorial representations of key concepts and the relationships between them (Kane & Trochim, 2007) and have been used in areas such as health care. Using this type of approach may allow evaluators to visually map out important risk factors, link various factors together, identify which factors are most relevant and start to establish the relationship between risk factors and risk for sexual violence.

## 4.5. Limitations

The main limitations of this study have generally been described in the above discussion, but will be briefly summarized here.

With respect to the training, the online format used was a new approach that had not previously been used for RSVP training. In particular, the online format resulted in little to no interaction between participants and the trainer, minimizing the opportunity for participants to ask questions or seek clarification from the trainer. Further, the training had limited focus on the risk factors and corresponding definitions, instead relying on participants to review the definitions provided in the manual. Assessment feedback was provided for each assessment submitted, but in some cases participants completed multiple assessments prior to obtaining feedback. The feedback provided an important means for clarifying item definitions and general assessment issues and proceeding through cases without feedback might have impacted the quality of assessments overall.

The cases selected for the study were actual cases, edited to a reasonable length (most case files were originally hundreds of pages). However, completing RSVP assessments using only file information is not ideal. As recommended by the measure's authors, information should be collected from multiple sources including interviews with the offender. The use of file-only cases limits the generalizability of the findings to real world settings. Also, the variability of the cases included was limited as they were all moderate and high risk cases.

Pre-investigation power analyses were completed to ensure there was sufficient power to interpret findings of the exploratory analyses for the case formulation similarity

ratings. The power analyses indicated six evaluators completing 10 cases were required, at a minimum, which is what was used in this study. The results from the case formulation analyses indicated that the power analysis was accurate for this component and there was adequate power to detect moderate effect sizes. However, obtaining a greater sample size would have been useful for improving the power for the reliability analyses.

The current study combined Presence Past and Recent ratings into a single item, disregarding the timeframe for the item and taking a more broad approach of *Was this risk factor ever a serious problem?* While it is true that this approach does not reflect the intended administration of the RSVP, which encourages evaluators to distinguish between risk factors that were present historically (i.e., more than a year prior to assessment) and those that are present recently (i.e., within the year prior to assessment), the decision to take this approach was based on the cases used in this study. For all 10 cases used in this study, the offender had been institutionalized for the entire year prior to assessment. As a result, offenders were not presented with the same opportunity to demonstrate behaviours and attitudes relevant to certain risk factors. For example, the five items within the Sexual Violence History domain were coded *No* for Recent ratings across all cases. Reliability analyses would be difficult to conduct with items that demonstrated such low variability. With that said, future research should investigate reliability for Past and Recent ratings separately.

## 4.6. Conclusion

The current study adds to the direct empirical literature of the RSVP. Although agreement was lower than expected, a moderate level of interrater reliability was obtained. There are a number of potential explanations for this finding and future research should address issues encountered in this study. Importantly, this is one of the first studies to extend beyond examining risk factors to investigating the process of case formulation. Case formulation, developing case scenarios, and developing management strategies, are integral aspects of the risk assessment process. To date, this area has been neglected in research and it is critical that the risk assessment field start to investigate and address this issue. The results pertaining to formulation are promising:

this study demonstrated evaluators provided the same case information can reliably develop some aspects of formulation, including scenarios, management strategies, types of violence, victims, level of harm, victim safety plans, and treatment strategies. There is a need for continued attention on this issue, including research examining the best approach for teaching this important skill. With that said, the current findings provide an important first step in evaluating case formulation and provide new avenues for future research to assess and evaluate case formulation.

# References

American Psychological Association Task Force on the Assessment of Competence in Professional Psychology, (2006). *APA Task Force on the Assessment of Competence in Professional Psychology: Final report*. Retrieved April 15, 2013 from http://www.apa.org/education/grad/competency-revised.pdf

Andrews, D. A. & Bonta, J. (1995). *The Level of Service Inventory-Revised*. Toronto, Canada: Multi-Health Systems.

Andrews, D. A. & Bonta, J (2003). *The psychology of criminal conduct* (3rd ed.). Cincinnati, OH: Anderson.

Andrews, D. A. & Bonta, J. (2006). *The psychology of criminal conduct* (4th ed.). Cincinnati, OH: Anderson.

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

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

Archer, R. P., Buffington-Vollum, J. K., Vauter Stredny, R., & Handel, R. W. (2006). A survey of psychological test use patterns among forensic psychologists. *Journal of Personality Assessment, 87*, 84-94. doi: 10.1207/s15327752jpa8701_07

Bartko, J. (1976). On various intraclass correlation reliability coefficients. *Psychological Bulletin, 83*, 762-765. doi: 10.1037/0033-2909.83.5.762

Bartosh, D. L, Garby, T., Lewis, D. & Gray S. (2003). Differences in the predictive validity of actuarial risk assessments in relations to sex offender type. *International Journal of Offender Therapy and Comparative Criminology, 47*, 422-438. doi: 10.1177/0306624X03253850

Bickle, A. (2008). The dangerous offender provisions of the Criminal Justice Act 2003 and their implications for psychiatric evidence in sentencing violent and sexual offenders. *The Journal of Forensic Psychiatry & Psychology, 19*, 603-619. doi: 10.1080/14789940801912341

Boer, D. P., Hart, S. D., Kropp, P. R., & Webster, C. D. (1997). *Manual for the Sexual Violence Risk-20: Professional guidelines for assessing risk of sexual violence*. Vancouver, British Columbia: British Columbia Institute on Family Violence and Mental Health, Law, & Policy Institute, Simon Fraser University.

British Psychological Society Professional Practice Board (2008). *Generic professional practice guidelines*, 2nd ed. Leicester: British Psychological Society. Retrieved April 15, 2013 from http://www.bps.org.uk/sites/default/files/documents/generic_professional_practice_guidelines.pdf

Chermack, T. J. & Lynham, S. A. (2002). Definitions and outcome variables of scenario planning. *Human Resources Development Review, 1*, 366-383. doi: 10.1177/1534484302013006

Chermack, T. J. & van der Merwe, L. (2003). The role of constructivist learning in scenario planning. *Futures, 35*, 445-460. doi: 10.1016/S0016-3287(02)00091-5

Code of Virginia, § 37.2-903 (2013).

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Criminal Code, R.S.c. C-46 (1985). Retrieved from http://laws-lois.justice.gc.ca/PDF/C-46.pdf

Daffern, M., Jones, L., Howells, K., Shine, J., Mikton, C., & Tunbridge, V. (2007). Editorial. Refining the definition of Offence Paralleling Behavior. *Criminal Behavior and Mental Health, 17*, 265-273. doi: 10.1002/cbm.671

Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). *HCR-20$^{V3}$ Assessing Risk for Violence. User Guide*. Burnaby, BC: Mental Health, Law, and Policy Institute, Simon Fraser University.

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. doi: 10.1177/009385402236735

Doyle, M. & Dolan, M. (2002). Violence risk assessment: Combining actuarial and clinical information to structure clinical judgments for the formulation and management of risk. *Journal of Psychiatric and Mental health Nursing, 9*, 649-657. doi: 10.1046/j.1365-2850.2002.00535.X

Doyle, D. J., & Ogloff, J. R. P. (2009). Calling the tune without the music: A psycho-legal analysis of Australia's post-sentence legislation. *Australian and New Zealand Journal of Criminology, 42*, 179 – 203. doi: 10.1375/acri.42.2.179

Dvoskin, J. A. & Heilbrun, K. (2001). Risk assessment and release decision-making: Toward resolving the great debate. *Journal of the American Academy of Psychiatry and Law, 26*, 6-10.

Eells, T. D. (2007). Psychotherapy case formulation: History and current status. In T.D. Eells (Ed.), *Handbook of Psychotherapy Case Formulation* (2nd edition, pp. 3-32). New York: Guilford Press.

Eells, T. D. & Lombart, K. G. (2011). Theoretical and evidence-based approaches to case formulation. In P. Sturmey & M. McMurran (Eds.), *Forensic Case Formulation* (pp. 3-32). West Sussex, UK: John Wiley & Sons, Ltd.

Epperson, D. L., Kaul, J. D., & Hesselton, D. (1998). *Final report on the development of the Minnesota Sex Offending Screening Tool-Revised (MnSOST-R)*. St. Paul: Minnesota Department of Corrections.

Epperson, D. L., Kaul, J. D., Huot, S., Hesselton, D., & Alexander, W. (2000). *Minnesota Sex Offender Screening Tool-Revised* (MnSOST-R). St. Paul, MN: Minnesota Department of Corrections.

Fernando, I., Cohen, M., & Henskens, F. (2012). Pattern-based formulation: A methodology for psychiatric case formulation. *Australasian Psychiatry, 20*, 121-126. doi: 10.1177/1039856212437255

First, M. B. & Hanlon, R. L. (2008). Use of DSM paraphilia diagnoses in Sexually Violent Predator commitment cases. *Journal of the American Academy of Psychiatry and the Law, 36*, 443-454.

Fleiss, J. L. (1981). *Statistical Methods for Ratios and Proportions*, 2nd ed. New York: John Wiley & Sons, Inc.

Florida Administrative Register & Florida Administrative Code. Code 65E-25.001

Furby, L., Weinrott, M. R. & Blackshaw, L. (1989). Sex offender recidivism: A review. *Psychological Bulletin, 105*, 3 – 30. doi: 10.1037/0033-2909.105.1.3

Gresswell, D. M. & Hollin, C. (1992). Towards a new methodology for making sense of case material: An illustrative case involving attempted multiple murder. *Criminal Behaviour and Mental Health, 2*, 329-341.

Grove, W. M. & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical-statistical controversy. *Psychology, Public Policy, and Law, 2*, 293-323. doi: 10.1037/1076-8971.2.2.293

Haggard-Grann, U. (2007). Assessing violence risk: A review and clinical recommendations. *Journal of Counseling and Development, 85*, 294-302. doi: 10.1002/j.1556-6678.2007.tb00477.x

Hanson, R. K. (1997). *The development of a brief actuarial scale for sex offender recidivism* (User Report No. 1997-04). Ottawa, ON: Department of the Solicitor General of Canada.

Hanson, R. K. (1998). What do we know about sex offender risk assessment? *Psychology, Public Policy, and Law, 4*, 50-72. doi: 10.1037/1076-8971.4.1-2.50

Hanson, R. K., (2000). Twenty years of progress in violence risk assessment. *Journal of Interpersonal Violence, 20*, 212-217. doi: 10.1177/0886260504267740

Hanson, R. K. & Bussière, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology, 66*, 348-362. doi: 10.1037/0022-006X.66.2.348

Hanson, R. K., Harris, A. J. R., Scott, T., & Helmus, L. (2007). *Assessing the risk of sexual offenders on community supervision: The Dynamic Supervision Project* (Corrections Research User Report No. 2007-05). Ottawa, ON: Public Safety Canada.

Hanson, R. K. & Morton-Bourgon, K. (2004). *Predictors of sexual recidivism: An updated meta-analysis 2004-02*. Ottawa: Public Works and Government Services Canada.

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

Hanson, R. K. & Thornton, D. (1999). *Static-99: Improving actuarial risk assessment for sex offenders*. Ottawa: Solicitor General of Canada.

Hanson, R. K. & Thornton, D. (2003). *Notes on the development of Static-2002*. (Corrections Research User Report No. 2003-01). Ottawa, ON: Department of the Solicitor General of Canada.

Hare, R. D. (1991). *The Hare Psychopathy Checklist-Revised* (1st ed.). Toronto, ON, Canada: Multi-Health Systems.

Hare, R. D. (2003). *Manual for the Revised Psychopathy Checklist* (2nd ed.). Toronto, ON, Canada: Multi-Health Systems.

Harris, G. T., Rice, M. E., & Cormier, C. A. (2002). Prospective replication of the Violence Risk Appraisal Guide in predicting violence recidivism among forensic patients. *Law and Human Behavior, 26*, 377-394. doi: 10.1023/A:1016347320889

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

Hart, S. D. (2008, July). *The Structured Professional Judgment approach to violence risk assessment: Core principles of SPJ*. Paper presented at the annual meeting of the International Association of Forensic Mental Health Services, Vienna, Austria.

Hart, S. D. (2008). Risk assessment: Sexual violence and the role of paraphilia. In A. J. R. Harris & C. A. Page (Eds.), *Sexual homicide and paraphilias: The Correctional Service of Canada's Experts Forum 2007*. Ottawa: Correctional Service of Canada.

Hart, S. D. (2009). Evidence-based assessment of risk for sexual violence. *Chapman Journal of Criminal Justice, 1*, 143-165.

Hart, S. D. (2010). Foreword. In M. Daffern, L. Jones, & J. Shine (Eds.) *Offence Paralleling Behavior. A Case Formulation Approach to Offender Assessment and Intervention*. West Sussex, UK: Wiley & Sons Ltd.

Hart, S. D. (2012). *The RSVP. Clinical-Forensic Assessment of Risk for Sexual Violence: Risk Assessment Training Course*. Vancouver, BC: ProActive ReSolutions.

Hart, S. D., & Boer, D. P. (2010). Structure professional judgment guidelines for sexual violence risk assessment: The Sexual Violence Risk-20 (SVR-20) and Risk for Sexual Violence Protocol (RSVP). In R. K. Otto and K. S. Douglas (Eds.) *Handbook of violence risk assessment*. New York, NY: Routledge/Taylor & Francis Group. P. 269-294.

Hart, S. D., Cox, D., & Hare, R. (1995). *The Hare Psychopathy Checklist Screening Version (PCL:SV)*. Toronto, ON, Canada: Multi-Health Systems.

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

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

Hart, S. D. & Logan, C. (2011). Formulation of violence risk using evidence-based assessments: The Structured Professional Judgment Approach. In P. Sturmey & M. McMurran (Eds). *Forensic Case Formulation*. 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*, 60-65. doi: 10.1192/bjp.190.5.s60

Hart, S., Sturmey, P., Logan, C., & McMurran, M. (2011). Forensic case formulation. *International Journal of Forensic Mental Health, 10*, 118-126. doi: 10.1080/14999013.2011.577137

Haynes, S. N. & O'Brien, W. H. (Eds.) (2000). *Principles of behavioural assessment: A functional approach to psychological assessment*. New York, NY: Plenum/Kluwer.

Haynes, S. N. & Williams, A. E. (2003). Case formulation and design of behavioural treatment programs. Matching treatment mechanisms to causal variables for behaviour problems. *European Journal of Psychological Assessment, 19*, 164-174. doi: 10.1027//1015-5759.19.3.164

Heilbrun, K., O'Neill, M. L., Strohman, L. K., Bowman, Q., & Philipson, J. (2000). Expert approaches to communicating violence risk. *Law and Human Behavior, 24*, 137-148. doi: 10.1023/A:1005435005404

Heilbrun, K., Yasuhara, K., & Shah, S. (2009). Violence risk assessment tools: Overview and critical analysis. In R. K. Otto & K. Douglas (Eds.), *Handbook of Violence Risk Assessment Tools*. New York, NY: Routledge/Taylor & Francis Group.

Hoge, R. D. & Andrews, D. A. (2006). *Youth Level of Service/Case Management Inventory: User's manual*. North Tonawanda, NY: Multi-Health Systems.

Jackson, K. & Healey, J. (2008, July). *Concurrent validity of the RSVP vis-à-vis the SVR-20, Static-99, Static-2002, and SORAG*. Paper presented at the annual meeting of the International Association of Forensic Mental Health Services, Vienna, Austria.

Janus, E. S. (2003). Legislative responses to sexual violence: An overview. *Annals of New York Academy of Sciences, 989*, 247 – 264.

Janus, E. & Prentky, R. (2003). Forensic use of actuarial risk assessment with sex offenders: Accuracy, admissibility and accountability. *American Criminal Law Review, 40*, 1443 – 1499.

Jones, L. F. (1997). Developing models for managing treatment integrity and efficacy in a prison based TC: the Max Glatt Centre. In E. Cullen, L. Jones, & R. Woodward (Eds.). *Therapeutic Communities for Offenders*. Chichester: Wiley.

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*. Retrieved June 14 from http://www.tandfonline.com.proxy.lib.sfu.ca/doi/full/10.1080/13552600.2013.767948#.Ubz32etQ00c. doi: 10.1080/13552600.2013.767948

Kane, M. & Trochim, W. M. K. (2007). *Concept mapping for planning and evaluation*. Thousand Oaks, CA: Sage.

Kendjelic, E. M. & Eells, T. D. (2007). Generic psychotherapy case formulation training improves formulation quality. *Psychotherapy: Theory, research, practice, training, 44*, 66-77. doi: 10.1037/0033-3204.44.1.66

Klaver, J., Watt, K., Kropp, P. R., & Hart, S. D. (2002, August). *Actuarial assessment of risk for sexual violence*. Paper presented at the annual meeting of the American Psychological Association, Chicago, IL.

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. (2008). *Guidelines for Stalking Assessment and Management (SAM) User manual*. Vancouver, Canada: ProActive ReSolutions Inc.

Kropp, P. R., Hart, S. D., Webster, C. D., & Eaves, D. (1994). *Manual for the spousal assault risk assessment guide*. Vancouver, BC: British Columbia Institute on Family Violence.

Kuyken, W. (2006). Evidence-based case formulation: Is the emperor clothed? In N. Tarrier (Ed.), *Case formulation in cognitive behaviour therapy: The treatment of challenging and complex clinical cases* (pp 12-35). London: Brunner-Routledge.

Larcombe, W. (2012). Sex offender risk assessment: The need to place recidivism research in the context. *Violence Against Women, 18*, 482 – 501. doi: 10.1177/1077801212452249

Levenson, J. S. (2004). Reliability of sexually violent predator civil commitment criteria in Florida. *Law and Human Behavior, 28*, 357-368. doi: 10.1023/B:LAHU.0000039330.22347.ad

Litwack, T. R. (2001). Actuarial versus clinical assessments of dangerousness. *Psychology, Public Policy, and Law, 7*, 409-443. doi: 10.1037/1076-8971.7.2.409

Lidz, C. W., Mulvey, E. P. & Gardner, W. (1993). The accuracy of predictions of violence to others. *JAMA: Journal of the American Medical Association, 269*, 1007-1111. doi: 10.1001/jama.1993.03500080055032

McGraw, K. O. & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*, 30-46. doi: 10.1037/1082-989X.1.1.30

McNiel, D. E. & Binder, R. L. (1995). Correlates of accuracy in the assessment of psychiatric inpatients' risk of violence. *American Journal of Psychiatry, 152*, 901-906.

Miller, H. A., Amenta, A. E., & Conroy, M. A. (2005). Sexually violent predator evaluations: Empirical evidence, strategies for professionals, and research directions. *Law and Human Behavior, 29*, 29-54. doi: 10.1007/s10979-005-1398-y

Monahan, J. (1984). The prediction of violent behavior: Toward a second generation of theory and policy. *American Journal of Psychiatry, 141*, 10-15.

Mumma, G. H. (2011). Current issues in case formulation. In P. Sturmey & M. McMurran (Eds.) *Forensic Case Formulation*. Chichester, UK: John Wiley & Sons, Ltd.

Nichols, D. P. (1998). *Choosing an Intraclass Correlation Coefficient*. Retrieved on June 1, 2013 from http://www.ats.ucla.edu/stat/spss/library/whichicc.htm.

Otto, R. K. & Douglas, K. S. (Eds.). (2010) *Handbook of violence risk assessment*. New York, NY: Routledge/Taylor & Francis Group.

Perrault, S. & Brennan, S. (2010). Criminal victimization in Canada, 2009. *Juristat, 30*, 1-33.

Persons, J. B. & Bertagnolli, A. (1999). Inter-rater reliability of cognitive-behavioral case formulations of depression: A replication. *Cognitive Therapy and Research, 23*, 271-283. doi: 10.1023/A:1018791531158

Persons, J. B., Mooney, K. A., & Padesky, C. A. (1995). Interrater reliability of cognitive-behavioral case formulations. *Cognitive Therapy and Research, 19*, 21-34. doi: 10.1007/BF02229674

Quinsey, V. L., Harris, G. T., Rice, M., & Cormier, C. A. (1998). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.

Quinsey, V. L., Harris, G., Rice, M., & Cormier, C. A. (2006). *Violent offenders: Appraising and managing risk* (2nd ed.). Washington, DC: American Psychological Association.

Shrout, P. E. & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing reliability. *Psychological Bulletin, 86*, 420-428. doi: 10.1037/0033-2909.86.2.420

Solicitor General Canada (2001). *High-Risk Offenders. A Handbook for Criminal Justice Professionals*. http://www.publicsafety.gc.ca/res/cor/rep/_fl/2001-05-hnd-bk-eng.pdf (retrieved April 27, 2013).

Statistics Canada (2013). *Crimes, by type of violation, and by province and territory* (CANSIM table 252-0051 and Catalogue no. 85-002-X). Retrieved from http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/legal50a-eng.htm

Sturmey, P. (2010). Case formulation in forensic psychology. In M. Daffern, L. Jones, & J. Shine (Eds.) *Offence Paralleling Behavior. A Case Formulation Approach to Offender Assessment and Intervention*. West Sussex, UK: Wiley & Sons Ltd.

Sutherland, A. A., Johnstone, L., Davidson, K. M., Hart, S. D., Cooke, D. J., Kropp, P. R…Stocks, R. (2012). Sexual violence risk assessment: An investigation of the interrater reliability of professional judgments made using the Risk for Sexual Violence Protocol. *International Journal of Forensic Mental Health, 11*, 119-133. doi: 10.1080/14999013.2012.690020

Thornton, D., Mann, R., Webster, S., Blud, L., Travers, R., Friendship, C., & Erikson, M. (2003). Distinguishing and combining risks for sexual and violent recidivism. *Annals of New York Academy of Sciences, 989*, 225-235.

Vess, J. & Skelton, A. (2010). Sexual and violent recidivism by offender type and actuarial risk: Reoffending rates for rapists, child molesters and mixed-victim offenders. *Psychology, Crime & Law, 16*, 541-554. doi: 10.1080/10683160802612908

Vincent, G. M., Maney, S. M., & Hart, S. D. (2009). The use of actuarial risk assessment instruments in sex offenders. In F. M. Saleh, A. J. Grudzinskas, J. M. Bradford, & D. J. Brodsky (Eds). *Sex Offenders: Identification, Risk Assessment, Treatment, and Legal Issues*. New York, NY: Oxford University Press.

Ward, T. & Brown, M. (2004). The good lives model and conceptual issues in offender rehabilitation. *Psychology, Crime & Law, 10*, 243-257. doi: 10.1080/10683160410001662744

Ward, T. & Stewart, C. A. (2003). The treatment of sex offenders: Risk management and good lives. *Professional Psychology: Research and Practice, 34*, 353-360. doi: 10.1037/0735-7028.34.4.353

Watt, K. A., Hart, S. D., Wilson, C. M., Guy, L., & Douglas, K. S. (2006, March). *An evaluation of the Risk for Sexual Violence Protocol (RSVP) in high risk offenders: Interrater reliability and concurrent validity*. Paper presented at the annual meeting of the American Psychology-Law Society, St. Petersburgh, FL.

Watt, K. A. & Jackson, K. (2008, July). *Interrater and structural reliabilities of the Risk for Sexual Violence Protocol (RSVP)*. 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). Vancouver, Canada: Mental Health, Law, & Policy Institute, Simon Fraser University.

Worling, J. R. & Curwen, T. (2001). *Estimate of Risk Adolescent Sexual Offense Recidivism (ERASOR) Version 2.0*. Toronto, ON: SAFE-T Program, Thistletown Regional Centre, Ontario Ministry of Community and Social Services.

# Appendices

# Appendix A.

# Recruitment Information

Dear Madam/Sir,

CONCEPT, Consolidated Continuing Education and Professional Training, offers an on-line training program in sex offender risk assessment using the Risk for Sexual Violence Protocol or RSVP (Hart et al., 2003). The RSVP is a set of structured professional judgment guidelines published by the Mental Health, Law, and Policy Institute at Simon Fraser University.

The training has two components. The first component is didactic. It involves a review of fundamental issues in the structured professional judgment approach to risk assessment and the administration of the RSVP. The second component is practical. It involves administration of the RSVP for 6 cases. Individualized feedback is provided for each completed case.

The training requires about 25 hours to complete: 3-4 hours for the first component (Part 1 on CONCEPT website), and approximately 20 hours (~ 3 hours per case) for the second component (Part 2 on CONCEPT website). Those who complete the training receive a certificate of competence as well as 25 hours of Continuing Education credits. CONCEPT is approved by the American Psychological Association (APA), the Canadian Psychological Association (CPA) to sponsor continuing education for psychologists, and approved by the Canadian Association of Threat Assessment Professionals (CATAP) to sponsor Continuing education for its members. Continuing Education credits granted by CONCEPT may also be recognized by other organizations.

Those who wish to complete the training should have some specialized training or experience in either sex offender risk assessment or in the use of structured professional judgment guidelines for violence risk assessment (and preferably in both).* The training is available only in English.

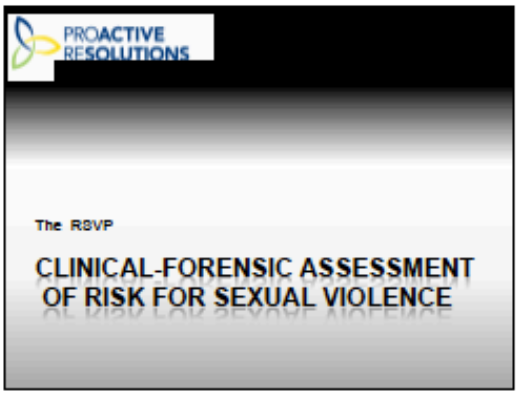
The standard cost of the training is USD $500, but there is a special opportunity for people to complete the training **FREE**. As part of her doctoral research in the Department of Psychology at Simon Fraser University, Catherine Wilson is conducting an evaluation of the reliability and validity of the RSVP. **People who complete the RSVP training through CONCEPT and who agree to have their training materials used in the evaluation research are eligible to have 100% of their fees reimbursed.** This opportunity is only available for a limited time and for a limited number of people.

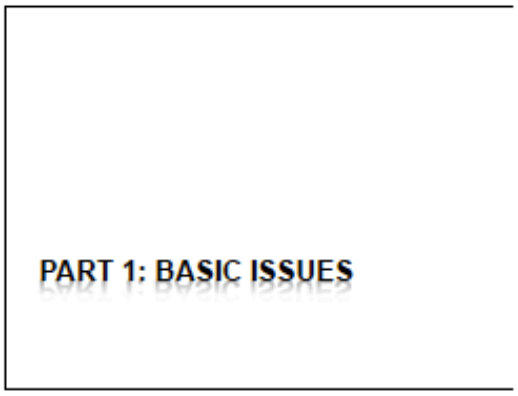
For further details about this special training opportunity, please visit the CONCEPT website (http://secure.concept-ce.com) or email Catherine Wilson ( ). Catherine's doctoral research is supervised by Professor Stephen Hart at Simon Fraser University ( ).

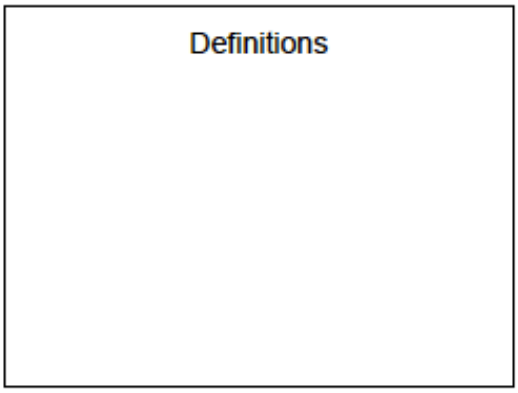
* This training may be appropriate for senior doctoral level students near completion of their degree and who meet the minimum inclusion requirements. Please contact directly to discuss options for participating in the training.

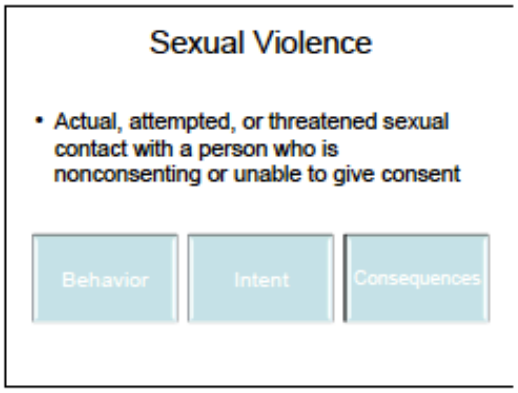
## Appendix B.

## RSVP Training Course Slides

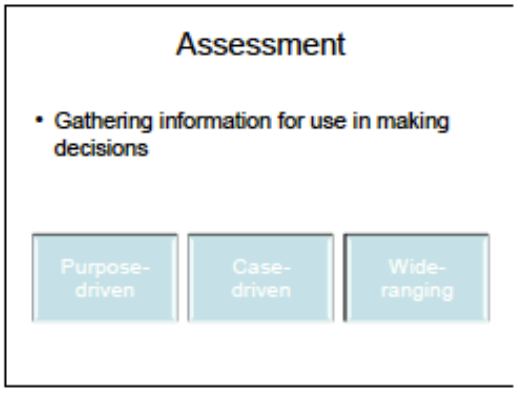






## Management

- Taking action to gain or maintain control

Active | Strategic | Tactical

## Synthesis

**Risk assessment**

- Process of speculating about risks posed by a perpetrator

**Risk management**

- Process of mitigating the risks posed by a perpetrator

## Goals

**Prevention**

- Guide and coordinate activities

**Accountability**

- Improve transparency and consistency of decisions

**Protection**

- Infringement of rights, exposure to liability

## Combining Risk Factors

## Two Basic Approaches

| Discretionary | Non-Discretionary |
|---|---|
| • Information weighted and combined according to the evaluator's judgment<br>• AKA informal, intuitive, impressionistic | • Information weighted and combined using fixed and explicit rules<br>• AKA actuarial, algorithmic, mechanical |

## Problems: Discretionary

**Capricious**

- Lacks accountability

**Imprecise**

- Lacks detail concerning risks posed

**Inconstant**

- May change rapidly

## Problems: Non-Discretionary

**Arbitrary**

- Lacks flexibility

**Rigid**

- Simplistic definition of risk

**Fixed**

- May be insensitive to change

## Challenge

- How do we combine the strengths of the discretionary and non-discretionary approaches?

## Credits

This training program was developed by Stephen D. Hart, ProActive ReSolutions Inc., with the assistance of Cathy M. Wilson, Simon Fraser University, and Patricia A. Zapf, CONCEPT (Consolidated Continuing Education and Professional Training).



www.proactive-resolutions.com

PROACTIVE RESOLUTIONS

The RSVP

# CLINICAL-FORENSIC ASSESSMENT OF RISK FOR SEXUAL VIOLENCE

## PART 2: STRUCTURED PROFESSIONAL JUDGMENT

## The SPJ Approach

## SPJ

- Relies on evidence-based guidelines to structure the exercise of professional discretion

| | | |
|---|---|---|
| Guide prevention via planning | Reflect best practice | Specify basic risk factors |
| Don't restrict scope | No scoring rules | Good for monitoring change |

## SPJ Guidelines

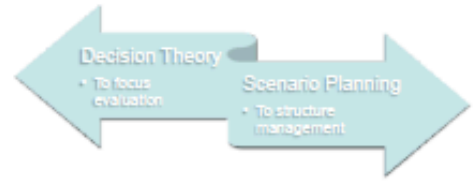

## General Method

1. Identify and gather relevant information
2. Determine presence of risk factors
3. Determine relevance of risk factors
4. Develop primary scenarios of violence
5. Develop case management plans
6. Communicate findings

## Limitations

- Judgments require expertise and training
  - Better suited for specialists than generalists
- Use requires resources
  - No such thing as "quick" risk assessment

## Formulation

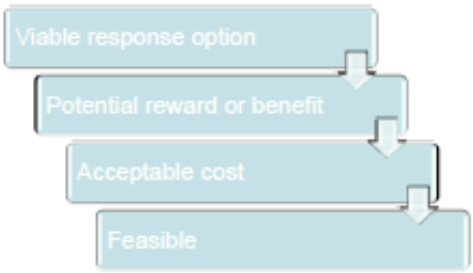



## Decision Theory

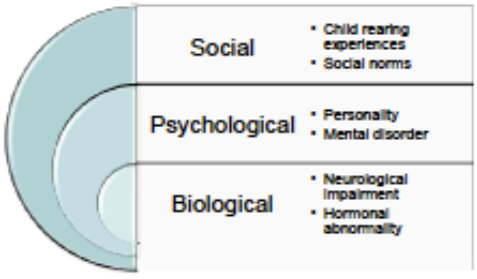
- The proximal cause of sexual violence is a choice or decision
  - Most decisions are not to commit sexual

| Decisions are: | Decisions are not: |
|---|---|
| • Goal-directed<br>• Planful | • Self-conscious<br>• Rational |

## Thinking About Sexual Violence




## Influences on Decisions



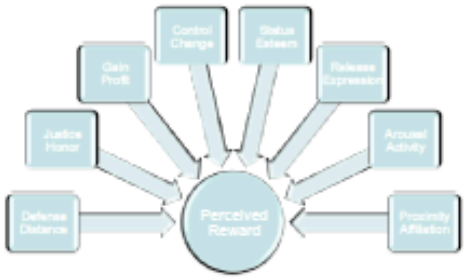

## Nature of Risk Factors

- Risk factors influence decisions

## Motivators

- Which factors drove decisions to act violently?
- Which factors increased the perceived gains or benefits of violence?
- What was the person trying to accomplish by acting violently?

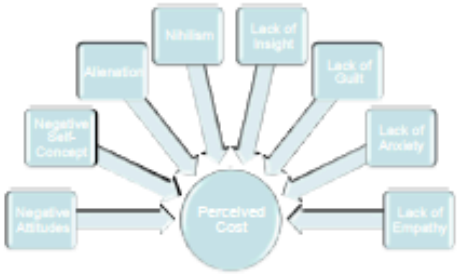
## Motivators




## Disinhibitors

- Which factors facilitated decisions to act violently?
- Which factors decreased the perceived costs or negative consequences of violence?
- How did the person overcome internal and external inhibitions against acting violently?

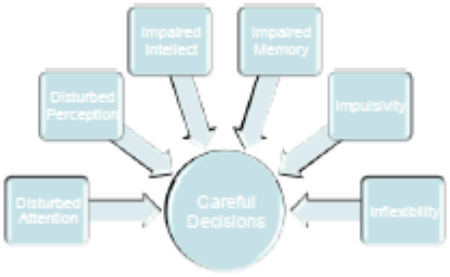
## Disinhibitors




## Destabilizers

- Which factors disturbed decisions to act violently?
- Which factors generally impaired or disrupted the person's ability to think rationally about violence?
- How clear, explicit, and systematic were the goals and plans for acting violently?

## Destabilize



## Simplify and Revise

- What is the simplest set of risk factors that explains why the person committed violence against specific people at specific times and in specific contexts?
- Is it necessary to add any other factors to the formulation?
- Can any factors be dropped from the formation?
- Does the formulation account for all the violence perpetrated by the person, or are multiple formulations required?

## Implications

**Assess decisions**

- Why might the perpetrator decide to commit sexual violence?

**Manage decisions**

- How do we encourage the perpetrator to decide not to commit sexual violence?

## Scenario Planning

- A specific approach to planning used under conditions of complexity and unbounded uncertainty
  - Public safety, health, engineering, industry, military
- Goal is to speculate systematically about "possible futures"
  - Avoid tunnel vision
  - Strive for desired outcomes
  - Avoid undesired outcomes

## Characteristics

**Directly guides action**

- Focuses on what should be done

**Based on prudence**

- Goes beyond abstract knowledge and technical skills

**Relies on qualitative, narrative reasoning**

- Uses formulation, rather than formulas

## Advantages

**Does not require prediction**

- Considers what might happen, not what will happen

**Does not require certainty**

- Evaluates action in light of current knowledge

**Encourages causal, systemic thinking**

- Identifies controlling factors

## Method

**Develop scenarios**

What are the "possible futures"?

**Evaluate scenarios**

Which "possible futures" most deserve attention?

**Use scenarios to plan**

Develop strategies and tactics to maximize success

## Developing Scenarios

| Repeat | Twist | Escalation | Improvement |
|---|---|---|---|
| • Consider all past violence, not just most recent | • Change in motivation, victimology, behavioral topography | • Including lethal or "worst case" | • Including desistence or "best case" |

## Evaluating Scenarios

| Plausibility | Utility | Reliability |
|---|---|---|
| • Information anchors should be relevant, comprehensive, and credible | • Should guide development of specific risk management plans | • Different people should develop similar scenarios |

## Using Scenarios

| Strategies | Tactics | Logistics |
|---|---|---|
| • Identify potentially feasible and effective procedures for each scenario | • Determine how to implement strategies in light of real-world constraints | • Organize the various agencies and people involved to maximize effectiveness |

## Credits

This training program was developed by Stephen D. Hart, ProActive ReSolutions Inc., with the assistance of Cathy M. Wilson, Simon Fraser University, and Patricia A. Zapf, CONCEPT (Consolidated Continuing Education and Professional Training).



www.proactive-resolutions.com

PROACTIVE RESOLUTIONS

The RSVP

# CLINICAL-FORENSIC ASSESSMENT OF RISK FOR SEXUAL VIOLENCE

## PART 3: THE RSVP

## Overview

## Uses and Users

For use with (alleged) perpetrators
- Adults, male or female, ≥ 18 yrs

At any stage of legal proceedings
- Pre-arrest, -trial, -sentence; pre- and post-release

For use by diverse professionals
- With expertise in assessment and sexual violence

## Administration

- Case information
- Presence of risk factors
- Relevance of risk factors
- Scenarios of sexual violence
- Management strategies
- Conclusory opinions

## Step 1: Information

Identify and gather information reasonably necessary for assessment

Document information reviewed and any important missing information

## Sexual Deviation

- Thoughts
- Images
- Fantasies
- Urges
- Physiological arousal
- Behaviour

## Step 2: Presence

- Consider standard and case specific risk factors
- Summarize judgments about presence, past and recent

## Risk Factors

- 22 basic risk factors in five domains
  - Sexual violence history
  - Psychological adjustment
  - Mental disorder
  - Social adjustment
  - Manageability
- Plus "Other considerations"

## Sexual Violence History

1. Chronicity of sexual violence
2. Diversity of sexual violence
3. Escalation of sexual violence
4. Physical coercion in sexual violence
5. Psychological coercion in sexual violence

## Psychological Adjustment

6. Extreme minimization or denial of sexual violence
7. Attitudes that support or condone sexual violence
8. Problems with self-awareness
9. Problems with stress or coping
10. Problems resulting from child abuse

## Mental Disorder

11. Sexual deviance
12. Psychopathic personality disorder
13. Major mental illness
14. Problems with substance use
15. Violent or suicidal ideation

## Social Adjustment

16. Problems with intimate relationships
17. Problems with non-intimate relationships
18. Problems with employment
19. Non-sexual criminality

## Manageability

20. Problems with planning
21. Problems with treatment
22. Problems with supervision

## Coding Presence

- Goal is to help summarize opinions regarding presence, based on the information reviewed

| | |
|---|---|
| Yes | Evidence the risk factor is or has been present |
| Possibly/ Partially | Possible or partial evidence the risk factor is or has been present |
| No | No evidence the risk factor is or has been present |

## Step 3: Relevance

Consider standard and case specific risk factors

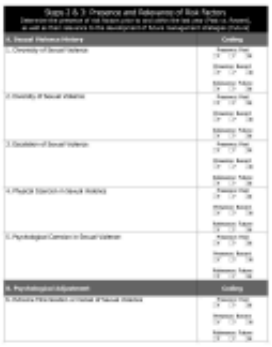
Summarize judgments about relevance to risks posed and management

## Relevance: Causal Roles

## Coding Relevance

- Goal is to help summarize opinions regarding presence, based on the information reviewed

| | |
|---|---|
| Yes | Evidence the risk factor is relevant to risks posed or risk management |
| Possibly/ Partially | Possible or partial evidence the risk factor is relevant to risks posed or risk management |
| No | No evidence the risk factor is relevant to risks posed or risk management |



## Step 4: Scenarios

Always consider past behavior

Consider possible development or evolution of offense behavior

## Scenarios

| Repeat | Twist | Escalation | Improvement |
|---|---|---|---|
| • Consider all past violence, not just most recent | • Change in motivation, victimology, behavioral topography | • Including lethal or "worst case" | • Including desistence or "best case" |

## Step 5: Management

Identify and target relevant risk factors

Specify management strategies and tactics

## Strategies

| Monitoring | Supervision | Treatment | Victim Safety Planning |
|---|---|---|---|
| • Surveillance or repeated assessment | • Imposition of controls or restriction of freedoms | • Rehabilitation, including further assessment | • Enhancement of security resources for identifiable targets |

## Step 6: Opinions

Communicate summary judgments clearly and simply

Identify important limitations on opinions

## Issues

| Case Prioritization | Serious Physical Harm | Immediate Action | Other Risks Indicated | Case Review |
|---|---|---|---|---|
| • Degree of effort or intervention required | • Risk for life-threatening violence | • Urgent intervention taken or needed | • Other risks that should be assessed | • Date for regular review and triggers for emergency review |

## Next Steps

## Next Steps

- Study the RSVP manual
  - Including the Sample Case in Appendix C
- Complete Part 4 of the training program
  - Reviews the Sample Case
- Complete the Evaluation
- Move on to the Practice Cases!

## Credits

This training program was developed by Stephen D. Hart, ProActive ReSolutions Inc., with the assistance of Cathy M. Wilson, Simon Fraser University, and Patricia A. Zapf, CONCEPT (Consolidated Continuing Education and Professional Training).



www.proactive-resolutions.com

PROACTIVE RESOLUTIONS

The RSVP

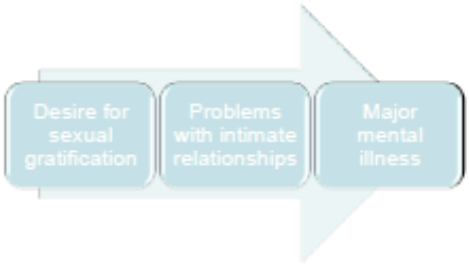
# CLINICAL-FORENSIC ASSESSMENT OF RISK FOR SEXUAL VIOLENCE

## Decision Theory

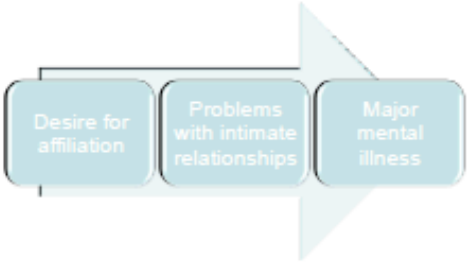
- Risk factors influence decisions


## Motivators


## Root Cause Analysis

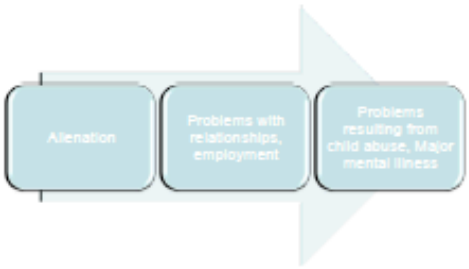



## Root Cause Analysis

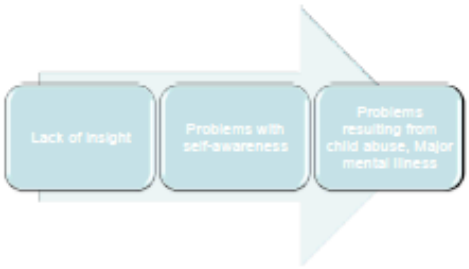



## Disinhibitors

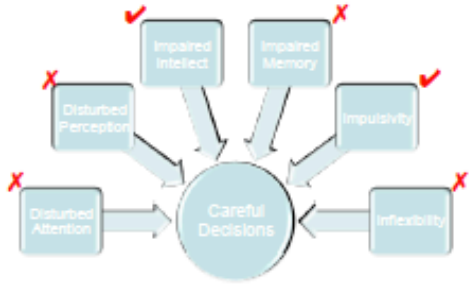
## Root Cause Analysis




## Root Cause Analysis

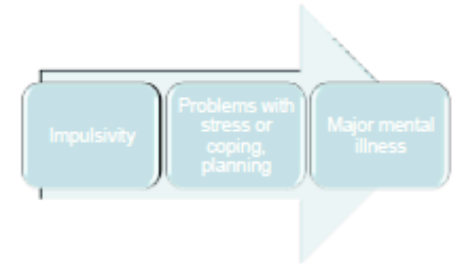



## Destabilize




## Root Cause Analysis


## Root Cause Analysis




## Summary

- Motivators: Desire for gratification, intimacy
- Disinhibitors: Alienation, lack of insight
- Destabilizers: Impaired intellect, impulsivity
- Root causes
  - Major mental illness
  - Problems resulting from child abuse
- Key unknown
  - Sexual deviance

## Scenario Planning

- Scenarios
- Management plans

## Credits

This training program was developed by Stephen D. Hart, ProActive ReSolutions Inc., with the assistance of Cathy M. Wilson, Simon Fraser University, and Patricia A. Zapf, CONCEPT (Consolidated Continuing Education and Professional Training).



www.proactive-resolutions.com

# Appendix C.

# Training Course Part 1 Questionnaire

### RSVP Specialized Training Course – Course Evaluation

Please complete the following course evaluation after you have completed Parts 1 through 4 of the training course and reviewed the RSVP Manual. All information required to complete this evaluation is provided in the training slides and/or the manual.

After you have completed and submitted the evaluation, your responses will be reviewed and individualized feedback will be provided. You will then be provided your six (6) practice cases to complete RSVP assessments.

1. What is the definition of “sexual violence”?
2. How should “sexual violence risk assessment” be defined?
3. What are the three primary goals of violence risk assessment?
4. What are the 6 elements (i.e., purpose and structure) of the structured professional judgment approach to violence risk assessment?
5. Identify and describe the 6 steps for completing a violence risk assessment, including the administration of the RSVP.
6. What are the main limitations of the structured professional judgment approach to violence risk assessment?
7. What are the three main ways that risk factors may influence decisions?
8. What are the three main advantages of scenario planning?
9. Describe the three main steps in the method for Scenario Planning.
10. Describe who can use the RSVP, when, and who the RSVP can be used with.
11. What is the convention for the coding the presence of risk factors using the RSVP?
12. What is the convention for coding the relevance of risk factors using the RSVP?
13. What four types of scenarios should always be considered when developing scenarios of future sexual violence on the RSVP?
14. What four major approaches/strategies should be included when developing risk management planning?
15. What information should be included and communicated in your final summary judgment?

# Appendix D.

# RSVP Worksheet

| **Step 1: Case Information**<br>Gather relevant background information |
|---|
| Identifying Information |
| Case:<br><br>Completed by:<br><br>Date completed: |

| **Steps 2 & 3: Presence and Relevance of Risk Factors**<br>Determine the presence of risk factors prior to and within the last year (Past vs. Recent), as well as their relevance to the development of future management strategies (Future) | |
|---|---|
| **A. Sexual Violence History** | **Coding**<br>**Yes, Possibly, No** |
| 1. Chronicity of Sexual Violence | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |
| 2. Diversity of Sexual Violence | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |
| 3. Escalation of Sexual Violence | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |
| 4. Physical Coercion in Sexual Violence | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |
| 5. Psychological Coercion in Sexual Violence | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |

| B. Psychological Adjustment | Coding |
|---|---|
| 6. Extreme Minimization or Denial of Sexual Violence | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |
| 7. Attitudes That Support or Condone Sexual Violence | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |
| 8. Problems With Self-Awareness | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |
| 9. Problems With Stress or Coping | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |
| 10. Problems Resulting From Child Abuse | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |

| C. Mental Disorder | Coding |
|---|---|
| 11. Sexual Deviance ☐*Definite* ☐*Provisional* | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |
| 12. Psychopathic Personality Disorder ☐*Definite* ☐*Provisional* | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |
| 13. Major Mental Illness ☐*Definite* ☐*Provisional* | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |
| 14. Problems With Substance Use ☐*Definite* ☐*Provisional* | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |
| 15. Violent or Suicidal Ideation ☐*Definite* ☐*Provisional* | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |

| D. Social Adjustment | Coding |
|---|---|
| 16. Problems With Intimate Relationships | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |
| 17. Problems With Non-Intimate Relationships | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |
| 18. Problems With Employment | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |
| 19. Non-Sexual Criminality | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |

| E. Manageability | Coding |
|---|---|
| 20. Problems With Planning | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |
| 21. Problems With Treatment | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |
| 22. Problems With Supervision | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |

| F. Other Considerations | Coding |
|---|---|
| Specify: | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |
| Specify: | Presence: Past ☐<br>Presence: Recent ☐<br>Relevance: Future ☐ |

| **Formulation of Sexual Violence** |
| --- |
| Provide an account of the person's sexual violence by identifying primary risk factors and their causal roles<br><br>Motivators (factors that increased the perceived likelihood of gains or benefits of violence)<br>Disinhibitors (factors that decreased the perceived likelihood of costs or negative consequences of violence)<br>Destabilizers (factors that impaired, disturbed, or disrupted the process of decision making) |

| Step 4: Risk Scenarios<br>Identify and describe the most plausible scenarios of future sexual violence; revise Step 3 as required | | | |
|---|---|---|---|
| | **Scenario #1** | **Scenario #2** | **Scenario #3** |
| **Nature**<br>What kind of sexual violence is the person likely to commit?<br><br>Who are the likely victims?<br><br>What is the likely motivation — that is, what might the person be trying to accomplish? | | | |
| **Severity**<br>What would be the psychological harm to victims?<br><br>What would be the physical harm to victims?<br><br>Is there a chance that the sexual violence might escalate to serious or life-threatening violence? | | | |
| **Imminence**<br>How soon might the person engage in sexual violence?<br><br>Are there any warning signs that might signal that the risk is increasing or imminent? | | | |
| **Frequency/Duration**<br>How often might the sexual violence occur — once, several times, frequently?<br><br>Is the risk chronic or acute (i.e., time-limited)? | | | |
| **Likelihood**<br>In general, how frequent or common is this type of sexual violence?<br><br>Based on this person's history, how likely is it that this type of sexual violence will occur? | | | |

| Step 5: Case Management<br>Recommend strategies for managing sexual violence risk; revise Steps 3 and 4 as required | | | |
|---|---|---|---|
| | **Scenario #1** | **Scenario #2** | **Scenario #3** |
| **Monitoring**<br>What is the best way to monitor warning signs that the risks posed by the person may be increasing?<br><br>What events, occurrences, or circumstances should trigger a re-assessment of risk? | | | |
| **Treatment**<br>What treatment or rehabilitation strategies could be implemented to manage the risks posed by the person?<br><br>Which deficits in psychosocial adjustment are high priorities for intervention? | | | |
| **Supervision**<br>What supervision or surveillance strategies could be implemented to manage the risks posed by the person?<br><br>What restrictions on activity, movement, association, or communication are indicated? | | | |
| **Victim Safety Planning**<br>What steps could be taken to enhance the security of any potential victims?<br><br>How might the physical security or self-protective skills of potential victims be improved? | | | |
| **Other Considerations**<br>What events, occurrences, or circumstances might increase or decrease risk?<br><br>What else might be done to manage risk? | | | |

| Step 6: Summary Judgments<br>Document judgments regarding overall risk | | |
|---|---|---|
| Opinion | Coding | Comments |
| **Summary Risk Rating**<br>What level of effort or intervention will be required to prevent further sexual violence?<br><br>To what extent is this opinion limited in light of information that is unclear, unavailable, or missing? | High/Urgent ☐<br>Moderate/Elevated ☐<br>Low/Routine ☐ | |
| **Serious Physical Harm**<br>What is the risk that the sexual violence will involve or escalate into serious or life-threatening physical harm?<br><br>To what extent is this opinion limited in light of information that is unclear, unavailable, or missing? | High ☐<br>Moderate ☐<br>Low ☐ | |
| **Immediate Action Required**<br>Does the person pose any imminent risks?<br><br>What preventive steps were or should be taken immediately? | High ☐<br>Moderate ☐<br>Low ☐ | |
| **Other Risks Indicated**<br>Is there evidence that the person poses other risks, such as general violence, suicide, or self-harm?<br><br>Should the person be evaluated for other risks? | Yes ☐<br>Possibly ☐<br>No ☐ | |
| **Case Review**<br>When should the case be scheduled for routine review (re-assessment)?<br><br>What circumstances should trigger a special review (re-assessment)? | Date for review:<br><br>_____ | |

# Appendix E.

# Summary of 10 Case Files

| Case<br>Violence Type | Victim | Mental Health |
|---|---|---|
| **1 "Atkinson"** | | |
| 1) Assault with Intent to Commit Sexual Abuse<br>2) Harassment x 3<br>3) Harassment | 1) Adult Female (stranger)<br>2) Same victim as #1<br>3) Adult female (Correctional Officer) | Paraphilia – Sexual sadism and exhibitionism (debated)<br>Antisocial Personality Disorder<br>Alcohol Abuse<br>Cannabis Abuse |
| **2 "Crosby"[a]** | | |
| 1) Assault & Battery (attempted rape)<br>2) Statutory Rape<br>3) Sexual Assault (rape)<br>4) Kidnapping (for purpose of rape)<br>5) Kidnapping & Wire Fraud<br>6) Kidnapping (for purpose of rape)<br>7) Kidnapping & 1st degree Forcible Sexual Assault | 1) Adult female<br>2) Adolescent female<br>3) Adolescent female<br>4) Adolescent female (17yo)<br>5) Adolescent female (13yo)<br>6) Adult female (20yo)<br>7) Adolescent females (16 & 18 yo) | Alcohol Abuse |
| **3 "Halko"** | | |
| 1) Sexual Assault (rape)<br><br>Non-adjudicated:<br>A) Sexual assault (victim unable to assist in investigation)<br>B) Burglary & Attempted Rape (Pled to reduced charge of Residential Burglary) | 1) Adult female (20 year old female, met night of offense)<br><br>Non-adjudicated:<br>A) Adult female (22 year old female, stranger)<br>B) Adult female (70 year old, stranger) | Alcohol dependence<br>Cocaine dependence<br>Cannabis abuse<br>Antisocial Personality Disorder (Debated) |
| **4 "Hasman"** | | |
| 1) Indecent Exposure x 2<br>2) Sexual Abuse (2nd degree) x 2<br>3) Sexual Assault<br>4) Attempt to touch | 1) Adolescent females<br>(14 & 17 yo strangers)<br>2) Minor females<br>(5 & 6 years old; niece and stranger)<br>3) Adolescent female<br>(16 yo, stranger?)<br>4) Adult males<br>(at correctional institution | None |

| Case<br>Violence Type | Victim | Mental Health |
|---|---|---|
| **5 "Hatch"** | | |
| 1) Sodomy (Juvenile offense) | 1) Adolescent male | Alcohol Dependence |
| 2) Sexual Abuse (3rd degree) | 2) Adolescent female (13 yo step-daughter) | Amphetamine and Cocaine Dependence |
| 3) Sexual Abuse (3rd degree) | 3) Adult female | |
| **6 "Hughes"** | | |
| 1) Sexual Abuse (3rd degree, rape) | 1) Adolescent male (15yo stranger) | Historical diagnosis of Pedophilia |
| 2) Sexual Abuse (3rd degree, fondling) | 2) Adult male (mentally retarded) | |
| 3) Assault with intent to Commit Sexual Abuse Causing No Bodily Harm | 3) Adolescent male | Query Paraphilia NOS (Hebephilia) |
| 4) Assault with Intent to Commit Sexual Abuse | 4) Adult male (mentally retarded) | Antisocial Personality Disorder |
| 5) Prostitution x 3 | 5) Adolescent male and adult males | |
| 6) Assault with Intent to Commit Sexual Abuse | 6) Adult male | |
| **7 "Magnusson"** | | |
| 1) Sexual Assault | 1) Adolescent female (ex-girlfriend) | Substance abuse |
| 2) Harassment | 2) Same as 1 | |
| 3) Uttering threats | 3) Adolescent female | |
| Pending charges:<br>A) Child Pornography | Pending charges:<br>A) Adolescent females (under age of 15) | |
| **8 "Markoff"** | | |
| 1) Sexual Assault | 1) Adult female (met night of offense) | None |
| 2) Sexual Assault x2 | 2) Adult female prostitutes | |
| **9 "Stowe"** | | |
| 1) Lascivious Acts with Child | 1) Minor female (6yo cousin) | Pedopohilia |
| 2) Sexual Abuse | 2) Adolescent female (15yo in context of "relationship" – he was28) | PTSD (related to sexual abuse as a child) |
| Non-adjudicated Offences (admitted by offender): | Non-adjudicated Offences (admitted by offender): | Anxiety<br>Alcohol and Cannabis Dependence |
| A) Lascivious Acts with Child | A) Minor male (brother of victim 1) | |
| B) Sexual Abuse | B) Minor female (daughter of girlfriend) | |
| C) Sexual Assault | C) 100+ adolescent females (met at parties) | |
| D) Sodomy (juvenile offence) | D) Minor male (neighbor) | |
| E) Sexual Abuse | E) Minor females (13, 10 and 6 yo nieces) | |
| **10 "Tate"** | | |
| 1) Sexual Assault | 1) Adolescent male (intellectually disabled) | Mild mental retardation |
| 2) Sexual Assault | 2) Adult female (acquaintance) | Schizophrenia (Undifferentiated, chronic) |
| 3) Sexual Misconduct | 3) Adult female (mentally ill and "vulnerable") | |

[a] Physical health condition – Scleroderma: Limit to physical mobility

# Appendix F.

# Case Feedback

**Advanced RSVP Training Evaluation**

**Name of Participant:**

**Case:**

**Step 1: Correct rating for presence (PAST) of risk factors?**

| Risk Factor | Correct | 1 Off | 2 Off |
|---|---|---|---|
| RSVP1 | ☐ | ☐ | ☐ |
| RSVP2 | ☐ | ☐ | ☐ |
| RSVP3 | ☐ | ☐ | ☐ |
| RSVP4 | ☐ | ☐ | ☐ |
| RSVP5 | ☐ | ☐ | ☐ |
| RSVP6 | ☐ | ☐ | ☐ |
| RSVP7 | ☐ | ☐ | ☐ |
| RSVP8 | ☐ | ☐ | ☐ |
| RSVP9 | ☐ | ☐ | ☐ |
| RSVP10 | ☐ | ☐ | ☐ |
| RSVP11 | ☐ | ☐ | ☐ |
| RSVP12 | ☐ | ☐ | ☐ |
| RSVP13 | ☐ | ☐ | ☐ |
| RSVP14 | ☐ | ☐ | ☐ |
| RSVP15 | ☐ | ☐ | ☐ |
| RSVP16 | ☐ | ☐ | ☐ |
| RSVP17 | ☐ | ☐ | ☐ |
| RSVP18 | ☐ | ☐ | ☐ |
| RSVP19 | ☐ | ☐ | ☐ |
| RSVP20 | ☐ | ☐ | ☐ |
| RSVP21 | ☐ | ☐ | ☐ |
| RSVP22 | ☐ | ☐ | ☐ |

**Summary:**

**Step 2: Correct rating for presence (RECENT) of risk factors?**

| Risk Factor | Correct | 1 Off | 2 Off |
|---|---|---|---|
| RSVP1 | ☐ | ☐ | ☐ |
| RSVP2 | ☐ | ☐ | ☐ |
| RSVP3 | ☐ | ☐ | ☐ |
| RSVP4 | ☐ | ☐ | ☐ |
| RSVP5 | ☐ | ☐ | ☐ |
| RSVP6 | ☐ | ☐ | ☐ |
| RSVP7 | ☐ | ☐ | ☐ |
| RSVP8 | ☐ | ☐ | ☐ |
| RSVP9 | ☐ | ☐ | ☐ |
| RSVP10 | ☐ | ☐ | ☐ |
| RSVP11 | ☐ | ☐ | ☐ |
| RSVP12 | ☐ | ☐ | ☐ |
| RSVP13 | ☐ | ☐ | ☐ |
| RSVP14 | ☐ | ☐ | ☐ |
| RSVP15 | ☐ | ☐ | ☐ |
| RSVP16 | ☐ | ☐ | ☐ |
| RSVP17 | ☐ | ☐ | ☐ |
| RSVP18 | ☐ | ☐ | ☐ |
| RSVP19 | ☐ | ☐ | ☐ |
| RSVP20 | ☐ | ☐ | ☐ |
| RSVP21 | ☐ | ☐ | ☐ |
| RSVP22 | ☐ | ☐ | ☐ |

**Summary:**

### Step 3: Correct rating for relevance of risk factors?

| Risk Factor | Correct | 1 Off | 2 Off |
|---|---|---|---|
| RSVP1 | ☐ | ☐ | ☐ |
| RSVP2 | ☐ | ☐ | ☐ |
| RSVP3 | ☐ | ☐ | ☐ |
| RSVP4 | ☐ | ☐ | ☐ |
| RSVP5 | ☐ | ☐ | ☐ |
| RSVP6 | ☐ | ☐ | ☐ |
| RSVP7 | ☐ | ☐ | ☐ |
| RSVP8 | ☐ | ☐ | ☐ |
| RSVP9 | ☐ | ☐ | ☐ |
| RSVP10 | ☐ | ☐ | ☐ |
| RSVP11 | ☐ | ☐ | ☐ |
| RSVP12 | ☐ | ☐ | ☐ |
| RSVP13 | ☐ | ☐ | ☐ |
| RSVP14 | ☐ | ☐ | ☐ |
| RSVP15 | ☐ | ☐ | ☐ |
| RSVP16 | ☐ | ☐ | ☐ |
| RSVP17 | ☐ | ☐ | ☐ |
| RSVP18 | ☐ | ☐ | ☐ |
| RSVP19 | ☐ | ☐ | ☐ |
| RSVP20 | ☐ | ☐ | ☐ |
| RSVP21 | ☐ | ☐ | ☐ |
| RSVP22 | ☐ | ☐ | ☐ |

**Summary:**

## Step 4: Formulation of Sexual Violence (Motivators, Disinhibitors, Destabilizers)

**Summary:**

**Documented 2 risk scenarios?**

| | | | |
|---|---|---|---|
| 2 Risk Scenarios | Y | ? | N |

**Identification of plausible, useful, and consensual scenarios?**

| | | | |
|---|---|---|---|
| Plausible | Y | ? | N |
| Useful | Y | ? | N |
| Consensual | Y | ? | N |

**Summary:**

## Step 5: Case Management

**Documented 2 management plans?**

| | | | |
|---|---|---|---|
| 2 Management Plans | Y | ? | N |

**Identification of strategies and tactics?**

| | | | |
|---|---|---|---|
| Strategies | Y | ? | N |
| Tactics | Y | ? | N |

**Summary:**

## Step 6: Correct coding of conclusory opinions

| Conclusory Opinions | Correct | 1 Off | 2 Off |
|---|---|---|---|
| Case Prioritization | ☐ | ☐ | ☐ |
| Serious Physical Harm | ☐ | ☐ | ☐ |
| Immediate Action Required | ☐ | ☐ | ☐ |
| Other Risks Indicated | ☐ | ☐ | ☐ |
| Case Review | Within acceptable range: Yes ☐ No☐ | | |

# Appendix G.

## Case Formulation Similarity Coding Form

Please read the following sections and complete the similarity ratings for each: Case formulations, Case Scenarios and Case Management Strategies. For the Case Scenario and Case Management sections, consider all of the information provided across all scenarios and management strategies (i.e., do not be concerned about columns).

Ratings should be based on overall global judgments for each section. The scale is 1 to 10, where 1 = Not similar at all, ~4 = different in many important respects, but similar in a few, ~7 = similar in many important respects, but different in a few, 10 = very similar.

**Section A: Case Formulation**

1. Please read the *case formulation* page and rate the *overall similarity* between the two provided cases.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|

1 - Not similar at all | 4 - Different in many important respects, but similar in a few | 7 - Similar in many important respects, but different in a few | 10 - Very Similar

1a) Rate the *similarity of the Motivators* provided in the case formulation (Note: consider both the specific motivator provided as well as any explanation included).

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|

1 - Not similar at all | 4 - Different in many important respects, but similar in a few | 7 - Similar in many important respects, but different in a few | 10 - Very Similar

1b) Rate the *similarity of the Disinhibitors* provided in the case formulation (Note: consider both the specific disinhibitor provided as well as any explanation included).

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|

1 - Not similar at all | 4 - Different in many important respects, but similar in a few | 7 - Similar in many important respects, but different in a few | 10 - Very Similar

1c) Rate *the similarity of the Destabilizers* provided in the case formulation (Note: consider both the specific destabilizer provided as well as any explanation included).

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|

1 - Not similar at all | 4 - Different in many important respects, but similar in a few | 7 - Similar in many important respects, but different in a few | 10 - Very Similar

**Section B: Case Scenarios**

2) Please review the *case scenarios* provided and rate the *overall similarity*.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|

1 - Not similar at all | 4 - Different in many important respects, but similar in a few | 7 - Similar in many important respects, but different in a few | 10 - Very Similar

2a) Rate the *similarity of the type of the violence* described in the scenarios

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|

1 - Not similar at all | 4 - Different in many important respects, but similar in a few | 7 - Similar in many important respects, but different in a few | 10 - Very Similar

2b) Rate the *similarity of the victims* provided in the scenarios

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|

1 - Not similar at all | 4 - Different in many important respects, but similar in a few | 7 - Similar in many important respects, but different in a few | 10 - Very Similar

2c) Rate the *similarity of the level of harm* (psychological and physical) provided in the scenarios.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|

1 - Not similar at all | 4 - Different in many important respects, but similar in a few | 7 - Similar in many important respects, but different in a few | 10 - Very Similar

2d) Rate the *similarity of the imminence of violence* provided in the scenarios.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|

1 - Not similar at all | 4 - Different in many important respects, but similar in a few | 7 - Similar in many important respects, but different in a few | 10 - Very Similar

2e) Rate the *similarity of the frequency of violence* provided in the scenarios.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|

1 - Not similar at all | 4 - Different in many important respects, but similar in a few | 7 - Similar in many important respects, but different in a few | 10 - Very Similar

2f) Rate the *similarity of the likelihood of violence* provided in the scenarios.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|

1 - Not similar at all | 4 - Different in many important respects, but similar in a few | 7 - Similar in many important respects, but different in a few | 10 - Very Similar

**Section C: Case Management Strategies**

3. Rate the overall similarity of the management strategies provided.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|

1 - Not similar at all | 4 - Different in many important respects, but similar in a few | 7 - Similar in many important respects, but different in a few | 10 - Very Similar

3a) Rate the similarity of monitoring strategies provided.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|

1 - Not similar at all | 4 - Different in many important respects, but similar in a few | 7 - Similar in many important respects, but different in a few | 10 - Very Similar

3b) Rate the similarity of treatment strategies provided.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|

1 - Not similar at all | 4 - Different in many important respects, but similar in a few | 7 - Similar in many important respects, but different in a few | 10 - Very Similar

3c) Rate the similarity of supervision strategies provided.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|

1 - Not similar at all | 4 - Different in many important respects, but similar in a few | 7 - Similar in many important respects, but different in a few | 10 - Very Similar

3d) Rate the similarity of victim safety planning strategies provided.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|

1 - Not similar at all | 4 - Different in many important respects, but similar in a few | 7 - Similar in many important respects, but different in a few | 10 - Very Similar

# Appendix H.

# Tables for Mean Scores and Standard Deviations for Presence, Relevance and Domain Scores Per Case

**Table H1.**
**Mean Scores and Standard Deviations for Presence and Relevance Risk Factors in the Sexual Violence History Domain**

| Case (*n*) | Chronicity | | Diversity | | Escalation | | Physical Coercion | | Psychological Coercion | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *M* (*SD*) | | *M* (*SD*) | | *M* (*SD*) | | *M* (*SD*) | | *M* (*SD*) | |
| | Pres. | Rel. | Pres. | Rel. | Pres. | Rel. | Pres. | Rel. | Pres. | Rel. |
| Case 1 (9) | 1.89 (.33) | 1.78 (.44) | 1.67 (.71) | 1.33 (.71) | 0.44 (.88) | 0.78 (.97) | 2.00 (.00) | 1.78 (.44) | 1.11 (.93) | 1.00 (.87) |
| Case 2 (11) | 2.00 (.00) | 1.91 (.30) | 0.91 (.83) | 0.91 (.83) | 1.18 (.87) | 1.18 (.87) | 2.00 (.00) | 2.00 (.00) | 0.91 (.94) | 0.82 (.98) |
| Case 3 (9) | 1.78 (.44) | 1.00 (.71) | 0.78 (.83) | 0.44 (.53) | 0.33 (.71) | 0.56 (.73) | 1.89 (.33) | 1.67 (.71) | 1.44 (.88) | 1.22 (.83) |
| Case 4 (12) | 1.00 (1.00) | 1.00 (.77) | 1.36 (.81) | 1.55 (.52) | 1.45 (.69) | 1.27 (.65) | 0.18 (.40) | 0.18 (.40) | 1.73 (.47) | 1.50 (.52) |
| Case 5 (9) | 1.22 (.97) | 1.11 (.93) | 1.78 (.44) | 1.89 (.33) | 1.44 (.53) | 1.11 (.60) | 1.11 (1.05) | 1.00 (1.00) | 2.00 (.00) | 1.44 (.53) |
| Case 6 (6) | 2.00 (.00) | 1.83 (.41) | 0.67 (.82) | 0.67 (1.03) | 0.50 (.84) | 1.00 (.89) | 2.00 (.00) | 1.83 (.41) | 1.33 (1.03) | 1.33 (1.03) |
| Case 7 (14) | 1.14 (.95) | 1.21 (.80) | 0.64 (.75) | 0.79 (.80) | 1.07 (.92) | 1.21 (.80) | 2.00 (.00) | 1.86 (.36) | 1.50 (.86) | 1.36 (.84) |
| Case 8 (7) | 1.14 (.90) | 1.29 (.49) | 0.57 (.79) | 0.43 (.54) | 1.86 (.38) | 1.86 (.38) | 1.86 (.38) | 2.00 (.00) | 0.29 (.49) | 0.29 (.49) |
| Case 9 (12) | 2.00 (.00) | 2.00 (.00) | 1.83 (.39) | 1.67 (.65) | 1.17 (.84) | 1.08 (.79) | 0.67 (.89) | 0.42 (.67) | 2.00 (.00) | 1.75 (.45) |
| Case 10 (14) | 2.00 (.00) | 1.93 (.27) | 1.57 (.65) | 1.50 (.65) | 0.64 (.84) | 0.57 (.76) | 2.00 (.00) | 1.64 (.63) | 1.93 (.27) | 1.79 (.43) |
| Overall Mean | 1.62 (.73) | 1.52 (.69) | 1.22 (.83) | 1.17 (.81) | 1.01 (.87) | 1.04 (.81) | 1.55 (.80) | 1.40 (.84) | 1.49 (.84) | 1.49 (.81) |

*Note.* Pres. = Presence; Rel. = Relevance

**Table H2.**
**Mean Scores and Standard Deviations for Presence and Relevance Risk Factors in the Psychological Adjustment Domain**

| Case (*n*) | Minimization/ Denial | | Attitudes Supportive of Sexual Violence | | Problems with Self-Awareness | | Problems with Stress/Coping | | Problems from Child Abuse | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *M* (*SD*) | | *M* (*SD*) | | *M* (*SD*) | | *M* (*SD*) | | *M* (*SD*) | |
| | Pres. | Rel. | Pres. | Rel. | Pres. | Rel. | Pres. | Rel. | Pres. | Rel. |
| Case 1 (9) | 2.00 (.00) | 1.78 (.44) | 2.00 (.00) | 1.78 (.44) | 2.00 (.00) | 1.78 (.44) | 1.89 (.33) | 1.78 (.44) | 1.33 (.71) | 1.00 (.71) |
| Case 2 (11) | 1.73 (.47) | 1.64 (.51) | 1.55 (.69) | 1.27 (.65) | 1.82 (.41) | 1.91 (.30) | 1.73 (.65) | 1.36 (.67) | 1.64 (.67) | 1.45 (.69) |
| Case 3 (9) | 1.78 (.67) | 1.33 (.71) | 2.00 (.00) | 1.33 (.50) | 1.89 (.33) | 1.67 (.50) | 1.89 (.33) | 2.00 (.00) | 1.56 (.73) | 1.00 (.71) |
| Case 4 (12) | 0.82 (.98) | 0.64 (.67) | 0.72 (.65) | 0.64 (.67) | 1.73 (.65) | 1.45 (.69) | 1.82 (.60) | 1.73 (.65) | 0.36 (.50) | 0.18 (.40) |
| Case 5 (9) | 1.11 (.93) | 0.56 (.53) | 1.22 (.83) | 0.67 (.71) | 1.56 (.73) | 0.89 (.60) | 1.78 (.44) | 1.44 (.73) | 0.78 (.67) | 0.22 (.44) |
| Case 6 (6) | 2.00 (.00) | 1.83 (.41) | 1.50 (.84) | 1.33 (.82) | 2.00 (.00) | 1.80 (.45) | 1.00 (1.00) | 0.60 (.89) | 0.80 (.84) | 0.80 (.84) |
| Case 7 (14) | 2.00 (.00) | 1.79 (.43) | 2.00 (.00) | 1.86 (.36) | 1.86 (.54) | 1.79 (.58) | 1.93 (.27) | 1.79 (.43) | 1.43 (.85) | 1.21 (.80) |
| Case 8 (7) | 2.00 (.00) | 1.86 (.38) | 1.29 (.49) | 1.43 (.54) | 1.57 (.79) | 1.43 (.79) | 1.00 (.89) | 0.83 (.75) | 0.14 (.38) | 0.14 (.38) |
| Case 9 (12) | 1.67 (.49) | 1.08 (.79) | 1.92 (.29) | 1.25 (.62) | 1.58 (.52) | .75 (.62) | 1.83 (.58) | .92 (.67) | 2.00 (.00) | 1.58 (.51) |
| Case 10 (14) | 1.79 (.43) | 1.71 (.47) | 1.07 (.73) | 1.00 (.78) | 1.93 (.27) | 1.93 (.27) | 1.93 (.27) | 1.79 (.43) | 1.93 (.27) | 1.79 (.43) |
| Overall Mean | 1.68 (.65) | 1.41 (.71) | 1.53 (.69) | 1.25 (.71) | 1.79 (.50) | 1.54 (.66) | 1.76 (.57) | 1.50 (.69) | 1.30 (.83) | 1.03 (.82) |

Note. Pres. = Presence; Rel. = Relevance

**Table H3.**
**Mean Scores and Standard Deviations for Presence and Relevance Risk Factors in the Mental Disorder Domain**

| Case (*n*) | Sexual Deviance | | Psychopathic Personality Disorder | | Major Mental Illness | | Problems with Substance Use | | Violent or Suicidal Ideation | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *M* (*SD*) | | *M* (*SD*) | | *M* (*SD*) | | *M* (*SD*) | | *M* (*SD*) | |
| | Pres. | Rel. | Pres. | Rel. | Pres. | Rel. | Pres. | Rel. | Pres. | Rel. |
| Case 1 (9) | 2.00 (.00) | 1.89 (.33) | 2.00 (.00) | 1.89 (.33) | 0.44 (.73) | 0.33 (.71) | 2.00 (.00) | 1.44 (.53) | 2.00 (.00) | 1.67 (.50) |
| Case 2 (11) | 1.36 (.81) | 1.36 (.81) | 0.40 (.84) | 0.40 (.84) | 0.82 (.87) | 0.73 (.79) | 2.00 (.00) | 1.91 (.30) | 1.73 (.47) | 1.55 (.69) |
| Case 3 (9) | 1.33 (.71) | 0.67 (.50) | 1.67 (.71) | 1.67 (.71) | 0.44 (.53) | 0.44 (.53) | 2.00 (.00) | 2.00 (.00) | 0.33 (.71) | 0.11 (.33) |
| Case 4 (12) | 1.00 (.63) | 1.00 (.63) | 0.64 (.81) | 0.55 (.69) | 1.64 (.81) | 1.55 (.82) | 0.00 (.00) | 0.09 (.30) | 1.91 (.30) | 1.64 (.67) |
| Case 5 (9) | 1.22 (.67) | 1.11 (.60) | 0.43 (.79) | 0.29 (.49) | 1.00 (1.00) | 1.00 (1.00) | 1.11 (1.05) | 1.11 (.93) | 1.33 (.71) | 0.89 (.78) |
| Case 6 (6) | 2.00 (.00) | 1.83 (.41) | 0.80 (.84) | 0.80 (.84) | 0.17 (.41) | 0.17 (.41) | 0.50 (.55) | 0.50 (.55) | 0.67 (.82) | 0.67 (.82) |
| Case 7 (14) | 1.21 (.80) | 1.21 (.80) | 1.57 (.65) | 1.57 (.65) | 0.79 (.80) | 0.86 (.77) | 2.00 (.00) | 2.00 (.00) | 1.79 (.43) | 1.64 (.50) |
| Case 8 (7) | 1.14 (.38) | 1.29 (.49) | 0.60 (.89) | 0.60 (.89) | 0.00 (.00) | 0.00 (.00) | 1.29 (.76) | 1.14 (.90) | 1.29 (.76) | 1.14 (.69) |
| Case 9 (12) | 2.00 (.00) | 1.92 (.29) | 0.18 (.41) | 0.18 (.41) | 1.67 (.65) | 1.08 (.67) | 2.00 (.00) | 1.08 (.67) | 2.00 (.00) | 0.33 (.49) |
| Case 10 (14) | 1.43 (.65) | 1.43 (.65) | 0.85 (.69) | 0.85 (.69) | 2.00 (.00) | 1.93 (.27) | 1.57 (.76) | 1.07 (.62) | 1.86 (.36) | 1.64 (.63) |
| Overall Mean | 1.45 (.67) | 1.36 (.69) | 0.95 (.88) | 0.91 (.86) | 1.02 (.91) | 0.92 (.86) | 1.51 (.82) | 1.27 (.80) | 1.58 (.70) | 1.19 (.82) |

Note. Pres. = Presence; Rel. = Relevance

**Table H4.**
**Mean Scores and Standard Deviations for Presence and Relevance Risk Factors in the Social Adjustment Domain**

| Case (*n*) | Problems with Intimate Relationships | | Problems with Non-intimate Relationships | | Problems with Employment | | Non-sexual Criminality | |
|---|---|---|---|---|---|---|---|---|
| | *M* (*SD*) | | *M* (*SD*) | | *M* (*SD*) | | *M* (*SD*) | |
| | Presence | Relevance | Presence | Relevance | Presence | Relevance | Presence | Relevance |
| Case 1 (9) | 2.00 (.00) | 1.89 (.33) | 2.00 (.00) | 2.00 (.00) | 2.00 (.00) | 1.78 (.44) | 2.00 (.00) | 1.56 (.53) |
| Case 2 (11) | 2.00 (.00) | 1.73 (.47) | 1.73 (.65) | 1.64 (.65) | 1.73 (.47) | 1.36 (.51) | .27 (.65) | .36 (.67) |
| Case 3 (9) | 1.89 (.33) | 1.89 (.33) | 2.00 (.00) | 1.56 (.73) | 2.00 (.00) | 1.78 (.44) | 2.00 (.00) | 1.56 (.53) |
| Case 4 (12) | 2.00 (.00) | 2.00 (.00) | 2.00 (.00) | 2.00 (.00) | 1.73 (.65) | 1.64 (.67) | 2.00 (.00) | 1.73 (.47) |
| Case 5 (9) | 1.89 (.33) | 1.78 (.44) | 1.56 (.73) | 1.56 (.73) | 1.78 (.44) | 1.33 (.71) | 2.00 (.00) | 1.44 (.53) |
| Case 6 (6) | 2.00 (.00) | 1.33 (.82) | 2.00 (.00) | 1.33 (.82) | 2.00 (.00) | 1.67 (.52) | 1.50 (.84) | .50 (.55) |
| Case 7 (14) | 2.00 (.00) | 2.00 (.00) | 1.79 (.43) | 1.64 (.50) | 2.00 (.00) | 1.79 (.43) | 1.93 (.27) | 1.71 (.47) |
| Case 8 (7) | 1.00 (.82) | .71 (.95) | .29 (.49) | .14 (.38) | .14 (.38) | .00 (.00) | .29 (.49) | .00 (.00) |
| Case 9 (12) | 1.92 (.29) | 1.67 (.49) | 1.92 (.29) | 1.17 (.84) | 1.67 (.49) | 1.00 (.74) | 2.00 (.00) | .50 (.52) |
| Case 10 (14) | 1.93 (.27) | 1.93 (.27) | 1.93 (.27) | 1.79 (.43) | 1.79 (.58) | 1.64 (.63) | 1.43 (.65) | .93 (.62) |
| Overall Mean | 1.89 (.37) | 1.75 (.53) | 1.76 (.57) | 1.54 (.71) | 1.73 (.60) | 1.44 (.71) | 1.58 (.75) | 1.08 (.78) |

**Table H5.**
**Mean Scores and Standard Deviations for Presence and Relevance Risk Factors in the Manageability Domain**

| Case (*n*) | Problems with Planning | | Problems with Treatment | | Problems with Supervision | |
|---|---|---|---|---|---|---|
| | *M* (*SD*) | | *M* (*SD*) | | *M* (*SD*) | |
| | Presence | Relevance | Presence | Relevance | Presence | Relevance |
| Case 1 (9) | 1.89 (.33) | 1.78 (.44) | 1.89 (.33) | 1.56 (.73) | 2.00 (.00) | 1.67 (.71) |
| Case 2 (11) | 2.00 (.00) | 1.91 (.30) | 1.64 (.67) | 1.55 (.69) | 1.45 (.82) | 1.45 (.82) |
| Case 3 (9) | 1.78 (.44) | 1.56 (.53) | 1.78 (.44) | 1.78 (.44) | 2.00 (.00) | 1.67 (.50) |
| Case 4 (12) | 1.82 (.40) | 1.73 (.47) | 1.82 (.60) | 1.09 (.70) | 2.00 (.00) | 1.55 (.52) |
| Case 5 (9) | 1.67 (.71) | 1.22 (.67) | 1.56 (.88) | .89 (.60) | 2.00 (.00) | 1.11 (.33) |
| Case 6 (6) | 1.17 (.75) | 1.33 (.82) | 1.83 (.41) | 1.67 (.52) | 1.83 (.41) | 1.67 (.52) |
| Case 7 (14) | 1.93 (.27) | 1.86 (.36) | 2.00 (.00) | 1.86 (.36) | 2.00 (.00) | 2.00 (.00) |
| Case 8 (7) | .57 (.79) | .57 (.79) | .40 (.89) | .60 (.89) | .25 (.50) | .50 (.58) |
| Case 9 (12) | 1.92 (.29) | .92 (.90) | 2.00 (.00) | .67 (.49) | 1.83 (.58) | .75 (.62) |
| Case 10 (14) | 2.00 (.00) | 2.00 (.00) | 1.93 (.27) | 1.93 (.27) | 2.00 (.00) | 1.93 (.27) |
| Overall Mean | 1.75 (.55) | 1.55 (.68) | 1.77 (.58) | 1.41 (.71) | 1.83 (.51) | 1.51 (.66) |

# Appendix I.

# Tables for Percent Agreement between Participant and Gold Standard Ratings of Presence and Relevance Scores and Summary Judgments per Case

**Table I1.**
**Percent Agreement between Participant Ratings (*n* = 9) and Expert Ratings for RSVP Items, Domain and Total Scores, and Summary Judgments for Case 1**

| | % Agreement | |
|---|---|---|
| | **Presence Rating** | **Relevance Rating** |
| 1. Chronicity of Sexual Violence | 89 | 78 |
| 2. Diversity of Sexual Violence | 78 | 33 |
| 3. Escalation of Sexual Violence | 78 | 44 |
| 4. Physical Coercion in Sexual Violence | 100 | 67 |
| 5. Psychological Coercion in Sexual Violence | 56 | 44 |
| **Mean Agreement - Sexual Violence History Domain** | 80 | 53 |
| 6. Extreme Minimization/Denial of Sexual Violence | 100 | 78 |
| 7. Attitudes that Support or Condone Sexual Violence | 100 | 78 |
| 8. Problems with Self-Awareness | 100 | 78 |
| 9. Problems with Stress or Coping | 89 | 78 |
| 10. Problems Resulting from Child Abuse | 56 | 56 |
| **Mean Agreement - Psychological Adjustment Domain** | 94 | 74 |
| 11. Sexual Deviance | 100 | 89 |
| 12. Psychopathic Personality Disorder | 100 | 89 |
| 13. Major Mental Illness | 67 | 78 |
| 14. Problems with Substance Use | 100 | 44 |
| 15. Violent or Suicidal Ideation | 100 | 78 |
| **Mean Agreement - Mental Disorder Domain** | 93 | 76 |
| 16. Problems with Intimate Relationships | 100 | 78 |
| 17. Problems with Non-Intimate Relationships | 100 | 100 |
| 18. Problems with Employment | 100 | 78 |
| 19. Non-sexual Criminality | 100 | 56 |
| **Mean Agreement - Social Adjustment Domain** | 100 | 78 |
| 20. Problems with Planning | 89 | 89 |
| 21. Problems with Treatment | 89 | 78 |
| 22. Problems with Supervision | 100 | 89 |
| **Mean Agreement - Manageability Domain** | 93 | 85 |
| **Overall Mean Agreement** | 91 | 72 |
| **Summary Judgments** | | |
| Case Prioritization | 78 | |
| Level of Harm | 22 | |
| Imminence | 56 | |

**Table I2.**
**Percent Agreement between Participant Ratings (*n* = 11) and Expert Ratings for RSVP Items, Domain and Total Scores, and Summary Judgments for Case 2**

| | % Agreement | |
|---|---|---|
| | **Presence Rating** | **Relevance Rating** |
| 1. Chronicity of Sexual Violence | 100 | 91 |
| 2. Diversity of Sexual Violence | 36 | 36 |
| 3. Escalation of Sexual Violence | 46 | 46 |
| 4. Physical Coercion in Sexual Violence | 100 | 100 |
| 5. Psychological Coercion in Sexual Violence | 46 | 55 |
| **Mean Agreement - Sexual Violence History Domain** | 66 | 66 |
| 6. Extreme Minimization/Denial of Sexual Violence | 73 | 64 |
| 7. Attitudes that Support or Condone Sexual Violence | 64 | 36 |
| 8. Problems with Self-Awareness | 82 | 91 |
| 9. Problems with Stress or Coping | 82 | 46 |
| 10. Problems Resulting from Child Abuse | 73 | 55 |
| **Mean Agreement - Psychological Adjustment Domain** | 75 | 58 |
| 11. Sexual Deviance | 27 | 27 |
| 12. Psychopathic Personality Disorder | 82 | 82 |
| 13. Major Mental Illness | 46 | 46 |
| 14. Problems with Substance Use | 100 | 91 |
| 15. Violent or Suicidal Ideation | 27 | 27 |
| **Mean Agreement - Mental Disorder Domain** | 56 | 55 |
| 16. Problems with Intimate Relationships | 100 | 73 |
| 17. Problems with Non-Intimate Relationships | 82 | 73 |
| 18. Problems with Employment | 73 | 36 |
| 19. Non-sexual Criminality | 82 | 73 |
| **Mean Agreement - Social Adjustment Domain** | 84 | 64 |
| 20. Problems with Planning | 100 | 91 |
| 21. Problems with Treatment | 73 | 64 |
| 22. Problems with Supervision | 64 | 64 |
| **Mean Agreement - Manageability Domain** | 79 | 73 |
| **Overall Mean Agreement** | 71 | 62 |
| **Summary Judgments** | | |
| Case Prioritization | 91 | |
| Level of Harm | 73 | |
| Imminence | 64 | |

**Table I3.**
**Percent Agreement between Participant Ratings (*n* = 9) and Expert Ratings for RSVP Items, Domain and Total Scores, and Summary Judgments for Case 3**

| | % Agreement | |
|---|---|---|
| | **Presence Rating** | **Relevance Rating** |
| 1. Chronicity of Sexual Violence | 78 | 56 |
| 2. Diversity of Sexual Violence | 44 | 56 |
| 3. Escalation of Sexual Violence | 78 | 56 |
| 4. Physical Coercion in Sexual Violence | 89 | 78 |
| 5. Psychological Coercion in Sexual Violence | 67 | 44 |
| **Mean Agreement - Sexual Violence History Domain** | 71 | 58 |
| 6. Extreme Minimization/Denial of Sexual Violence | 89 | 44 |
| 7. Attitudes that Support or Condone Sexual Violence | 100 | 33 |
| 8. Problems with Self-Awareness * | 89 | 67 |
| 9. Problems with Stress or Coping | 89 | 100 |
| 10. Problems Resulting from Child Abuse | 67 | 56 |
| **Mean Agreement - Psychological Adjustment Domain** | 87 | 60 |
| 11. Sexual Deviance | 44 | 67 |
| 12. Psychopathic Personality Disorder | 78 | 78 |
| 13. Major Mental Illness | 56 | 56 |
| 14. Problems with Substance Use | 100 | 100 |
| 15. Violent or Suicidal Ideation | 11 | 11 |
| **Mean Agreement - Mental Disorder Domain** | 58 | 62 |
| 16. Problems with Intimate Relationships | 89 | 89 |
| 17. Problems with Non-Intimate Relationships | 100 | 67 |
| 18. Problems with Employment | 100 | 78 |
| 19. Non-sexual Criminality | 100 | 56 |
| **Mean Agreement - Social Adjustment Domain** | 97 | 73 |
| 20. Problems with Planning | 78 | 56 |
| 21. Problems with Treatment | 78 | 78 |
| 22. Problems with Supervision | 100 | 67 |
| **Mean Agreement - Manageability Domain** | 85 | 67 |
| **Overall Mean Agreement** | 78 | 63 |
| **Summary Judgments** | | |
| Case Prioritization | 44 | |
| Level of Harm | 56 | |
| Imminence | 56 | |

**Table I4.**
**Percent Agreement between Participant Ratings (*n* = 12) and Expert Ratings for RSVP Items, Domain and Total Scores, and Summary Judgments for Case 4**

| | % Agreement | |
|---|---|---|
| | **Presence Rating** | **Relevance Rating** |
| 1. Chronicity of Sexual Violence | 46 | 27 |
| 2. Diversity of Sexual Violence | 55 | 55 |
| 3. Escalation of Sexual Violence | 36 | 55 |
| 4. Physical Coercion in Sexual Violence | 82 | 82 |
| 5. Psychological Coercion in Sexual Violence | 27 | 55 |
| **Mean Agreement - Sexual Violence History Domain** | 49 | 55 |
| 6. Extreme Minimization/Denial of Sexual Violence | 55 | 46 |
| 7. Attitudes that Support or Condone Sexual Violence | 55 | 46 |
| 8. Problems with Self-Awareness | 9 | 36 |
| 9. Problems with Stress or Coping | 91 | 82 |
| 10. Problems Resulting from Child Abuse | 64 | 82 |
| **Mean Agreement - Psychological Adjustment Domain** | 55 | 58 |
| 11. Sexual Deviance | 64 | 64 |
| 12. Psychopathic Personality Disorder | 55 | 55 |
| 13. Major Mental Illness | 82 | 73 |
| 14. Problems with Substance Use | 100 | 91 |
| 15. Violent or Suicidal Ideation | 91 | 73 |
| **Mean Agreement - Mental Disorder Domain** | 78 | 71 |
| 16. Problems with Intimate Relationships | 100 | 100 |
| 17. Problems with Non-Intimate Relationships | 100 | 100 |
| 18. Problems with Employment | 82 | 18 |
| 19. Non-sexual Criminality | 100 | 73 |
| **Mean Agreement - Social Adjustment Domain** | 96 | 73 |
| 20. Problems with Planning | 82 | 73 |
| 21. Problems with Treatment | 91 | 55 |
| 22. Problems with Supervision | 100 | 55 |
| **Mean Agreement - Manageability Domain** | 91 | 61 |
| **Overall Mean Agreement** | 71 | 63 |
| **Summary Judgments** | | |
| Case Prioritization | 73 | |
| Level of Harm | 73 | |
| Imminence | 27 | |

**Table I5.**
**Percent Agreement between Participant Ratings (*n* = 9) and Expert Ratings for RSVP Items, Domain and Total Scores, and Summary Judgments for Case 5**

| | % Agreement | |
|---|---|---|
| | **Presence Rating** | **Relevance Rating** |
| 1. Chronicity of Sexual Violence | 56 | 44 |
| 2. Diversity of Sexual Violence | 78 | 89 |
| 3. Escalation of Sexual Violence | 44 | 67 |
| 4. Physical Coercion in Sexual Violence | 56 | 44 |
| 5. Psychological Coercion in Sexual Violence | 100 | 44 |
| **Mean Agreement - Sexual Violence History Domain** | 67 | 58 |
| 6. Extreme Minimization/Denial of Sexual Violence | 44 | 56 |
| 7. Attitudes that Support or Condone Sexual Violence | 44 | 44 |
| 8. Problems with Self-Awareness | 67 | 67 |
| 9. Problems with Stress or Coping | 78 | 56 |
| 10. Problems Resulting from Child Abuse | 56 | 22 |
| **Mean Agreement - Psychological Adjustment Domain** | 58 | 49 |
| 11. Sexual Deviance | 56 | 22 |
| 12. Psychopathic Personality Disorder | 11 | 22 |
| 13. Major Mental Illness | 44 | 44 |
| 14. Problems with Substance Use | 56 | 44 |
| 15. Violent or Suicidal Ideation | 44 | 44 |
| **Mean Agreement - Mental Disorder Domain** | 42 | 35 |
| 16. Problems with Intimate Relationships | 89 | 78 |
| 17. Problems with Non-Intimate Relationships | 67 | 67 |
| 18. Problems with Employment | 78 | 44 |
| 19. Non-sexual Criminality | 100 | 44 |
| **Mean Agreement - Social Adjustment Domain** | 84 | 58 |
| 20. Problems with Planning | 78 | 56 |
| 21. Problems with Treatment | 78 | 67 |
| 22. Problems with Supervision | 100 | 89 |
| **Mean Agreement - Manageability Domain** | 85 | 71 |
| **Overall Mean Agreement** | 65 | 52 |
| **Summary Judgments** | | |
| Case Prioritization | 11 | |
| Level of Harm | 56 | |
| Imminence | 22 | |

**Table I6.**
**Percent Agreement between Participant Ratings (*n* = 6) and Expert Ratings for RSVP Items, Domain and Total Scores, and Summary Judgments for Case 6**

| | % Agreement | |
|---|---|---|
| | **Presence Rating** | **Relevance Rating** |
| 1. Chronicity of Sexual Violence | 100 | 83 |
| 2. Diversity of Sexual Violence | 50 | 67 |
| 3. Escalation of Sexual Violence | 67 | 33 |
| 4. Physical Coercion in Sexual Violence | 100 | 83 |
| 5. Psychological Coercion in Sexual Violence | 67 | 67 |
| **Mean Agreement - Sexual Violence History Domain** | 77 | 67 |
| 6. Extreme Minimization/Denial of Sexual Violence | 100 | 83 |
| 7. Attitudes that Support or Condone Sexual Violence | 67 | 50 |
| 8. Problems with Self-Awareness | 83 | 67 |
| 9. Problems with Stress or Coping | 17 | 67 |
| 10. Problems Resulting from Child Abuse | 50 | 50 |
| **Mean Agreement - Psychological Adjustment Domain** | 63 | 63 |
| 11. Sexual Deviance | 100 | 83 |
| 12. Psychopathic Personality Disorder | 50 | 50 |
| 13. Major Mental Illness | 83 | 83 |
| 14. Problems with Substance Use | 50 | 50 |
| 15. Violent or Suicidal Ideation | 33 | 50 |
| **Mean Agreement - Mental Disorder Domain** | 63 | 63 |
| 16. Problems with Intimate Relationships | 100 | 50 |
| 17. Problems with Non-Intimate Relationships | 100 | 50 |
| 18. Problems with Employment | 100 | 33 |
| 19. Non-sexual Criminality | 17 | 50 |
| **Mean Agreement - Social Adjustment Domain** | 79 | 46 |
| 20. Problems with Planning | 50 | 33 |
| 21. Problems with Treatment | 83 | 67 |
| 22. Problems with Supervision | 83 | 67 |
| **Mean Agreement - Manageability Domain** | 72 | 56 |
| **Overall Mean Agreement** | 70 | 60 |
| **Summary Judgments** | | |
| Case Prioritization | 17 | |
| Level of Harm | 50 | |
| Imminence | 33 | |

**Table I7.**
**Percent Agreement between Participant Ratings (*n* = 14) and Expert Ratings for RSVP Items, Domain and Total Scores, and Summary Judgments for Case 7**

| | % Agreement | |
|---|---|---|
| | **Presence Rating** | **Relevance Rating** |
| 1. Chronicity of Sexual Violence | 50 | 43 |
| 2. Diversity of Sexual Violence | 36 | 43 |
| 3. Escalation of Sexual Violence | 43 | 43 |
| 4. Physical Coercion in Sexual Violence | 100 | 86 |
| 5. Psychological Coercion in Sexual Violence | 71 | 57 |
| **Mean Agreement - Sexual Violence History Domain** | 60 | 54 |
| 6. Extreme Minimization/Denial of Sexual Violence | 100 | 79 |
| 7. Attitudes that Support or Condone Sexual Violence | 100 | 86 |
| 8. Problems with Self-Awareness | 93 | 86 |
| 9. Problems with Stress or Coping | 93 | 79 |
| 10. Problems Resulting from Child Abuse | 64 | 36 |
| **Mean Agreement - Psychological Adjustment Domain** | 90 | 73 |
| 11. Sexual Deviance | 36 | 36 |
| 12. Psychopathic Personality Disorder | 29 | 64 |
| 13. Major Mental Illness | 36 | 36 |
| 14. Problems with Substance Use | 100 | 100 |
| 15. Violent or Suicidal Ideation | 79 | 64 |
| **Mean Agreement - Mental Disorder Domain** | 56 | 60 |
| 16. Problems with Intimate Relationships | 100 | 100 |
| 17. Problems with Non-Intimate Relationships | 79 | 36 |
| 18. Problems with Employment | 100 | 79 |
| 19. Non-sexual Criminality | 93 | 71 |
| **Mean Agreement - Social Adjustment Domain** | 93 | 72 |
| 20. Problems with Planning | 93 | 86 |
| 21. Problems with Treatment | 100 | 86 |
| 22. Problems with Supervision | 100 | 100 |
| **Mean Agreement - Manageability Domain** | 98 | 91 |
| **Overall Mean Agreement** | 77 | 68 |
| **Summary Judgments** | | |
| Case Prioritization | 71 | |
| Level of Harm | 79 | |
| Imminence | 29 | |

**Table I8.**
**Percent Agreement between Participant Ratings (n = 7) and Expert Ratings for RSVP Items, Domain and Total Scores, and Summary Judgments for Case 8**

| | % Correct Across all Cases | |
|---|---|---|
| | **Presence Rating** | **Relevance Rating** |
| 1. Chronicity of Sexual Violence | 43 | 71 |
| 2. Diversity of Sexual Violence | 29 | 43 |
| 3. Escalation of Sexual Violence | 86 | 86 |
| 4. Physical Coercion in Sexual Violence | 86 | 100 |
| 5. Psychological Coercion in Sexual Violence | 71 | 71 |
| **Mean Agreement - Sexual Violence History Domain** | 63 | 74 |
| 6. Extreme Minimization/Denial of Sexual Violence | 100 | 86 |
| 7. Attitudes that Support or Condone Sexual Violence | 71 | 57 |
| 8. Problems with Self-Awareness | 71 | 57 |
| 9. Problems with Stress or Coping | 29 | 43 |
| 10. Problems Resulting from Child Abuse | 86 | 86 |
| **Mean Agreement - Psychological Adjustment Domain** | 71 | 66 |
| 11. Sexual Deviance | 86 | 71 |
| 12. Psychopathic Personality Disorder | 71 | 71 |
| 13. Major Mental Illness | 100 | 100 |
| 14. Problems with Substance Use | 43 | 29 |
| 15. Violent or Suicidal Ideation | 43 | 57 |
| **Mean Agreement - Mental Disorder Domain** | 69 | 66 |
| 16. Problems with Intimate Relationships | 29 | 29 |
| 17. Problems with Non-Intimate Relationships | 71 | 86 |
| 18. Problems with Employment | 86 | 100 |
| 19. Non-sexual Criminality | 71 | 100 |
| **Mean Agreement - Social Adjustment Domain** | 64 | 79 |
| 20. Problems with Planning | 57 | 57 |
| 21. Problems with Treatment | 86 | 71 |
| 22. Problems with Supervision | 86 | 71 |
| **Mean Agreement - Manageability Domain** | 76 | 66 |
| **Overall Mean Agreement** | 68 | 70 |
| **Summary Judgments** | | |
| Case Prioritization | 43 | |
| Level of Harm | 14 | |
| Imminence | 14 | |

**Table I9.**
**Percent Agreement between Participant Ratings (*n* = 12) and Expert Ratings for RSVP Items, Domain and Total Scores, and Summary Judgments for Case 9**

| | % Correct Across all Cases | |
|---|---|---|
| | **Presence Rating** | **Relevance Rating** |
| 1. Chronicity of Sexual Violence | 100 | 100 |
| 2. Diversity of Sexual Violence | 83 | 75 |
| 3. Escalation of Sexual Violence | 33 | 42 |
| 4. Physical Coercion in Sexual Violence | 17 | 67 |
| 5. Psychological Coercion in Sexual Violence | 100 | 75 |
| **Mean Agreement - Sexual Violence History Domain** | 67 | 72 |
| 6. Extreme Minimization/Denial of Sexual Violence | 67 | 42 |
| 7. Attitudes that Support or Condone Sexual Violence | 92 | 33 |
| 8. Problems with Self-Awareness * | 58 | 58 |
| 9. Problems with Stress or Coping | 92 | 58 |
| 10. Problems Resulting from Child Abuse | 100 | 42 |
| **Mean Agreement - Psychological Adjustment Domain** | 82 | 47 |
| 11. Sexual Deviance | 100 | 92 |
| 12. Psychopathic Personality Disorder | 83 | 83 |
| 13. Major Mental Illness | 75 | 58 |
| 14. Problems with Substance Use | 100 | 58 |
| 15. Violent or Suicidal Ideation | 100 | 67 |
| **Mean Agreement - Mental Disorder Domain** | 92 | 72 |
| 16. Problems with Intimate Relationships | 92 | 67 |
| 17. Problems with Non-Intimate Relationships | 92 | 33 |
| 18. Problems with Employment | 67 | 50 |
| 19. Non-sexual Criminality | 100 | 50 |
| **Mean Agreement - Social Adjustment Domain** | 88 | 50 |
| 20. Problems with Planning | 92 | 33 |
| 21. Problems with Treatment | 100 | 67 |
| 22. Problems with Supervision | 92 | 58 |
| **Mean Agreement - Manageability Domain** | 95 | 53 |
| **Overall Mean Agreement** | 83 | 59 |
| **Summary Judgments** | | |
| Case Prioritization | 42 | |
| Level of Harm | 67 | |
| Imminence | 33 | |

**Table I10.**
**Percent Agreement between Participant Ratings (*n* = 14) and Expert Ratings for RSVP Items, Domain and Total Scores, and Summary Judgments for Case 10**

| | % Correct Across all Cases | |
|---|---|---|
| | **Presence Rating** | **Relevance Rating** |
| 1. Chronicity of Sexual Violence | 100 | 93 |
| 2. Diversity of Sexual Violence | 64 | 57 |
| 3. Escalation of Sexual Violence | 57 | 57 |
| 4. Physical Coercion in Sexual Violence | 100 | 71 |
| 5. Psychological Coercion in Sexual Violence | 93 | 79 |
| **Mean Agreement - Sexual Violence History Domain** | 83 | 71 |
| 6. Extreme Minimization/Denial of Sexual Violence | 79 | 71 |
| 7. Attitudes that Support or Condone Sexual Violence | 50 | 29 |
| 8. Problems with Self-Awareness | 93 | 93 |
| 9. Problems with Stress or Coping | 93 | 79 |
| 10. Problems Resulting from Child Abuse | 93 | 79 |
| **Mean Agreement - Psychological Adjustment Domain** | 82 | 70 |
| 11. Sexual Deviance | 43 | 43 |
| 12. Psychopathic Personality Disorder | 50 | 50 |
| 13. Major Mental Illness | 100 | 93 |
| 14. Problems with Substance Use | 71 | 64 |
| 15. Violent or Suicidal Ideation | 86 | 71 |
| **Mean Agreement - Mental Disorder Domain** | 70 | 64 |
| 16. Problems with Intimate Relationships | 93 | 93 |
| 17. Problems with Non-Intimate Relationships | 93 | 79 |
| 18. Problems with Employment | 86 | 71 |
| 19. Non-sexual Criminality | 43 | 64 |
| **Mean Agreement - Social Adjustment Domain** | 79 | 77 |
| 20. Problems with Planning | 100 | 100 |
| 21. Problems with Treatment | 93 | 93 |
| 22. Problems with Supervision | 100 | 93 |
| **Mean Agreement - Manageability Domain** | 98 | 95 |
| **Overall Mean Agreement** | 81 | 74 |
| **Summary Judgments** | | |
| Case Prioritization | 79 | |
| Level of Harm | 64 | |
| Imminence | 43 | |