UNIVERSITY OF BIRMINGHAM

MANAGING SEXUAL OFFENDERS IN THE COMMUNITY

by

MARGUERITE LOUISE DONATHY

A thesis submitted to the Centre for Forensic and Criminological Psychology

of the University of Birmingham

for the degree of ForenPsyD

Centre for Forensic and Criminological Psychology

School of Psychology

College of Life and Environmental Sciences

The University of Birmingham

October 2018

Abstract

The aim of the current thesis was to explore current methods being utilised to manage sexual offenders in the community. Specifically, it centred on assessing risk by combining actuarial data and dynamic risk factors, together with protective factors. Emphasis was placed on case formulation with a view to preparing holistic risk assessments that are tailored to individual offenders.

The introduction (chapter one) provides an overview of the evolution of risk assessments with sexual and violent offenders and considers the changing role of police officers managing offenders in the community within the context of the sex offender register and Multi-Agency Protection Panel Arrangements (MAPPA). The chapter goes on to describe the current research, introducing a relatively new SPJ tool, the Active Risk Management System (ARMS) and setting out the aims of this thesis.

Chapter two is an updated systematic review evaluating the effectiveness of risk assessment tools for predicting sexual recidivism in adult male offenders. It describes the employment of Actuarial Risk Assessment Instruments (ARAIs) and the biases inherent within them. Difficulties in defining recidivism were explored, together with the variability in risk probabilities. Within 41 quality assessed studies that evaluated the predictive validity of 16 static and 9 dynamic risk assessment tools, this review revealed largely moderate predictive accuracy. Only three of 25 tools consistently demonstrated a large effect size (AUC > .714). These were STATIC-2002R, VASOR-2 and SRA-FV. Most risk assessment tools within the current review demonstrated moderate predictive accuracy (18 of 25 tools). However, four of the assessment tools were low in predictive accuracy (RRASOR, STABLE-2000, SVR-20 and SARN-TMA). The findings in

relation to previous reviews are discussed, including the strengths and limitations of the methodology. The chapter concludes that whilst actuarial assessments can be helpful, they are not best used independently and ought to form part of a wider assessment that includes case formulation and scenario planning. Examples include the employment of the Risk for Sexual Violence Protocol (RSVP) and the Structured Assessment of Protective Factors (SAPROF).

Chapter three aimed to investigate the usefulness of ARMS, by conducting an exploratory study ascertaining the views of police officers (*N=22*) employing the tool in their daily work with sex offenders (phase one) and gathering their opinions with regard to the ARMS training they received (*N=31*, phase two). Descriptive statistics were derived from Likert scale questionnaires developed by the author and qualitative data, collected through focus groups and group interviews were transcribed and analysed employing thematic analysis. Nine basic themes were identified and within these, a number of organising themes. The global theme of 'risk reduction' was identified. The results suggested that police officers employing ARMS in their daily work would like to see some changes to the tool, with common complaints about time needed to complete assessments. Although aspects of the training were criticized, most participants considered that employing ARMS in their daily work improved their confidence in assessing sexual offenders, particularly when training incorporated information about psychosexual development. Phase three was developed with a view to exploring whether dynamic risk and protective factors (based on structured professional judgement) altered perceived risk when combined with RM2000. Data were collected from a regional police force (*N=434*). Initial factor analyses were conducted, followed by Multidimensional Scaling. The results indicated that rather than a set of eleven

individual factors, ARMS may best be understood as four components, namely, sexual risk, protective factors, hostile attitude and relationship status. Whilst the results have important practice implications for police officers managing offenders in the community, additional longitudinal research will be required with regard to the effectiveness of ARMS.

As RM2000 is incorporated within ARMS assessments, chapter four provides a critique of this tool. The chapter concludes that whilst RM2000 can be a helpful adjunct to risk assessment, it is important for practitioners to take an approach that combines these results with structured professional judgement, protective factors and a thorough case formulation.

The thesis concludes with a general discussion regarding the importance of a tailored approach to sexual offender assessment and treatment.

Acknowledgements

I would like to thank several members of staff at UoB for their support/guidance over the years. In particular, my academic supervisors, Dr Jessica Woodhams and Dr Darren Bishopp, who encouraged me to persevere when challenges arose. To Sue Hanson (former Course Administrator) who encouraged me in my studies and firmly believed in me. She showed such patience and warmth. I would also like to thank my examiners, Dr Zoe Stephenson and Dr Jayne Allam for valued feedback regarding modifications to the final thesis.

I would like to thank a number of individuals of various ranks and positions within the police force (and associated organisations) for their support in accessing the relevant data, together with the team of police officers whose patience shone through when technical challenges arose.

I would also like to thank the practice supervisors and their organisations for providing me with placements that undoubtedly increased my knowledge but also significantly enhanced my confidence.

Special thanks go to Professor Daniel T Wilcox and his wife, Helen Wilcox MBE, both of whom are inspirations and without whom none of this would have been achievable. They have supported me emotionally and financially and, like Sue Hanson, have always believed in my abilities.

To my husband, Mark Taylor, for tolerating a largely absent wife for most of our married life and for his unfaltering encouragement and confidence in me. Finally, to my family and friends whom I have missed spending valuable time with, I look forward to recapturing precious moments and building happy memories.

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1

Introduction

Introduction

The nature of risk assessments in the world of forensic psychology has evolved over several decades. There has been a 'generational' transition from unstructured clinical judgement to actuarial assessment to structured professional judgement. There is a dearth of literature regarding the history of risk assessment, particularly with sexual and violent offenders and a brief summary is provided in this introduction.

The first 'generation' of risk assessment, unstructured clinical judgement (UCJ), was favoured some decades ago. This involved practitioners collating all of the information they had about offenders and making a subjective determination (based on experience and intuition) as to whether the individual posed a risk (Quinsey, Harris, Marnie, Rice & Cormier, 2006). However, research demonstrated that UCJ was inadequate in determining risk dependent upon the decision maker, the information available and the specificity of the decision. Indeed, these kinds of approaches to risk assessment were found to perform little better than chance (Harris, Rice, Quinsey, & Cornier, 2015). False positive rates (those at low risk were deemed to be high risk) were also found to be high (Elwood, 2016) and this likely related to an overestimation of risk by clinicians who erred on the side of caution (Harris, & Tough, 2004).

As such, a second generation of risk assessments began to develop, Actuarial Risk Assessment Instruments (ARAIs). ARAIs are based on algorithms that combine and weight information to 'predict' recidivism, though these tools also have several criticisms, including a lack of information about the type and severity of risk (this is explored further in chapter four). Nonetheless, these tools have some benefits to them, including: training in their use is relatively simple; they can be used by non-clinicians; they are quick and easy to complete (dependent upon having relevant information

available); and they offer a risk category that can inform intensity of monitoring and/or treatment required, depending upon the nature of the assessment. For example, police personnel have historically employed ARAIs to inform the frequency of their scheduled visits to convicted sex offenders residing in the community, whilst probation staff, psychologists and forensic mental health workers would normally employ them to decide the duration/intensity of treatment.

The third generation of risk assessments involved structured professional judgement (SPJ), a framework for drawing together static, acute dynamic and stable dynamic risk. It is more of a collaborative approach to risk assessment that considers what *might* happen in the future rather than what *will* happen in the future (Wilkes & Barker, 2014). SPJs rely on collateral information, interviews with the offender, self-report, psychometric assessment and clinical judgement. However, a regular criticism of these third generation risk assessments is that they are often more time consuming and resource demanding than ARAIs. Nonetheless, in assessing risk, practitioners have a 'duty' to take the time and effort required to produce a good quality and meaningful assessment of risk.

The fourth generation of risk assessment (which continues to grow in the field of forensic psychology) combines SPJ and case formulation. It is based on static and dynamic risk, as well as protective factors. It appears to be a more holistic approach that takes into account what is likely to increase risk for a particular offender, together with considering what factors might decrease risk or increase desistance. A good example of this kind of approach, and one which is widely used, specifically within the field of forensic psychology, is the Risk for Sexual Violence Protocol (RSVP: Hart, Kropp, Laws, Klaver, Logan, & Watt, 2003), in combination with the Structured

Assessment of Protective Factors (SAPROF: de Vogel, de Ruiter, Bouman, & de Vries Robbé, 2012).

The RSVP allows for scenario planning such that clinicians are able to provide details about the nature, severity, imminence, frequency/duration and likelihood of risk. This information is gathered through case details, the presence of risk factors, and the relevance of risk factors. Such detailed information allows the user to formulate risk management strategies that are relevant to the individual offender. The SAPROF contains 17 protective factors, 15 of which are dynamic, thus allowing the assessor to focus upon treatment goals and evaluate treatment effects.

Notably, this approach is also time consuming and, traditionally within the purview of chartered psychologists. However, the role of police officers has changed since the introduction of the sex offender register (Sex Offenders Act, 1997). The register was introduced to help police officers verify information in relation to convicted sex offenders and quickly identify suspects. Since its introduction, in addition to investigating sexual offences and apprehending offenders, police became responsible (alongside probation staff) for assessing risk and managing sexual offenders in the community.

In 2015/16, there were 104 registered sex offenders (RSOs) per 100,000 of the population in England and Wales (MoJ, 2016). This equates to almost 60,000 offenders and represents a year on year increase (92 per 100,000 in 2013/14 and 98 per 100,000 in 2014/15). More recent figures were not readily available, though there is a vast body of research suggesting that many sexual offences go unreported and/or undetected, which

may indicate that not all sexual offenders are apprehended and registered and many RSOs have more victims than official records document.

Whilst a number of registered offenders are incarcerated or hospitalised, a significant proportion are residing in the community and subject to probation/license and monitored by Multi-Agency Public Protection Arrangements (MAPPA). MAPPA is a process via which police, probation and prison services, together with other agencies, collaborate; sharing information to protect the public and manage sexual and violent offenders in the community. This sharing of information is vital as, according to Kewley, Larkin, Harkins and Beech (2017), almost all convicted sex offenders are released back into the community *at some point*, whilst many others do not receive custodial sentences.

Recidivism rates for sexual offenders vary between 1.1% (Friendship, Mann, & Beech, 2003) and 39% (de Vogel, de Ruiter, van Beek, & Mead, 2004). These variations occur in part because sex offenders are not a homogenous group and include sub-groups of offenders (e.g. rapists, child sexual abusers, internet offenders, indecent exposers). Recidivism rates may also vary according to victim gender, socioeconomic status, intelligence, and treatment effects. Nonetheless, it would appear that recidivism rates for at least some groups of sexual offenders are lower than for other, non-sexual offenders (Brown, 2011). Hanson & Thornton (2003) found that rapists were slightly higher risk of recidivism (16%) than child molesters (13%) but that intrafamilial child molesters were less likely (5%) to recidivate than extrafamilial child molesters (18%). This discrepancy between intrafamilial and extrafamilial child sex offenders will have skewed the overall recidivism rates. Importantly, the relatively low rate of intrafamilial recidivism likely relates to professional intervention, for example, the removal of the

perpetrator from households where children are living and limiting their access to related children.

Thus, whilst a large proportion of convicted offenders will not be reconvicted of sexual offences, a minority will go on to commit further offences, causing significant harm to their victims, as well as the wider community. With this in mind, it is critical that effective assessment and treatment efforts are invested with each offender to ensure that they are being sufficiently supported in the community to maintain offence-free lifestyles. Practitioners working with sexual offenders have a responsibility, not just to the public, but to the offenders themselves to employ effective methods in assessing and treating their needs.

Risk, need, responsivity

Risk assessments and effective management/treatment of sexual offenders ought to be based on the Risk-Need-Responsivity (RNR) principles wherein the central tenet is redirecting assessments away from purely static (historical) information to 'predict' risk, to a more holistic approach that takes into account dynamic (changing) factors that impact upon criminogenic needs, as well as the individual offender's relative strengths and challenges that may enhance or reduce effective intervention efforts. The 'risk' and 'needs' principles resulted from meta analyses of offender treatment programmes (Andrews, Bonta, & Hoge, 1990; Andrews & Bonta, 2006; Bonta & Andrews, 2007) and the 'responsivity' principle was derived from treatment methods that have been linked with reductions in recidivism.

The assessment of RNR (Andrews, Zinger, Hoge, Bonta, Gendreau, & Cullen, 1990; Andrews & Bonta, 2006; Bonta & Andrews, 2007) is critical in facilitating

rehabilitation amongst convicted sex offenders, enhancing desistance/promoting offence-free lifestyles and reducing victimisation. However, this needs to be balanced with increasingly limited resources. When these kinds of approaches are taken, reductions in recidivism have been observed (Hanson, Bourgon, Helmus, & Hodgson, 2009).

Within non-clinical settings, there appears to have been an emphasis on 'risk' assessment, to the detriment of the underlying clinical needs that increase risk or enhance an offender's potential for recidivism (Lussier & Davies, 2011). A good example is perhaps the STATIC-99 (Hanson & Thornton, 2000) wherein the following variables are scored to provide a 'risk score': age at release; ever lived with someone for at least two years; index non-sexual violence; prior non-sexual violence; prior sex offences; prior sentencing; non-contact sex offences; unrelated victims; stranger victims; and male victims. Whilst these features have been associated with recidivism (Hanson & Morton-Bourgon, 2009), there is no empirical evidence to suggest that they correlate with criminogenic needs for individual offenders as they are static in nature.

Whilst it could be argued that completing 'quick and dirty' risk assessments, frees up additional resources for police officers monitoring offenders, it should be noted that such risk assessments do not take account of numerous variables that may have led to the index offence, maintained offending behaviour or explained the specific level of risk posed by the individual. Relatedly, Beech, Wakeling, Szumski and Freemantle (2016) assert that there are difficulties in measuring dynamic risk. The authors conclude that only *some* tools within *some* samples can be used to measure dynamic risk, thus highlighting the need for tailored assessments.

Desistance

Bensel and Harris (2017) suggested that the concept of desistance (cessation of offending) is difficult to conceptualise. Nonetheless, they note that desistance is a "natural human process that is observed across a wide range of offenses and offender types" (p.97). Ward and Laws (2010) suggested that desistance is viewed as a moderator of risk, as opposed to evidence that the individual is no longer engaged in crime. They describe desistance as a process not an event. It is a journey that offenders take. Relatedly, in a study by Hanson, Harris, Helmus and Thornton (2014), it was reported that the longer offenders remain sex-offence free in the community, particularly high risk offenders, the less likely they are to recidivate.

Historically, assessment and treatment of sex offenders was based on a "deficits-focused approach" (Maruna & LeBel, 2003), such that practitioners and policymakers have focused on what factors will likely increase the potential risk of recidivism for convicted sex offenders and focused their treatment efforts on these 'deficits'. For example, the areas targeted in treatment have included cognitive distortions, victim empathy, problem solving skills and social skill deficits. Because there is a focus on 'deviant' attitudes and behaviour Mingus and Burchfield (2012) argued that sex offenders are labeled and thus stigmatized by society. This can be counterproductive to community reintegration, as offenders may withdraw in an effort to psychologically distance themselves from the stigma. This may also lead to a reduced investment by offenders into prosocial activities that may facilitate desistance (Hulley, 2016). However, Farmer, McAlinden, and Maruna (2016) also describe "shame management" (an internalisation of stigma) as a protective cognition for offenders that may increase

desistance. These opinions further advocate the employment of comprehensive risk assessment and treatment planning through the use of case formulations.

Strengths-Based Approaches

Strength-based approaches began developing some years ago with the introduction of the Good Lives Model (Ward, 2002). Ward defined good lives as lives that are "beneficial and fulfilling for individuals" (p.514). Ward promoted a focus on equipping offenders to live offence-free lifestyles by achieving personal, prosocial goals. Since that time, there has been a plethora of literature endorsing the utility of strengths-based approaches in the treatment of sexual offenders. Unfortunately, there have not been any outcome studies to measure the efficacy of these kinds of approaches (Marshall, Marshall, Serran, & O'Brien, 2011; Marshall, Marshall, & Olver, 2017).

Protective Factors

According to Thornton (2013), protective factors are "social or psychological factors that make recidivism less likely" (p.64). De Vries Robbé, de Vogel, Koster and Bogaerts (2015) asserted that the incorporation of protective factors may provide a more "balanced view" with regard to assessing risk, offering positive treatment goals and high quality treatment and decision making. Furthermore, Cording and Beggs Christofferson (2017) asserted, "The measurement of protective factors is necessary to ensure that risk is not over-estimated and that strengths are instead incorporated into our current assessment of the risk and needs of individuals" (p.46), though importantly, they suggest that the inclusion of protective factors within assessments still remains in its infancy. In defining protective factors, the authors suggested that the term protective factors could be separated out into two alternative factors, promotive (e.g. religion) and

buffering factors (e.g. above average intelligence). They posited that promotive factors relate to low rates of re-offending amongst convicted offenders or low onset of offending in non-offenders, whilst suggesting that buffering factors may act to indirectly lower offence probability. Importantly, this is not true for all sexual offenders. For example, above average intelligence for some offenders might facilitate their ability to access victims through subtle grooming behaviours and the manipulation of others.

Current Research

The current research set out to explore police officers' perceptions of the utility of a relatively new SPJ tool (phase one), the Active Risk Management System (ARMS), together with their perceptions about the training they received to allow them to utilise ARMS (phase two). As noted above, since the sex offender register was introduced, it has increasingly fallen to police officers to manage sexual offenders in the community and, according to Nash (2016), this signified a divergence from more conventional policing roles and led to some labelling of offender managers as "scum cuddlers" (p. 412). This may relate to the hatred and disgust sex offenders evoke in non-offenders (Darjee & Russell, 2012). It may also relate to group mentality, insofar as Nash highlights, that officers working with sexual offenders need to have a different set of professional and personal qualities to effectively engage offenders in ongoing risk management. Their counterparts, who may not have these skill sets, may distance themselves as a means of self-protection.

Historically, police have had an investigative role in identifying crime and bringing the offenders to justice. Offender managers working with sexual offenders in the community require the skills to form and maintain consistent working relationships

with a view to continually monitoring risk and keeping the public safe. As such, they are required to undertake specific training as defined by the College of Policing.

On the basis of the empirical research supporting strengths-based approaches and the inclusion of protective factors in risk assessments, Blandford, Farmer, Mann, Scott, and Jarvis (2013) developed ARMS. ARMS provides a framework to assist the user in drawing together information from various sources including case files, colleagues, intelligence sources, offender self-disclosure and police observations. It brings together information in a format that was aimed to assist officers working in multi-agency settings. When completing ARMS, the assessor is required to apply priority ratings to relevant factors. It was developed with the aim of facilitating early identification of high priority cases so that multi-agency arrangements could be made and relevant resources allocated at an early stage in order to more effectively manage sex offenders in the community (Blandford, Farmer, Mann, Scott, & Jarvis, 2013).

ARMS incorporates the Risk Matrix 2000 (RM2000: Thornton et al., 2003), an actuarial risk instrument widely used by police, prison and probation services, to assess the level of risk of future sexual and violent convictions. The RM2000 (as well as other ARAIs) has been widely criticised (c.f. Cooke & Michie, 2014). A critique of RM2000 is provided in chapter four of this thesis.

ARMS was piloted with a small number of officers and following initial findings, training began to be rolled out nationally with ongoing evaluation. The original research did not set out to assess the tool itself but to report a consultation with those using the tool during the pilot phase that gathered their perspectives about its usefulness. There have been a number of evaluations of ARMS since its development (please refer to chapter three).

The research evolved over the course of two and a half years and a third phase involved the collection of data from completed ARMS assessments via a national police force electronic database.

Aims/hypotheses

The aim of the current thesis was to systematically evaluate the predictive validity of contemporary risk assessment tools employed with adult male sex offenders. The research element was aimed at exploring police officers' experience of using ARMS and their perceptions about the training they received. It also aimed to explore whether the incorporation of ARMS in any way altered risk levels as identified by the RM2000 when officers applied priority ratings based on structured professional judgement. This is because they would have the opportunity to take into account additional information that is known about/shared by the offender (dynamic factors) rather than simply historical information. As the RM2000 is incorporated within ARMS, a critique of this measure was also conducted.

Chapter 2

The effectiveness of risk assessment tools for predicting sexual recidivism in adult male offenders: An updated systematic literature review

Abstract

Purpose: The current study aimed to explore the effectiveness of risk assessment tools for predicting sexual recidivism in adult male offenders. The study replicates earlier work by Tully, Chou and Browne (2013).

Background: The utility of risk assessment tools employed with sex offenders has been the subject of much research and discussion over the past few decades. The current review aimed to systematically evaluate studies that have assessed the predictive validity of a number of tools that have been utilised with adult male sexual offenders.

Method: Six electronic databases were searched and two experts in the field were contacted to identify relevant studies. Inclusion criteria were applied and the included studies were quality assessed by the author, with a second assessor rating 20% of the studies, employing pre-defined criteria (replicating the Tully et al review). The remaining studies were then subjected to data extraction and synthesis.

Results: Electronic searches of six databases yielded 4991 hits. Within these, 4370 irrelevant hits and 409 duplicates were excluded. The 43 publications included in the earlier Tully et al study were also excluded, so that only new, updated evidence was incorporated. In addition, 128 studies did not meet the inclusion criteria. 41 studies were retained for quality assessment/data synthesis.

Conclusions: Most risk assessment tools within the current review demonstrated moderate predictive accuracy (18 of 25 tools). Three tools demonstrated large effect sizes (STATIC-2002R, VASOR-2 and SRA-FV). However, four of the assessment tools were low in predictive accuracy (RRASOR, STABLE-2000, SVR-20 and SARN-TMA). As such, practitioners may need to be more selective in the tools that they use. Although the current review did not incorporate the RSVP, many practitioners use this

tool in their daily work, alongside the SAPROF. Whilst these tools can be time consuming, it is imperative that sufficient time, energy and resources are allocated to the preparation of effective risk assessments that fully inform the treatment and management of sex offenders in order to avoid future sexual victimisation.

Background

The assessment and treatment of sex offenders is an area of research that has received much attention over several decades, with various theories and models being introduced, reviewed and, in some cases, disposed of or revised substantially. The history and evolution of forensic risk assessment with adult male sexual offenders is well-documented (Beech, Fisher & Thornton, 2003; Boer, Beech, Ward, Craig, & Rettenberger et al, 2017). Over time, there have been several generations of risk assessments (Wilkes, & Barker, 2014) with a move from unstructured clinical judgement (Quinsey, Harris, Rice, & Cormier, 1998) to actuarial assessment, to structured professional judgement (SPJ), and last, but by no means least, a combination of SPJ and case formulation (Andrews & Bonta, 2010; Thornton et al, 2003).

Amongst professionals working with sex offenders, there continues to be much debate about best practice in effectively assessing risk and identifying accurate treatment needs in this population (Wilcox, 2017). This may be because many risk assessments fail to take account of the underlying difficulties that lead to offending, whilst others blur the boundaries between risk and treatment needs (Vess, 2011). The introduction of actuarial risk assessment instruments (ARAIs) has led to numerous research studies and professional publications about the utility of these tools and their ability to 'predict' recidivism. Notably, however, many of these studies have been produced by the developers of such tools and may therefore be subject to 'developer

bias'. In addition, the predictive accuracy of these tools has been estimated based on large samples of sexual offenders and therefore cannot be applied to individual offenders, thus leading to a high level of error within individual assessments. When these tools are employed with individuals, often practitioners are inclined to indicate that the individual in question has a certain percentage risk of re-offending in the future and placing them in a risk category, low, medium, high, very high. Often, this will determine the level of treatment the individual receives, together with the level of monitoring.

There are difficulties with the way that recidivism is defined across studies and therefore the way risk is estimated. Whilst some define recidivism as a sexual reconviction, others use sexual re-offending (without conviction) or any reoffending (without a sexual element) and this has led to variable estimates of risk probability. In the most recent study of recidivism with prison-based treated sexual offenders (Mews, Di Bella & Purver, 2017), analysts at the Ministry of Justice concluded that the core sex offender treatment programme (SOTP) has no effect on recidivism, with treated sex offenders' sexual re-offending rates at 10%, whilst the comparison group had sexual re-offending rates of 8%.

Tully, Chou and Browne (2013) carried out the first systematic review of the predictive validity of risk assessment tools with sex offenders. They found that only two of the 15 tools evaluated within 43 studies they reviewed had large effect sizes, namely, the Violence Risk Scale: Sexual Offender version (Wong, Olver, Nicholaichuk, & Gordon, 2000) and the Structured Risk Assessment (Thornton, 2002). A large effect size is considered to have an Area Under the Curve (AUC) value of 0.714 and above (Rice & Harris, 2005).

Aims and Objectives

The aim of the current systematic review was to evaluate the predictive accuracy of risk assessment tools employed with adult male sex offenders. The assessment (and treatment) of sexual offenders appears to be an area that is ever-evolving and, in some cases, practitioners appear to be favouring previously used methods of assessment, as opposed to more recent tools, based on developing theories and models. It will be noted that Tully et al, (2013) have already produced a systematic review addressing the predictive accuracy of risk assessments tools. However, there are no clear guidelines as to when, and how, to conduct an updated systematic review (Moher et al, 2008). Furthermore, since a period of time had elapsed since Tully et al conducted their review (2011), and their subsequent publication (2013) of the results, it was considered prudent to provide an update. The current review replicates Tully's previous review.

METHOD

PRISMA guidelines (Moher, Liberati, Tetzlaff, & Altman, 2009) were followed in conducting this systematic review.

Sources of Literature

The databases that were searched replicated those selected by Tully, Chou and Browne (2013). Electronic databases were all searched on 20 December 2016 and included OVID: Psychinfo (2011-week 2 December 2016); OVID: Medline (2011-week 2 December 2016); OVID: Embase (2011-week 2 December 2016); Web of Science (Science Citation Index Expanded (SCI-EXPANDED); Social Sciences Citation Index (SSCI); Arts and Humanities Citation Index (A&HCI); Conference Proceedings Citation Index — Science (CPCI-S); Conference Proceedings Citation Index — Social Science and Humanities (CPCI-SSH); 2011–2016); Proquest: ASSIA

Search Terms

The following search terms were applied to all databases:

(sex* offen* /rape/paedophilia)

AND

(risk assessment/recidivism/names of known sex offender risk tools). The following tools were selected as they are the most widely utilised in the assessment of sexual offenders: RM2000; SVR-20; STATIC-99; RRASOR; MnSOST; SORAG; RSVP; SARN; SRA; STATIC-2002.

AND

(predict*/validity/area under curve/sensitivity/specificity/measurement/accuracy)

The full search syntax can be found at appendix A.

Study Selection

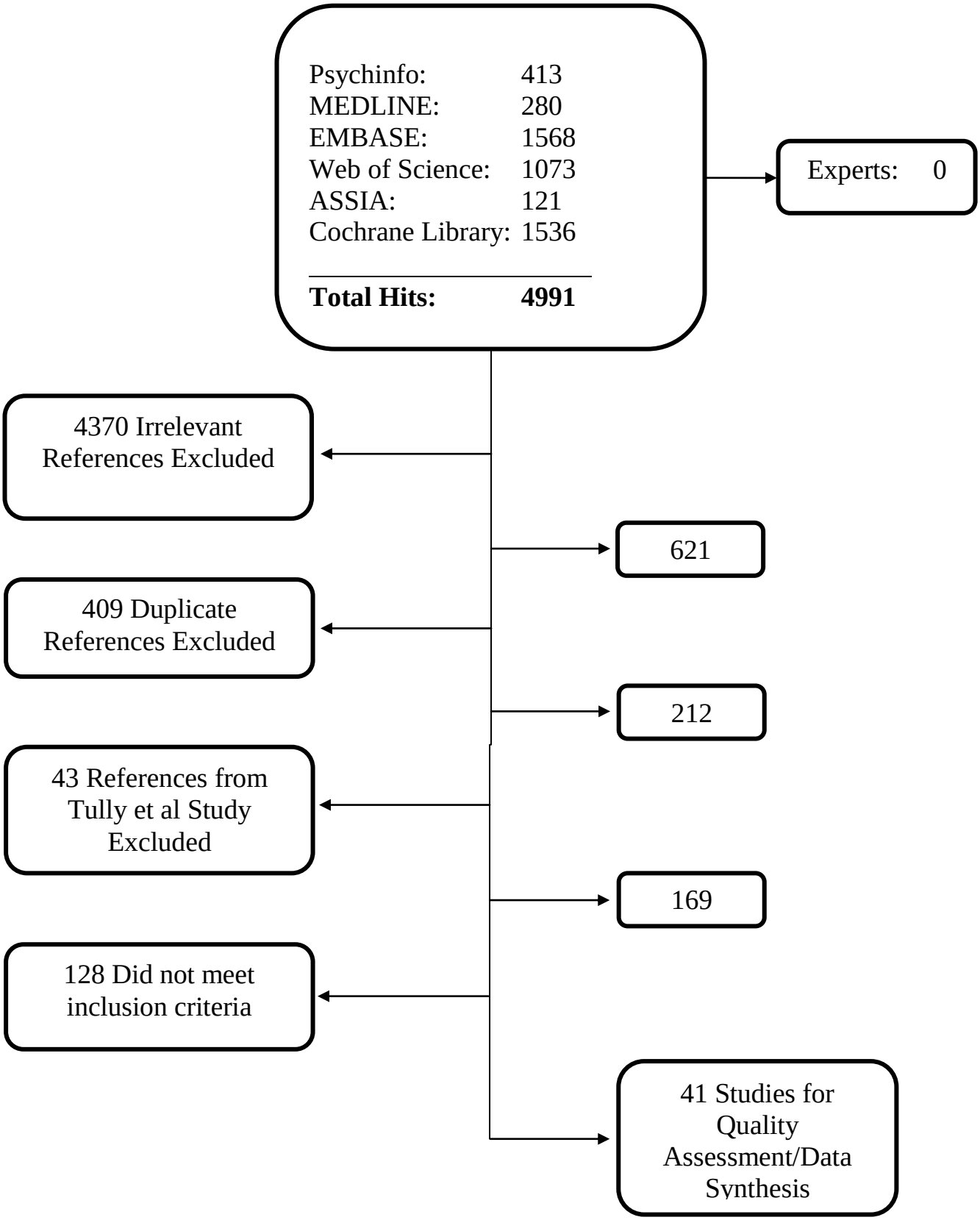
Inclusion and exclusion criteria were applied based on the PICO (Box 1), employing predefined inclusion and exclusion forms (appendix B). Studies to be reviewed had to meet all of the inclusion criteria. Excluded studies can be found at appendix C.

Population: Adult male sexual offenders.
Exposure: Sex offender risk assessment tool (designed to assess risk in sexual offenders aged 18 or over, including actuarial tools, SPJ, and a combination of both).
Outcome: Sexual reoffending, reconviction or recidivism.
Study type: Case control or cohort.
Language: English.
Date of publication: 2011 onwards.
Exclusion: Opinion papers, editorials

Box 1

The population was restricted to adult males (18 years and over) where a risk assessment tool (both static and dynamic) designed to be used with sex offenders had been evaluated. Case control or cohort studies were to be included with a publication date of 2011 onwards. Non-English studies were excluded, as were opinion papers and editorials.

Figure 1. *Systematic literature review search strategy*



Quality Assessment

Quality assessment was conducted in two phases, following Tully et al's methodology.

Threshold Criteria

The minimum threshold criteria for included studies was a clear description of both the risk assessment tool and the outcome measure. Furthermore, appropriate statistical analysis of the predictive power of the tool needed to be incorporated.

Quality Assessment Forms

Quality assessment forms were replicated from Tully et al's study, which were based on the Critical Appraisal Skills Programme (CASP 2004, 2006). See appendices D and E.

The first author assessed the methodological quality of all included studies, with a second reviewer (a forensic psychologist in training) independently assessing the quality of 20% of the studies to aid the consistency of the assessment process. 'Quality' was measured based on a number of factors, namely, objectives, selection, measurement, attrition and results. Intra-class correlation coefficient (ICC) of 0.704 (single measure) was achieved between the two assessors, demonstrating 'good' inter-rater reliability (Cicchetti, 1994). However, an ICC of >0.75 would have been desirable (Fleiss, 1986). The differences in rating were discussed and centred on the interpretation of whether the results of the studies could be generalised.

Data Extraction

The data extraction proforma utilised by Tully et al, was employed for the current review, with relevant data being extracted from the included studies prior to synthesis. In some cases, information could not be extracted due to ambiguity.

Table 1. *Results of static sex offender risk assessment tools*

Risk Tool	Reference	Sample Size	Follow-Up	Location	Recidivism Rate	Quality Score	AUC for Sexual Recidivism
STATIC-99 (Hanson & Thornton, 2000)	Eher et al (2012)	263	6.4 years	Austria	Unclear	34	0.71-0.75
	Eher et al (2015)	261	6.28 years	Austria	Unclear	36	0.67
	Helmus et al (2012a)	8390	8.2 years	Various	12.4% Overall 11.1% 5 years 16.6% 10 years	36	0.713 5 years 0.706 10 years
	Hill et al (2012)	90	10.22 years	Germany	28.9%	36	0.56
	Montana et al (2012)	337	16.05 years	UK	6.2%	26	0.672
	Olver & Wong (2011)	321	10 years	Canada	Charges 22% LRLC 24%LRHC 43% HRLC 27% HRHC Convictions 14% 16% 36% 24%	36	0.66-0.67 low risk 0.64-0.65 moderate risk 0.57-0.66 moderate high 0.55-0.56 high risk
	Olver et al (2014)	676	6.31 years	Canada	6.2%	34	0.71-0.78
	Olver et al (2014a)	267	18.2 years	Canada	27.3%	40	0.53-0.54**
	Olver et al (2014b)	539	15.5 years	Canada & New Zealand	22.4%	42	0.71
	Parent et al (2012)	414	5 years	USA	4.9%-29.1%	36	0.70
	Rettenberger et al (2013)	1077	6.35 years	Austria	6.6%	40	0.73
	Smallbone & Rallings (2013)	399	29 months	Australia	4.8%	29	0.81
	Smid et al (2014)	397	145 months	Netherlands	14.1%	42	0.72 5 year 0.73 10 year
	Turner et al (2016)	277	5.55 CSA-W 5.65 CSA-E 5.79 CSA-I	Austria	13.5% CSA-W 25.8% CSA-E 2.4% CSA-I	34	CSA 0.83 CSA-W 0.78 CSA-E 0.79 CSA-I ANY RECIDIVISM 0.65

	Varela et al (2013)	1911	4.85 years white 4.89 black 4.58 latino	USA	Unclear	34	White 0.57 Black 0.58 Latino 0.59
	Woodrow & Bright (2011)	117	45 months	Australia	8.5%	29	0.679-0.718**
STATIC-99R (Helmus, 2009)*	Brouillette-Alarie & Proulx (2013)	711	5 years	Canada	12.8% mixed 11% women 12.6% children	31	0.73 0.73 women 0.77 children
	Hanson et al (2014)	7740	8.2 years	Various 21 samples	11.9% mixed 2.9% low risk 8.5% mod risk 24.2% high risk	34	Survival analysis**
	Hanson et al (2015)	768	7.4 years	Canada	10.8%	40	0.728
	Helmus et al (2012a)	8390	8.2 years	Various	12.4% Overall 11.1% 5 years 16.6% 10 years	36	0.720 5 years 0.710 10 years
	Helmus et al (2015)	410	7.5 years	Canada	8.8%	32	0.769-0.771
	Lee & Hanson (2016)	947	7.4 years	Canada & USA	1.80%	36	0.577 [.519, .635]**
	Lehmann et al (2013)	940	9 years	Germany	7.53%	27	0..68-0.69
	Looman et al (2012)	272	6.7 years	Canada	15.4%	34	0.70
	Nunes et al (2013)	462	6909.33 days	Canada	23.,2%	30	0.519 [.446, .595]**
	Rettenberger et al (2013)	1077	6.35 years	Austria	6.6%	40	0.71
	Smid et al (2014)	397	145 months	Netherlands	14.1%	42	0.74 5 year 0.74 10 year
	Smid et al (2016)	266	148 months	Netherlands	15%	35	0.78 0.83 untreated 0.66 inpatient
	Thornton & Knight (2015)	480 @ 5years 391 @ 10years	5 & 10 years	USA	19.2% @ 5years 23.3% @ 10years	41	0.686
	Varela et al (2013)	1911	4.85 yrs white 4.89 black 4.58 latino	USA	Unclear	34	White 0.59 Black 0.65 Latino 0.57

RRASOR (Hanson, 1997)	Lehmann et al (2013)	940	9 years	Germany	7.53%	27	0.58-0.60
	Smid et al (2014)	397	145 months	Netherlands	14.1%	42	0.68 5 year 0.69 10 year
RM2000/S (Thornton et al, 2003)	Barnett et al (2012)	3402	3 years	England and Wales	4.9-11%	34	0.60/S Treated
	Helmus et al (2015a)	710	7.7 years	Canada	13.7%	32	0.685-0.695
	Parent et al (2012)	414	5 years	USA	4.9%-29.1%	36	0.65 /S
	Smid et al (2014)	397	145 months	Netherlands	14.1%	42	0.72 5 year 0.71 10 year
	Thornton & Knight (2015)	480 @ 5years 391 @ 10years	5 & 10 years	USA	19.2% @ 5years 23.3% @ 10years	41	0.665
STATIC-2002 (Hanson & Thornton, 2003)	Helmus et al (2012a)	2609	8.2 years	Various	12.4% Overall 11.1% 5 years 16.6% 10 years	36	0.709 5 years 0.700 10 years
	Parent et al (2012)	414	5 years	USA	4.9%-29.1%	36	0.68
	Smid et al (2014)	397	145 months	Netherlands	14.1%	42	0.76 5 years 0.75 10 years
STATIC-2002R (Hanson & Thornton, 2003)*	Ennis et al (2016)	345	2 years	Canada	Cluster 1 8.7% Cluster 2 10.7% Cluster 3 16%	35	0.62-0.69
	Hanson et al (2015)	768	7.4 years	Canada	10.8%	40	0.731
	Helmus et al (2012a)	2609	8.2 years	Various	12.4% Overall 11.1% 5 years 16.6% 10 years	36	0.713 5 years 0.699 10 years
	Helmus et al (2015)	410	7.5 years	Canada	8.8%	32	0.773-0.780
	Lee & Hanson (2016)	947	7.4 years	Canada & USA	1.80%	36	0.588 [.520, .656]**
	Lehmann et al (2013)	940	9 years	Germany	7.53%	27	0.67-0.69
	Smid et al (2014)	397	145 months	Netherlands	14.1%	42	0.77 5 year 0.75 10 year
MnSOST-R (Epperson et al, 1998)	Parent et al (2012)	414	5 years	USA	4.9%-29.1%	36	0.69
SACJ-Min	Smid et al (2014)	397	145 months	Netherlands	14.1%	42	0.69 5 years

(Thornton, 1997)							0.71 10 years
STABLE-2007 (Hanson et al, 2007)*	Eher et al (2012)	263	6.4 years	Austria	Unclear	34	0.67-0.71
	Hanson et al (2015)	768	7.4 years	Canada	10.8%	40	0.689
	Helmus et al (2012)	597	3.4 years	Canada	11.4% aboriginal 7.3% non-aboriginal	34	0.529 aboriginal 0.701 non-aboriginal
	Helmus et al (2015)	410	7.5 years	Canada	8.8%	32	0.684-0.709
	Helmus et al (2015a)	710	7.7 years	Canada	13.7%	32	COMBINED WITH RMS 0.694-0.709
	Lee & Hanson (2016)	947	7.4 years	Canada & USA	1.80%	36	0.649 [.585, .714]**
	Tamatea (2014	245	6.4 years	New Zealand	Unclear	33	0.66
SORAG (Quinsey et al, 2006)	Eher et al (2012)	263	6.4 years	Austria	Unclear	34	0.72-0.79
	Fedoroff et al (2016)	112	31.88 months	Canada	Unclear	33	0..70 Learning Disabled
	Parent et al (2012)	414	5 years	USA	4.9%-29.1%	36	0.67
	Smid et al (2014)	397	145 months	Netherlands	14.1%	42	0.63 5 years 0.64 10 years
	Turner et al (2016)	277	5.55 CSA-W 5.65 CSA-E 5.79 CSA-I	Austria	13.5% CSA-W 25.8% CSA-E 2.4% CSA-I	34	CSA 0.77 CSA-W 0.76 CSA-E 0.74 CSA-I ANY RECIDIVISM 0.66
STABLE-2000 (Hanson & Harris, 2004)*	Eher et al (2012)	263	6.4 years	Austria	Unclear	34	0.62
	Helmus et al (2015)	410	7.5 years	Canada	8.8%	32	0.575-0.599
	Lee & Hanson (2016)	947	7.4 years	Canada & USA	1.80%	36	0.612 [.547, .677]**
	Hanson et al (2015)	768	7.4 years	Canada	10.8%	40	0.661
BARR-2002R (Babchishin et al, 2015)*	Lee & Hanson (2016)	947	7.4 years	Canada & USA	1.80%	36	0.557 [.490, .624]**

SSPI* (Seto & Lalumier, 2001)*	Eher et al (2015)	261	6.28 years	Austria	Unclear	36	0.71
	Helmus et al (2015)	410	7.5 years	Canada	8.8%	32	0.621-0.641
	Lee & Hanson (2016)	947	7.4 years	Canada & USA	1.80%	36	0.553 [.470, .637]**
	Nunes et al (2013)	462	6909.33 days	Canada	23.2%	30	0.69 [.62, .76]**
VASOR-2* (McGrath & Hoke, 1994)	McGrath et al (2012)	759	5 years	Canada		37	0.74
	McGrath et al (2014)	1581	5 years	Canada & USA	8.6%	37	0.74
BARS* (Olver et al, 20130)	Nicholaichuk et al (2014)	2158	12 years	Canada	12.6%	30	0.67-0.73 older offenders 0.65-0.66 younger offenders
ASRS* (Skelton & Vess, 2008)	Tamatea (2014	245	6.4 years	New Zealand	Unclear	33	0.65

* Tool not included in Tully et al's study
** Original statistics converted to AUC values

Table 2. *Results of dynamic sex offender risk assessment tools*

Risk Tool	Reference	Sample Size	Follow-Up	Location	Recidivism Rate	Quality Score	AUC for Sexual Recidivism
SVR-20 (Boer et al, 1997)	de Vries Robbé et al (2015)	83	3-24	Netherlands	2% 1 year 7% 3 year 19% 15 year	40	0.63 3 years 0.58 15 years
	Hill et al (2012)	90	10.22 years	Germany	28.9%	36	0.52
	Smid et al (2014)	397	145 months	Netherlands	14.1%	42	0.53 5 years 0.48 10 years
	Turner et al (2016)	277	5.55 CSA-W 5.65 CSA-E 5.79 CSA-I	Austria	13.5% CSA-W 25.8% CSA-E 2.4% CSA-I	34	CSA 0.75 CSA-W 0.77 CSA-E 0.73 CSA-I ANY RECIDIVISM 0.73
SAPROF (de Vogel et al, 2012)*	de Vries Robbé et al (2015)	83	15.1 years	Netherlands	2% 1 year 7% 3 year 19% 15 year	40	0.76 3 years 0.71 15 years
	Turner et al (2016)	277	5.55 CSA-W 5.65 CSA-E 5.79 CSA-I	Austria	13.5% CSA-W 25.8% CSA-E 2.4% CSA-I	34	CSA 0.52 CSA-W 0.53 CSA-E 0.58 CSA-I ANY RECIDIVISM SAPROF 0.64
VRS: SO (Wong et al, 2000)	Beggs & Grace (2011)	218	12.24 years	New Zealand	13.3%	38	0.70
	Eher et al (2015)	261	6.28 years	Austria	Unclear	36	0.76
	Olver & Wong (2011)	321	10 years	Canada	Charges 22% LRLC 24%LRHC 43% hrlc 27% hrhc Convictions 14% 16% 36% 24%	36	0.69-0.73 low risk 0.66-0.70 moderate risk 0.64-0.65 moderate high 0.60-0.65 high risk
	Olver et al (2014)	676	6.31 years	Canada	6.2%	34	0.66-0.69
	Olver et al (2014a)	267	18.2 years	Canada	27.3%	40	Pre-TX 0.55 [.46, .65]** Post-TX 0.63 [.54, .71]**

	Olver et al (2014b)	539	15.5 years	Canada & New Zealand	22.4%	42	0.73
	Olver et al (2016)	668	10.2 years	Canada	10.4%	34	0.68-0.74
SRA-FV (Thornton & Knight, 2015)*	Thornton & Knight (2015)	480 @ 5years 391 @ 10years	5 & 10 years	USA	19.2% @ 5years 23.3% @ 10years	41	0.727
SGAS (Hogue, 1994)*	Beggs & Grace (2011)	218	12.24 years	New Zealand	13.3%	38	0.66
SOTIPS (McGrath et al, 2011)*	McGrath et al (2012)	759	3 years	Canada	4.6%	40	0.61-0.72
MOLEST & RAPE Scales (Bumby, 1996)*	Nunes et al (2016)	146	7.59 MOLEST 7.53 RAPE	Canada	19.7 MOLEST 18.8% RAPE	24	MOLEST Scale Pre-TX 0.53 [.41, .64] Post-TX 0.50 [.39, .62] RAPE Scale Pre-TX 0.53 [.41, .64] Post-TX 0.53 [.41, .64]
CPORT (Seto & Eke, 2015)*	Seto & Eke (2015)	266	5 years	Canada	11%	37	0.74 0.63 for pornography offenders only
SARN-TMA (Thornton, 2002)*	Tully et al (2015)	496	2 & 4 Years	UK	5.6% 2 years 16.8% 4 years	37	0.59 2 years 0.57 4 years

* Tool not included in Tully et al's study

** Original statistics converted to AUC values

Results

Description of Studies

The initial search yielded 4991 hits, 4370 of which were irrelevant. A further 409 duplicates were removed and the 43 studies included in Tully et al's original study were removed so that only new evidence was incorporated. Of the 169 studies left, 128 did not meet the inclusion criteria, leaving 41 studies for inclusion in the review.

Characteristics

On inspection, 40,544 participants were included in the 41 studies reviewed. However, many of these studies utilised the same data, resulting in an overlap of participants. For example, some of the studies by Helmus and colleagues took data from a previous study conducted by Hanson, Harris, Scott and Helmus (2007) and the studies conducted by Olver and colleagues utilised data from the Clearwater Sex Offender Program across several studies. Therefore, it is likely that some of these studies utilised the same cases and, as such, an accurate count of participants was difficult to decipher.

Most of the studies came from a Canadian sample (*N=16*). Other samples included Austria (*N=4*), Netherlands (*N=3*), USA (*N=3*), New Zealand (*N=2*), Germany (*N=2*), UK (*N=2*) and Australia (*N=2*). Some of the samples were mixed with England and Wales (*N=1*), Canada and New Zealand (*N=2*), and Canada and USA (*N=1*). Two studies had various samples incorporated (Hanson et al, 2014; Helmus et al, 2012a). This was due to the fact that the studies were reviews of previous studies.

Within the 41 studies, 25 tools were reviewed (see Tables 1 and 2), 16 of which were static tools, namely:-

- STATIC-99 (16 studies, 16,683 participants, mixed treated/untreated, inpatient/incarcerated/community samples with one study looking at sexual murderers, another at diagnosed paedophiles and another at catholic priests)

- STATIC-99R (14 studies, 24,771 participants, mixed treated/untreated inpatient/incarcerated/community samples, including sexually violent predators)
- RRASOR (2 studies, 1,337 participants, discharged inpatients/correctional facilities including those who offended against children, adults and mixed); RM2000/S (5 studies, 5,403 participants, community samples treated and untreated inpatient and incarcerated)
- STATIC-2002 (3 studies, 3,420 participants, mixed treated/untreated discharged inpatient/incarcerated including sexually dangerous offenders)
- STATIC-2002R (7 studies, 6,416 participants, treated/untreated inpatients/outpatients)
- MnSOST-R (1 study, 414, participants, sexually dangerous offenders)
- SACJ-Min (1 study, 397 participants, treated and untreated contact sexual offenders)
- STABLE-2007 (7 studies, 3,940 participants, six of the seven studies included participants who were being supervised in the community, with one study incorporating child sexual offenders only)
- SORAG (5 studies, 1,463 participants, mixed samples discharged inpatient/incarcerated, community sample, discharged child sex offenders and a sample of intellectually disabled offenders)
- STABLE-2000 (4 studies, 2,388 participants, offenders supervised in the community)
- BARR-2002/R (1 study, 947 participants, community supervised offenders)
- SSPI (4 studies, 2,080 participants, two samples community supervised offenders, one sample of diagnosed paedophiles and one sample of incarcerated offenders)
- VASOR-2 (2 studies, 2,468 participants, community treated offenders)

- BARS (1 study, 2,158 participants, sample who had reached warrant expiry date)
- ASRS (1 study, 245 participants, incarcerated and community sample).

The remaining nine studies evaluated dynamic risk tools as follows:-

- SVR-20 (4 studies, 847 participants, mixed discharged inpatients/incarcerated offenders, one study incorporated released child sex offenders and one included catholic priests)
- SAPROF (2 studies, 360 participants, discharged inpatients and released child sexual offenders)
- VRS: SO (7 studies, 2,950 participants, mixed samples of treated/untreated incarcerated/community samples including specifically one sample of treated child sex offenders and one sample of diagnosed paedophiles)
- SRA-FV (1 study, 480 participants, released sexually dangerous offenders)
- SGAS (1 study, 218 participants, treated and released child sex offenders)
- SOTIPS (1 study, 759 participants, in treatment offenders)
- MOLEST & RAPE Scales (1 study, 146 participants, incarcerated treated offenders
- CPORT (1 study, 266 participants, child pornography offenders)
- SARN-TMA (1 study, 496 participants, released offenders who had started SOTP in prison)

The included studies used various measures of outcome, with the most prevalent being a new conviction (*N=11*). Some however, measured recidivism as any new charge and/or conviction (*N=9*). Re-offense or problematic sexual behaviour was the outcome utilised in three studies and re-arrest in a further two. One study defined recidivism as where there is an identifiable victim and another where there was sexual contact and/or use of child pornography. Sexually motivated behaviour including breaches (*N=3*) and any new charge including breaches (*N=3*) were also used to define recidivism. A further study described

recidivism as any breach or re-imprisonment. Two of the studies were reviews and therefore included various definitions of recidivism across samples and four of the studies failed to specify what the outcome measure was. The studies were included in the current review and these variances taken into account during the quality assessment phase, impacting upon the overall quality score.

All of the included studies were cohort, so the quality assessment form for the case control was not required but is included in the appendices for reference.

Quality

A large proportion of the studies (*N=37*, 90.2%) had in excess of 100 participants. Sample sizes ranged from 83 (de Vries Robbé et al, 2015) to 8,390 (Helmus et al, 2012a), with a mean sample size of 989. Follow up periods varied, with 7 studies (17.1%) having a follow-up period of more than two years, 22 (53.7%) more than five years and 12 (29.3%) more than ten years. In the majority of studies, the risk assessment tools were clearly described (*N=40*, 97.6%), together with the outcome measure (*N=33*, 80.5%).

In more than half of the studies (*N=23*, 56.1%) predictive outcome was clearly stated and missing information dealt with appropriately (*N=22*, 53.7%). Inter-rater reliability (kappa >0.80) was only reported in 16 (39.0%) of the studies and concurrent validity in 18 (43.9%). A large proportion (*N=18*, 44.0%) of the studies were unclear as to whether the rater was blind to the outcome.

The quality assessment proforma produced by Tully et al and employed within the current review provides a total quality assessment score of 48. The current review produced scores ranging from 24 to 42, with a mean score of 35. The previous systematic review conducted in 2011, revealed scores ranging from 16 to 48, with a mean of 34. Higher scores equated to better quality studies and the authors considered that the general overall quality was “good”. The quality of the studies included in the current review was similar.

Descriptive data synthesis

The most widely evaluated tool within this review was the STATIC-99 (*N=16*), with the STATIC-99R, being the second most evaluated (*N=14*). Many of the tools (*N=11*, 42.3%) only had one evaluation and the remainder ranged between two and seven evaluations.

STATIC-99 was included in 16 studies with AUC values between 0.56 (Hill et al, 2012) and 0.81 (Smallbone & Rallings, 2013). Seven studies reported large effects sizes ranging between 0.71 and 0.83 (Brouillette-Alarie & Proulx, 2013; Eher, Matthes, Schilling, Haubner-MacLean, & Rettenberger, 2012; Eher, Olver, Heurix, Schilling, & Rettenberger, 2015; Olver, Nicholaichuk, & Wong, 2014; Rettenberger, Haubner-Maclean, & Eher, 2013; Smid, Kamphuis, Wever, & Van Beek, 2014; Turner, Rettenberger, Yoon, Klein, Eher, & Briken, 2016), though the variability amongst the studies, brought the mean score down to a moderate effect size with a mean of 0.65.

STATIC-99R was the next most prevalent tool to be evaluated, with 14 studies, reporting AUCs from 0.60 (Verala et al, 2013) to 0.77 (Helmus et al, 2015), with six studies reporting large effect sizes (Brouillette-Alarie & Proulx, 2013; Hanson, Helmus, & Harris, 2015; Helmus, Thornton, Hanson, & Babchishin, 2012; Helmus, Ciardha & Seto, 2015; Smid, Kamphuis, Wever, & Van Beek, 2014; Smid, Kamphuis, Wever, & Van Beek, 2016). The overall mean of the included studies was within the moderate range, slightly below what is required for a large effect size (AUC = 0.712). One study (Hanson et al, 2014) reported upon risk ratios using survival analysis and it was not possible to compute any effect sizes as only the summary data was available.

RRASOR was reviewed in two studies producing AUC values of 0.59 (Lehmann et al, 2013) and 0.685 (Smid et al, 2014) with a mean value of 0.638. The two studies that evaluated RRASOR reported effect sizes from small to medium.

RM2000/S was included within five studies, with reported AUC values of between 0.60 (Barnett et al, 2012) and 0.715 (Smid et al, 2014) with a mean AUC of 0.664. Only one of the studies reported a large effect size (Smid et al, 2014).

Three studies evaluated STATIC-2002, reporting AUC values between 0.68 (Parent et al, 2012) and 0.755 (Smid et al, 2014), with a mean of 0.713. Smid, Kamphuis, Wever, & Van Beek (2014) were the only ones to report large effect sizes.

STATIC-2002R was included in seven studies, with AUC values between 0.59 (Lee & Hanson, 2016) and 0.777 (Helmus et al, 2015), with a mean AUC of 0.699. Three of these studies reported large effect sizes (Hanson, Helmus, & Harris, 2015; Helmus, Ciardha & Seto, 2015; Smid et al, 2014).

Seven studies evaluated STABLE-2007, with AUC values ranging from 0.615 (Helmus et al, 2012) to 0.702 (Helmus et al, 2015a), with a mean AUC of 0.671. Only one of these studies reported a large effect size (Hanson, Helmus, & Harris, 2015).

SORAG was evaluated in five studies, all with AUC values. These ranged from 0.635 (Smid et al, 2014) to 0.755 (Eher et al, 2012), with a mean score of 0.699. Two of the studies reported large effect sizes (Eher et al, 2012; Turner, Rettenberger, Yoon, Klein, Eher, & Briken, 2016).

Of four studies evaluating STABLE-2000, AUC values from 0.587 (Helmus et al, 2015) to 0.661 (Hanson et al, 2015) were reported upon. The mean AUC value was 0.619 with only one study demonstrating a large effect size (Hanson, Helmus, & Harris, 2015).

SSPI was incorporated into four studies with AUC values between 0.553 (Lee & Hanson, 2016) (Eher et al, 2015). These studies demonstrated small to moderate effect sizes.

Two studies set out to evaluate VASOR-2 (McGrath et al, 2012 and 2014). Both studies described a large effect size of 0.74.

A number of tools were only included in one study, namely MnSOST-R 1 (Parent et al, 2012); SACJ-Min; BARS (Nicholaichuk et al, 2014); ASRS (Tamatea, 2014). These studies all reported moderate effect sizes, ranging from 0.65 to 0.70.

Of the 16 static risk assessment tools evaluated, only two demonstrated a large effect size (based on mean AUCs), STATIC-2002R (AUC = 0.718) and VASOR-2 (AUC = 0.74). With the exception of two tools (RRASOR and STABLE-2000), the remaining assessments achieved moderate effect sizes (ranging from 0.65 to 0.713). However, when we examine each of the studies individually for large effect sizes, the following prevalence is noted, STATIC-99, 7 of 16 (44%); STATIC-99R, 6 of 14 (43%); RM2000/S, 3 of 7 (43%); STABLE-2007, 1 of 7 (14%); and SORAG, 2 of 5 (40%).

In relation to dynamic risk tools, the most evaluated was VRS: SO, included in seven studies. AUC values ranged from 0.55 (Olver et al, 2014a) to 0.76 (Eher et al, 2015), with a mean AUC of 0.69. Two studies demonstrated large effect sizes (Eher et al, 2015; Olver et al, 2014b), and two demonstrated medium to large effect sizes (Olver & Wong, 2011; Olver et al, 2016).

Four studies evaluated SVR-20. AUC values between 0.505 (Smid et al, 2014) and 0.745 (Turner et al, 2016) were reported, with a mean value of 0.594. Only one of the studies described a large effect size (Turner et al, 2016).

SAPROF was included within two studies, with AUCs ranging from 0.52 (Turner et al, 2016) to 0.71 (de Vries Robbé et al, 2015), with a mean value of 0.596. Based on these studies, SAPROF demonstrated small to medium effect sizes.

The remaining dynamic risk assessment tools were only included in one study each. They demonstrated small (RAPE and MOLEST Scales, Nunes et al, 2016; SARN-TMA, Tully et al, 2015), moderate (SGAS, Beggs & Grace, 2011; SOTIPS, McGrath et al, 2012; CPORT, Seto & Eke, 2015) and large effect sizes (SRA-FV, Thornton & Knight, 2015).

Of the 10 dynamic risk assessment tools, only one achieved a large effect size based on the mean AUC, SRA-FV (AUC = 0.727). However, when we explore the studies individually, one study reported a large effect size for SVR-20 (25%), another two for VRS: SO (29%) and one for SRA-FV (100%). More studies for the SRA-FV may reduce the effect size. Two further studies demonstrated small effect sizes for SVR-20 and SARN-TMA, whilst the remaining tools produced moderate effect sizes (ranging from 0.65 to 0.707).

The three tools that demonstrated large effect sizes (STATIC-2002R; VASOR-2; SRA-FV) were not significantly different from the other studies in terms of their quality score or the mean sample size. However, both studies on the VASOR-2 and the one study on the SRA-FV may have been subject to developer bias and artificially elevated due to limited studies. It was not clear why the STATIC-2002R performed as well.

Table 3. *Study Methodology*

	N = 41			
	YES	PARTIAL	NO	UNCLEAR
Large N (n>100)	37 (90.2%)	0	4 (9.8%)	0
Follow up >2 years	7 (17.1%)	0	0	0
Follow up >5 years	21 (51.2%)	0	0	0
Follow up >10 years	13 (31.7%)	0	0	0
Risk tool clearly described	40 (97.6%)	1 (2.4%)	0	0
Outcome measure clearly described	33 (80.5%)	1 (2.4%)	7 (17.1%)	0
Rater blind to outcome	16 (39.0%)	2 (4.9%)	5 (12.2%)	18 (44.0%)
Inter-rater reliability kappa >0.80	16 (39.0%)	4 (9.8%)	20 (48.8%)	1 (2.4%)
Missing information dealt with appropriately	22 (53.7%)	4 (9.8%)	3 (7.3%)	12 (29.3%)
Concurrent validity discussed	18 (43.9%)	1 (2.4%)	22 (53.7%)	0
Predictive validity clearly stated	23 (56.1%)	1 (2.4%)	17 (41.5%)	0

Table 4. *Quality scores, sample sizes and AUC values*

Risk Tool	No. of Studies	Mean quality Score (range)	Mean sample Size (range)	Mean AUC (range)	No. of studies With AUC>0.714	No. of studies With AUC<0.714	No. of included studies by at least one developer of the tool
STATIC-99	16	35.3 (26-42)	1026 (90-8390)	0.70 (0.58-0.81)	7	7	2
STATIC-99R	14	35.3 (27-42)	1769 (266-8390)	0.71 (0.52-0.76)	6	5	4
RRASOR	2	34.5 (27-42)	669 (397-940)	0.64 (0.59-0.69)	0	2	0
RM2000/S	5	37.0 (32-42)	1081 (397-3402)	0.67 (0.60-0.72)	1	4	2
STATIC-2002	3	38.0 (36-42)	1140 (397-2609)	0.71 (0.68-0.76)	1	2	1
STATIC-2002R	7	35.4 (27-42)	917 (345-2609)	0.71 (0.59-0.78)	3	4	4
MnSOST-R	1	36.0 (n/a)	414 (n/a)	0.69 (n/a)	0	1	0
SACJ-Min	1	42.0 (n/a)	397 (n/a)	0.70 (0.69-0.71)	0	1	0
STABLE-2007	7	34.4 (32-40)	563 (245-947)	0.67 (0.62-0.70)	0	5	5
SORAG	5	35.8 (33-42)	293 (112-414)	0.70 (0.64-0.76)	2	3	0
STABLE 2000	4	35.5 (32-40)	597 (263-947)	0.62 (0.59-0.66)	0	3	3
BARR-2002R	1	36.0 (n/a)	947 (n/a)	0.55 (n/a)	0	1	1
SSPI	4	33.5 (30-36)	520 (261-947)	0.65 (0.55-0.71)	0	4	0
VASOR-2	2	37.0 (n/a)	1234 (887-1581)	0.76 (0.74-0.77)	2	0	2
BARS	1	30.0 (n/a)	2158 (n/a)	0.68 (n/a)	0	1	1
ASRS	1	33.0 (n/a)	245 (n/a)	0.65 (n/a)	0	1	0
SVR-20	4	38.0 (34-42)	212 (83-397)	0.60 (0.52-0.75)	1	3	0
SAPROF	2	37.0 (34-40)	180 (83-277)	0.66 (0.57-0.74)	1	1	1
VRS: SO	7	37.1 (34-42)	421 (218-676)	0.69 (0.55-0.76)	2	5	6
SRA-FV	1	41.0 (n/a)	480 (n/a)	0.73 (n/a)	1	0	1
SGAS	1	38.0 (n/a)	218 (n/a)	0.66 (n/a)	0	1	0
SOTIPS	1	40.0 (n/a)	759 (n/a)	0.67 (0.61-0.72)	0	1	1
MOLEST & RAPE	1	24.0 (n/a)	146 (n/a)	0.52 (n/a)	0	1	0
CPORT	1	37.0 (n/a)	266 (n/a)	0.69 (0.63-0.74)	0	1	1
SARN-TMA	1	37.0 (n/a)	496 (n/a)	0.58 (0.57-0.59)	0	1	0

Table 5: *Studies with Large Effect Sizes*

Tool	AUC	Study
STATIC-99	0.73-0.77	Brouillette-Alarie & Proulx (2013)
	0.75	Eher, Matthes, Schilling, Haubner-MacLean, & Rettenberger (2012)
	0.75	Eher, Olver, Heurix, Schilling, & Rettenberger (2015)
	0.71-0.78	Olver, Nicholaichuk, & Wong (2014)
	0.73	Rettenberger, Haubner-Maclean, & Eher (2013)
	0.72-0.73	Smid, Kamphuis, Wever, & Van Beek (2014)
	0.78-0.83	Turner, Rettenberger, Yoon, Klein, Eher, & Briken (2016)
STATIC-99R	0.73-0.77	Brouillette-Alarie & Proulx (2013)
	0.788	Hanson, Helmus, & Harris (2015)
	0.720	Helmus, Thornton, Hanson, & Babchishin (2012)
	0.771	Helmus, Ciardha & Seto (2015)
	0.74	Smid, Kamphuis, Wever, & Van Beek (2014)
	0.78	Smid, Kamphuis, Wever, & Van Beek l (2016)
RM2000S	0.71-0.72	Smid, Kamphuis, Wever, & Van Beek (2014)
STATIC-2002	0.75-0.76	Smid, Kamphuis, Wever, & Van Beek (2014)
STATIC-2002R	0.747	Hanson, Helmus, & Harris, (2015)
	0.780	Helmus, Ciardha & Seto (2015)
	0.75-0.77	Smid, Kamphuis, Wever, & Van Beek (2014)
STABLE-2007	0.784	Hanson, Helmus, & Harris, (2015)
SORAG	0.72	Eher, Matthes, Schilling, Haubner-MacLean, & Rettenberger (2012)
	0.74-0.77	Turner, Rettenberger, Yoon, Klein, Eher, & Briken (2016)
STABLE-2000	0.784	Hanson, Helmus, & Harris (2015)
VASOR-2	0.74-0.77	McGrath, Lasher & Cumming (2012), McGrath, Lasher, Cumming, Langton & Hoke (2014)
SVR-20	0.73-0.77	Turner, Rettenberger, Yoon, Klein, Eher, & Briken (2016)
VRS: SO	0.76	Eher, Olver, Heurix, Schilling, & Rettenberger (2015)
	0.69-0.73*	Olver & Wong (2011)
	0.73	Olver, Christofferson, Grace, & Wong (2014)
	0.68-0.74*	Olver, Klepfisz, Stockdale, Kingston, Nicholaichuk, & Wong (2016)
SRA-FV	0.727	Thornton & Knight (2015)

*Effect size: medium to large

Discussion/Conclusions

The current review examined 41 studies that evaluated 25 risk assessment tools designed to be employed with adult male sex offenders. Of the 25 tools, only three demonstrated consistent predictive validity. The STATIC-2002R, the VASOR-2 and the SRA-FV were all found to have large effect sizes. Of the remaining 22 tools, 18 had a moderate effect size and four had a low effect size. The selected studies varied in their overall methodological quality with none of them reaching the upper limit on the quality assessment form. Indeed, the highest score was 42 across two studies (Smid et al, 2014; Olver et al, 2014b) with a low score of 24 for one study (Nunes et al, 2016). There was also a significant amount of overlap in relation to participants. The definitions of outcome measures within the current review were variable. Ideally, recidivism should be measured based on reconviction, though (as noted elsewhere) this underestimates reoffending. Nonetheless, this could account for some of the variance across these studies, together with different populations (age, culture, child sexual abusers, adult abusers, non-contact offenders), treatment effects and follow-up periods.

Relation to other reviews and research

The current review aimed to replicate that of Tully et al's previous systematic review. Whilst the methodology was very similar, the current review produced some differences. Specifically, the number of tools that were incorporated within the studies increased from 15 to 25, with a number of tools included in this study that were not within the original review, namely, STATIC-99R, STATIC-2002R, STABLE-2000, BARR-2002R, SSPI, VASOR-2, BARS, ASRS, SAPROF, SRA-FV, SGAS, SOTIPS, MOLEST & RAPE Scales, CPORT, SARN-TMA. All of the studies in the current review were cohort studies, whereas the original review contained two case control studies. None of the studies in the current review were in the original review and this was intentional. The search criteria was set to 2011

onwards, and all of the studies reviewed by Tully et al excluded. Similarities were noted in terms of the quality of the studies within the two reviews, though it should be noted that the ICC in the original study (0.911) was stronger than for the current study (0.704). This may relate to differences in the stage of career development between the primary researcher and second rater.

In a large meta-analysis of 118 predictive studies (*N=45*,398), Hanson and Morton-Bourgon (2009) identified that actuarial risk measures were more accurate at predicting future reconviction than unstructured professional judgement. However, the effect sizes across these tools varied from 0.67 to 0.97 demonstrating effect sizes that ranged from moderate to large. Tully and colleagues expressed the view that a moderate effect size was justified in aiding defensible decision making. Nonetheless, there is increasing concern that whilst actuarial risk measures have a place in assessing sexual offenders, they are only able to identify a percentage of offenders who will re-offend but not which offenders these will be.

Ward and Beech (2009) assert that whilst actuarial risk assessments offer descriptors associated with risk factors, they do not provide underlying theoretical explanations. Therefore, their utility in case formulation is limited and does not translate sufficiently enough from the research field to the practice of psychology. It is therefore important that these tools are employed alongside multiple other tools and sources of information.

Strengths and limitations of the current review

The focus of the review was on adult male sexual offenders only and did not include female or juvenile offenders. Furthermore, it did not take account of type of offence (contact v non-contact), victim type (child, adult, both) or ethnicity. Due to the nature of systematic literature reviews, this chapter provides a snapshot of the studies conducted between 2011 and December 2016. The earlier work by Tully et al (2013) was intentionally excluded from this review so that comparisons could be made between the results. However, the results may

have been different had Tully's data been included. There is therefore potential for a further, more longitudinal review to be undertaken.

The majority of studies included within this review reported AUC values and where these were not reported, the statistics were converted to reflect an AUC value. Whilst this was done due to time constraints and to improve consistency within the review, it should be noted that the external validity of AUC in predictive studies is considered by some (Hand, 2009) to be 'fundamentally flawed'. Importantly, the AUC derives from regression estimates and thus assumes a simple linear relationship between risk factors and outcomes. At best, the external validity of an AUC value in predicting an outcome measure is questionable and the confidence intervals tend to be low and dependent upon the error rate within each of the samples. Despite this, AUC values are regularly employed and they are appropriate for reporting relative predictive accuracy, though not absolute predictive accuracy (Helmus & Babchshin, 2017).

Conclusions and recommendations

Based on the current review, the predictive validity of numerous measures currently being employed with adult male sex offenders is variable. Only three of the 25 tools evaluated (12%) consistently demonstrated a large effect size in terms of predicting recidivism. The remaining 88.5% have variable effect sizes according to the evaluations conducted to date. Whilst these tools continue to have some utility in the assessment of sex offenders, questions will be raised as to what it is practitioners can take from such tools. When we look at 'predicting' recidivism, we firstly need to consider what each of these terms mean. Elwood (2016) argues that we have problems with the term prediction because we apply frequencies or frequentist probabilities to individuals when they can only be applied at the group level. Furthermore, Helmus and Babchishin (2017) argue that predictions are used to diagnose, whereas the risk assessment tools currently employed with sex offenders are

developed to be prognostic tools, that is a diagnosis is pre-existing, whereas a prognosis is the "likelihood of a future event that has not yet happened" (p.9).

Secondly, the published research in relation to recidivism rates varies considerably with regard to how recidivism is defined, with some studies using official reconviction data and others referring to breaches and self-report. In any event, reconviction rates seriously underestimate re-offending (Falshaw, Friendship, & Bates, 2003).

Interestingly, risk assessment tools appear to be progressing with many developers taking into account protective factors as suggested within the Risk-Need-Responsivity (RNR) model (Andrews, Bonta, & Hoge, 1990). It has been argued that the combination of risk and protective factors is one of the most major advances in recent years in the field of violence risk assessment (de Vries Robbé, de Vogel, & Douglas, 2013). In the current climate, many practitioners prefer to employ the Risk for Sexual Violence Protocol (RSVP: Hart et al, 2003) alongside the Structured Assessment of Protective Factors (SAPROF: de Vogel et al, 2012). The RSVP is a comprehensive management plan that the user can prepare and implement, as well as evaluate at regular intervals and repeatedly update (Logan, 2016). The tool is employed internationally and places much emphasis on psychological risk factors to aid case management plans and inform treatment targets (Hart & Boer, 2010). The SAPROF focuses on protective factors and has been found to be a statistically significant predictor of future violence and sexual violence (de Vries Robbé, de Vogel, Koster, & Bogaerts, 2015).

It is imperative that risk assessments take account of individual differences, with a focus on protective/desistance factors and case formulation and that sufficient energy and resources are allocated accordingly if these assessments are to inform, as fully as possible, treatment planning and thereby aim to reduce further victimisation. As such, robust methods of evaluating risk assessment tools will be required so that practitioners can have confidence in their utility as part of the broader aim of formulating risk and identifying

appropriate treatment and protective factors. When treatment targets are well formulated and outcomes are positive, e.g. desistance, this will inevitably suggest (at least for those considered medium, high and very high risk), that the predictive accuracy of the risk assessment tool was limited. Nonetheless, it should also be noted that actuarial risk assessment tools are seen as particularly helpful when it comes to sentencing, offering practitioners the opportunity to weigh-up the intensity of treatment that an offender should engage with. This is particularly true for low risk offenders, where recidivism rates are lower. Therefore, treatment resources can be allocated to higher risk offenders (Wakeling, Mann & Carter, 2012). Post-sentencing, treatment needs for low risk offenders may increase based on assessment of dynamic factors, though undoubtedly, an assessment of static risk offers a helpful way forward at the sentencing stage.

Chapter 3

Active Risk Management System (ARMS): An Exploration of its Usefulness

Abstract

This research set out to explore the Active Risk Management System (ARMS) recently developed to provide a national standard for the dynamic assessment of risk associated with sexual offenders being managed in the community. The ARMS framework incorporates both risk and protective factors into the assessment, combining the outcome with actuarial risk assessment (RM2000) to arrive at a General Level of Risk Management and Risk Management Plan. The research took place over a period of almost two and a half years with data being analysed both qualitatively and quantitatively. There were three phases to the evaluation. Phase one was aimed at evaluating the experience of police officers who had piloted the tool employing Likert-scale questionnaires and conducting a focus group. Phase two was aimed at evaluating the training being provided to offender managers and consisted of Likert-scale questionnaires and qualitative interviews. Phase three involved the quantitative analysis of a set of ARMS assessments completed by police officers managing sexual offenders in the community.

Whilst some drawbacks were identified during the initial findings of phase one, the overall results were largely positive. Similarly, phase two identified areas for development, though the general consensus appeared to be that participants would like to see ARMS rolled out nationally and would recommend the tool to their colleagues. Whilst this research did not set out to evaluate ARMS itself, it was aimed at exploring what offender managers' experiences were in terms of the training they received to employ the tool in their daily duties and their views with regard to national roll-out. The participants expressed support for national roll-out, though at phase two of the evaluation, caution was advised regarding the critical role of ongoing evaluation of ARMS.

Phase three was aimed at exploring whether or not the inclusion of dynamic risk and protective factors alongside a measure of actuarial risk altered the overall perceived risk

posed by sexual offenders being managed in the community, through the employment of ARMS. A sample of 434 adult male offenders' assessments were accessed via an identified police force's electronic database, ViSOR, and analysed using SPSS Version 22 (IBM, 2013). A number of analyses were undertaken including factor analysis and multidimensional scaling to explore the underlying structure of the tool and identify potential subscales. It was hypothesised that the inclusion of dynamic risk and protective factors would alter the perceived risk levels from actuarial assessment. The results revealed that of 11 dynamic risk factors, eight could be loaded onto two facets or components, namely sexual risk and protective factors. Two variables (hostility and relationship status) did not load onto these components and there was some ambiguity with regard to a further variable (opportunity to offend) as this item appeared to load onto both components. However, it may be helpful to retain these items as providing qualitative information to offender managers. Employing structured professional judgement to adjust actuarial risk scores revealed some increases in perceived risk, though notably, for very high risk offenders, decreases in perceived risk were noted. The current study may have practical implications for police officers managing sexual offenders in the community and how they direct resources. However, it will be important for officers managing sexual offenders in the community to incorporate a thorough case formulation for each of the offenders they manage to ensure that specific information regarding type of offence, imminence, frequency and victim ~~type~~ to inform management and treatment plans as fully as possible.

Introduction

Historically, individuals working with sexual offenders, in particular from an assessment point of view, have focused much attention on risk factors, including age at commencement of risk, previous sexual convictions, emotional congruence with children, sexually deviant interests/fantasies, pro-offending thinking and distorted attitudes. There

have been numerous journal articles and book chapters on the evolution of risk assessment with sexual offenders and, yet, there continues to be much debate about best practice. The various changes over time have been described in earlier chapters of this thesis. A good overview can be found in Laws and O'Donohue (2016).

Increasingly, professionals involved in managing risk of sexual (and/or violent) offending are looking to incorporate protective factors within their risk management plans. This can be seen within the National Health Service, where the Structured Assessment of Protective Factors (SAPROF: de Vogel, de Ruiter, Bouman, & de Vries Robbé, 2012) is increasingly employed within risk assessments relating to violence. Additionally, the Short-Term Assessment of Risk and Treatability (START: Webster, Martin, Brink, Nicholls, & Middleton, 2004) is utilised with violent offenders and the Assessment of Risk and Manageability of Individuals with Developmental and Intellectual Limitations who Offend-Sexually (ARMIDILO-S: Boer et al, 2012) is employed with offenders who have learning disabilities. Boer (2013) opined that the incorporation of protective factors into risk assessment is "essential to effective community integration" (p.8). In addition, Akerman, Craig, and Beech (2015) argue that combining protective factors with actuarial and dynamic factors, together with a detailed case formulation, affords the offender the opportunity to 'tell their story' in a way that takes account of historical factors and current clinical influences, thereby tailoring risk assessment and treatment planning to the specific individual.

In an effort to address both risk and protective factors for police officers managing sex offenders in the community, Blandford, Farmer, Mann, Scott and Jarvis (2013) developed the Active Risk Management System (ARMS). ARMS incorporates the Risk Matrix 2000 (RM2000: Thornton et al, 2003), an actuarial risk assessment tool widely used by police, prison and probation services, to assess the level of risk of future sexual and violent recidivism (please refer to chapter four for a critique of RM2000). An overview of ARMS

was presented in chapter one. The relevant factors to which police officers apply priority ratings are listed in table six.

Table 6: *ARMS factors*

Risk Items	**Protective Factors**
Opportunity to offend	Commitment to desist
Sexual preoccupation	Intimate relationship
Offence related sexual interests	Employment or positive routine
Emotional congruence with children	Social investment – 'giving something back'
Hostile orientation	Pro-social network
Poor self-management	
Anti-social influences	

The risk factors were derived from empirically supported factors (Mann, Hanson, & Thornton, 2010), whilst the protective factors were developed with a focus on the desistance literature (Kewley & Blandford, 2017) with the authors highlighting that the empirical literature in relation to protective factors was limited at the time the tool was developed. The tool was developed with the aim of facilitating early identification of high priority cases so that multi-agency arrangements and relevant resources could be allocated at an early stage in order to more effectively manage sex offenders in the community (Blandford et al, 2013). The developers devised three case studies for training purposes. Each of the case studies were provided both in written format (a case summary) and visual/audio format (DVDs depicting offender managers engaging with and assessing offenders in their home environments – the offenders were actors). The developers then scored all three case studies employing the ARMS framework and reached agreement about risk. The case studies included a low risk, medium risk and high risk offender and the authors described them as the 'Gold Standard' for measuring inter-rater reliability for future training in ARMS. However, there were insufficient assessments completed during the training phase to allow for statistical analysis using intra-class correlation coefficients (Nicholls & Webster, 2014). It

should also be noted that the cases were artificial and the ratings were produced by the developers of the tool, thereby potentially introducing developer bias wherein the 'gold standard' was based on their subjective views of a tool they developed, without the support of scientific evidence or validation data. Furthermore, whilst it has been reported that the team consisted of both academics and practitioners with expertise in assessing, treating and managing sexual offenders (Kewley & Blandford, 2017), it was unclear as to the level of expertise within the cohort regarding the development of risk assessment tools.

Barnett (2013) conducted a study with probation (N=19) and police (N=22) officers who completed training in ARMS to evaluate inter-rater reliability. She found that whilst there was some variability between various factors embedded within ARMS, more than two thirds of the participants achieved consistency on the low risk case study and 83% correctly classified the high risk case study set against the 'Gold Standard'. Whilst at face value these figures appear rather high, taking account of the small sample sizes, there was, in fact, quite a high level of disagreement amongst participants.

Furthermore, participants found the medium risk case study more difficult to rate and tended to overestimate risk in this scenario. Fifty percent (50%) of police officers and 80% of probation officers 'misclassified' this case study as high risk. The developers of ARMS defined medium risk as "a stable balance between risk and protective factors. There are no grounds for an active investigation or immediate action but periodic engagement with the offender will be necessary to reduce the risk further and to enable ongoing review of the assessment" (p.2). Whereas, high risk was defined as, "no evidence of an imminent likelihood of offending but continued exposure to current risk factors could lead to the commission of a further offence. The case will require further investigation and action aimed at reducing the risk. This will involve regular engagement with the offender until further assessment indicates the level of risk has reduced" (p.2). Therefore, the discrepancy

highlighted by Barnett (2013) suggests that the assessment of dynamic risk and protective factors, at least amongst police and probation officers, remains extremely subjective. Barnett concluded that further evaluation was needed in a number of different areas, including the training of personnel, the quality of the case studies and the varying perspectives of personnel from different backgrounds. The training may also have benefited from the employment of 'real life' cases so that comparisons could be made between the individual's recorded risk levels (assessed by professionals involved in their case management) and the perception of those doing the training.

Subsequently, Blandford (2014), conducted a mid-pilot evaluation study exploring the successes and challenges of ARMS in the six different police force areas where it was being piloted. Blandford found that whilst 51% of ARMS evaluations corresponded with RM2000 risk estimates, 42% of ARMS assessments identified offenders as lower risk than the RM2000. Seven percent (7%) demonstrated increased risk. However, the sample size was small (*N=66*) and likely skewed by two pilot areas; one where all offenders (*N=3*) were reduced from very high to medium; and one where all offenders (*N=6*) were reduced from high to medium. These results may be indicative of procedural changes rather than an actual decrease in risk. It would appear that the incorporation of dynamic risk factors in these two areas was a new phenomenon. However, it may also reflect the subjectivity of the assessments, as well as highlighting how much officers had invested in the success of ARMS. Nonetheless, on the basis of Blandford's findings, it was noted that scheduled visits by offender managers to sexual offenders were revised. Importantly, the one area that had previously incorporated dynamic risk into their assessments for some time prior to the development of ARMS, demonstrated very little change. Indeed, there was only one change (*N=9*) and this related to an increase in risk. Again, it should be noted that this was a small scale study.

From this study came further recommendations for the developers. In particular, participants suggested further training, streamlining of the tool, the importance of time management and quality assurance. Despite these recommendations, there appeared to be a focus upon how best the tool could be 'marketed across the Responsible Authority'. It should be noted that at this stage of the pilot, it would perhaps have been important to focus on developing the tool further, rather than efforts to promote it more widely, although it is recognised that widespread dissemination of the tool would allow for larger scale studies to be undertaken.

Towards the end of the initial pilot, Nicholls and Webster (2014) carried out a small scale qualitative evaluation. They observed clear differences between how police officers and probation officers rated each of the case studies. They also observed disagreement between the developers of the tool in relation to the case studies. The authors concluded that a further, larger study was required.

Research Questions

ARMS was developed with a view to eventually employing this measure with every registered sex offender (RSO) residing in the community and being managed by their local police force. It was deemed important to ascertain how helpful frontline staff found the tool, as well as the training they received in order to employ it in their daily activities. These research questions were time-limited as the agenda was to roll-out ARMS in the short-term. However, the intention was always to expand upon this smaller scale research and explore the utility of ARMS. Specifically, ascertaining whether the application of structured professional judgement to actuarial assessment, employing ARMS, influence the perceived, subjective level of risk (based on priority ratings) amongst police officers managing sexual offenders in the community.

Qualitative Method

Participants

Pre-selected participants (*N*=22) who had been piloting ARMS were invited to complete Likert scale questionnaires. Participants were police officers of both genders and a range of ages who were responsible for the community management of sexual offenders. Fourteen participants completed questionnaires, giving a 36% attrition rate.

For the focus group, six volunteers from the above pool of participants put themselves forward. All six individuals represented a different pilot area.

During phase two, 31 offender managers attending training consented to complete Likert scale questionnaires. For this phase, there was no attrition. At the conclusion of the training, participants were interviewed in their allocated groups regarding their experiences of training.

Measure

The questionnaires were produced by the researcher and were based on a Likert scale. They can be found at appendices H and I. For the focus group, an interview schedule was developed (appendix J).

Procedure

During phase one, participants were invited to complete questionnaires about their experiences of using ARMS within the course of their duties. The second element of phase one involved a focus group held during March 2014. The first author facilitated this focus group, providing guidance as to the kinds of issues that required consideration. As per best practice, the researcher aimed to achieve flexibility through open ended enquiry. Focus group discussions were audio recorded.

Phase two was conducted during June 2014, at which time training was being delivered to a number of sex offender managers.

Ethics

Ethical approval for phases one and two was sought and approved (09/01/14) through the Ethical Committee for the University of Birmingham, prior to any data collection. Informed consent was provided by each of the participants prior to the research being conducted. Participants were provided with an information leaflet (appendix K) explaining the nature of their participation and all provided informed consent (appendix L).

Analysis

Descriptive statistics from the questionnaires were produced through SPSS Version 22 (IBM, 2013) and the focus group and group interviews were transcribed and analysed employing *Thematic Analysis*. Braun and Clarke's (2006) checklist was employed to ensure a thorough analysis was adhered to. The researcher immersed herself into the data, beginning with transcribing focus group interviews from phase one and individual responses from phase two before re-reading all material. Initial codes were then generated by annotating the transcripts with handwritten notes, following which codes were collated into potential themes. Themes were reviewed using mind mapping software (XMind: V3.4.1.201401221918, 2013) to generate a thematic map and refined to produce definitive themes and a final thematic map (Appendix N).

Quantitative Method

Participants

There were no active participants involved in this phase of the research. A sample of 434 ARMS assessments for male adults convicted of sexual offences was gathered through the UK Police Force's electronic database, ViSOR. The offenders' ages ranged between 18 and 84, with a mean age of 39.58 (SD = 14.67). The average number of non-sexual convictions amongst offenders was 1.70 (Range = 0-54; SD = 4.89). The range for sexual convictions was slightly smaller (1-49), though due to the variability amongst offenders, the

mean was higher at 4.30, with a standard deviation of 5.68 (SPSS output can be found in Appendix M). Data were collected on 435 ARMS assessments. However, it transpired that one of the assessments related to an individual who had not been convicted of a sexual offence but for whom intelligence suggested he was a Person Posing a Risk to Children (PPRC). The RM2000 was not designed for use with non-convicted individuals and, as such, the data from the assessment for this individual was excluded.

Procedure

Following strict vetting procedures by the relevant police authority, the researcher was permitted access to the electronic system ViSOR, a secure and confidential database where police, probation and prison staff share information regarding risk assessment and risk management about individual offenders. The data was coded on-site at an identified police station and no identifying information was recorded. The data collected included age at time of assessment, sexual convictions, non-sexual convictions, RM2000 risk categories (sexual, violent and combined scales), individual ratings on ARMS factors and overall ARMS priority ratings.

Ethics

Ethical approval for phase three was sought and approved (27/05/16) through the Ethical Committee for the University of Birmingham, prior to any data collection. Furthermore, the research was supported by the National Lead for the management of sexual and violent offenders for the Association of Chief Police Officers and her team (25/01/16).

Analysis

The third phase involved statistical analysis of the outcome of ARMS assessments with community managers of sexual offenders. Secondary analyses were conducted using SPSS incorporating factor analysis and multidimensional scaling.

Qualitative Results

During *Thematic Analysis,* nine (9) basic themes (lower order premises) were identified as follows: Time; Resources; Training; Structure; Previous Experience; Current Practice; Usefulness; Defensibility; Psychosexual Development. Within several of these basic themes, a number of organising themes (abstract principles) were identified. For ease of reference, the results have been displayed in terms of each of the basic themes with organising themes and descriptive statistics presented where appropriate. The global theme (superordinate theme) was identified as Risk Reduction. These themes are represented in a final thematic map at Appendix N.

Theme 1: Time

The first theme identified related to the amount of time required to complete ARMS. As will be noted throughout, this resonates with other SPJs, such that practitioners employing SPJ tools need to allocate sufficient time and resources to facilitate an accurate, individually tailored assessment of risk. A number of factors were relevant to this theme including the practicalities associated with completing the tool, novelty, other work commitments and longer-term gains. Each of these will be addressed in order.

Duration

There appeared to be some disagreement amongst respondents with regard to the length of time it takes to complete ARMS, such that 34% of respondents (*N=18*) neither agreed nor disagreed to the statement, "ARMS can be completed in a reasonable amount of time". Of the remaining respondents, there was a fairly even split, with 32% of participants responding affirmatively and 34% disagreeing (see Figure 2).

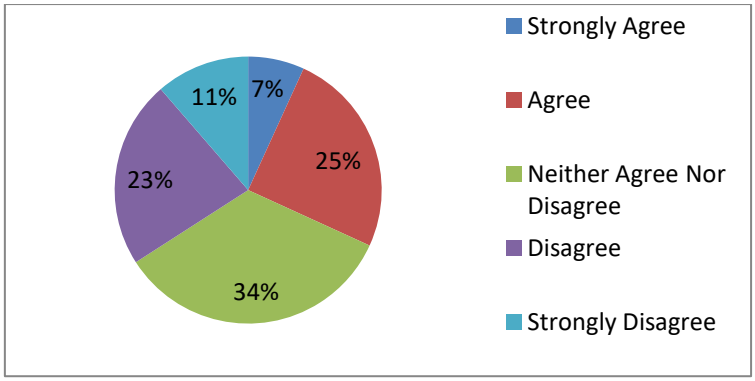


Figure 2. Percentage of Responses to the statement:

"ARMS can be completed in a reasonable amount of time".

Relatedly, the following comments were identified: "It's . . . very time consuming"; "If you're talking high and very high risk and I'd go along with that and say . . . three hours writing that up". There was a suggestion that this was inevitable: "The time consuming thing err . . . it is time consuming, there's no way of getting around that". For those involved in phase two of the evaluation, time restrictions were also a concern: "I'd like to start using it but it's going to add lots of time"; "I'm unsure how it's going to fit into my daily work".

Novelty

However, novelty appeared to impact upon the amount of time it takes to complete ARMS and one participant suggested:-

I think the first one that I did probably took the best part of a day . . . I've now got it down to, like you say, maybe three hours and, and I'm quite happy doing that you know it's not an entire morning

There was some agreement between participants that the time taken to complete ARMS reduces over time, as the user becomes more familiar with the tool:"I think you've gotta accept when you first do it, particularly if you're a force that've come straight from the Risk Matrix to that . . . you're gonna have to take a hit and you're gonna get behind to start with" and:-

> The next plan is much quicker because you're only really discussing that or new things you found out on the visit it, it, you don't have to go through the whole lot again so your first initial visit certainly three hours err but the, the more you, certainly high risk as well because you're seeing them more often erm it gets quicker and quicker and quicker

There is an underlying assumption that in order to complete SPJs, the practitioner needs to be an expert in the field (Neal & Grisso, 2014). Experts differ from novices, insofar as they have extensive knowledge and training that is far broader than the problem with which they are being presented (Eysenck & Keane, 2010). There continues to be much debate about whether experts have innate abilities e.g. a special talent, or whether expertise can be achieved through 'deliberate practice' (Plomin, Shakeshaft, McMillan, & Trzaskowski, 2014): "The more you do it the better you get at it and we're down … I mean I can err we can cer, certainly the low risk we can probably knock out in about half an hour now".

Other Commitments

Participants also identified that their capacity for completing ARMS for offenders they manage is impacted upon by their other work commitments:"I've got to look at a thousand of these a year"; "It's so time consuming with the rest of our job to do".

It is not uncommon for people to resist organisational change, specifically when it relates to a shift in strategic direction, as this inevitably involves more work for the individuals involved. Psychologists, forensic mental health workers and therapists regularly refer to 'readiness to change' with clients and patients. Indeed, they often subscribe to theoretical models and psychometric testing in gathering information regarding the participant's motivation to change. At times, this may be overlooked with

organisational/procedural changes and have negative consequences for those involved. Research has found that involving staff in the decision making processes, allows them to feel more in control of the situation and less resistant to change (McKay, Kuntz, & Näswall, 2013).

> . . . with the sheer volume of work and you know, you know, obligations, I've got and commitments. I don't get out [in the community] that often but what we've done is that we've did it with about 15 to 20 offenders . . . at a time when our staffing levels were horrendous

The comments identified did not appear to be direct opposition to incorporating ARMS into their daily workload by participants. Rather, it appeared that the focus group afforded a forum whereby they could share their concerns about the additional work these changes have involved for them.

Longer-Term Benefits

However, there appeared to be a general consensus that by initially investing the time and effort, ARMS produces longer term gains: "It's an investment"; "We have got less staff than we had before and far more offenders than we ever had before and we're probably doing better now than we did with the old system".

According to Hill and Jackson (2016), people are prepared to make short-term sacrifices on the basis that their efforts will result in longer-term benefits. This was particularly evidenced in the use of the word "investment" but also implicit when reference was made to "the old system". Almost certainly, the most important gain that was identified for police officers was lower numbers of high or very high risk offenders: "Over time it'll get quicker, and not only that . . . I absolutely guarantee you we'll get lower high and very high risk [offenders]"; "We certainly gained. I did a, did a high [risk] well, what, and he's actually . . . could've gone down to a low [risk]"; "We're not wasting time visiting very high

risk [offenders] who have been very high risk for 15 years . . . they're not very high risk . . . so we're now visiting them probably twice a year as opposed to every month"; "The return [from the initial investment] you'll get it back by bucketfuls".

It is of interest that participants appeared to place value upon reducing the levels of risk for offenders, though this appeared to relate to the longer-term benefits for the offender managers, e.g. a reduced workload, than it did for the offenders themselves. This was particularly notable in the comment about "wasting time". Nonetheless, participants considered that reducing risk was motivational for some offenders:-

> Where it's reduced and the offender's been really chuffed that you know, really pleased to know of that and agrees with, *'I'm really glad you, you know, you've listened to me, and can see that I've done x, y and z and it says in the evaluation that I'm not high, I'm actually low, you know, you're quite happy to reduce me down to a medium'*, so in, in the discussion it's good that he's taken ownership for that but it's been quite a confidence booster for them as well

Research (Leue, Brocke, & Hoyer, 2008) has found that the impact of 'rewarding' sexual offenders differs between various subgroups. It was reported, that sexual offenders respond more to rewarding stimuli than non-offenders, though this was greater for sex offenders with impulse control disorders than for paraphilic offenders who spend greater amounts of time and energy planning their offending. As such, it will be important for police officers to recognise the function of the offending in order to assess whether being informed of a reduced risk has any significant bearing on desistance. A further word of caution in relation to these issues might relate to complacency. If an offender is advised that risk has reduced, they may not put in place as many 'safeguards' as they had previously to prevent re-offending.

I can understand that, you could get some, *'I've done all this, and you're still coming to see me every 28 days, is there no recognition for the fact that I done essentially'*, they would say, *'everything you wanted me to do'* and I think yeah you're right actually, if they've then done this and you can say well actually '*congratulations we're now gonna come and see you a little bit less*' There was also a suggestion that not only will completion of ARMS reduce risk but with that, workloads may decrease (over time):-

> My view on it is, and I, I've been doing it a long time is that erm you have gotta bring, your [management] is gonna have to be prepared to take a hit initially that, that you're gonna be more busy, you're gonna get behind your visits and err the pay off for that over time, it, it'll take about a year won't it . . . it will save you time
>
> "Eventually, once we take that hit, it will reduce the time, irrespective of how many [offenders] you manage".
>
> We're now reducing our workload because of err you know if someone has a very high you know it really is very high and they'll probably be Monday to Friday on it, just working on it constantly, on that one offender all week 'cause they've got the time to do that because we've got less of em

It would appear that police officers managing sex offenders in the community place a significant amount of importance on reducing risk amongst the offenders they manage and this is not surprising given their vital roles in protecting the public. Whilst there was specific reference to reducing workloads, there was also reference to ensuring good quality assessments: "They've gotta look at the quality as well, do you wanna visit five hundred sex offenders and do a, a superficial job where one of them's gonna go wrong".

Theme 2: Resources

Participants were in disagreement with regard to the level of paperwork required for ARMS, with 34% believing it is manageable and 43% disagreeing (see Figure 3). The remaining 23% remained ambivalent (*N=12*).

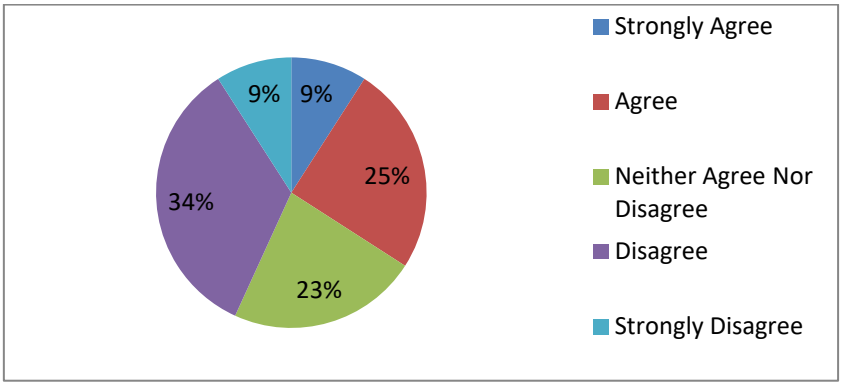


Figure 3. Percentage of Responses to the statement:

"The level of paperwork required for ARMS is manageable in the course of my daily duties".

As noted elsewhere, the introduction of ARMS and its impact on current commitments had relevance to the resources available to participants: "At a time when our staffing levels were horrific"; "It didn't fit in with what we've got at the moment"; "I think you need to look at how many offenders you visit, 'cause the headline figure … is 120 but they don't visit anyone who's got probation, [location] don't visit loads of mediums anymore".

A criminal justice joint inspection report (Bridges & O'Connor, 2010) found generally good practice by police and probation staff, particularly with regard to restrictive efforts. However, it was noted that improvements were needed in the channels of communication between police and probation services. Collaboration with other professionals was identified by two participants with one noting: "The DS or the assessment liaison officer would complete that and let me know that it's on the shared drive and then I'd look into it"; and another:

We've done them together [with probation officers], and that was the best, that, that was the best outcome and erm it's interesting to interview together and then it's interesting, so [name] had gone away and err, written his evaluation and then we had another written together

Sharing the workload may serve to improve communications, adhere to MAPPA standards and, as a result, improve management and treatment efforts with sexual offenders in the community.

Theme 3: Training

Importance

The level of training that was received appears to be a critical feature of the positive responses received in relation to this research. Ninety-eight percent (98%) of respondents (*N=52*) believed that the training they received in ARMS was sufficient for them to incorporate ARMS into their daily duties (see Figure 4).

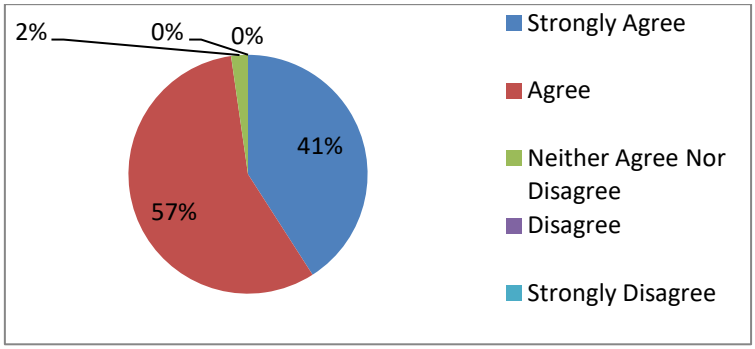


Figure 4. Percentage of responses to the statement:

"The training I received in ARMS was sufficient for me to use it in my daily duties"

Indeed, the training was considered a critical aspect to the roll-out of ARMS across police, prison and probation services: "The training for me is everything, if we get that wrong, it ain't gonna work"; "I found that it equipped me, as, as much as I needed to be . . . I knew what I was doing really once I'd had that training"' "If we had the money every offender manager should have [this training]. The training currently is dated".

Also of relevance, was the expertise of the trainers:-

> We were lucky, we had [the trainer] because he knew what he was talking about but if you have a trainer that thinks he knows what he's talking about and doesn't put it across as, as well, then I think you might have a watered down version

One participant noted, "I didn't want to do the training but I've thoroughly enjoyed it now". Another stated, "I've done the MOSOVO course and was involved in the pilot, so was fully aware of what to expect". Despite more than half of the participants reporting that they would change some aspect of the training (see Figure 6), 98% considered that the training gave them sufficient confidence in completing ARMS (see Figure 5) with the offenders they manage:-

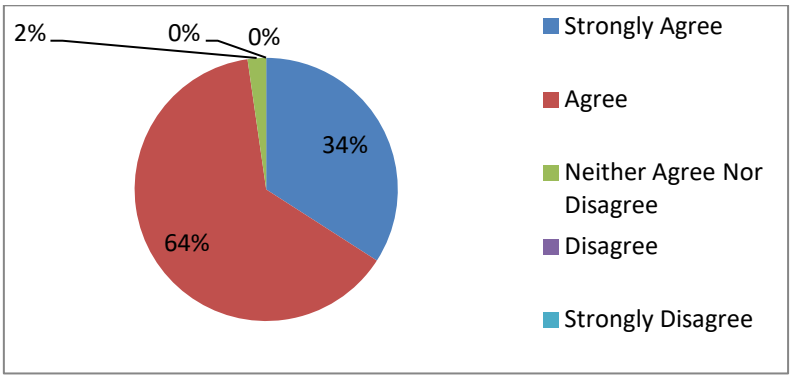


Figure 5. Percentage of responses to the statement:

"I am confident in completing ARMS"

Training

Nonetheless, training in relation to managing sexual offenders within the community was considered vital:-

> The old CEOP stuff but it, it's in here, it's at the back of your mind and you kind've understand where you're coming from with offender management, so when you do this training, you're quarter of the way there already aren't you, you know about the factors, you know what they're talking about

Furthermore, participants considered that having the right staff was another serious consideration:-

> Some of that would come back to selection I think, who are they putting in these teams and how long are you staying here . . . if you wanna do this for public protection then this is you for a minimum of five years and this is your path to doing it
>
> The reality of it is that they need a, as a minimum, they need a ViSOR course as a minimum, before they even speak to a sex offender otherwise they're gonna come back err, it only takes one to go wrong doesn't it, *what training did you have, well no they just sent me out*

As well as experience:-

> You need to convince them [the managers] . . . you've only gotta go on the course, do a couple of these, get your head around it, based on experience and you think *you know what that's pretty good*, but it's higher up, you, you need [managers] on board

Improvements

Sixty-four percent (*N=34*) of participants would change some aspect of the training:-

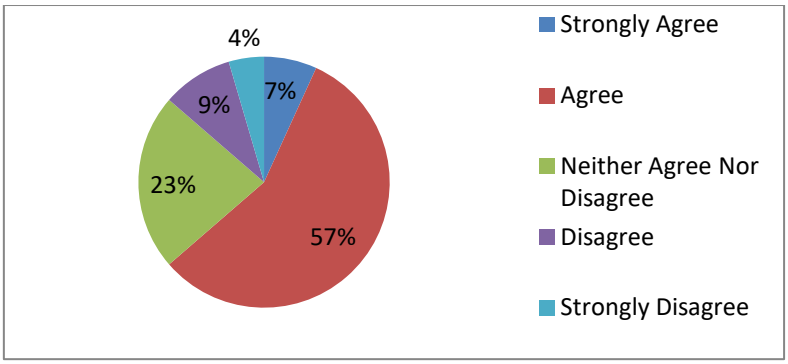


Figure 6. Percentage of responses to the statement:

"There are aspects of the training that I would change"

Given that ARMS has been rolled out nationally across all police and probation areas, the trainers recognised that they would need to have a 'train the trainers' element (Yarber et al, 2015) added to their model so as to manage time efficiently. These kinds of training techniques have been found effective in various settings, including mental health services (Limm et al, 2015). However, for participants, there was perhaps some concern that the training would be 'diluted' if this approach were approved.

> I think it would be really helpful if one of the, the people on the team [developers] had come out, sat at my desk, or sat at your desk, and said *right operationally this is what we've got* and that's, and I think [the trainer] was gonna do

Theme 4: Structure

An important finding in this research was that some of the offender managers believed ARMS required modification, both in the structure of the assessment and the way in which it is recorded.

Layout

Participants considered that the layout of the ARMS paperwork needed restructuring:-

> It's really crucial I think 'cause if, if, where it'll fall down, if you've identified a high priority there and you haven't got it in the plan, what you've done about it that's gonna go spectacularly wrong. If you identify it, you've gotta do something about it and so, the first bit's fantastic, I think the last bit needs tweaking

> You're opening up and thinking right who have we got here, and it only mentions, it mentions their name, it mentions a few things and then it goes on to opportunity and I think well, who is this person, wha, wha, what, have they offended, who do they live with

> You've gotta split that up so the second one down is the case summary, that needs to be at the beginning . . . summary could go at the beginning, you've then got

> your risk assessment, and then at the end, I think if you get [inaudible] and incorporate it into that bit . . . and then just right at the end on your risk management plan, is just [inaudible] strategy and your actions

Content

The group were directed to the item content within ARMS and discussions surrounding each of the factors was encouraged. Whilst the general consensus appeared to be that the various subheadings were helpful, there was also concern that staff might not collaborate with other agencies when completing ARMS:; "If you give people a template they stick to it and what you're getting is what the offender said"; "There's no section in it for probation services, housing, social care, you know, again, on ours we've got a section for every agency"; and that there can be a degree of repetition amongst the various sections: "I think that you, you do get offen, offender managers who will duplicate the same thing again but it's usually down to the fact that they've not read the section and understand it properly".

The critical role of multi-agency working has been highlighted elsewhere and has, in the past, been a slight criticism for police and probation services, such that it will be important for the tool to be revised so that it is a 'catch all' in terms of the various agencies.

One of the ARMS factors in particular arose during discussions, Factor 11: Social Investment – 'Giving Something Back': "They think it's charity work but it's not about charity work, it's do they do anything just to kee, to give something back . . . it wasn't just about charity work, it's about helping out";

> That aspect, in fact that was the only one that we have an issue, sort, sort of trying to convince the sergeants and the inspector about some, some of them, we've all got them, will just sit there all day watching Jeremy Kyle won't they

This item was included to capture any contribution the offender is making to society, their community, and families. The rationale for incorporating this factor was that

contributing something positive can enhance self-esteem and agency and lead to more positive, prosocial relationships. All of which may increase the potential for desistance.

Recording

Participants expressed ambivalence with regard to how information is recorded once it is captured:-

> The problem in [location] was that our systems we use are very different . . . I found that the real issue was, was ViSOR. I didn't know where to put it, certainly not within the structure that we have to do business, so although I found it extremely useful and a really good tool, it didn't fit in with what we've got at the moment so I think we would need to change quite considerably the way that we go at it

Even for the most recently trained cohort, this appeared to be a concern: "I'm unsure how it's going to work on ViSOR".

Notwithstanding, 71% of respondents expressed the view that ARMS is user-friendly (see Figure 7) and there appears to be a suggestion that the developers are taking on board feedback: "I think this actually it's like it's giving more structure to what has been evolving".

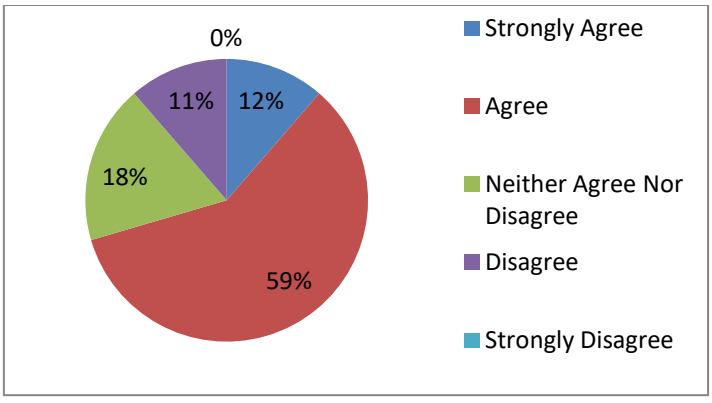


Figure 7. Percentage of responses to the statement:

"ARMS is user friendly"

Purpose of ARMS

However, there were discussions with regard to whether ARMS is a risk assessment or a management plan:-

> "I think the important question for the team . . . is that the plan, or is that an assessment, 'cause it does say plan [this was changed during the course of training] on the front of it, it doesn't say it's an assessment".

Any assessment of risk, irrespective of the mode, should be done with a view to informing risk management and treatment. However, for police officers, they are not tasked with addressing treatment needs and it would therefore suggest that ARMS is a joint risk assessment and management tool.

Theme 5: Previous Experience

ARMS

The degree to which various managers have been exposed to ARMS and been able to use it within the course of their daily duties was variable: "I'd had an, an erm like an overview of what ARMS was proposing to be erm for quite some months so I would often refer to that in doing any assessments that I was doing"; "I was only able, I was only able to do the half dozen that were suggested for the pilot".

Sexual Matters

Focus group participants considered that previous experience of working within the field of sexual offending was critical to the training, for example:-

> The training was really good. I just think it's a little bit dependent on your previous experience as well . . . if you're gonna ask him about his sexual preoccupation and they sort've said well *'how am I gonna ask him how many times a day he you know plays with himself, I can't ask an offender that'* and you're not are

you, you're not gonna just sit there and say right, *my name's [participant] how many times a day* . . . your staff have gotta have the experience to be able to do that

Theme 6: Current Practice

At the time of data collection, just over half (55%) of participants (*N=29*) were using ARMS with every offender they manage (see Figure 8).

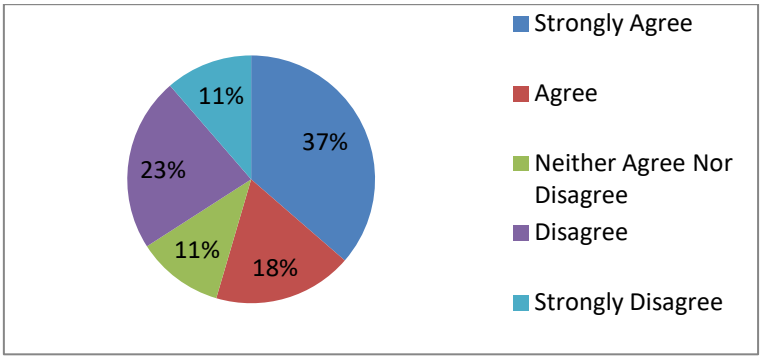


Figure 8. Percentage of responses to the statement:

"I use ARMS with all the offenders I manage"

With the national roll-out, it would be anticipated that these figures will have increased significantly, as the tool is now employed within a standardised approach to assessing and managing sexual offenders in the community.

RM 2000

The utility of the RM2000 (Thornton, 2007) was discussed within the context of current practice and participants appeared to express the view that RM2000 can overestimate risk, thereby resulting in the inappropriate deployment of resources:-

> The Risk Matrix is static isn't it so, err, these aren't accurate figures by the way, but, but it's along the lines of if you've got a hundred people in a room, 75 of them are not gonna reoffend, 25 of them in there are, so, it doesn't tell you which ones, you know so you've gotta put the dynamic on there to say which one of those hundred, my 25 do I need to concentrate on

That's so much more of an indicator, it's felt like that there, that says he's very high risk and you're thinking I'm not having that, that's just ridiculous, or you've got ones that say low and you're thinking I'm just not happy with the dynamics of [that]

So you do your risk matrix, he's coming out as very high, right we need to do a visit now to see if he really is or not, whether he's one of those who is . . . we never, certainly not with the Risk Matrix, but even with some of the old systems, we never had that, if it went wrong . . . it's just another factor isn't it, you've spoken to probation, you've been and visited err you've spoken to relatives . . . you're risk matrix score, you're not gonna ignore it"

Participants appeared to be advocating the use of structured professional judgement, over unstructured clinical judgement:

We have him, 25 of age, on drugs, no job, no partner, comes out as very high risk, and then three years later they've got a job, their wife's, their, their girlfriend's pregnant, they're not on the drugs, Risk Matrix, very high . . . hold on

Theme 7: Usefulness

Overall, the general feedback regarding the usefulness of ARMS was positive: "I found it extremely useful and a really good tool . . . It's something that we want"; "I think it's superb and it's about time we had something that makes it that clear".

One participant in particular, reflected on their experience of using ARMS, stating, "It focuses, it gives you structure . . . it sort of tells you, gives you a bit of a guide . . . so … I think it's really good".

In managing resources:-

The force policy plan is that because they're very high risk they will be monitored by visits every month … that, that's part of your plan and if, and then your

last line and your catch all is any changes to be brought to the attention of the supervisor

Now we've got, there is a problem with this particular risk factor and this is what we're gonna do about it, and as soon as you do that, you start to think *well right I'm gonna need the help of the beat team or I'm gonna need, you know, we're gonna have to go nick him or whatever else there is,* an, and it's there written out

Despite some uncertainties about ARMS, as referenced above, 78% of respondents (*N=41*) would recommend ARMS to their colleagues (see Figure 9).

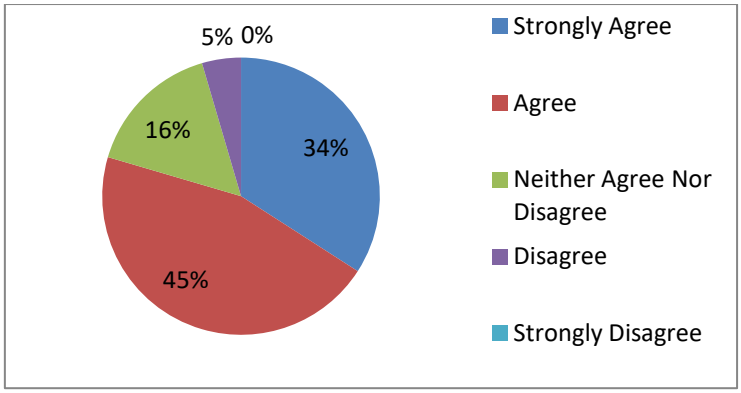


Figure 9. Percentage of responses to the statement:

"I would recommend ARMS to my colleagues"

89% would like to continue using ARMS within the course of their duties (see Figure 10).

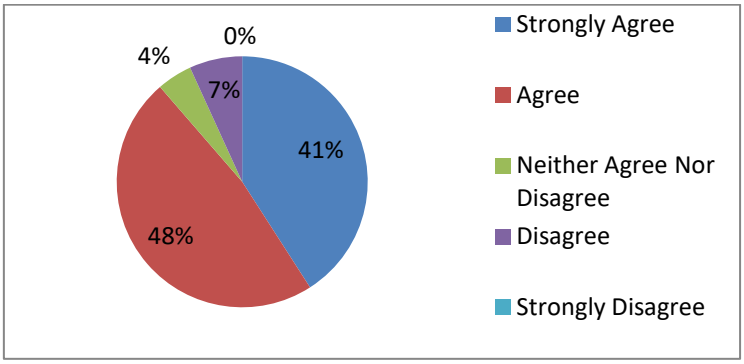


Figure 10. Percentage of responses to the statement:

"I would like to continue using ARMS during the course of my daily duties"

Theme 8: Defensibility

Importantly, participants expressed confidence that ARMS would provide a good framework for justifying decisions they make:-

> I'm 100% confident to stand there, once you've done one of these, in fact I'm more confident . . . I'm happy to stand there and say, *because of this, there's my evidence for what I did, this is why I did what I did* . . . and I think that gives, gives you confidence anyway

> *This is a really difficult decision, I didn't know what to do, but based on the evidence here what I'm gonna do is this*, yeah, and if they see that you're virtually [inaudible] proof, you know even if you've taken 'cause they talk about erm good decisions, with a bad outcome, you know that's what they talk about to us . . . to me that is everything you've got there to back you up

Eighty-seven percent (87%) of participants (*N=46*) believed that ARMS had improved their confidence in assessing sex offenders in the community (see Figure 11).

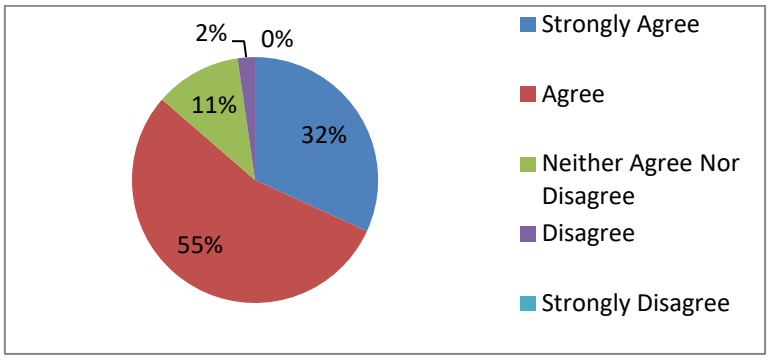


Figure 11. Percentage of responses to the statement:

"ARMS has improved my confidence in assessing sex offenders"

In addition, when combining the data from phases one and two, 91% of participants would like to see ARMS rolled out nationally (see Figure 12).

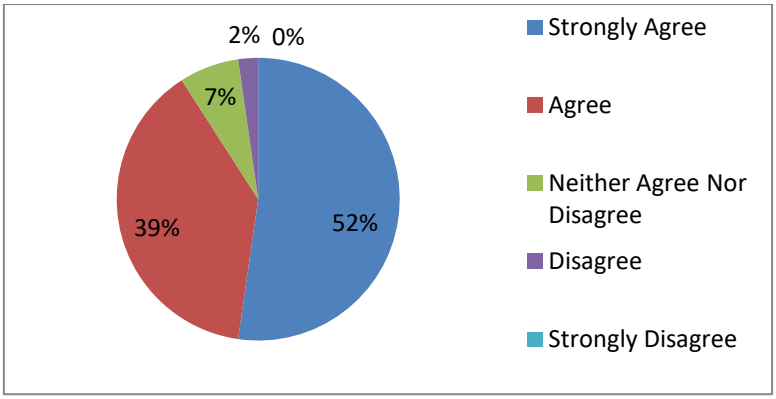


Figure 12. Percentage of responses to the statement:

"I would like to see ARMS rolled out nationally"

Resistance to Reducing Risk

Notably, participants anticipated that they might be met with resistance from their superiors in terms of lowering risk levels:-

> If they see that the person was a very high and then [participant] turns up and says *I want to make him a medium* it's gonna be so much easier for that boss to say *actually I'm not happy just leave him as he was*

> I've done a few that've gone high to low on this and I'm still managing them at medium 'cause the boss won't have it that you can drop from high to low and I've got to review it in another year again

Theme 9: Psychosexual Development

As noted above, the initial findings from the data highlighted some concerns about how offender managers address sensitive sexual matters with offenders. As such, an additional day was added to the training to assist offender managers in understanding psychosexual development with the aim of enabling a more confident approach to addressing sexual matters with the offenders they manage. An Independent Social Work Consultant was recruited to deliver this training. The consultant had specific experience in the fields of sexual trauma, psychosexual therapy, sexual offending and child protection.

The presentation *'Understanding Normal Sexual Functioning to Assess Sexual Deviancy'* was delivered interactively with participants who engaged well with the presenter's style. A number of areas were covered including the various stages of psychosexual development, sexual health, differences between genders, the impact of age and illness upon sexual health and psychosexual dysfunction. The presenter spent time exploring with participants different sexual pathways, sexual desire, pornography and sexual aggression, as well as various paraphilias. She presented a biopsychosocial model of sexuality within the presentation. The overriding theme of this training related to a recognition that sexual offenders are often meeting 'normal' needs through deviant behaviours and that professionals managing these offenders in the community should apply their knowledge of 'normal' sexual functioning when discussing sensitive issues with offenders.

Five organising themes were identified within this basic theme of psychosexual development: the level of training; the relevance of the training; the format; improvements that could be made; and the presenter's incorporation of humour.

Level of Training

> Participants commented that the level of the training was appropriate: "[The presenter] gauged the audience well"; "[The presenter] made it personal but not creepy, cringy or condescending".

However, one participant expressed the view that one element of the training was difficult to follow as they stated, "[The presenter] lost me at a couple of points on the brain structure". Nonetheless, the general consensus appeared to be that the training was helpful:

> "[The training] was very beneficial".

Notably, although the presenter was only responsible for delivering training on the final day of a three day event, she attended all three days and this was considered useful as

she was able to refer back to previous information that had been presented:"[The presenter] was referring back to previous days [training], making the case study relevant".

Relevance

Seventy percent (70%) of the participants believed that some offender managers struggle to discuss sensitive sexual issues with offenders and 87% expressed the view that training was required. In this context it was noted: "[The training] helps us understand where the offending came from, where we should be starting, going back to get insights into risks of re-offending"; "Officers in our role [receive] half a day's in-house training, specialising. We're assessing all this sexual risk but know nothing about it".

Ninety-three percent (93%) of participants expressed the view that their knowledge of psychosexual development had improved through the training delivered on day three andhe training appeared to instill confidence in offender managers in discussing sexual issues with offenders, with 91% reporting this to be the case. In particular, the training was considered as having: "Helped with the way that you word the questions".

One participant referred directly to a specific aspect of the training noting, "Courtship and grooming was really good". This related to the presenter contextualising deviant sexual behaviours (grooming) meeting 'normal' sexual needs (courtship). Relatedly, 80% of participants agreed that the training would help them to feel less anxious about discussing sexual matters and 91% felt better equipped to do so. It was noted: "I think (the presenter) covered pretty much the issues we talk about"; "(About) how we will be able to speak to them, so they don't get aroused, but even if they are, how to address it now".

Ninety-seven (97%) of participants expressed the view that day three of the training enhanced their overall training experience and believed it was relevant to the previous two days training: "The input was really useful. It's not an area that has ever really been explored before".

Nearly all of the participants (97%) concluded that day three should be incorporated within future training events.

Format

Some participants considered that the training agenda was disorganised and that day three might have been better placed elsewhere within the itinerary: “Should have been on day one, some of day two and then ARMS training”and that additional training with regard to psychosexual development may have been appropriate: “A two day input would have been helpful. [The presenter has] got a lot more to tell us”.

Improvements

In relation to these comments, participants identified where improvements could be made.

They also commented upon the content of the presentation: “A little more [information needed] on data around the abuse [offenders] tell you they’ve received”; “Touching on younger people and brain development, more information needed. . . practical information about how significant it is”.

Humour

The presenter reported that many people working with sex offenders have also experienced negative/traumatic sexual experiences and how it is important to be aware of the possible impact, with the training potentially raising some anxieties. Participants were encouraged to do whatever they needed to do, to feel safe in the training. The humour within the training was aimed at addressing and reducing the anxieties staff have regarding discussing sexual behaviours and assessing how far from the norm an offender is thinking and behaving. This style of engagement appeared to work well: “Very good, well presented, funny and relevant”; “Very good, light hearted. [The presenter] kept everybody’s attention”.

Quantitative Results

Initial analysis was conducted to ensure that the data set (*N=434)* was appropriate for factor analysis. The results were significant (p<.000, KMO = .90). A principal component analysis (PCA) was therefore conducted to identify patterns within the data and highlight similarities and differences (Bartholomew, Steel, Moustaki, & Galbraith, 2008). The results suggested three possible components, although one item (hostility) did not make complete theoretical sense as it loaded more heavily onto protective factors than it did sexual risk factors. Research (Firestone, Nunes, Moulden, Broom, & Bradford, 2005) has found that hostility is significantly associated with recidivism amongst intra and extra-familial child molesters. However, there was no association between hostility and recidivism amongst rapists and mixed offenders. The current study did not differentiate between the types of offenders, thus, it was unclear why hostility loaded onto protective factors in the PCA.

Furthermore, relationship loaded onto a factor of its own and there was a degree of ambiguity in relation to the opportunity to offend factor as it appeared to load onto both the protective factors and the sexual risk components (appendix O).

The item associated with 'opportunity' to offend loaded onto both protective and risk factors, leading to some ambiguity with regard to this item and perhaps inferring factorial complexity or a lack of specificity (Kline, 2009). Notably, within the data set, the only variable that was nominal was the relationship variable. However, one could argue that opportunity to offend is also nominal in that the opportunity is either there or it is not (the offender falls into one category or the other). Whilst it was removed from the model, this item is important in terms of imminence of risk and will be of specific interest for the police (and probation) officers managing offenders in the community because although there may be some common elements in offenders' decisions to desist from re-offending, there may be other features that are idiosyncratic to the particular offender (Wakeling, Webster, Moulden,

& Marshall, 2007). Furthermore, when present, 'opportunity to offend' may, in some circumstances, represent a step towards 'overcoming external inhibitors' (Finkelhor, 1984).

Of the eight items, five demonstrated an 'excellent' loading where a cut-off of .71 is employed (Tabachnick & Fidell, 2007). Whilst there are only a few loadings on each of the factors, this is considered acceptable as the sample size was greater than 300 (Stevens, 2002).

Of interest, the developers of ARMS incorporated social influences into both risk and protective factors, separating them out as anti-social influences (risk) and pro-social network (protective). The current analysis revealed that social influences loaded onto protective factors rather than risk. This will likely be specific to the sample population as de Vries Robbé, Mann, Maruna and Thornton (2015) correctly point out, these features can co-exist for any offender at any time in their lives.

In order to examine the results further, particularly the ordinal variables pertaining to risk (low, medium, high and very high), Multidimensional Scaling (MDS: Lingoes & Guttman, 1973) was employed using an ordinal coefficient appropriate for the data. MDS is a family of techniques which offer a non-linear alternative to factor analysis by representing the relationships between variables as points in geometric space (Bishopp, 2003). Clusters of variables are considered to represent distinct facets in the same way that cluster of variables represent factors in factor analysis. The data was analysed using specialist software that is not available in SPSS (HUDAP, ver 8 Reuven: Barak, Shiloh, & Amar, 2005). The specific MDS analysis employed was a smallest space analysis (Guttman, 1968). The fit of the model within MDS is determined via a coefficient of alienation (COA). The COA for the solution was 0.05 which is very good and indicates a very low stress in the model. A model would be considered to fit less well and have a higher stress index (Shye, 1978) with a threshold of between 0.15 and 0.2. The MDS is provided in appendix P.

The orientation of the relationship variable clearly skewed the rest of the structure and was therefore removed. Factor analysis and MDS support the structure as having two main components and independent relationship and hostility variables. This four component model would suggest that relationship status and hostile orientation do not sit neatly within either risk factors or protective factors in this model. They are, nonetheless, important considerations for practitioners managing sex offenders in the community and may influence an offender's motivation and commitment to treatment. Thus, from a quantitative perspective whilst ARMS could discount these two variables (together with opportunity to offend) in order to streamline the tool, they likely hold important qualitative information that will assist offender managers in terms of managing dynamic risk factors. These items may also represent a further content domain that could be developed in future research. When considering relationships and hostility, it should be noted that they can represent both risk and protective factors. For instance, a relationship can be a settling factor for some offenders wherein their sexual and emotional needs are being met. However, where dysfunctional relationships exist, this may serve to increase risk. Furthermore, for preferential paedophiles, the existence of an adult intimate relationship may serve only to mask the offender's true desires. With regard to hostility, Firestone, Nunes, Moulden, Broom, & Bradford (2005), found that overall it was correlated with sexual (and violent) recidivism. However, they also noted that the correlation was stronger for child sexual abusers than for rapists. Therefore, offender managers will benefit from taking account of these features within case formulations.

The two solid components, sexual risk and protective factors, explain over 60% of the variance across the incorporated items. The results indicate that ARMS would be better considered as a set of underlying facets or scales, rather than simply a total score or individual items. These facets might therefore be better considered as four distinct

components, namely, sexual risk, protective factors, hostile attitude (to management) and relationship status.

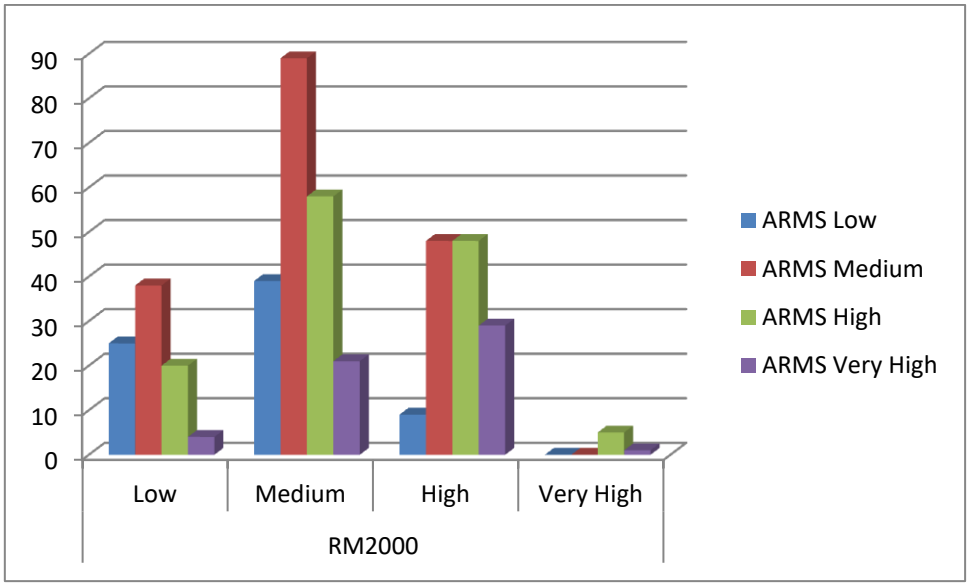
The data was explored in a number of ways to examine how ARMS items were related to one another and to the final ARMS rating and the RM2000 scores. This revealed a high level of agreement between items in terms of correlation (r= .815, p .029).

Table 7. *ARMS v RM2000*

		ARMS				Total
		Low	Medium	High	Very high	
RM2000/S	Low	25	39	9	0	73
	Medium	38	89	48	0	175
	High	20	58	48	5	131
	Very high	4	21	29	1	55
Total		87	207	134	6	434

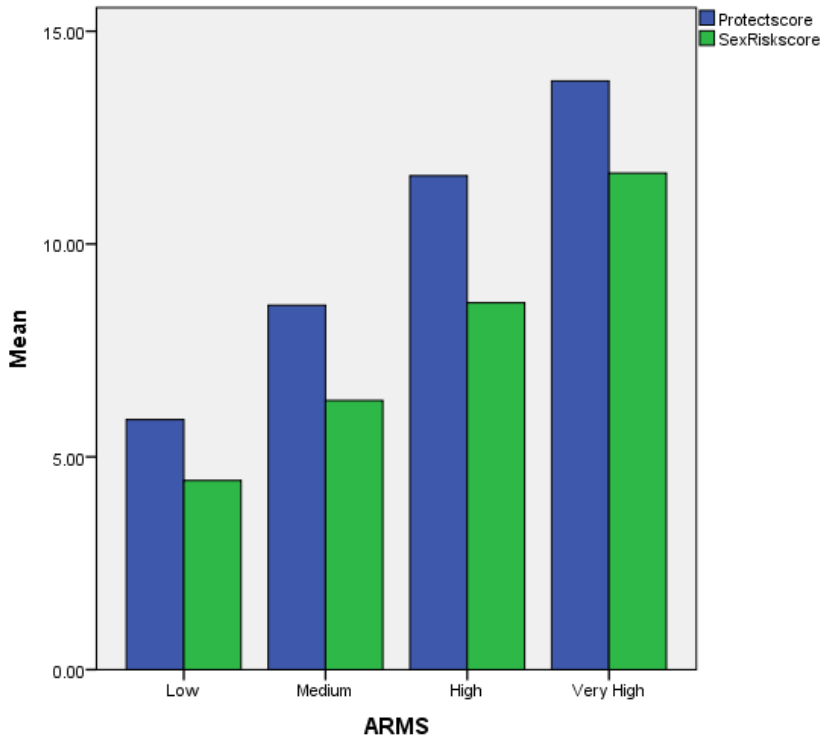
In 163 (37.6%) of the cases, RM2000 scores and ARMS priority ratings were similar and much of the agreement falls within the medium risk band (N=89, 54%). We can also see that ARMS ratings reduced perceived risk in 170 (39%) of the cases and increased perceived risk in 101 (23%) of the cases. These results are represented visually in Figure 13 below:

Figure 13. *ARMS v RM2000*



Using the items combined from the factor analysis and excluding the two items that did not predict (regression), two empirical scales were created from the items, namely protective factors and sexual risk using MDS, which maps factors which share similar characteristics closer in geometric space. The reliability was very good for both protective factors (Cronbach's a 0.85) and sexual risk (Cronbach's a 0.82).

Figure 14. *Protective factors v sexual risk*



An inspection of each of the risk categories revealed differences between the RM2000 scores and ARMS priority ratings. Within the low category, the numbers increased from 73 to 87 on RM2000 and ARMS respectively. For the medium risk category, the number increased from 175 to 207 and for the high category, from 131 to 134. For low, medium and high categories, it would appear that practitioners perceive that RM2000 fails to correctly identify risk based on actuarial assessment, suggesting that for these three categories, dynamic risk and protective factors have little impact upon actuarial data. However, for the

very high risk category the number decreased from 55 (12.7%) to 6 (1.4%) suggesting that when dynamic factors are taken into account, the perceived number of very high risk offenders is somewhat less than identified with actuarial assessment. However, it should be noted that these alterations are made based on subjective judgements and until further outcome data is available, these risk estimates cannot be considered accurate as subjective judgements are little better than chance in identifying risk.

Nonetheless, this finding will no doubt be of interest to police and probation officers managing offenders in the community. If these priority ratings are employed in making decisions about the level of monitoring being offered with the current sample, officers would see an increase of 91 more visits per year to offenders falling within the low to high categories. However, they would see a decrease of 588 visits per year to very high risk offenders. This equates to 49 less visits per month in the region where data was gathered, a huge demand on resources. Nonetheless, if action is taken in the absence of scientific evidence regarding the utility of ARMS, this could have serious consequences for individual offenders and the public. Importantly, police detection rates for sexual offences tend to be around 29% (Smith, Taylor, & Elkin, 2013). Given that sexual offences account for a small proportion of recorded crime, around 1% (MoJ, 2013), together with the fact that many sexual offences go undetected or unreported, perhaps a more practical and valuable allocation of resources could be focused on understanding what leads to sexual offending in the first place, with a greater focus on prevention, particularly at a primary and secondary level, e.g. before offending has taken place (Smallbone, Marshall, & Wortley, 2008) but also at a tertiary level, to prevent re-offending and re-victimisation.

Discussion

Qualitative Studies

Initially this research was focused on evaluating the small pilot study of police officers employing ARMS (phase one) and the training delivered to offender managers (phase two). The numbers of participants involved was small. Further, although some descriptive statistics have been highlighted within this report, the data has largely been analysed from a qualitative perspective. Whilst this approach allows for the capture of rich data, this kind of analysis also has a strong researcher effect, such that there will always be an element of subjectivity.

This small study provided some constructive feedback to the developers of ARMS with suggestions, in particular, that the layout of the paperwork would benefit from some modifications. Notably, the initial findings identified that some offender managers were uncomfortable discussing sensitive sexual issues with some of the offenders they manage and the developers took this on board and incorporated relevant training. There are indications that the developers have responded well to feedback and that they will continue to benefit from ongoing evaluation of ARMS. Rigorous evaluation over time is highly recommended, particularly since the current research was based on a small study focusing upon the experiences of those piloting ARMS and those engaged in training. Nonetheless, the findings tend to suggest that the decision to roll-out ARMS nationally is likely to prove positive in terms of improving consistency and qualitative and quantitative outcomes for the practice of managing sexual offenders in the community. Certainly, the findings from this study suggested that offender managers who are already using ARMS and those that have been trained more recently are confident in using ARMS, would recommend this tool to their colleagues and supported national roll-out.

Quantitative Study

The results revealed that ARMS may show some potential as a structured tool with both the factor analysis and MDS indicating two clear factors that are readily interpreted. The remaining items within the tool, namely opportunity to offend, relationship status and hostile orientation, perhaps require further development as additional themes.

Officers managing sexual offenders in the community allocate resources based on the identified levels of risk. Historically, risk has been based purely on the actuarial results of RM2000. The number of visits for assessing and monitoring offenders vary from once a year (low risk), bi-annually (medium risk), quarterly (high risk) and monthly (very high risk). The current study revealed that where ARMS is employed, the risk level for 37.6% (*N=163*) of the total sample (*N=434*) remained the same as measured by RM2000. However, 39.2% (*N=170*) revealed a decrease in perceived level of risk and 23.3% (*N=101*) produced increased risk levels. Officers managing sexual offenders in the community perceive, based on their priority ratings on ARMS, that in more than half of cases (62.5%), RM2000 fails to correctly identify risk. As can be seen, for the vast majority, this is an overestimation of risk. These results suggest that officers managing sex offenders in the community appear to place more weight on the specific offender's current circumstances than past behaviour. An overview of the impact of these results on allocated resources can be found at appendix Q.

Based on the changes in perceived risk when ARMS is applied to individual offenders, police officer visits for the purpose of assessing and monitoring offenders in their care could reduce from 1606 visits per year to 1109, 497 visits less per year, a 31% reduction in allocated resources. This is in one region alone and the allocated police hours nationally could be significantly reduced. However, these visits are carried out by police officers for a number of reasons. Primarily, their role is to protect the public and ensure that offenders are complying with their imposed conditions. They are also constantly gathering intelligence and

may also have a role in preventing or disrupting offending behaviour. Whilst the outcome of this research may have practical utility for offender managers and may be welcome news for operational staff, it should be noted that the current research did not evaluate the efficacy of ARMS and to reduce visits without systematic evaluation of the tool, may lead to offenders not being sufficiently monitored and thereby expose the public to risk.

It will be important for longitudinal research to be conducted in the future, taking account of actual recidivism rates. This would afford the opportunity for a comparison of predictive accuracy between ARMS and RM2000. The current research would suggest that ARMS is a helpful tool in assessing risk. However, follow-up studies will be required in terms of determining its accuracy. At this stage, it is unclear whether ARMS outperforms other measures employed in assessing sexual risk.

Furthermore, the current study did not take account of the different kinds of offences (contact, non-contact, child/adult victims, crossover, etc.) committed by each of the 434 sexual offenders. In assessing future risk potential, and formulating management plans, these issues may have significant bearing and will require further research. An additional limitation relates to the officers conducting the assessments. Information concerning their rank, years in the force, experience of working with sex offenders and relevant training was not collected for the current study, though will likely prove important in terms of the level of knowledge and expertise required for carrying out these kinds of assessments.

As noted earlier, these officers require specific skill sets in order to work with sexual offenders. The role of police has changed over the last twenty years and frontline staff are now responsible for managing sexual and violent offenders in the community. This includes monitoring offenders but also contributing to their community reintegration. Relevant training will continue to be important.

Summary and Conclusions

The current research was a mixed design of qualitative and quantitative analysis containing three phases. Phase one focused on the perceived usefulness of ARMS from the perspective of police officers employing the tool in their daily duties and phase two explored the training being provided to frontline staff to enable them to employ the tool. Overall, participants considered that ARMS was a helpful tool which, with some modifications, would increase their confidence in assessing and managing sexual offenders in the community. Participants supported the national roll-out of ARMS and this was implemented prior to the conclusion of this research.

The quantitative analysis focused on assessing the variables within ARMS, together with its overall association with the RM2000. The results revealed that the 11 dynamic risk factors incorporated within ARMS are best considered as four distinct scores or components: sexual risk; protective factors; hostility; and relationship status. By incorporating dynamic risk into the risk assessment, police officers managing sexual offenders in the community can appropriately focus their attention on those offenders who continue to present with high levels of risk. Risk assessments need to be tailored to the individual and by incorporating dynamic factors, ARMS goes some way to achieving this. However, it is still based on 11 items and therefore does not present an holistic view of the individual offender's risk, needs and responsivity. As identified throughout this thesis, thorough case formulations for individual offenders is vital.

The current study revealed that when dynamic risk factors are incorporated into the assessment of sexual offenders being managed in the community, risk levels are (on average) deemed to be lower, than purely through actuarial assessment. However, these results are based on a snapshot, as well as subjective determinations by police officers. It will be

important for longitudinal research to be conducted to evaluate whether the attributed risk levels are consistent with actual recidivism rates.

Chapter 4

A Critique of the RM2000: Sexual, Violent and Combined Recidivism Scales

Introduction

One of the roles of the forensic psychologist is to assess risk so that offenders can be managed appropriately both within the community and in custodial/inpatient settings. Over the last twenty-five years, risk assessment has evolved based on empirical research (please refer to chapter one for an overview). Each of the four approaches have their opponents and proponents, though this critique offers an overview of a second generation risk tool, the Risk Matrix 2000 (RM2000), in terms of its practical utility, its scientific properties and its use in research studies. The RM2000 was specifically chosen for critique because the research element of the doctoral thesis relating to ARMS incorporates the RM2000 into the risk assessment methods and contributes to the assessors' understanding of risk for specific offenders.

Assessment of Sexual and Violent Risk:

A number of tools have been developed over time to assess risk of sexual and violent offending. These include the Rapid Risk Assessment for Sex Offence Recidivism (RRASOR: Hanson, 1997); the Sex Offender Risk Appraisal Guide (SORAG: Quinsey, Harris, Rice & Cormier, 1998), the Violence Risk Appraisal Guide (VRAG: Quinsey, Harris, Rice & Cormier, 1998); the Minnesota Sex Offender Screening Tool – Revised (MnSOST-R: Epperson, Kaul & Hesselton, 1998); Structured Anchored Clinical Judgment (SACJ: Hanson & Thornton, 2000); and the STATIC-99 (Hanson & Thornton, 2000). However, it has been asserted that none of these tools take account of non-sexual violence with sexual offenders (Thornton et al, 2003). Rather, they classify risk in relation to sexual recidivism or overall violent recidivism (to include sexual recidivism) in terms of its statistical likelihood. In a 10 year follow up study of convicted sex offenders released from prison in England and Wales, Thornton and Travers (1991) found that approximately one fifth ($1/5^{th}$) of offenders were reconvicted of sexual offences, whilst a further fifth ($1/5^{th}$) were reconvicted of non-sexual

violent offences, thus arguing that sexual offenders pose two distinct types of risk (sexual and non-sexual violence) that existing tools failed to assess.

This research led to the revision of the SACJ, resulting in the SACJ-Min (Grubin, 1998), a brief classification algorithm which included two steps, the first of which contained five items relating to static risk factors associated with criminal history. These factors were incorporated on the basis of the popular maxim that 'past behaviour best predicts future behaviour'. However, these items do not take into account important dynamic factors that may reduce risk. The second stage contained four aggravating factors which, based on published literature, were associated with increased risk of sexual recidivism. However, a meta-analysis (Hanson & Bussière, 1998) identified that the SACJ-Min did not take into account important correlates of recidivism, namely age and prior sexual offences (Thornton et al, 2003). Thornton and colleagues therefore developed the Risk Matrix 2000 (RM2000).

RM2000 Overview

The RM2000 (Thornton et al, 2003) was developed as a 'quick' alternative to Clinical Judgement (Barnett, Wakeling & Howard, 2010), a method regularly employed in police, prison and probation services in the UK during the 1990s (Ireland & Craig, 2011). It has since become one of the most widely used tools for sex offender assessment within prison and probation services in England and Wales (Barnett, Wakeling, & Howard, 2010). It is also approved by MAPPA, not only within prison and probation services but also within the police force (Ministry of Justice, 2012). The RM2000 is readily available for anyone to use, although the authors assert that it is the responsibility of the user/organisation to take account of all available information about the person they are assessing, together with relevant legal, policy, professional, organisational and clinical frameworks. Furthermore, official training in the use of the RM2000 is recommended.

The RM2000 was designed to be used with males over the age of 18 who have been convicted of at least one sexual offence (after the age of 16). The most recent version of the tool is the RM2000.10/SVC (Thornton, 2010). The tool contains three scales that examine risk of recidivism for sexual violence, non-sexual violence and combined risk, e.g. sexual or non-sexual violence. Each scale is described in more detail below.

Risk for Sexual Recidivism (RM2000/S)

This scale has two steps in evaluating risk. The first step is based on three factors that have been identified in the empirical literature as associated with risk, namely, age, number of sexual appearances and number of criminal appearances (Thornton, 2007). Step two relates to 'aggravating factors'. Aggravating factors are considered present when any of the four items are identified as relevant. The four factors are male victim, stranger victim, non-contact offence and the perpetrator being single, i.e. has not been in a relationship with another adult for at least two years. Where aggravating factors are present, they can increase the initial risk category. Where two or three factors are endorsed as present, the risk category is raised one category, e.g. medium to high but where all four aggravating factors are present, the category goes up by two, e.g. from medium to very high.

Risk for Violent Recidivism (RM2000/V)

The RM2000/V scale contains three items, namely, age at commencement of risk, number of sentencing appearances or formal police cautions for non-sexual violent offence and any conviction for burglary. Points from each of these three items are accrued to give an overall score that assigns the category in which the offender is placed, low, medium, high or very high.

Risk for Sexual and Violent Recidivism Combined (RM2000/C)

The final scale affords the opportunity to combine the results from the sexual and violence scales to evaluate the risk of recidivism for either type of offending. According to

the authors, this decision was taken on the basis that the two former scales have equal weighting, e.g. their individual ability to 'predict' risk being "equally effective" (p.8). However, where this occurs, there are likely to be errors of estimation as the weighting is based on variables, rather than individual cases.

By assigning categories, assessors employing the RM2000 are able to compare offenders against predefined recidivism rates based on convicted offenders with various characteristics (Thornton et al, 2003). It is a statistically-driven tool that relies on relationships between various factors found within the empirical literature to be associated with recidivism. However, like many of its predecessors, the RM2000 does not take into account dynamic or protective factors, a fact the author readily acknowledges.

In conducting this critique, reliability and validity issues in relation to the RM2000 are considered, that is, does the measure perform consistently over time and does it measure what it was intended to measure. ?

Reliability

A number of studies have considered inter-rater reliability using the sexual scale of the RM2000 (RM2000/S) wherein raters achieve similar results to one another when independently assessing an offender in terms of sexual risk (Grubin, 2011; Wakelin, Mann, & Milner, 2011). To date, the violence scale (RM2000/V) and combined scale (RM2000/C) have not been assessed for inter-rater reliability. In this regard, because items are factual, one would expect a high level of inter-rater reliability, though this would be dependent upon the rater having received appropriate training in the use of this tool (Pike, 2015) and being familiar with and taking account of validity, reliability and normative samples when assessing offenders (BPS, 2012). It would also be subject to the different raters having the same information available to them (Barnett et al, 2010), though generally, research suggests high levels of inter-rater reliability (Kingston et al, 2008).

Validity

Face Validity

Whilst the RM2000 claims to predict the likelihood of reconviction for sexual and violent offending, it is not offender specific. That is, it places offenders into categories of low, medium, high and very high risk, with a percentage of these offenders being reconvicted of a subsequent sexual or violent offence. These categories are based on static risk and aggravating factors, without offering much meaning as to what led to the offending behaviour. Furthermore, within each of the categories, the RM2000 does not identify which individuals will be reconvicted. Indeed, at an individual level, Hart, Michie and Cook (2007) assert that all ARAIs are "virtually meaningless" (p.63). When an RM2000 is completed for an individual, the assessor is making a determination of the risk they pose of further offending behaviour. The RM2000 does not do this. It simply classifies the individual based on the behaviour of others. Furthermore, whilst the RM2000 sets out to predict future sexual and violent crime, with a further category combining the two, the author acknowledges that the violence subscale should not be used to predict recidivism but should only be used for treatment purposes. To his credit, Thornton (2007) highlights the limitations of the RM2000 within the scoring guide. He cautions that the reconviction rates are subject to sampling error, that reconviction rates may vary across jurisdictions and over time, and that reconviction represents an underrepresentation of reoffending.

Concurrent Validity

In terms of concurrent validity, Craig, Brown and Stringer (2004) found that the RM2000 sexual scale correlated highly with the STATIC-99 and the Sexual Violence Risk-20 (SVR-20: Boer, Wilson, Gauthier & Hart, 1997). Kingston et al (2008) also found that the RM2000 correlated well with STATIC-99 and SORAG.

Predictive Validity

In assessing predictive validity (as will be seen thus far), a number of researchers have employed the AUC statistic. The author (Thornton, 2007) reported the following AUC values for the RM2000:-

Table 8. *RM2000 AUC values*

Scale	Type of Reconviction	N	Follow Up (years)	AUC
S	Sexual	647	2	0.77
S	Sexual	429	16-19	0.75
V	Non-Sexual Violence	647	2	0.85
V	Non-Sexual Violence	311	10	0.78
V	Non-Sexual Violence	423	16-19	0.80
C	Sexual and Other Violence	276	Mean 3.7	0.81
C	Sexual and Other Violence	406	16-19	0.74

However, there have been various AUC values reported, particularly as the measure applies to different populations. A further study was conducted by Tully, Chou and Browne (2015) with a focus on the sexual scale of the RM2000. Ten studies are described with AUC values ranging from 0.58 (Intellectually Disabled offenders) to 0.77 (treatment completers). Further studies included internet offenders, female offenders and subtypes of male sexual offenders, e.g. rapists and child molesters (see table 9 below).

It will be noted that these various studies produced an average AUC of 0.67 (range 0.58-0.77), whereas the author of the tool reported an average AUC of 0.79 (range 0.74 to 0.85), taking the effect size from moderate to large. Whilst the above study demonstrated a moderate effect size for the RM2000, once again, it can only be applied to the sample as a whole and does not provide any form of 'prediction' as to whether the specific individual will re-offend. In view of the differences between Tully and Browne's results and those of Thornton, it should be questioned whether developer bias has had an impact on these results.

Table 9. *From Tully and Browne (2015)*

Study	Sample Size (N)	Follow-Up	Location	Recidivism Rate	AUC	Sub-Group
Wilcox, Beech, Markall and Blacker (2009)	27	76 months	UK	30%	0.58	Intellectually disabled
Wakeling, Howard and Barnett (2011)	1,344	2 years	UK	3.1%	0.67	Internet offenders
Kingston, Yates, Firestoone, Babchishin and Bradford (2008)	192	11.4 years	Canada	3.8% 2yr 9.4% 5yr	0.65	Contact offenders including rapists and child molesters
Craissati, Bierer and South (2011)	221	9 years 2 months	UK	11%	0.71	Child molesters
	80			16%	0.64	Rapists
Bengtson (2008)	304	16.2 years	Denmark	28%	0.65	Rapists and child molesters
	160			27%	0.61	Rapists
	144			29%	0.71	Child molesters
Barnett, Wakeling and Howard (2010)	4,946	4 years	UK	5.5%	0.68	Treated and untreated, mixed offence types and mixed victim ages
Parent, Guay and Knight (2011)	275	5 years	USA	15.1%	0.72	Child molesters
	174			16.9%	0.63	Rapists
	54			4.5%	0.60	Both adult and child victims
Looman and Abracen (2010)	419	7.1 years	Canada	12.9%	0.66	Rapists and child molesters
Thornton et al (2003)	647	3.7 years	UK	2.6%	0.77	Treatment completers
	429	19 years		27.7%	0.75	Prison discharges
Grubin (2008, 2011)	1,029	5 years	Scotland	10.8%	0.73	Prison discharges

It should also be noted that this method of evaluating ARAIs has been described as misleading due to insufficient transparency within the data and a large margin of error in producing individual predictions (Cooke & Michie, 2014). They go on to question how this information can be admissible in judicial settings when they do not appear to meet the criteria for scientific rigour (Frye v United States, 1923; Daubert v Merrell Dow Pharmaceuticals, 1993). These assertions may have relevance to many of the risk assessments currently employed when working with sexual and violent offenders (please see chapter 2). In addition, Cooke and Michie (2014) advocate the use of SPJs, which assist in 'breaking down' risk in relation to type, imminence, frequency, victim type and management, rather than

providing estimates of future risk potential for a group of individuals rather than a particular offender (Craig, Browne, & Stringer, 2004). One such SPJ, the Risk of Sexual Violence Protocol (RSVP: Hart, Kropp, Laws, Klaver, Logan, & Watt, 2003) is a commonly used tool amongst forensic psychologists and is described in chapter one. It is often employed alongside the Structured Assessment of Protective Factors (SAPROF: de Vogel, de Ruiter, Bouman, & de Vries Robbé, 2012).

Content Validity

A number of factors have been found to be associated with sexual and non-sexual recidivism amongst sexual offenders (Hanson & Bussière, 1998) and in order to have content validity, the RM2000 needs to cover a representative sample of these factors. Mann, Hanson and Thornton (2010) identified sexual preference for children (under the age of 12) as a strong risk factor and whilst this is incorporated within the STATIC-2002 (Hanson & Thornton, 2003), the RM2000 manual does not indicate whether this factor was taken into consideration for future risk potential. Craig, Beech and Brown (2006) described four further items which correlate well with future sexual and violent recidivism which are not incorporated within the RM2000. These relate to a history of being in foster care, school maladjustment, past substance abuse and a history of employment difficulties/instability.

In any event, the RM2000 does not discriminate between individual offenders, taking account of the various treatment needs for different kinds of offenders, for example, child molesters versus rapists. Nonetheless, by categorising offenders, the RM2000 indicates which offenders are likely to require more intensive treatment and thus, guides professionals with regard to treatment mode and duration. Furthermore, it is recognised that some of the identified factors associated with recidivism, do not need to be targeted in treatment in order to reduce offending behaviour. Rather, it has been argued that sexual offending is influenced by multiple factors that can be the focus of treatment goals (Hanson & Yates, 2013).

Construct Validity

In developing RM2000, Thornton and his colleagues incorporated items based on past behaviour as a means of judging risk. None of the items took into consideration treatment needs or "psychologically meaningful risk factors" (Mann, Hanson, & Thornton, 2010). Sex Offender Treatment Programmes (SOTPs) tend to target not just risk, sexual deviancy and relapse prevention but also problem solving skills, emotional regulation and interpersonal effectiveness (Gray & Wilcox, 2015).

The three scales of the RM2000 have been cross-validated with various groups of offenders (Barnett, Wakeling & Howard, 2010), though the analysis that was conducted has been criticised by some as being misleading due to a lack of significant association between the seven factors identified within the RM2000. Cooke and Michie (2014) identified that on the sexual scale of the RM2000, only four of the seven factors demonstrated a significant association with one another, namely age, number of sexual appearances, stranger victim and non-contact offence. They reported that only two of the factors on the violence subscale were significantly associated. This point appears to be moot, given that the RM2000 was not designed to assess one relatively pure construct and different risk factors will be different for different types of offender.

As noted elsewhere, the author of the RM2000 acknowledges that the violence scale is not appropriate for 'predicting' risk of future violence, but rather was intended to inform treatment plans. Furthermore, in most circumstances, the author opines that RM2000 should not be used as a standalone tool but should be used in combination with dynamic risk assessment. It is described as a tool to direct practitioners with regard to the allocation of resources, including treatment and supervision (Thornton, 2010) which is guided by the likelihood of recidivism.

Appropriate Norms

The base rates provided by RM2000 emanate from a ten year follow-up study of convicted male sex offenders released from prisons in England and Wales during 1980 (Thornton & Travers, 1991). Thornton and colleagues validated the RM2000 on several samples by conducting a logistic regression to identify predictive factors in relation to reconviction and found moderate to high AUC values of between 0.74 and 0.85 (Thornton et al, 2003). However, RM2000 has been widely employed across a range of heterogeneous groups of offenders including internet offenders, female offenders, learning disabled offenders, as well as sub-groups of offenders, e.g. child molesters and rapists. The RM2000 was not normed for these various groups and therefore its predictive ability is likely to be compromised with such groups. Furthermore, the base rates provided by Thornton and colleagues, would vary across samples of offenders and the RM2000 does not therefore have the predictive accuracy across these various sub-groups.

Conclusion

Much research has been conducted regarding the RM2000 and the developer has highlighted that the tool ought to be used in conjunction with dynamic factors and functional analysis (Beech, Fisher, & Thornton, 2003). Thornton (2013) also acknowledges the importance of protective factors which he describes as "social or psychological factors that make recidivism less likely" (p.64). As such, he opines that the RM2000 should only be used as a standalone tool if the offender does not cooperate with the assessment process. For example, it would be difficult to conduct an assessment of dynamic risk and protective factors if the offender were unwilling to share information about their current circumstances. As such, practitioners would need to rely on official records to make a determination about risk. This would then enable them to appropriately direct resources such as levels of

supervision and treatment intensity. In these circumstances, the RM2000 may be viewed as a helpful screening measure.

Risk is individual-specific and takes into account dynamic risk and protective factors. The importance of protective factors and the impact this can have upon desistance, particularly for life-course persistent offenders (de Vries Robbé, Mann, Maruna, & Thornton, 2015), currently appears to be at the forefront of professional thinking. These factors are also considered important in relation to developing risk management plans (Bonta, 2002). Indeed, Boer (2013) opined that the incorporation of protective factors into risk assessment are "essential to effective community integration" (p. 8).

Efforts to address this shortfall are currently being made within police, prison and probation services through the national roll-out of ARMS developed by the National Offender Management Service (NOMS) and the Association of Chief Police Officers (ACPO). In the meantime, Winder, Lievesley, Elliott, Norman and Kaul (2015), assert that the RM2000 remains "one of the most robust static risk tools used to predict sexual and violent offending" (p.346).

This critique has highlighted the importance of individual risk assessment rather than a 'one size fits all' approach. It is notable that ARAIs categorising offenders as low, medium, high or very high risk, often stay with the offender even when protective factors become evident. Any risk assessment should be fluid and therefore subject to change. Risk assessments should also be focused on the management of risk, together with treatment plans rather than simply labelling offenders or putting them into categories. Indeed, there is evidence to suggest that when this happens, it can have a negative impact on social reintegration for offenders (Mingus & Burchfield, 2012). On the basis that the RM2000 is not considered by the Courts in Europe as a useful risk assessment tool, it would appear that practitioners are now preferring a combination of risk assessment methods to include

actuarial risk, structured professional judgement and protective factors to inform treatment planning and within the wider context of an individualised case formulation for each offender, for whom idiographic features may be most salient in their offending behaviour (Vess, Ward, & Collie, 2008).

Chapter 5

General Discussion

General Discussion

The emphasis of this thesis was on the assessment of risk with adult male sex offenders. The introduction (chapter one) explored the current status of risk assessments with reference to the various approaches employed over time and the different 'generations' of risk assessment. Risk assessment has moved from unstructured professional judgement, which was found to be little better than chance, through to the use of actuarial assessments, structured professional judgement (SPJ) and a combination of SPJ and actuarial tools. The prevalence of sexual offending was explored, together with difficulties in detection rates. The role of police officers managing offenders in the community was discussed within the context of the sex offender register and MAPPA. The thesis described a move from deficits-based approaches to strengths-based approaches to risk assessment and treatment planning. The introduction considered the incorporation of protective factors and highlighted the need to assess risk employing the Risk-Need-Responsivity (RNR) model and for assessments and interventions to be tailored to the individual. It highlighted the common use of the Risk of Sexual Violence Protocol (RSVP) and the Structured Assessment of Protection Factors (SAPROF), particularly amongst forensic psychologists. It also reported on the importance of thorough case formulations with sexual offenders. The introduction concluded with an overview of ARMS and set out the purpose of the current research.

Summary of Findings

The systematic review (chapter two) revealed that many of the risk assessment tools currently employed with adult male sex offenders have only moderate predictive validity. Furthermore, where large predictive validity was found, the results will likely have been subject to some developer bias. Nonetheless, when incorporated with an assessment of dynamic risk, case formulation and psychological input, these tools can be helpful in

determining risk. The chapter concluded with a discussion around whether practitioners working with offenders should be trying to 'predict' recidivism and problems inherent with the term itself. These issues were explored in relation to existing diagnoses versus prediction. Furthermore, it was argued that that tailored comprehensive assessments would be a more appropriate option than risk assessment tools alone. These kinds of assessments should be based on the RNR model and should include comprehensive formulations and functional analyses. Use of existing SPJ tools, including the RSVP and SAPROF were advocated, as these go further than simply classifying risk, breaking it down into nature, severity, imminence, frequency/duration and likelihood of risk and incorporating protective factors that include internal factors (e.g. intelligence, empathy, coping), motivational factors (e.g. work, attitudes towards authority, life goals) and external factors (e.g. social network, professional care, living circumstances). This reduces the amount of weight given to actuarial assessment and facilitates detailed psychological assessments that aid defensible legal and treatment decisions. Tarrier and Calam (2002) assert that in order to understand the "origins, development and maintenance" of any problematic behaviour, one needs to integrate theoretical psychological knowledge with appropriate information (p.312). Case formulation should be evidence-based, nomothetic knowledge (based in law) which can be applied to ideographic context (conceptual) including origins of symptoms; triggers and maintaining factors; sensitivity and specificity of treatment; all of which can contribute to tailored treatment (Sturmey & McMurran, 2011). Furthermore, formulations should be developed collaboratively and regularly tested and revised. Vess and Ward (2011) argued that in order to develop a dynamic and responsive risk management framework, detailed information regarding the offender and their life circumstances, together with the availability of supervision, support and treatment is required (p. 191). In conducting individual case formulations, practitioners ought to take into account specific vulnerabilities for the offender

(Darjee & Russell, 2012), including but not restricted to, issues such as mental health, learning disabilities, posttraumatic stress symptoms. These kinds of assessment approaches are regularly employed by forensic psychologists, though perhaps less so with police officers managing sexual offenders (please see chapter three).

Chapter three examined whether the incorporation of dynamic risk and protective factors into risk assessment methods currently employed by community managers of sexual offenders reduced the perceived risk posed by those offenders. Specifically, the research focused upon the introduction of a new tool, Active Risk Management System (ARMS), and the impact this had upon perceived risk, as assessed by police officers managing offenders in the community. It began with a description of the initial evaluation of ARMS training and highlighted recommended developments to the training being delivered, together with some adjustments to the tool itself. The research was a mixed design including Likert-scale questionnaires, focus groups and group interviews with officers engaging with the training. Overall, qualitative analysis of officers' views appeared to lend support to ARMS in terms of its usefulness and the confidence it instilled in officers. Thereafter, ARMS was rolled out nationally with police and probation services and this presented an opportunity to collect a large amount of assessment data from a regional police force. The chapter goes on to describe a large scale study, wherein the results of 434 ARMS assessments were collated and analysed with statistical software programmes, SPSS and Multidimensional scaling. Initial analysis revealed that two dynamic risk items, hostility and relationship status, appeared to operate independently from the remaining items. When these two items were removed from the analysis, the results highlighted two specific components across the remaining nine dynamic risk items incorporated within the ARMS assessment, namely, sexual risk and protective factors. Importantly, opportunity to offend loaded onto both of these components and this may relate to factorial complexity suggesting that the item is either too ambiguous or

too general. Nonetheless, opportunity to offend is critical in terms of the frequency, severity and impact of risk. Five of the items loaded onto protective factors, namely, employment, self-management, social investment, social influences and desistance. The remaining three items loaded onto sexual risk, namely sexual interests, preoccupation and emotional congruence with children. These all make theoretical sense.

There are clearly some limitations to the tool as all of the factors could be further elaborated, and further factors explored concerning offenders' attitudes towards the assessment, and indeed their attitudes towards offending. There are also questions as to whether offenders are likely to be honest about these issues with police officers in an authoritative role. For this reason, the continued incorporation of RM2000 will be vital. As noted elsewhere, often police officers have to rely on official records only and, in these circumstances, the RM2000 can offer an estimated level of risk which can then inform treatment and management plans. On ARMS, the item relating to social investment suggested that there is a large area untapped in terms of the social circumstances and support available to individual offenders which may promote desistance.

The results also demonstrated that when dynamic factors are weighed up against static risk, police officers' judgements about perceived levels of risk can be altered. In this study, this was true for almost two fifths of the assessments conducted. Whilst this may have an impact for operational staff when allocating resources, a word of caution is necessary. The current research did not evaluate the efficacy of the tool and it is yet unclear whether it is an appropriate resource. Despite this, ARMS has been rolled out nationally. Future research may support this decision. However, it may also prove to have been premature and, if deemed inadequate, may have caused considerable financial losses for the statutory authorities. More critically however, if the tool is found to be ineffective, some offenders may not have been adequately supervised in the community and this could potentially extend

to further victimisation. Furthermore, ARMS does not take into consideration, other motives or underlying sexual interests that may persist. Whilst officers receive specific training in working with sexual (and violent) offenders, their role continues to focus on "an investigative approach, underpinned by respectful skepticism" (College of Policing, 2017). The current research revealed that officers employing ARMS during the course of their work, highlighted a lack of specific training with regard to the complex nature and functional aspects related to sexual offending.

Chapter four focused specifically on the RM2000 and its utility in assessing sex offenders. It provided a critical review of the tool and considered issues associated with reliability and validity. The chapter offered recommendations regarding the incorporation of dynamic risk and protective factors, alongside actuarial measurement. Risk assessment employing the RM2000 as a standalone tool does not afford the assessor the opportunity to take into account individual differences amongst offenders or provide the opportunity to explore what initiated or maintained offending behaviour, all factors that contribute to effective treatment planning and community reintegration. There has been a lot of research in relation to this tool and many criticisms of it. However, it was never intended to be a standalone tool and the importance of dynamic and protective factors within the wider context of case formulation were discussed. Fluidity of risk assessments is also recommended, with alterations being made as and when changes occur in the life of the offender.

Theoretical and Practical Implications

Practitioners working with sexual offenders need to take a more holistic approach to assessing risk and formulating management plans that take account of the offender's history and their current circumstances, including social, environmental, familial, interpersonal, occupational, spiritual and self-concept aspects in their presentation. Whilst more research is

needed in terms of the predictive validity of ARMS, the current research suggests that there is still considerable variance in the way that different practitioners perceive risk and often, determinations are extremely subjective. Nonetheless, in combination with actuarial data, police officers employing ARMS may now have a better framework for the beginnings of a more thorough case formulation. Despite the fact that ARMS has not been scientifically validated, the tool has been rolled out nationally across police and probation services. Whilst there are indications that this may decrease the resources required for police officers managing sexual offenders in the community, it should be noted that significant resources have been allocated in terms of the roll-out and the training required for these officers. Yet, it is not clear whether ARMS is effective. Reducing monitoring efforts with sex offenders in the community, based on subjective judgements, may lead to catastrophic results, particularly as the current research revealed a trend in reducing very high risk offenders to medium risk. It is recognised that people have the capacity to change and this is no different for sex offenders. However, if an offender is deemed very high risk on the RM2000 and a police officer subsequently reduces his risk to medium, based on ARMS priority ratings, the offender would only receive monitoring visits twice a year, rather than monthly. Without external restrictions, this may lead some offenders to become complacent and to engage in risk-related behaviours. It may also leave some offenders feeling isolated and unsupported.

Thesis Limitations

The thesis was limited to considering assessment of adult male sexual offenders and did not take consideration of the assessment of other groups of offenders, for example, female offenders, juvenile offenders or offenders with special needs. Furthermore, it did not take account of various types of offences, namely contact v non-contact, child v adult and crossover offenders. These areas are particularly important given that the rates of recidivism for different kinds of offenders are markedly different. The research did not explore issues

associated with cultural differences, cognitive functioning or ethnicity. These too will be important factors to take into account as the recidivism rates for various groups of offenders vary (Skelton & Vess, 2008). When formulating a case, through sound psychological practice, all of these factors would be captured.

The qualitative analysis was a small study and subject to researcher bias. Whilst the quantitative study was much larger, it did not examine the utility of ARMS. Rather, it determined whether police officers managing offenders in the community perceived different levels of risk in offenders when incorporating ARMS factors, as compared with the RM2000 alone.

Conclusions

Over time, many non-clinical practitioners have favoured an actuarial approach to assessing risk. However, two systematic reviews (Tully et al, 2013 and chapter two of the current thesis) have demonstrated that many of these tools have only moderate predictive validity. As can be seen in chapter four, the RM2000 is no different. Perhaps the problem relates to how these tools are employed. In the world of psychology, practitioners do not, and should not, view any assessment tool, irrespective of what it is measuring, as a standalone resource. Rather, the results obtained from assessment tools need to be considered in the wider context of the individual being assessed. For example, if we consider the use of a personality inventory, any features identified through item endorsement need to be weighed with the respondent's known history, clinical presentation, interpersonal relationships, current situation, past trauma, support networks and so forth. Unfortunately, it is often the case that non-clinical practitioners rely heavily on the results of actuarial measures without taking account of these additional features. A comprehensive risk assessment including dynamic

risk and protective factors is surely warranted in order to safely protect the public and assist the offender in maintaining an offence-free lifestyle.

In terms of the current research, it should be noted that further, longitudinal research will be required regarding the effectiveness of ARMS. What the research has highlighted is support for practitioners to continue to utilise all of the resources at their disposal in order to produce a comprehensive risk assessment and to recognise that desistance can be achieved based on changing circumstances for the individual offender. The combination of actuarial assessment and structured professional judgement may offer a framework for taking a more holistic approach when working with sexual offenders, such as that afforded by RSVP and SAPROF. The importance of thorough case formulations was highlighted and advocated.

Whether professionals involved in the assessment and treatment of sexual offenders continue with the route of combining actuarial data with structured professional judgement remains to be seen. However, it is of critical importance that the risk assessment tools employed are 'fit for purpose' and that they are understood within the wider context of the offender, such that tailored assessment and treatment are conducted with specific relevance to the individual offender. For many years, practitioners have debated the 'one size fits all' approach and, more recently, research (Mews, Di Bella, & Purver, 2017) has demonstrated that manualised programmes such as the prison-based Sex Offender Treatment Programme (SOTP) have little effect on recidivism, with treated sex offenders more likely (10%) to recidivate, than their non-treated comparison group of offenders (8%).

Importantly, the current research highlighted that whilst police officers are expected to manage sexual offenders in the community, this is a diversion from their traditional roles, wherein they were responsible for investigating crime and detaining criminals. These changes came about with the introduction of the sex offender register and the requirement for

multi-agency public protection arrangements to be adhered to in order to protect the public and manage sexual and violent offenders in the community. It would appear that this has not been an easy transition, with many officers in the current study describing the implications these kinds of efforts have on their workloads and other time commitments, particularly since the introduction of ARMS. Nonetheless, there appeared to be a general consensus that ARMS was helpful in the course of their duties and many offered support for its national roll-out. It became evident that significant resources have been allocated to developing ARMS, providing training and rolling the tool out nationally. With time and further research, it may be that this option was premature, given that we do not know whether ARMS is effective. Furthermore, police officers may benefit from the additional expertise of forensic psychologists in producing risk assessment and management plans.

Ward and Salmon (2011) indicated that "each of us is obliged to think deeply about our responsibilities to sex offenders, victims, the community, and ourselves" (p.398) in working with sex offenders. Relatedly, effective community integration, and with this, an offence-free lifestyle are critical factors to achieve in working with sexual offenders. If practitioners managing offenders in the community are unable to achieve these kinds of results, then this will inevitably lead to more victims of sexual abuse.

References

References marked with an asterisk were included in the systematic review, those with two asterisks were excluded

Akerman, G., Craig, L.A., & Beech, A.R. (2015). Treating child sex abusers: A person-centered approach. In. D.T. Wilcox, T. Garrett and L. Harkins (Eds.). *Sex offender treatment: A case study approach to issues and interventions.* Chichester, West Sussex: John Wiley & Sons Ltd. ISBN: 978-1-118-67440-6

**Allen, B.P., & Pflugradt, D.M. (2014). An exploration of the latent constructs of the STATIC-99, *International Journal of Offender Therapy and Comparative Criminology, 58(11),* 1376-1388. DOI: 10.1177/0306624X13496046

Andrews, D.A., & Bonta, J. (2006). *The psychology of criminal conduct* (4th ed.). Newark, NJ: LexisNexis. ISBN: 1593453213 9781593453213

Andrews, D.A., & Bonta, J. (2010). *The psychology of criminal conduct (fifth edition).* New Providence, NJ: Matthew Bender & Company Inc. ISBN: 978-1-4224-6329-1

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., Zinger, I., Hoge, R.D., Bonta, J., Gendreau, P., & Cullen, F.T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology, 28,* 369-404. DOI: 10.1111/j.1745-9125.1990.tb01330.x

**Babchishin, K.M., Hanson, R., & Blais, J, (2016). Less is more: Using Static-2002R subscales to predict violent and general recidivism among sexual offenders, *Sexual Abuse: Journal of Research and Treatment, 28(3),* 187-217. DOI: 10.1177/1079063215569544

**Babchishin, K.M., Hanson, R.K., & Herman, C.A. (2011). The characteristics of online sex offenders: A meta-analysis, *Sexual Abuse: A Journal of Research and Treatment, 23(1),* 92-123. DOI: 10.1177/1079063210370708

**Babchishin, K.M., Nunes, K.L., Hermann, C.A., & Malcom, J.R. (2015). Implicit sexual interest in children: Does separating gender influence discrimination when using the Implicit Association Test? *Journal of Sexual Aggression, 21(2),* 194-208. DOI: 10.1080/13552600.2013.836575

**Baltieri, D.A., & Boer, D.P. (2015). Two clusters of child molesters based on impulsiveness, Revista Brasileira de Psiquiatria, 37(2), 139-145. DOI: 10.1590/1516-4446-2014-1568

**Banse, R., Koppehele-Gossel, J., Kistemaker, L.M., Werner, V.A., & Schmidt, A.F. (2013). Pro-criminal attitudes, intervention, and recidivism, *Aggression and Violent Behavior, 18(6),* 673-685. DOI: 10.1016/j.avb.2013.07.024

Barak, A., Shiloh, A., & Amar, L. (2005). An organizational grid of federated MOSIX clusters. Conference paper. DOI: 10.1109/CCGRID.2005.1558576

Barnett, G. (2013). *ARMS v2: Inter-Rater Reliability.* Unpublished Presentation.

Barnett, G.D., Wakeling, H.C., & Howard, P.D. (2010). An examination of the predictive validity of the Risk Matrix 2000 in England and Wales. *Sexual Abuse: Journal of Research and Treatment, 22,* 443-470. DOI: 10.1177/1079063210384274

*Barnett, G.D., Wakeling, H.C., Mandeville-Norden, R., & Rakestrow, J. (2012). How Useful Are Psychometric Scores in Predicting Recidivism for Treated Sex Offenders? *International Journal of Offender Therapy and Comparative Criminology, 56(3),* 420-446. DOI: 10.1177/0306624X11403125

Bartholomew, D.J., Steel, F., Moustaki, I., & Galbraith, J. (2008). *Analysis of Multivariate Social Science Data (2nd Edition).* Florida, USA: Taylor & Francis Group LLC. ISBN: 978-1-58488-960-1

Beech, A.R., Fisher, D.D., & Thornton, D. (2003). Risk assessment of sex offenders. *Professional Psychology: Research and Practice, 34,* 339-352. DOI: 10.1037/0735-7028.34.4.339

Beech, A.R., Wakeling, H.C., Szumski, F., & Freemantle, N. (2016). Problems in the measurement of dynamic risk factors in sexual offenders. *Psychology, Crime & Law, 22(1-2),* 68-83. DOI: 10.1080/1068316X.2015.1109095

*Beggs, S.M., & Grace, R.C. (2011). Treatment gain for sexual offenders against children predicts reduced recidivism: A comparative validity study. *Journal of Consulting and Clinical Psychology, 79(2),* 182-192. DOI: 10.1037/a0022900

**Bench, L.L., & Allen, T.D. (2013). Assessing sex offender recidivism using multiple measures: A longitudinal analysis, *The Prison Journal, 93(4),* 411-428. DOI: 10.1177/0032885513500765

Bensel, T., & Harris, D. (2017). Introduction for special issue on desistance from sexual offending, *Criminal Justice Studies, 30(2),* 97-100. DOI: 10.1080/1478601X.2017.1301918

Bishopp, D.F.C. (2003). *Dimensions of Sexual Aggression.* Thesis Submission: University of Surrey.

**Blais, J., & Bonta, J. (2014) Tracking and managing high risk offenders: A Canadian initiative, *Law & Human Behavior, 39(3),* 253-265. DOI: 10.1037/lhb0000109

Blandford, M., Farmer, M., Mann, R., Scott, J., & Jarvis, P. (2013). *Active Risk Management System (ARMS).* Crown Copyright.

Blandford, M. (2014). *ARMS. Active Risk Management System: Mid Pilot Evaluation Report*. Unpublished.

Blandford, M., & Sheppard, D. (2015). *Active Risk Management System (ARMS).* Conference Presentation: National Organisation for the Treatment of Abusers (NOTA). Dublin.

**Blasko, B.L., Jeglic, E.L., & Mercado, C.C. (2011). Are actuarial risk data used to make determinations of sex offender risk classification? An examination of sex offenders selected for enhanced registration and notification, *International Journal of Offender Therapy and Comparative Criminology, 55(5),* 676-692. DOI: 10.1177/0306624X10372784

**Boccaccini, M.T., Murrie, D.C., Mercado, C., Quesada, S., Hawes, S., Rice, A.K., & Jeglic, E.L. (2012). Implications of static-99 field reliability findings for score use and reporting, *Criminal Justice and Behavior, 39(1),* 42-58. DOI: 10.1177/0093854811427131

**Boccaccini, M.T., Turner, D.B., Murrie, D.C., Henderson, C.E., & Chevalier, C. (2013). Do Scores From Risk Measures Matter to Jurors? *Psychology Public Policy and Law, 19(2),* 259-269. DOI: 10.1037/a0031354

Boer, D.P. (2013). Some essential environmental ingredients for sex offender reintegration. *International Journal of Behavioral and Consultation Therapy, 8(3-4),* 8-11. DOI: 10.1037/h0100976

Boer, D.P., Beech, A.R., Ward, T., Craig, L.A., Rettenberger, M., Marshall, L., & Marshall, W.L. (2017). *The Wiley handbook on the theories, assessment, and treatment of sexual offending Vols 1-3.* Chichester, West Sussex: John Wiley & Sons, Ltd. ISBN: 978-1-118-57266-5

Boer, D.P., Haaven, J., Lambrick, F., Lindsay, W.R., McVilly, K.R., Sakdalan, J., & Frize,

M. (2012). Manual for the Assessment of Risk and Manageability of Individuals with Developmental and Intellectual Limitations who Offend Sexually (ARMIDILO-S). Author: Available online at www.armidilo.net.

Bonta, J. (2002). Offender risk assessment: Guidelines for selection and use. *Criminal Justice and Behavior, 29*(4), 355-379. DOI: 10.1177/0093854802029004002

Bonta, J., & Andrews, D.A. (2007). *Risk-need-responsivity model for offender assessment and rehabilitation* (User Report No. 2007-06). Ottawa, ON: Public Safety and Emergency Preparedness Canada. ISBN: 978-0-662-05049-0

Bonta, J., & Wormith, S.J. (2008). Risk and need assessment. In. G. McIvor & P. Raynor (Eds.). *Developments in social work with offenders.* (p.131-152). London: Jessica Kingsley Publishers. ISBN: 1843105381

Boyle, G.J. (1991). Does item homogeneity indicate internal consistency or item redundancy in psychometric scales? *Personality and Individual Differences, 12(3),* 291-294. DOI: 10.1016/0191-8869(91)90115-R

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3(2),* 77-101. ISBN: 1-478-0887

Bridges, A., & O'Connor, D. (2010). Restriction and rehabilitation: Getting the right mix. An inspection of the management of sexual offenders in the community. ISBN: 978-1-84099-388-2

**Briken, P., & Műller, J.L. (2014). Beurteilung der Schuldfähigkeit bei paraphiler Störung: Kann der Schweregrad der Störung mithilfe von Kriterien aus Prognoseinstrumenten erfasst werden? *Nervenarzt, 85,* 304-311. DOI: 10.1007/s00115-013-3901-x

British Psychological Society (2012). *Psychological Testing: A User's Guide.* Leicester, UK: BPS.

**Brouillette-Alarie, S., Babchishin, K.M., Hanson, R., & Helmus, L.M. (2016). Latent constructs of the Static-99R and Static-2002R: A three-factor solution, *Assessment, 23(1),* 96-111. DOI: 10.1177/1073191114568114

**Brouillette-Alarie, S., Hanson, R.K., Babchishin, K.M., & Benbouriche,M. (2014). De la prédiction à la compréhension : recension des dimensions psychologiques de la Statique-99*, Pratiques Psychologiques, 20,* 1-19. DOI: 10.1016/j.prps.2013.12.001

*Brouillette-Alarie, S., & Proulx, J. (2013). Predictive validity of the Static-99R and its dimensions. *Journal of Sexual Aggression, 19(3),* 311-328. DOI: 10.1080/13552600.2012.747630

Brown, S. (2011). *Treating sex offenders: An introduction to sex offender treatment programmes.* Oxon, UK: Routledge.

**Brown, S., Harkins, L., Beech, A.R. (2012). General and victim-specific empathy: Associations with actuarial risk, treatment outcome, and sexual recidivism, *Sexual Abuse: Journal of Research and Treatment, 24(5),* 411-430. DOI: 10.1177/1079063211423944

**Buttars, A., Huss, M.T., Brack, C. (2015). Sex offender risk assessment: A re-examination of the coffee can study, *International Journal of Law and Psychiatry, 42-43,* 31-36. DOI: 10.1016/j.ijlp.2015.08.004

**Camilleri, J.A., & Quinsey, V.L. (2011). Appraising the risk of sexual and violent recidivism among intellectually disabled offenders, *Psychology, Crime & Law, 17(1),* 59-74. DOI: 10.1080/10683160903392350

**Caudy, M.S., Durso, J.M., & Taxman, F.S. (2013). How well do dynamic needs predict recidivism? Implications for risk assessment and risk reduction, *Journal of Criminal Justice, 41(6),* 458-466. DOI: 10.1016/j.jcrimjus.2013.08.004

**Chevalier, C.S., Boccaccini, M.T., Murrie, D.C., & Varela, J.G. (2015). Static-99R reporting practices in sexually violent predator cases: Does norm selection reflect adversarial allegiance? *Law and Human Behavior, 39(3),* 209-218. DOI: 10.1037/lhb0000114

Cicchetti, D.V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6,* 284-290. DOI: 1040-3590/94/53.00

**Coid, J.W., Yang, M., Ullrich, S., Zhang, T.Q., Sizmur, S., Farrington, D., & Rogers, R. (2011). Most items in structured risk assessment instruments do not predict violence, *Journal of Forensic Psychiatry & Psychology, 22(1),* 3-21. DOI: 10.1080/14789949.2010.495990

Coid, J., Ying, M., Ullrich, S., Zhang, T., Roberts, A., Roberts, C., Rogers, R., & Farrington, D. (2007). Predicting and Understanding Risk of Re-offending: The Prisoner Cohort Study. Ministry of Justice: Research Summary 6.

Cooke, D.J., & Michie, C. (2010). Limitations of diagnostic precision and predictive utility in the individual case: A challenge for forensic practice. *Law and Human Behavior, 34(4),* 259-274. DOI: 10.1007/s10979-009-9176-x

Cooke, D.J., & Michie, C. (2014). The generalizability of the Risk Matrix 2000: On Model Shrinkage and the misinterpretation of the Area Under the Curve. *Journal of Threat Assessment and Management, 1(1),* 42-55. DOI: 10.1037/tam0000004

College of Policing (2017). Introduction to managing sexual offenders and violent offenders. Available from: https://www.app.college.police.uk/app-content/major-investigation-and-public-protection/managing-sexual-offenders-and-violent-offenders/#training

Cording, J.R., & Beggs Christofferson, S.M. (2017). Theoretical and practical issues for the measurement of protective factors. *Aggression and Violent Behavior, 32,* 45-54. DOI: 10.1016/j.avb.2016.12.007

**Corovic, J., Christianson, S.Ȧ., & Bergman, L.R. (2012). From crime scene actions in stranger rape to prediction of rapist type: Single-victim or serial rapist? *Behavioural Sciences and the Law, 30,* 764-781. DOI: 10.1002/bsl.2026

**Coyle, I.R. (2011). The cogency of risk assessments, *Psychiatry, Psychology and Law, 18(2),* 270-296. DOI: 10.1080/13218719.2010.543406

Craig, L., Beech, A.R., & Browne, K. (2006). Evaluating the predictive accuracy of sex offender risk assessment measures on UK samples: A cross-validation of the Risk Matrix 2000 scales. *Sexual Offender Treatment, 1,* 1-16. (Available online at http://www.sexual-offender-treatment.org/14.0.html).

Craig, L.A., Browne, K.D., & Stringer, I. (2004). Comparing sex offender risk assessment measures on a UK sample. *International Journal of Offender Therapy and Comparative Criminology, 48,* 7-27. DOI: 10.1177/0306624X03257243

**Craissati, J., & Blundell, R. (2013). A Community service for high-risk mentally disordered sex offenders: A follow-up study, *Journal of Interpersonal Violence, 28(6),* 1178-1200. DOI: 10.1177/0886260512468235

Critical Skills Appraisal Programme (2004). Making sense of evidence: 11 questions to help you make sense of a cohort study. Oxfordshire: Public Health Resource Unit.

Critical Skills Appraisal Programme (2006). Making sense of evidence: 12 questions to help you make sense of a cohort study. Oxfordshire: Solutions for Public Health.

**Dahle, K.P., Biedermann, J., Lehmann, R.J., & Gallasch-Nemitz, F. (2014). The development of the Crime Scene Behavior Risk measure for sexual offense recidivism, *Law and Human Behavior, 38(6),* 569-579. DOI: 10.1037/lhb0000088

Darjee, R., & Russell, K. (2012). What clinicians need to know before assessing risk in sexual offenders. *Advances in psychiatric treatment, 18,* 467-478. DOI: 10.1192/apt.bp.110.008094

Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579 (1993).

de Vogel, V., de Ruiter, C., Bouman, Y., & de Vries Robbé, M. (2012). *SAPROF. Guidelines for the Assessment of Protective Factors for Violence Risk. English version 2nd Edition.*Utrecht: Forum Educatief.

de Vogel, V., de Ruiter, C., van Beek, D., & Mead, G. (2004). Predictive validity of the SVR-20 and STATIC-99 in a Dutch sample of treated sex offenders. *Law and Human Behavior, 28(3),* 235-251. DOI: 10.1023/B:LAHU.0000029137.41974.eb

**de Vries Robbé, M., de Vogel, V., & Douglas, K.S. (2013). Risk factors and protective factors: A two-sided dynamic approach to violence risk assessment, *Journal of Forensic Psychiatry & Psychology, 24(4),* 440-457. DOI: 10.1080/14789949.2013.818162

*de Vries Robbé, M., de Vogel, V., Koster, K., & Bogaerts, S. (2015). Assessing protective factors for sexually violent offending with the SAPROF, *Sexual Abuse: A Journal of Research and Treatment, 27(1),* 51-70. DOI: 10.1177/1079063214550168

de Vries Robbé, M., de Vogel, Maruna, S., & Thornton, D. (2015). An exploration of protective factors supporting desistance from sexual offending. *Sexual Abuse: A Journal of Research and Treatment, 27(1),* 16-33. DOI: 10.1177/1079063214547582

**de Vries Robbé, M., de Vogel, V., Wever, E.C., Douglas, K.S., & Nijman, H.L. (2016). Risk and protective factors for inpatient aggression, *Criminal Justice and Behavior, 43(10),* 1364-1385. DOI: 10.1177/0093854816637889

de Vries Robbé, M., Mann, R.E., Maruna, S., & Thornton, D. (2015). An exploration of protective factors supporting desistance from sexual offending. *Sexual Abuse: A Journal of Research and Treatment, 27(1),* 16-33. DOI: 10.1177/107906321454758

**Dickson, S.R., Polaschek, D.L., & Casey, A.R. (2013). Can the quality of high-risk violent prisoners' release plans predict recidivism following intensive rehabilitation? A comparison with risk assessment instruments, *Psychology, Crime & Law, 19(4),* 371-389. DOI: 10.1080/1068316X.2011.640634

**Doyle, M., Carter, S., Shaw, J., & Dolan, M. (2012). Predicting community violence from patients discharged from acute mental health units in England, *Social Psychiatry & Psychiatric Epidemiology, 47(4),* 627-37. DOI 10.1007/s00127-011-0366-8

**Doyle, D.J., Ogloff, J.R.P., & Thomas, S.D.M. (2011). An analysis of dangerous sexual offender assessment reports: Recommendations for best practice, *Psychiatry Psychology and Law, 18(4),* 537-556. DOI: 10.1080/13218719.2010.499159

**Ducro, C., Pham, T.H., Saloppé, X., Chudzicck, L., & Réveillére, C. (2012). E´valuation du risque de re´ cidive d'agresseurs sexuels franc¸ais: Statique-99 et Sorag, *Annales Me´ dico-Psychologiques 170,* 106–108. DOI:10.1016/j.amp.2012.01.001

**Duwe, G. (2012). Using logistic regression modeling to predict sexual recidivism: The Minnesota Sex Offender Screening Tool-3 (MnSOST-3), *Sexual Abuse: Journal of Research and Treatment, 24(4),* 350-377. DOI: 10.1177/1079063211429470

**Duwe, G. (2013). Can Circles of Support and Accountability (COSA) work in the United States? Preliminary results from a randomized experiment in Minnesota, *Sexual Abuse: Journal of Research and Treatment, 25(2),* 143-165. DOI: 10.1177/1079063212453942

**Duwe, G., & Freske, P.J. (2012). Using logistic regression modeling to predict sexual recidivism: The Minnesota Sex Offender Screening Tool-3 (MnSOST-3), *Sexual*

Abuse: A Journal of Research and Treatment, 24(4), 350-377. DOI: 10.1177/1079063211429470

**Edens, J.F., Cox, J., Smith, S.T., DeMatteo, D., & Sorman, K. (2015). How reliable are Psychopathy Checklist-Revised scores in Canadian criminal trials? A case law review, *Psychological Assessment, 27(2),* 447-456. DOI: 10.1037/pas0000048

**Eher, R., Schilling, F., Hansmann, B., Pumberger, T., Nitschke, J., Habermeyer, E., & Mokros, A. (2016). Sadism and violent reoffending in sexual offenders, *Sexual Abuse: A Journal of Research and Treatment, 28(1),* 46-72. DOI: 10.1177/1079063214566715

*Eher, R., Matthes, A., Schilling, F., Haubner-MacLean, T., & Rettenberger, M. (2012). Dynamic risk assessment in sexual offenders using STABLE-2000 and the STABLE-2007: An investigation of predictive and incremental validity. *Sexual Abuse: A Journal of Research and Treatment, 24(1),* 5-28. DOI: 10.1177/1079063211403164

*Eher, R., Olver, M.E., Heurix, I., Schilling, F., & Rettenberger, M. (2015). Predicting reoffense in pedophilic child molesters by clinical diagnoses and risk assessment. *Law and Human Behavior, 39(6),* 571-580. DOI: 10.1037/lhb0000144

Elwood, R.W. (2016). Defining probability in sex offender risk assessment, *International Journal of Offender Therapy and Comparative Criminology, 60(16),* 1328-1941. DOI: 10.1177/0306624X15587912

*Ennis, L., Buro, K., & Jung, S. (2016). Identifying male sexual offender subtypes using cluster analysis and the Static-2002R. *Sexual Abuse: A Journal of Research and Treatment, 28(5),* 403-426. DOI: 10.1177/1079063214527481

Epperson, D.L., Kaul, J.D., & Hesselton, D. (1998). *Final Report of the Development of the Minnesota Sex Offender Screening Tool – Revised (MnSOST-R).* Presented at the

Annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers. Vancouver, Canada.

Eysenck, M.W., & Keane, M.T. (2010). *Cognitive Psychology: A Student's Handbook (6th Edition).* Hove, East Sussex: Psychology Press. ISBN: 978-1-84169-540-2

Falshaw, L., Friendship, C., & Bates, A. (2003). Sexual offenders – measuring reconviction, reoffending and recidivism. Crown Copyright. ISSN: 1473-8406

Farmer, M., McAlinden, A.M., & Maruna, S. (2016). Sex offending and situational motivation: Findings from a qualitative analysis of desistance from sexual offending. *International Journal of Offender Therapy and Comparative Criminology, 60(15),* 1756-1775. DOI: 10.1177/0306624X16668175

**Faust, E., Bickart, W., Renaud, C., & Camp, S. (2015). Child pornography possessors and child contact sex offenders: A multilevel comparison of demographic characteristics and rates of recidivism, *Sexual Abuse: A Journal of Research and Treatment, 27(5),* 460-478. DOI: 10.1177/1079063214521469

*Fedoroff, J., Richards, D., Ranger, R., & Curry, S. (2016). The predictive validity of common risk assessment tools in men with intellectual disabilities and problematic sexual behaviours. *Research in Developmental Disabilities, 57,* 29-38. DOI: 10.1016/j.ridd.2016.06.011

Firestone, P., Nunes, K.L., Moulden, H., Broom, I., & Bradford, J.M. (2005). Hostility and recidivism in sexual offenders. *Archives of Sexual Behavior, 34(3),* 277-283. DOI: 10.1007/s10508-005-3116-8

Fleiss, J.L. (1986). *Design and Analysis of Clinical Experiments*. New York: John Wiley & Sons. ISBN: 9780471349914

Friendship, C., Mann, R.E., & Beech, A.R. (2003). Evaluation of a national prison-based treatment programme for sexual offenders in England and Wales. *Journal of Interpersonal Violence, 18,* 744-759. DOI: 10.1177/0886260503253236

Frye v. United States, 293 F. 1013. (D. C. Cir. 1923). Gigerenzer, G. (2004). Mindless statistics. *Journal of Socio-Economics, 33,* 587–606.

**Fujita, G., Watanabe, K., Yokota, K., Suzuki, M., Wachi, T., Otsuka, Y., & Kuraishi, H. (2016). A multivariate model for analyzing crime scene information: Predicting stranger, solo offender, and money-oriented motive of Japanese homicides, *Homicide Studies, 20(4),* 295-320. DOI: 10.1177/1088767915613828

**Gannon, T.A., & O'Connor, A. (2011). The development of the Interest in Child Molestation Scale, *Sexual Abuse: A Journal of Research and Treatment, 23(4),* 474-493. DOI: 10.1177/1088767915613828

**Gardner, B.O., Boccaccini, M.T., Bitting, B.S., & Edens, J.F. (2015). Personality Assessment Inventory scores as predictors of misconduct, recidivism, and violence: A meta-analytic review, *Psychological Assessment, 27(2),* 534-544. DOI: 10.1037/pas0000065

**Garombo, M.F., Salvadori, S., Contarino, A., Castellino, N., Molinaro, S., Garofano, P., Molo, M.T., Veglia, F., & Rosso, C. (2016). Factorial analysis of risk matrix 2000 on an Italian sample of sexual offenders, *Journal of Sexual Aggression, 22(2),* 137-146. DOI: 10.1080/13552600.2015.1047908

**Giguére, G., & Lussier, P. (2016). Debunking the psychometric properties of the LS/CMI: An application of item response theory with a risk assessment instrument, *Journal of Criminal Justice, 46,* 207-218. DOI: 10.1016/j.jcrimjus.2016.05.005

Gottfredson, S.D., & Moriarty, L.J. (2006). Statistical risk assessment: Old problems and new applications. *Crime & Delinquency, 52(1),* 178-200. DOI: 10.1177/0011128705281748

**Gray, S.R., Abel, G.G., Jordan, A., Garby, T., Wiegel, M., & Harlow, N. (2015). Visual Reaction Time as a predictor of sexual offense recidivism, *Sexual Abuse: A Journal of Research and Treatment, 27(2),* 173-188. DOI: 10.1177/1079063213502680

Gray, R., & Wilcox, D.T. (2015). Standard offender assessment, supervision, and group Treatment. In. D.T. Wilcox, T. Garrett, & L. Harkins (Eds.). *Sex Offender Treatment: A Case Study Approach to Issues and Interventions.* Chichester, West Sussex: John Wiley & Sons Ltd. ISBN: 978-1-118-67440-6

Grubin, D. (1998). *Sex Offending Against Children: Understanding the Risk.* Police Research Series Paper 99, London: Home Office.

Grubin, D. (2008). Validation of Risk Matrix 2000 for Use in Scotland. Unpublished Report prepared for the Risk Management Authority, Paisley, UK.

**Grubin, D. (2011). A large-scale evaluation of risk matrix 2000 in Scotland, *Sexual Abuse: A Journal of Research and Treatment, 23(4),* 419-433. DOI: 10.1177/1079063211411309

**Guay, J.P. (2016). L'évaluation du risque et des besoins criminogènes à la lumière des données probantes: une étude de validation de la version franc¸ aise de l'inventaire de niveau de service et de gestion des cas – LS/CMI, *Revue européenne de psychologie appliquée 66,* 199–210. DOI: 10.1016/j.erap.2016.04.003

**Gutierrez, L., Wilson, H.A., Rugge, T., & Bonta, J. (2013). The prediction of recidivism with Aboriginal offenders: A theoretically informed meta-analysis, *Canadian Journal of Criminology and Criminal Justice, 55(1),* 55-99. DOI:10.3138/cjccj.2011.E.51

Guttman, L. (1968). A general nonmetric technique for finding the smallest coordinate space for a configuration of points. *Psychometrica, 33,* 469–506. DOI: 10.1007/BF02290164

**Hamilton, M. (2015). Risk-needs assessments: Constitutional and ethical challenges, *American Criminal Law Review, 52(2),* 231-291. DOI: 10.2139/ssrn.2506397

Hand, D.J. (2009). Measuring classifier performance: A coherent alternative to the area under the ROC curve. *Machine Learning, 77,* 103-123. DOI: 10.1007/s10994-009-5119-5

**Hannah-Moffat, K. (2016). A conceptual kaleidoscope: Contemplating 'dynamic structural risk' and an uncoupling of risk from need, *Psychology, Crime & Law, 22(1-2),* 33-46. DOI: 10.1080/1068316X.2015.1114115

Hanson, R.K. (1997). *The Development of a Brief Actuarial Risk Scale for Sexual Offense Recidivism*. (User Report 97-04). Ottawa, Canada: Department of the Solicitor General of Canada.

Hanson, R.K. (2006). Does static-99 predict recidivism among older sexual offenders? *Sexual Abuse: A Journal of Research and Treatment, 18,* 343-355. DOI: 10.1007/s11194-006-9027-y

**Hanson, R.K., Babchishin, K.M., Helmus, L., & Thornton, D. (2013). Quantifying the relative risk of sex offenders: Risk ratios for STATIC-99R, *Sexual Abuse: A Journal of Research and Treatment, 25(5),* 482-515. DOI: 10.1177/1079063212469060

Hanson, R.K., Bourgon, G., Helmus, L., & Hodgson, S. (2009). *A Meta-Analysis of the Effectiveness of Treatment for Sex Offenders: Risk, Need, and Responsivity.* Ottawa, Ontario, Canada: Public Safety Canada. ISBN: 978-1-100-11957-1

Hanson, R.K., & Bussiére, M.T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting Clinical Psychology, 66(2),* 348-362. DOI: 10.1037/0022-006X.66.2.348

*Hanson, R.K., Harris, A.J.R., Helmus, L., & Thornton, D. (2014). High-risk sex offenders may not be high risk forever. *Journal of Interpersonal Violence, 29(15),* 2792-2813. DOI: 10.1177/0886260514526062

Hanson, R.K., Harris, A.J.R., Scott, T.L., & Helmus, L. (2007). Assessing the risk of sexual offenders on community supervision: The dynamic supervision project. Public Safety Canada.

*Hanson, R.K., Helmus, L., & Harris, A.J.R. (2015). Assessing the risk and needs of supervised sexual offenders: A prospective study using STABLE-2007, Static-99R, and Static-2002R. *Criminal Justice and Behavior, 42(12),* 1205-1224. DOI: 10.1177/0093854815602094

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

Hanson, R.K., & Thornton, D.M. (2000). Improving risk assessments for sex offenders: A comparison of three actuarial scales. *Law and Human Behaviour, 24(1),* 119-136. DOI: 10.1023/A:1005482921333

Hanson, R. K., & Thornton, D. (2003). *Notes on the development of Static–2002* (Corrections Research User Report 2003–01). Ottawa, Ontario, Canada: Department of the Solicitor General of Canada.

Hanson, R.K., & Yates, P.M. (2013). Psychological treatment of sex offenders. *Current psychiatry reports. 15(3).* 348. DOI: 10.1007/s11920-012-0348

**Harkins, L., Howard, P. Barnett, G. Wakeling, H., & Miles, C. (2015). Relationships between denial, risk, and recidivism in sexual offenders, *Archives of Sexual Behavior, 44,* 157-166. DOI 10.1007/s10508-014-0333-z

**Harris, G.T., & Rice, M.E. (2015). Progress in violence risk assessment and communication: Hypothesis versus evidence, *Behavioral Sciences and the Law, 33,* 128-145. DOI: 10.1002/bsl.2157

Harris, G.T., Rice, M.E., Quinsey, V.L., & Cormier, C.A. (2015). *Violent offenders: Appraising and managing risk (3rd ed.)*. Washington, DC: American Psychological Association. ISBN: 978-1-4338-1901-8

Harris, A.J.R., & Tough, S. (2004). Should actuarial risk assessments be used with sex offenders who are intellectually disabled? *Journal of Applied Research in Intellectual Disabilities, 17,* 235-241. DOI: 10.1111/j.1468-3148.2004.00211

Hart, S.D., & Boer, D.P. (2010). Structured professional guidelines for sexual violence risk assessment: The Sexual Violence Risk-20 (SVR-20) and Risk for Sexual Violence Protocol (RSVP). In. R.K. Otto, & K.S. Douglas (Eds.). *Handbook of Violence Risk Assessment.* New York, USA: Taylor & Francis Group. ISBN: 978-1-138-87288-2

**Hart, S.D., & Cooke, D.J. (2013). Another look at the (Im-)precision of individual risk estimates made using actuarial risk assessment instruments, *Behavioral Sciences and the Law, 31(1),* 81-102. DOI: 10.1002/bsl.2049

Hart, S.D., Kropp, P.R., Laws, D.R., Klaver, J., Logan, C., & Watt, K.A. (2003). The Risk for Sexual Violence Protocol (RSVP): Structured professional guidelines for assessing risk of sexual violence. Burnaby, Canada: Mental Health, Law, & Policy Institute, Simon Fraser University.

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, 140(39),* 139-170. DOI: 10.1192/bjp.190.5.s60

**Hawes, S.W., Boccaccini, M.T., & Murrie, D.C. (2013). Psychopathy and the combination of psychopathy and sexual deviance as predictors of sexual recidivism: Meta-analytic

findings using the Psychopathy Checklist-Revised, *Psychological Assessment, 25(1),* 233-243. DOI: 10.1037/a0030391

**Hecker, J.E. (2014). Baby with the bath water: Response to Fanniff and Letourneau, *Sexual Abuse: A Journal of Research and Treatment, 26(5),* 395-400. DOI: 10.1177/1079063214525644

Helmus, L., & Babchishin, K.M. (2017). Primer on risk assessment and the statistics used to evaluate its accuracy, *Criminal Justice and Behavior, 44(1),* 8-25. DOI: 10.1177/0093854816678898

*Helmus, L., Babchishin, K.M., & Blais, J. (2012). Predictive accuracy of dynamic risk factors for Aboriginal and non-Aboriginal sex offenders: An exploratory comparison using STABLE-2007. *International Journal of Offender Therapy and Comparative Criminology, 56(6),* 856-876. DOI: 10.1177/0306624X11414693

Helmus, L., Babchishin, R., & Hanson, R. (2013). The predictive accuracy of the RM2000: A meta-analysis. *Sexual Offender Treatment, 8(2).* (Available online at: http://www.sexual-offender-treatment.org/125.html).

*Helmus, L., Ciardha C., & Seto, M.C. (2015). The Screening Scale for Pedophilic Interests (SSPI): construct, predictive, and incremental validity. *Law & Human Behavior, 39(1),* 35-43. DOI: 10.1037/lhb0000099

**Helmus, L., Hanson, R.K., Babchishin, K.M., & Mann, R.E. (2013). Attitudes supportive of sexual offending predict recidivism: A meta-analysis, *Trauma Violence & Abuse, 14(1),* 34-53. DOI: 10.1177/1524838012462244

**Helmus, L., Hanson, R.K., Thornton, D., Babchishin, K.M., & Harris, A.J.R. (2012). Absolute recidivism rates predicted by STATIC-99R and STATIC-2002R sex offender

risk assessment tools vary across samples: A meta-analysis, *Criminal Justice and Behavior, 39(9),* 1148-1171. DOI: 10.1177/0093854812443648

**Helmus, L., & Thornton, D. (2015). Stability and predictive and incremental accuracy of the individual items of STATIC-99R and STATIC-2002R in predicting sexual recidivism: A meta-analysis, *Criminal Justice and Behavior, 42(9),* 917-937. DOI: 10.1177/0093854814568891

*Helmus, L., Thornton, D., Hanson, R., & Babchishin, K.M. (2012a). Improving the predictive accuracy of Static-99 and Static-2002 with older sex offenders: Revised age weights. *Sexual Abuse: Journal of Research and Treatment, 24(1),* 64-101. DOI: 10.1177/1079063211409951

*Helmus, L., Hanson, R.K., Babchishin, K.M., & Thornton, D. (2015a). Sex offender risk assessment with the Risk Matrix 2000: Validation and guidelines for combining with the STABLE-2007. *Journal of Sexual Aggression, 21(2),* 136-157. DOI: 10.1080/13552600.2013.870241

**Hempel, I.S., Buck, N.M., Goethals, K.R., & van Marle, H.J. (2013). Unraveling sexual associations in contact and noncontact child sex offenders using the single category - implicit association test, *Sexual Abuse: A Journal of Research & Treatment, 25(5),* 444-60. DOI: 10.1177/1079063212464660

**Hendry, M.C., Douglas, K.S., Winter, E.A., Edens, J.F. (2013). Construct measurement quality improves predictive accuracy in violence risk assessment: An illustration using the Personality Assessment Inventory, *Behavioral Sciences & the Law, 31(4),* 477-493. DOI: 10.1002/bsl.2066

**Higgins, R.Q. (2012). The prediction of violent sexual recidivism among sexual offenders: A review of clinicians' and forensic professionals' predictions, *Dissertation Abstracts International: Section B: The Sciences and Engineering, 74(2-B(E)).*

Hill, P.L., & Jackson, J.J. (2016). The invest-and-accrue model of conscientiousness. *Review of General Psychology, 20(2),* 141-154. DOI: 10.1037/gpr0000065

*Hill, A., Rettenberger, M., Habermann, N., Berner, W., Eher, R., & Briken, P. (2012). The utility of risk assessment instruments for the prediction of recidivism in sexual homicide perpetrators. *Journal of Interpersonal Violence, 27(18),* 3553-3578. DOI: 10.1177/0886260512447570

**Hockenhull, J.C., Whittington, R., Leitner, M., Barr, W., McGuire, J., Cherry, M.G., Flentje, R., Quinn, B., Dundar, Y., & Dickson, R. (2013). A systematic review of prevention and intervention strategies for populations at high risk of engaging in violent behaviour: Update 2002-8, *Health Technology Assessment, 16(3),* 1-152. DOI: 10.3310/hta16030

**Howard, P.D., Barnett, G.D., Wakeling, H.C. (2015). Predicting nonsexual violent reoffending by sexual offenders: A comparison of four actuarial tools, *Legal and Criminological Psychology, 20(2),* 267-287. DOI:10.1111/lcrp.12027

Hulley, J. (2016). "While this does not in any way excuse my conduct . . .". The role of treatment and neutralizations in desistance from sexual offending. *International Journal of Offender Therapy and Comparative Criminology, 60(14),* 1776-1790.

IBM Corp (2013). IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.

Ireland, C., & Craig, L.A. (2011). Adult sexual offender assessment. In. D.P. Boer, R. Eher, L.A. Craig, M.H. Miner, & F., Pfafflin (Eds.). International Perspectives on the Assessment and Treatment of Sexual Offenders. Chichester, West Sussex: John Wiley & Sons Ltd. ISBN: 978-1-1190-4614-1

**Janka, C., Gallasch-Nemitz, F., Biedermann, J., & Dahle, K.P. (2012). The significance of offending behavior for predicting sexual recidivism among sex offenders of various age

groups, *International Journal of Law and Psychiatry, 35(3),* 159-164. DOI:10.1016/j.ijlp.2012.02.002

**Jung, S., Daniels, M.K., Friesen, M., & Ledi, D. (2012). An examination of convergent constructs among Level of Service measures and other measures, *Journal of Forensic Psychiatry & Psychology, 23(5-6),* 601-619. DOI: 10.1080/14789949.2012.732595

**Jung, S., Pham, A., & Ennis, L. (2013). Measuring the disparity of categorical risk among various sex offender risk assessment measure, *The Journal of Forensic Psychiatry and Psychology, 24(3),* 353-370. DOI: 10.1080/14789949.2013.806567

Kewley, S., & Blandford, M. (2017). The development of the active risk management system. *Journal of Criminal Psychology, 7(3),* 155-167. DOI: 10.1108/JCP-10-2016-0034

Kewley, S., Larkins, M., Harkins, L., & Beech, A.R. (2017). Restoring identity: The use of religion as a mechanism to transition between an identify of sexual offending to a non-offending identity. *Criminology & Criminal Justice, 17(1),* 79-96.

**Kiland, K.L. (2016). Ethnicity weighting adjustment for actuarial measure of sex offender assessment: STATIC-99R, *PROQUEST LLC.*

**Kingree, J.B., & Thompson, M. (2015). A comparison of risk factors for alcohol-involved and alcohol-uninvolved sexual aggression perpetration, *Journal of Interpersonal Violence, 30(9),* 1478-1492. DOI: 10.1177/0886260514540806

**Kingston, D.A., Olver, M.E., Harris, M., Wong, S.C.P., & Bradford, J.M. (2015). The relationship between mental disorder and recidivism in sexual offenders, *International Journal of Forensic Mental Health, 14(1),* 10-22. DOI: 10.1080/14999013.2014.974088

**Kingston, D.A., Yates, P.M., & Firestone, P. (2012). The self-regulation model of sexual offending: Relationship to risk and need, *Law and human behavior, 36(3),* 215-224. DOI: 10.1007/s10979-011-9287-z

Kingston, D.A., Yates, P.M., Firestone, P., Babchishin, K., & Bradford, J.M. (2008). Long term predictive validity of the Risk Matrix 2000: A comparison with the STATIC-99 and the Sex Offender Risk Appraisal Guide. *Sexual Abuse: A Journal of Research and Treatment, 20(4),* 466-484. DOI: 10.1177/1079063208325206

**Kingston, D.A., Yates, P.M., & Olver, M.E. (2014). The self-regulation model of sexual offending: Intermediate outcomes and posttreatment recidivism, *Sexual Abuse: A Journal of Research and Treatment, 26(5),* 429-449. DOI: 10.1177/1079063213495896

Klassen, D., & O'Connor, W. (1988). A prospective study of predictors of violence in adult male mental patients. *Law and Human Behavior, 12(2),* 143-158. DOI: 10.1007/BF01073122

**Kleban, H., Chesin, M.S., Jeglic, E.L., & Mercado, C.C. (2013). An exploration of crossover sexual offending, *Sexual Abuse-a Journal of Research and Treatment, 25(5),* 427-443. DOI: 10.1177/1079063212464397

Kline, P, (1999). *Handbook of Psychological Testing (Second Edition).* London: Routledge. ISBN: 0415211581

**Krauss, D.A., & Scurich, N. (2013). Risk assessment in the law: Legal admissibility, scientific validity, and some disparities between research and practice, *Behavioral Sciences and the Law, 31,* 215-229. DOI: 10.1002/bsl.2065

**Larcombe, W. (2012). Sex offender risk assessment: The need to place recidivism research in the context of attrition in the Criminal Justice System, *Violence Against Women, 18(4),* 482-501. DOI: 10.1177/1077801212452249

**Lasher, M.P., McGrath, R.J., & Cumming, G.F. (2015). Sex offender modus operandi stability and relationship with actuarial risk assessment, *Journal of Interpersonal Violence, 30(6),* 911-927. DOI: 10.1177/0886260514539757

Laws, D.R., & O'Donohue, W. (2016). *Treatment of sex offenders: Strengths and weaknesses in assessment and intervention.* London: Springer Publishing. DOI: 10.1007/978-3-319-25868-3

*Lee, S.C., & Hanson, R.K. (2016). Recidivism risk factors are correlated with a history of psychiatric hospitalization amongst sex offenders. *Psychological Services, 13(3),* 261-271. DOI: 10.1037/ser0000081

**Lee, A.F., Li, N.C., Lamade, R., Schuler, A., & Prentky, R.A. (2012). Predicting hands-on child sexual offenses among possessors of Internet child pornography, *Psychology, Public Policy, and Law, 18(4),* 644-672. DOI: 10.1037/a0027517

**Lehmann, R.J., Goodwill, A.M., Gallasch-Nemitz, F., Biedermann, J., & Dahle, K.P. (2013). Applying crime scene analysis to the prediction of sexual recidivism in stranger rapes, *Law and Human Behavior, 37(4),* 241-254. DOI: 10.1037/lhb0000015

**Lehmann, R.J., Goodwill, A.M., Hanson, R., & Dahle, K.P. (2016). Acquaintance rape: Applying crime scene analysis to the prediction of sexual recidivism, *Sexual Abuse: Journal of Research and Treatment, 28(7),* 679-702. DOI: 10.1177/1079063215569542

*Lehmann, R..J., Hanson, R.K., Babchishin, K.M., Gallasch-Nemitz, F., Biedermann, J., & Dahle, K.P. (2013). Interpreting multiple risk scales for sex offenders: evidence for averaging. *Psychological Assessment, 25(3),* 1019-24. DOI: 10.1037/a0033098

**Lewis, K., Olver, M.E., & Wong, S.C. (2013). The Violence Risk Scale: Predictive validity and linking changes in risk with violent recidivism in a sample of high-risk offenders

with psychopathic traits, *Assessment, 20(2),* 150-164. DOI: 10.1177/1073191112441242

Leue, A., Brocke, B., & Hoyer, J. (2008). Reinforcement sensitivity of sex offenders and non-offenders: An experimental and psychometric study of reinforcement sensitivity theory. *British Journal of Psychology, 99,* 361-378. DOI: 10.1348/000712607X228519

Limm, H., Heinmuller, M., Gundell, H., Liel, K., Seeger, K., Salman, R., & Angerer, P. (2015). Effects of a health promotion program based on a train-the-trainer approach on quality of life and mental health of long-term unemployed persons. *Biomed Research International.* DOI: 10.1155/2015/719327

Lingoes, J.C., & Guttman, L. (1973). *The Guttman-Lingoes nonmetric program series.* Ann Arbor, Michigan: Mathesis Press.

**Lofthouse, R.E., Lindsay, W.R., Totsika, V., Hastings, R.P., Boer, D.P., & Haaven, J.L. (2013). Prospective dynamic assessment of risk of sexual reoffending in individuals with an intellectual disability and a history of sexual offending behaviour, *Journal of Applied Research in Intellectual Disabilities, 26(5),* 394-403. DOI: 10.1111/jar.12029

Logan, C. (2016). Risk formulation: The new frontier in risk assessment and management. In. D.R. Laws, & W. O'Donohue (Eds.). *Treatment of sex offenders: Strengths and weaknesses in assessment and intervention.* London: Springer Publishing. DOI: 10.1007/978-3-319-25868-3

*Looman, J., Morphett, N.A., & Abracen, J. (2012). Does consideration of psychopathy and sexual deviance add to the predictive validity of the static-99R? *International Journal of Offender Therapy & Comparative Criminology, 57(8),* 939-65. DOI: 10.1177/0306624X12444839

Lussier, P., & Davies, G. (2011). Person-oriented perspective on sexual offenders, offending trajectories, and risk of recidivism: A new challenge for policymakers, risk assessors, and actuarial prediction? *Psychology, Public Policy, and Law, 17(4),* 530-561. DOI: 10.1037/a0024388

Mann, R.E., Hanson, R.K., & Thornton, D. (2010). Assessing risk for sexual recidivism: Some proposals on the nature of psychologically meaningful risk factors. *Sexual Abuse: A Journal of Research and Treatment, 22(2),* 191-217. DOI: 10.1177/1079063210366039

Marshall, L.E., & Marshall, W.L. (2017). Motivating sex offenders to enter and effectively engage in treatment. In. D.T. Wilcox, M.L. Donathy, R. Gray & C. Baim (Eds.). *Working with sex offenders: A guide for practitioners.* Abingdon, Oxon: Routledge. ISBN: 978-1-138-93354-5

Marshall, W.L., Marshall, L.E., & Olver, M.E. (2017). An evaluation of strength-based approaches to the treatment of sex offenders: A review. *Journal of Criminal Psychology, 7(3),* 221-228. DOI: 10.1108/JCP-04-2017-0021

Marshall, W.L., Marshall, L.E., Serran, G.A., & O'Brien, M.D. (2011). *Rehabilitating sexual offenders: A strength-based approach.* ISBN: 978-1-4338-0942-2

Maruna, S., & LeBel, T.P. (2003). Welcome home? Examining the "Reentry Court" concept from a strengths-based perspective. *Western Criminology Review, 4(2),* 91-107. ISSN: 1096-4886

**McDougall, C., Pearson, D.A., Willoughby, H., & Bowles, R.A. (2013). Evaluation of the ADViSOR project: Cross-situational monitoring of high-risk offenders in prison and the community, *Legal and Criminological Psychology, 18(2),* 205-228. DOI:10.1111/j.2044-8333.2012.02059.x

*McGrath, R.J., Lasher, M.P., & Cumming, G.F. (2012). The Sex Offender Treatment Intervention and Progress Scale (SOTIPS): Psychometric properties and incremental predictive validity with Static-99R. *Sexual Abuse: A Journal of Research and Treatment, 24(5),* 431-458. DOI: 10.1177/1079063211432475

*McGrath, R.J., Lasher, M.P., Cumming, G.F., Langton, C.M., & Hoke, S.E. (2014). Development of Vermont assessment of sex offender risk-2 (VASOR-2) reoffense risk scale. *Sexual Abuse: A Journal of Research and Treatment, 26(3),* 271-290. DOI: 10.1177/1079063213486936

McKay, K., Kuntz, J.R.C., & Näswall, K. (2013). The effect of affective commitment, communication and participation on resistance to change: The role of change readiness. *New Zealand Journal of Psychology, 42(1),* 55-66.

**McNally, M.R., & Fremouw, W.J. (2014). Examining risk of escalation: A critical review of the exhibitionistic behavior literature, *Aggression and Violent Behavior, 19,* 474-485. DOI: 10.1016/j.avb.2014.07.001

**McPhail, I.V., Hermann, C.A., & Fernandez, Y.M. (2014). Correlates of emotional congruence with children in sexual offenders against children: A test of theoretical models in an incarcerated sample, *Child Abuse and Neglect, 38(2),* 336-346. DOI: 10.1016/j.chiabu.2013.10.002

Mews, A., Di Bella, L, & Purver, M. (2017). *Impact evaluation of the prison-based Core Sex Offender Treatment Programme.* Ministry of Justice: Crown Copyright. ISBN: 978-1-84099-783-5

**Miller, C.S., Kimonis, E.R., Otto, R.K., Kline, S.M., & Wasserman, A.L. (2012). Reliability of risk assessment measures used in sexually violent predator proceedings, *Psychological Assessment, 24(4),* 944-953. DOI: 10.1037/a0028411

Mingus, W., & Burchfield, K.B. (2012). From prison to integration: Applying modified labeling theory to sex offenders. *Criminal Justice Studies, 25(1),* 97-109. DOI: 10.1080/1478601X.2012.657906

Ministry of Justice (MoJ: 2012). MAPPA Guidance: Version 4.

Ministry of Justice (MoJ: 2016). Multi-Agency Public Protection Arrangements Annual Report 2015/16. Ministry of Justice Statistics Bulletin.

Moffitt, T. (1993). 'Life-course persistent' and 'adolescent-limited' antisocial behaviour: A developmental taxonomy. *Psychological Review, 100,* 674-701. DOI: 10.1037/0033-295X.100.4.674

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Journal of Clinical Epidemiology, 62(10),* 1006–1012. DOI: 10.1016/j.jclinepi.2009.06.005

Moher, D., Tsertsvadze, A., Tricco, A., Eccles, M., Grimshaw, J., Sampson, M., & Barrowman, N. (2008). When and how to update systematic reviews. *Cochrane Database of Systematic Reviews, Issue 1.* DOI: 10.1002/14651858.MR000023.pub3

**Mokros, A., Gebhard, M., Heinz, V., Marschall, R.W., Nitschke, J., Glasgow, D.V., Gress, C.L.Z., & Laws, D.R. (2013). Computerized assessment of pedophilic sexual interest through self-report and viewing time: Reliability, validity, and classification accuracy of the Affinity Program, *Sexual Abuse: Journal of Research and Treatment, 25(3),* 230-258. DOI: 10.1177/1079063212454550

*Montana, S., Thompson, G., Ellsworth, P., Lagan, H., Helmus, L., & Rhoades, C.J. (2012). Predicting relapse for Catholic clergy sex offenders: The use of the Static-99. *Sexual Abuse: Journal of Research and Treatment, 24(6),* 575-590. DOI: 10.1177/1079063212445570

**Mooney, J.L., & Daffern, M. (2013). The Offence Analogue and Offence Reduction Behaviour Rating Guide as a supplement to violence risk assessment in incarcerated offenders, *The International Journal of Forensic Mental Health, 12(4),* 255-264. DOI: 10.1080/14999013.2013.867421

Nash, M.R. (2016). 'Scum cuddlers': Police offender managers and the sex offenders' register in England and Wales. *Policing and Society, 26(4),* 411-427. DOI: 10.1080/10439463.2014.842855

Neal, T.M.S., & Grisso, T. (2014). Assessment practices and expert judgment methods in forensic psychology and psychiatry: An international snapshot. *Criminal Justice and Behavior, 41(12),* 1406-1421. DOI: 10.1177/0093854814548449

**Neller, D.J., & Frederick, R.I. (2013). Classification accuracy of actuarial risk assessment instruments, *Behavioral Sciences & the Law, 31(1),* 141-153. DOI: 10.1002/bsl.2047

**Neller, D.J., & Petris, G. (2013). Sexually violent predators: Toward reasonable estimates of recidivism base rates, *Behavioral Sciences & the Law, 31(4),* 429-43. DOI: 10.1002/bsl.2072

*Nicholaichuk, T.P., Olver, M.E., Gu, D.Q., & Wong, S.C.P. (2014). Age, actuarial risk, and long-term recidivism in a national sample of sex offenders. *Sexual Abuse: A Journal of Research and Treatment, 26(5),* 406-428. DOI: 10.1177/1079063213492340

Nicholls, C..M., & Webster, S. (2014). *Sex offender management and dynamic risk: Pilot evaluation of the Active Risk Management System (ARMS).* National Offender Management Service (NOMS). Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/308159/sex-offender-management-and-dynamic-risk.pdf

**Nilsson, T., Wallinius, M., Gustavson, C., Anckarsater, H., & Kerekes, N. (2011). Violent recidivism: A long-time follow-up study of mentally disordered offenders, PLoS ONE, 6 (10): e25768. DOI: 10.1371/journal.pone.0025768

**Nunes, K.L., & Babchishin, K.M. (2012). Construct validity of Stable-2000 and Stable-2007 scores, *Sexual Abuse: A Journal of Research and Treatment, 24(1),* 29-45. DOI: 10.1177/1079063211404921

*Nunes, K.L., Hermann, C.A., Malcom, J.R., & Lavoie, K. (2013). Childhood sexual victimization, pedophilic interest, and sexual recidivism. *Child Abuse & Neglect, 37(9),* 703-711. DOI: 10.1016/j.chiabu.2013.01.008

*Nunes, K.L., Pettersen, C., Hermann, C.A., Looman, J., & Spape, J. (2016). Does change on the MOLEST and RAPE scales predict sexual recidivism? *Sexual Abuse: A Journal of Research and Treatment, 28(5),* 427-447. DOI: 10.1177/1079063214540725

Nunnally, J.C., & Bernstein, I.H. (1994). *Psychometric Theory (Third Edition).* New York, USA: McGraw-Hill Inc. ISBN: 0-07-047849-X

*Olver, M.E., Christofferson, S.M.B., Grace, R.C., & Wong, S.C.P. (2014b). Incorporating change information into sexual offender risk assessments using the Violence Risk Scale-Sexual Offender version. *Sexual Abuse-a Journal of Research and Treatment, 26(5),* 472-499. DOI: 10.1177/1079063213502679

**Olver, M.E., Kingston, D.A., Nicholaichuk, T.P., & Wong, S.C. (2014). A psychometric examination of treatment change in a multisite sample of treated Canadian federal sexual offenders, *Law & Human Behavior, 38(6),* 544-59. DOI: 10.1037/lhb0000086

*Olver, M.E., Klepfisz, G., Stockdale, K.C., Kingston, D.A., Nicholaichuk, T.P., & Wong, S.C. (2016). Some notes on the validation of VRS-SO static scores. *Journal of Sexual Aggression, 22(2),* 147-160. DOI: 10.1080/13552600.2015.1116625

*Olver, M.E., Nicholaichuk, T.P., Kingston, D.A., & Wong, S.C. (2014). A multisite examination of sexual violence risk and therapeutic change. *Journal of Consulting & Clinical Psychology, 82(2),* 312-24. DOI: 10.1037/a0035340

**Olver, M.E., Nicholaichuk, T.P., Gu, D., & Wong, S.C.P. (2013). Sex offender treatment outcome, actuarial risk, and the aging sex offender in Canadian corrections: A long-term follow-up, *Sexual Abuse: Journal of Research and Treatment, 25(4),* 396-422. DOI: 10.1177/1079063212464399

*Olver, M.E., Nicholaichuk, T.P., & Wong, S.C. (2014a). The predictive and convergent validity of a psychometric battery used to assess sexual offenders in a treatment programme: An 18-year follow-up. *Journal of Sexual Aggression, 20(2),* 216-239. DOI: 10.1080/13552600.2013.816791

*Olver, M.E., & Wong, S.C.P. (2011). A comparison of static and dynamic assessment of sexual offender risk and need in a treatment context. *Criminal Justice and Behavior, 38(2),* 113-126. DOI: 10.1177/0093854810389534

**Olver, M.E., & Wong, S. (2011a). Predictors of sex offender treatment dropout: Psychopathy, sex offender risk, and responsivity implications, *Psychology Crime & Law, 17(5),* 457-471. DOI: 10.1080/10683160903318876

**Olver, M.E., & Wong, S.C.P. (2013). Treatment programs for high risk sexual offenders: Program and offender characteristics, attrition, treatment change and recidivism, *Aggression and Violent Behavior, 18,* 579-591. DOI: 10.1016/j.avb.2013.06.002

**Ostermann, M., & Salerno, L.M. (2016). The validity of the Level of Service Inventory-Revised at the intersection of race and gender, *The Prison Journal, 96(4),* 554-575. DOI: 10.1177/0032885516650878

*Parent, G., Guay, J.P., & Knight, R.A. (2012). Can we do better? The assessment of risk of recidivism by adult sex offenders. *Criminal Justice and Behavior, 39(12),* 1647-1667. DOI: 10.1177/0093854812451680

**Pettersen, C., Nunes, K.L., & Cortoni, F. (2016). Does the factor structure of the Aggression Questionnaire hold for sexual offenders? *Criminal Justice and Behavior, 43(6),* 811-829. DOI: 10.1177/0093854815618692

Pike, R. (2015). *Risk Matrix 2000 (RM2000).* Presentation: University of Birmingham, Spring.

Plomin, R., Shakeshaft, N.G., McMillan, A., & Trzaskowski, M., 2014). Nature, nurture, and expertise. *Intelligence, 45( ),* 46-59. DOI: 10.1016/j.intell.2013.06.008

**Quesada, S.P., Calkins, C., & Jeglic, E.L. (2014). An Examination of the interrater reliability between practitioners and researchers on the Static-99, *International Journal of Offender Therapy and Comparative Criminology, 58(11),* 1364-1375. DOI: 10.1177/0306624X13495504

**Quinn, R., Miles, H., & Kinane, C. (2013).The validity of the Short-Term Assessment of Risk and Treatability (START) in a UK medium secure forensic mental health setting, *International Journal of Forensic Mental Health, 12(3),* 215-224. DOI: 10.1080/14999013.2013.832714

Quinsey, V.L., Harris, G.T., Rice, M.E., & Cormier, C.A. (1998). *Violent Offenders: Appraising and Managing Risk.* Washington, DC: American Psychological Society. ISBN: 978-1-55798-495-1

Quinsey, V.L., Harris, G.T., Rice, M.E., & Cormier, C.A. (2006). *Violent Offenders: Appraising and Managing Risk* (2nd ed.). Washington, DC: American Psychological Association. ISBN: 978-1-59147-343-5

**Ragusa-Salerno, L.M., Ostermann, M., & Thomas, S.S. (2013). Does the LSI-R have utility for sex offenders? *Criminal Justice and Behavior, 40(9),* 952-969. DOI: 10.1177/0093854813481667

**Rettenberger, M., & Eher, R. (2013). Actuarial risk assessment in sexually motivated intimate-partner violence, *Law and Human Behavior, 37(2),* 75-86. DOI: 10.1037/b0000001

*Rettenberger, M., Haubner-Maclean, T., & Eher, R. (2013). The contribution of age to the Static-99 risk assessment in a population-based prison sample of sexual offenders. *Criminal Justice and Behavior, 40(12),* 1413-1433. DOI: 10.1177/0093854813492518

Rice, M.E., & Harris, G.T. (2005). Comparing effect sizes in follow-up studies: ROC area, Cohen's d, and r, *Law and Human Behavior, 29,* 615-620. DOI: 10.1007/s10979-005-6832-7

**Rice, M.E., & Harris, G.T. (2014). What does it mean when age is related to recidivism among sex offenders? *Law and Human Behavior, 38(2),* 151-161. DOI: 10.1037/lhb0000052

**Rice, M.E., Harris, G.T., & Lang, C. (2013). Validation and revision to the VRAG and SORAG: The Violence Risk Appraisal Guide-Revised (VRAG-R), *Psychological Assessment, 25(3),* 951-965. DOI: 10.1037/a0032878

**Rocque, M., & Plummer-Beale, J. (2014). In the eye of the beholder? An examination of the inter-rater reliability of the LSI-R and YLS/CMI in a correctional agency, *Journal of Criminal Justice, 42(6),* 568-578. DOI: 10.1016/j.jcrimjus.2014.09.011

Rogers, R. (2000). The uncritical acceptance of risk assessment in forensic practice. *Law and Human Behavior, 24(5),* 595-605. DOI: 10.1023/A:1005575113507

**Romine, R.E., Miner, M.H., Poulin, D., Dwyer, S.M., & Berg, D. (2012). Predicting reoffense for community-based sexual offenders: An analysis of 30 years of data,

Sexual Abuse: A Journal of Research and Treatment, 24(5), 501-514. DOI: 10.1177/1079063212446514

**Rossegger, A., Gerth, J., Seewald, K., Urbaniok, F., Sing, J.P., & Endrass, J. (2013). Current obstacles in replicating risk assessment findings: A systematic review of commonly used actuarial instruments, *Behavioral Sciences and the Law, 31,* 154-164. DOI: 10.1002/bsl.2044

**Ryan, J., Wilson, N., Kilgour, G., & Reynolds, N. (2014). Balancing efficiency and accuracy: Guidelines for the psychological assessment of extended supervision orders for child sex offenders in New Zealand, *Psychiatry, Psychology and Law, 21(1),* 139-157. DOI: 10.1080/13218719.2013.793154

**Sandler, J.C., Freeman, N.J., Farrelland, P., & Seto, M.C. (2013). Predicting institutional sexual misconduct by adult male sex offenders, *Criminal Justice and Behavior, 40(10),* 1131-1148. DOI: 10.1177/0093854813485574

**Schmidt, A.F., Gykiere, K., Vanhoeck, K., Mann, R.E., & Banse, R. (2014). Direct and indirect measures of sexual maturity preferences differentiate subtypes of child sexual abusers, *Sexual Abuse: A Journal of Research and Treatment, 26(2),* 107-128. DOI: 10.1177/1079063213480817

**Schmucker, M., & Losel, F. (2015). The effects of sexual offender treatment on recidivism: An international meta-analysis of sound quality evaluations, *Journal of Experimental Criminology, 11(4),* 597-630. DOI 10.1007/s11292-015-9241-z

*Seto, M.C., & Eke, A.W. (2015). Predicting recidivism among adult male child pornography offenders: Development of the Child Pornography Offender Risk Tool (CPORT). *Law and Human Behavior, 39(4),* 416-429. DOI: 10.1037/lhb0000128

**Seto, M.C., & Fernandez, Y.M., (2011). Dynamic risk groups among adult male sexual offenders, *Sexual Abuse: A Journal of Research and Treatment, 23(4),* 494-507. DOI: 10.1177/1079063211403162

**Seto, M.C., Hanson, R.K., & Babchishin, K.M. (2011). Contact sexual offending by men with online sexual offenses, *Sexual Abuse: A Journal of Research and Treatment, 23(1),* 124-145. DOI: 10.1177/1079063210369013

Sex Offenders Act (1997). London, UK: The Stationery Office Limited. ISBN: 0-10-545197-5

Shye, S. (1978), Theory Construction and Data Analysis in the Behavioral Sciences. San Francisco, USA: Jossey-Bass Inc. ISBN: 0875893791

**Singh, J.P., Grann, M., & Fazel, S. (2011). A comparative study of violence risk assessment tools: A systematic review and metaregression analysis of 68 studies involving 25,980 participants, *Clinical Psychology Review, 31(3),* 499-513. DOI:10.1016/j.cpr.2010.11.009

**Singh, J.P., Grann, M., Lichtenstein, P., Langstrom, N., & Fazel, S. (2012). A novel approach to determining violence risk in schizophrenia: Developing a stepped strategy in 13,806 discharged patients, *Plos One, 7(2): e31727.* DOI: 10.1371/journal.pone.0031727

Skelton, A., & Vess, J. (2008). Risk of sexual recidivism as a function of age and actuarial risk. *Journal of Sexual Aggression, 14(3),* 199-209. DOI: 10.1080/13552600802267098

Smallbone, S., Marshall, W.L., & Wortley, J. (2008). *Preventing child sexual abuse: Evidence, policy and practice.* Abington, Oxon: Routledge. ISBN: 978-1-84392-222-3

*Smallbone, S., & Rallings, M. (2013). Short-term predictive validity of the Static-99 and Static-99-R for indigenous and nonindigenous Australian sexual offenders. *Sexual*

Abuse: Journal of Research and Treatment, 25(3), 302-316. DOI: 10.1177/1079063212472937

**Smid, W.J., Kamphuis, J.H., Wever, E.C., & Van Beek, D. (2013). Treatment referral for sex offenders based on clinical judgment versus actuarial risk assessment: Match and analysis of mismatch, *Journal of Interpersonal Violence, 28(1),* 2273-2289. DOI: 10.1177/0886260512475313

*Smid, W.J., Kamphuis, J.H., Wever, E.C., & Van Beek, D.J. (2014). A comparison of the predictive properties of nine sex offender risk assessment instruments. *Psychological Assessment, 26(3),* 691-703. DOI: 10.1037/a0036616

*Smid, W.J., Kamphuis, J.H., Wever, E.C., & Van Beek, D.J. (2016). A quasi-experimental evaluation of high-intensity inpatient sex offender treatment in the Netherlands. *Sexual Abuse: A Journal of Research and Treatment, 28(5),* 469-485. DOI: 10.1177/1079063214535817

**Smid, W.J., Kamphuis, J.H., Wever, E.C., & Verbruggen, M.C.F.M. (2015). Risk levels, treatment duration, and drop out in a clinically composed outpatient sex offender treatment group, *Journal of Interpersonal Violence, 30(5),* 727-743. DOI: 10.1177/0886260514536276

Smith, K., Taylor, P., & Elkin, M. (2013). Crimes detected in England and Wales 2012/13 (Second Edition). Crown Copyright. ISBN: 978-1-78246-167-8

Steadman, H.J. (1980). The right not to be a false positive: Problems in the application of the dangerousness standard. *Psychiatric Quarterly, 32,* 84-99. DOI: 10.1007/BF01071879

Stevens, J.P. (2002). *Applied Multivariate Statistics for the Social Sciences (4th Edition).* Mahwah, NJ: Lawrence Erlbaum Associates. ISBN: 978-0805837773

**Stinson, J.D. (2016). Predictors of treatment noncompletion in a sample of inpatient sex offenders with serious mental illness, *Psychiatric Services, 67(1),* 43-48. DOI: 10.1176/appi.ps.201400415

**Storey, J.E., Watt, K.A., Jackson, K.J., & Hart, S.D. (2012). Utilization and implications of the Static-99 in practice, *Sexual Abuse: A Journal of Research and Treatment, 24(3),* 289-302. DOI: 10.1177/1079063211423943

Sturmey, P., & McMurran, M. (2011). *Forensic Case Formulation.* Chichester, West Sussex: John Wiley & Sons Ltd. ISBN: 978-0-470-68394-1

Tabachnick, B.G., & Fidell, L.S. (2007). *Using Multivariate Statistics (Fifth Edition).* Pearson Education Inc. ISBN: 0205459382

*Tamatea, A.J, (2014). Predictive validity of the STABLE-2007: A New Zealand study. *Sexual Abuse in Australia and New Zealand, 6(1),* 57-71. ISSN: 0833-8488

Tarrier, N., & Calam, R. (2002). New developments in cognitive-behavioural case formulation. Epidemiological, systemic and social context: An integrative approach. *Behavioural and Cognitive Psychotherapy, 30(3),* 311-328. DOI: 10.1017/S1352465802003065

Thornton, D. (2002). Constructing and testing a framework for dynamic risk assessment. *Sexual Abuse: A Journal of Research and Treatment, 14(2),* 139-153. DOI: 10.1177/107906320201400205

Thornton, D. (2007). *Scoring Guide for Risk Matrix 2000.9/SVC*. Unpublished Manual.

Thornton, D. (2010). *Scoring guide for Risk Matrix 2000.10/SVC*. London, England: Ministry of Justice.

Thornton, D. (2013). Implications of our developing understanding of risk and protective factors in the treatment of adult male sexual offenders. *International Journal of Behavioral Consultation and Therapy, 8(3-4),* 62-65. ISSN: 1555–7855

*Thornton, D., & Knight, R.A. (2015). Construction and validation of SRA-FV need assessment. *Sexual Abuse: Journal of Research and Treatment, 27(4),* 360-375. DOI: 10.1177/1079063213511120

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 New York Academy of Sciences, 989,* 225–235. DOI: 10.1111/j.1749-6632.2003.tb07308.x

Thornton, D., & Travers, R. (1991). A longitudinal study of the criminal behavior of convicted sexual offenders. *Proceedings of the 1991 Prison Psychologist's Conference.* Scarborough, UK: HM Prison Service.

**Trinh, B.V. (2011). A replication study of the Domestic Violence Risk Appraisal Guide (DVRAG), Dissertation Abstracts International: Section B: The Sciences and Engineering, 72(6-8), 3775.

Tully, R.J., Chou, S., & Browne, K.D. (2013). A systematic review on the effectiveness of sex offender risk assessment tools in predicting sexual recidivism of adult male sex offenders. *Clinical Psychology Review, 33,* 287-316. DOI: 10.1016/j.cpr.2012.12.002

**Tully, R.J., & Browne, K.D. (2015). Appraising the risk matrix 2000 static sex offender risk assessment tool, *International Journal of Offender Therapy and Comparative Criminology, 59(2),* 211-224. DOI: 10.1177/0306624X13508928

*Tully, R.J., Browne, K.D., & Craig, L.A. (2015). An examination of the predictive validity of the structured assessment of risk and need-treatment needs analysis (SARN-TNA) in England and Wales. *Criminal Justice and Behavior, 42(5),* 509-528. DOI: 10.1177/0093854814553096

**Turner, D., Rettenberger, M., Lohmann, L., Eher, R., & Briken, P. (2014). Pedophilic sexual interests and psychopathy in child sexual abusers working with children, *Child Abuse & Neglect, 38,* 326-335. DOI: 10.1016/j.chiabu.2013.07.019

*Turner, D., Rettenberger, M., Yoon, D., Klein, V., Eher, R., & Briken, P. (2016). Risk assessment in sexual abusers working with children. *Sexual Abuse: A Journal of Research and Treatmen, 28(6),* 572-596. DOI: 10.1177/1079063214564390

**van Leeuwen, M.L., van Baaren, R.B., Chakhssi, F., Loonen, M.G.M., Lippman, M., & Dijksterhuis, A. (2013). Assessment of implicit sexual associations in non-incarcerated pedophiles, *Archives of Sexual Behavior, 42,* 1501-1507. DOI 10.1007/s10508-013-0094-0

**Varela, J.G., Boccaccini, M.T., Cuervo, V.A., Murrie, D.C., & Clark, J.W. (2014). Same score, different message: Perceptions of offender risk depend on Static-99R risk communication format, *Law and Human Behavior, 38(5),* 418-427. DOI: 10.1037/lhb0000073

*Varela, J.G., Boccaccini, M.T., Murrie, D.C., Caperton, J.D., & Gonzalez, E. (2013). Do the Static-99 and Static-99R perform similarly for white, black, and Latino sexual offenders? *International Journal of Forensic Mental Health, 12(4),* 231-243. DOI: 10.1080/14999013.2013.846950

Vess, J., & Ward, T. (2011). Sexual offences against children. In. P. Sturmey, & M. McMurran (Eds.). Forensic Case Formulation. Chichester, West Sussex: John Wiley & Sons Ltd. ISBN: 978-0-470-68394-1

Vess, J., Ward, T., & Collie, R. (2008). Case formulation with sex offenders: An illustration of individualized risk assessment. *The Journal of Behavior Analysis of Offender and Victim Treatment and Prevention, 1(3),* 284-293. DOI: 10.1037/h0100450

**Vitacco, M.J., Erickson, S.K., Kurus, S., & Apple, B.N. (2012). The role of the Violence Risk Appraisal Guide and Historical, Clinical, Risk-20 in US Courts: A case law survey, *Psychology, Public Policy and Law, 18(3),* 361-391. DOI: 10.1037/a0025834

**Wakeling, H., Beech, A.R., & Freemantle, N. (2013). Investigating treatment change and its relationship to recidivism in a sample of 3773 sex offenders in the UK, *Psychology, Crime & Law, 19(3),* 233-252. DOI: 10.1080/1068316X.2011.626413

Wakeling, H.C., Mann, R.E., & Carter, A.J. (2012). Do low-risk sexual offenders need treatment? *The Howard Journal of Criminal Justice, 51(3),* 286-299.

**Wakeling, H.C., Mann, R.E., & Milner, R.J. (2011). Interrater reliability of risk matrix 2000/s, *International Journal of Offender Therapy and Comparative Criminology, 55(8),* 1324-1337. DOI: 10.1177/0306624X11386933

**Walters, G.D. (2016). Predicting recidivism with the Criminal Sentiments Scale: A meta-analysis of a putative measure of criminal thought content, *Criminal Justice and Behavior, 43(9),* 1159-1172. DOI: 10.1177/0093854816649004

**Walters, G.D., Deming, A., & Casbon, T. (2015). Predicting recidivism in sex offenders with the Psychological Inventory of Criminal Thinking Styles (PICTS), *Assessment, 22(2),* 167-177. DOI: 10.1177/1073191114539384

Ward, T. (2002). Good lives and the rehabilitation of sexual offenders: Promises and problems. *Aggression and Violent Behavior, 7(5),* 513-528. DOI: 10.1016/S1359-1789(01)00076-3

Ward, T., & Beech, A.R. (2009). An integrated theory of sexual offending. *Aggression and Violent Behavior, 11,* 44-63. ISSN: 1359-1789

Ward, T., & Laws, D.R. (2010). Desistance from sex offending: Motivating change, enriching practice. *International Journal of Forensic Mental Health, 9(1),* 11-23. DOI: 10.1080/1499011003791598

Ward, T., & Salmon, K. (2011). The ethics of care and treatment of sex offenders. *Sexual Abuse: A Journal of Research and Treatment, 23(3),* 397-413. DOI: 10.1177/1079063210382049

Watson, J.B. (1913). Psychology from the standpoint of a behaviorist. *The Psychological Review, 20,* 158-177. ISBN**:** 9781230371382

Webster, C. D., Martin, M. L., Brink, J., Nicholls, T. L., & Middleton, C. (2004). *Short-Term Assessment of Risk and Treatability (START).* St. Josephs Healthcare, Hamilton and British Columbia Mental Health and Addiction Services.

Wilcox, D.T. (2017). Preface In. D.T. Wilcox, M.L. Donathy, R. Gray, & C. Baim (Eds.). *Working with sex offenders: A guide for practitioners.* Abingdon, Oxon: Routledge. ISBN: 978-1-138-93354-5

Wilcox, D.T., & Gray, R. (2017). Developing more effective interviewing skills with men who commit sexual offences. In. D.T. Wilcox, M.L. Donathy, R. Gray, & C. Baim (Eds.). *Working with sex offenders: A guide for practitioners.* Abingdon, Oxon: Routledge. ISBN: 978-1-138-93354-5

Wilkes, V., & Barker, R. (2014). *Assessing Risk of Violence.* Presentation: University of Birmingham, Autumn.

**Wilson, R J., Abracen, J., Looman, J., Picheca, J.E., & Ferguson, M. (2011). Pedophilia: An evaluation of diagnostic and risk prediction methods, *Sexual Abuse: A Journal of Research and Treatment, 23(2),* 260-274. DOI: 10.1177/1079063210384277

Winder, B., Lievesley, R., Elliott, H.J., Norman, C., & Kaul, A. (2015). Understanding the journeys of high-risk male sex offenders voluntarily receiving medication to reduce their sexual preoccupation and/or hypersexuality. In. D.T. Wilcox, T. Garrett, and L. Harkins (Eds.). *Sex Offender Treatment: A Case Study Approach to Issues and*

Interventions. Chichester, West Sussex: John Wiley & Sons Limited. ISBN: 978-1-118-67440-6

**Woessner, G., & Schwedler, A. (2014). Correctional treatment of sexual and violent offenders: Therapeutic change, prison climate, and recidivism, *Criminal Justice and Behavior, 41(7),* 862-879. DOI: 10.1177/0093854813520544

Wong, S., Olver, M.E., Nicholaichuk, T.P., & Gordon, A. (2000). The Violence Risk Scale: Sexual Offender version. Saskatoon, SK, Canada: Regional Psychiatric Centre and University of Saskatchewan.

*Woodrow, A.C., & Bright, D.A. (2011). Effectiveness of a sex offender treatment programme: A risk band analysis. *International Journal of Offender Therapy and Comparative Criminology, 55(1),* 43-55. DOI: 10.1177/0306624X09352162

Workman, K. (2015). Is this the Dawning of the Age of Surveillance? - Monitoring offenders in New Zealand. *Journal of New Zealand Studies,* 69-86. ISSN: 2324-3740

**Wormith, J., Hogg, S., & Guzzo, L. (2012). The predictive validity of a general risk/needs assessment inventory on sexual offender recidivism and an exploration of the professional override, *Criminal Justice and Behavior, 39(12),* 1511-1538. DOI: 10.1177/0093854812455741

Yarber, L., Brownson, C.A., Jacob, R.R., Baker, E.A., Jones, E., Baumann, C., Deshpande, A.D., Gillespie, K.N., Scharff, D.P., & Brownson, R.C. (2015). Evaluating a train-the-trainer approach for improving capacity for evidence-based decision making in public health. *BMC Health Services Research, 15,* 547.

Appendices

Appendix A: Search Syntax

OVID: Psychinfo (2011 to December Week 2, 2016, completed on 20 December 2016)

1. sex$ offen$.mp. [mp=title, abstract, heading, word, table of contents, key concepts, original title, tests and measures] (10,681)

2. exp Sex Offenses/(31,519)

3. exp Child Abuse/or exp Sexual Abuse/or sex$ abuse.mp. (44,480)

4. exp Rape/or exp Acquaintance Rape/or rape.mp. (8,689)

5. sex$ assault.mp. (4,946)

6. exp Pedophilia/or child molest$.mp. (2,062)

7. paedophilia.mp. or exp Paraphilias/ (7,192)

8. exp Risk Assessment/or exp Risk Management/or risk.mp. (294,181)

9. exp Recidivism/or reconviction.mp. (5,039)

10. reoffen$.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests and measures] (1,126)

11. rm2000.mp. (19)

12. exp “Clinical Judgment (Not Diagnosis)”/or exp Decision Making/or actuarial.mp./or exp Prediction/ (158,135)

13. risk matrix 2000.mp. (88)

14. svr-20.mp. (41)

15. static-99.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests and measures] (335)

16. static-2002.mp. [mp=title, abstract, headingword, table of contents, key concepts, original title, tests and measures] (47)

17. rrasor.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests and measures] (43)

18. MnSOST-R.mp. [mp=title, abstract, headingword, table of contents, key concepts, original title, tests and measures] (28)

19. SORAG.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests and measures] (37)

20. RSVP.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests and measures] (508)

21. RISK FOR SEXUAL VIOLENCE PROTOCOL.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests and measures] (20)

22. SARN.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests and measures] (4)

23. SRA.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests and measures] (277)

24. STRUCTURED RISK ASSESSMENT.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests and measures] (118)

25. STRUCTURED PROFESSIONAL JUDGEMENT.mp. (28)

26. exp Judgment/or judgement.mp. (28,193)

27. prediction.mp. or exp Prediction/(63,861)

28. predictive validity.mp. or exp Statistical Validity/ (20,328)

29. exp Measurement/or exp “Predictability (Measurement)”/or exp Statistical Measurement/or measurement.mp. (132,442)

30. statistical analysis.mp. or exp Statistical Analysis/ (87,925)

31. test validity.mp. or exp Test Validity/(63,720)

32. test reliability.mp. or exp Test Reliability/(45,179)

33. specificity.mp. (30,451)

34. sensitivity.mp. (83,577)

35. accuracy.mp. (57,558)

36. area under curve.mp. (199)

37. auc.mp. (1,952)

38. 1 or 2 or 3 or 4 or 5 or 6 or 7 (58,029)

39. 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or

20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 (457,013)

40. 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 (394,251)

41. 38 and 39 and 40 (1,294)

42. limit 41 to yr="1980–Current" (1,292)

43. limit 42 to (adulthood <18+ years> and male) (413)

OVID: MEDLINE (2011 to December week 1 2016, completed on 20 December 2016)

1. sex$ offen$.mp. [mp=protocol supplementary concept, rare disease supplementary concept, title, original title, abstract, name of substance word, subject heading word, unique identifier] (8,385)

2. exp Sex Offenses/(21,755)

3. exp Child Abuse, Sexual/or sex$ abuse.mp. or exp Child Abuse/ (32,138)

4. rape.mp. or exp Rape/ (9,541)

5. sex$ assault.mp. (3,489)

6. exp Pedophilia/or child molest$.mp. (1,099)

7. exp Paraphilias/or paedophilia.mp. (5,243)

8. risk.mp. or exp Risk/or exp Risk Management/or exp Risk Assessment/(2,007,787)

9. reconviction.mp. (80)

10. exp Recurrence/or recidivism.mp. (172,390)

11. reoffen$.mp. [mp=protocol supplementary concept, rare disease supplementary concept, title, original title, abstract, name of substance word, subject heading word, unique identifier] (317)

12. rm2000.mp. (6)

13. exp Actuarial Analysis/or actuarial.mp. (38,799)

14. risk matrix 2000.mp. (16)

15. svr-20.mp. (22)

16. static-99.mp. (78)

17. static-2002.mp. (5)

18. rrasor.mp. (16)

19. MnSOST-R.mp. (7)

20. SORAG.mp. (16)

21. RSVP.mp. (513)

22. RISK FOR SEXUAL VIOLENCE PROTOCOL.mp. [mp=protocol supplementary concept, rare disease supplementary concept, title, original title, abstract, name of substance word, subject heading word, unique identifier] (0)

23. SARN.mp. [mp=protocol supplementary concept, rare disease supplementary concept, title, original title, abstract, name of substance word, subject heading word, unique identifier] (10)

24. SRA.mp. [mp=protocol supplementary concept, rare disease supplementary concept, title, original title, abstract, name of substance word, subject heading word, unique identifier] (1,255)

25. STRUCTURED RISK ASSESSMENT.mp. (57)

26. statistical measurement.mp. (45)

27. exp Judgment/or judgement.mp. (24,111)

28. exp “Predictive Value of Tests”/or predictive validity.mp. (186,199)

29. statistical validity.mp. (264)

30. measurement.mp. (496,615)

31. statistical measurement.mp. (45)

32. exp Data Interpretation, Statistical/or statistical analysis.mp. (129,891)

33. test validity.mp. or exp “Sensitivity and Specificity”/(354,291)

34. accuracy.mp. (265,853)

35. area under curve.mp. or exp Area Under Curve/(36,646)

36. auc.mp. (43,128)

37. 1 or 2 or 3 or 4 or 5 or 6 or 7 (50,170)

38. 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 (2,190,822)

39. 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 (1,292,984)

40. 37 and 38 and 39 (690)

41. limit 40 to yr=“1980–Current” (689)

42. limit 41 to (male and “all adult (19 plus years)”) (280)

OVID: EMBASE (2011 to 2016 December 19, completed 20 December 2016)

1. exp sexual crime/or sex* offen*.mp. or exp sexual abuse/(29,734)

2. rape.mp. or exp rape/(10,797)

3. sex* assault.mp. (5,116)

4. exp pedophilia/or exp child sexual abuse/or p*dophil*.mp. or exp sexual deviation/(12,269)

5. risk.mp. or exp risk/or exp recurrence risk/or exp risk management/or exp risk assessment/(2,930,991)

6. reconviction.mp. or exp recidivism/(3,386)

7. reoffen*.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword] (489)

8. rm2000.mp. or exp rating scale/(108,735)

9. actuarial.mp. (23,106)

10. risk matrix 2000.mp. (22)

11. svr-20.mp. (39)

12. static-99.mp. (97)

13. static-2002.mp. (5)

14. rrasor.mp. (16)

15. exp “prediction and forecasting”/or MnSOST-R.mp. (1,600)

16. exp psychological test/or SORAG.mp. (153,864)

17. RSVP.mp. (635)

18. RISK FOR SEXUAL VIOLENCE PROTOCOL.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword] (1)

19. SARN.mp. (10)

20. SRA.mp. (1,917)

21. exp psychological aspect/or STRUCTURED RISK ASSESSMENT.mp. (476,602)

22. STRUCTURED PROFESSIONAL JUDGEMENT.mp. (26)

23. judgement.mp. or exp decision making/(295,624)

24. prediction.mp. or exp prediction/(445,544)

25. predictive validity.mp. or exp predictive validity/(10,564)

26. exp measurement/or measurement.mp. (2,340,023)

27. statistical analysis.mp. or exp statistical analysis/(2,112,231)

28. test validity.mp. (540)

29. exp reliability/or test reliability.mp. (160,963)

30. exp “sensitivity and specificity”/or specificity.mp. (728,684)

31. sensitivity.mp. (1,144,447)

32. accuracy.mp. or exp accuracy/(646,478)

33. area under curve.mp. or exp area under the curve/(118,863)

34. auc.mp. (73,188)

35. 1 or 2 or 3 or 4 (41,065)

36. 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 (4,091,396)

37. 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 (5,717,556)

38. 35 and 36 and 37 (3,904)

39. limit 38 to yr=“1980–Current” (3,889)

40. limit 39 to (male and adult <18 to 64 years>) (1,568)

Web of Science (Science Citation Index Expanded (SCI-EXPANDED); Social Sciences Citation Index (SSCI); Arts and Humanities Citation Index (A&HCI); Conference Proceedings Citation Index — Science (CPCI-S);
Conference Proceedings Citation Index — Social Science and Humanities (CPCI-SSH); 2011–2016, completed on 20 December 2016)

1. TS=(sex offen* OR “child abuse” OR sex* abuse OR paedophilia OR rape OR sex* assault OR paraphili*) Databases=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH Timespan=2011-2016 Lemmatization=On (21,176).

2. TS=(“risk assessment” OR “risk management” OR risk OR recidivism OR reconviction OR actuarial OR “professional judgement” OR prediction OR RM2000 OR “risk matrix 2000” OR SVR-20 OR RSVP OR STATIC-2002 OR STATIC-99 OR MnSOST-R OR SARN OR “STRUCTURED RISK ASSESSMENT”) Databases=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH Timespan=2011-2016 Lemmatization=On (1,239,617).
3. TS=(“predictive validity” OR “statistical validity” OR measurement OR “statistical measurement” OR “statistical analysis” OR validity OR reliability OR specificity OR sensitivity OR accuracy OR “area under curve”) Databases=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH Timespan=2011-2016 Lemmatization=On (1,630,827).
4. #3 AND #2 AND #1 Databases=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH timespan=2011-2016 Lemmatization=On (1,073).

Applied Social Sciences Index and Abstracts (ASSIA; 2011–2016, completed on 20 December 2016)
all (“sex offend*” OR “sex* abuse” OR paedophilia OR rape OR “sex* assault” OR paraphili*) (3,733)
AND
all (“risk assessment” OR “risk management” OR risk OR recidivism OR reconviction OR actuarial OR “professional judgement” OR prediction OR RM2000 OR “risk matrix 2000” OR SVR-20 OR RSVP OR STATIC-2002 OR STATIC-99 OR MnSOST-R OR SARN OR “STRUCTURED RISK ASSESSMENT”) (79,512)
AND
all (“predictive validity” OR “statistical validity” OR measurement OR “statistical measurement” OR “statistical analysis” OR validity OR reliability OR specificity OR sensitivity OR accuracy OR “area under curve”) (50,561)

AFTER Dec 2011 (121)

Cochrane Central (2011–2016, completed on 20 December 2016)

"sex* offen*" OR rape OR "sex* assault" in Title, Abstract or Keywords and risk OR prediction OR "risk assessment" OR actuarial OR structured in Title, Abstract or Keywords and effectiveness OR "area under curve" OR specificity OR sensitivity in Title, Abstract or Keywords, from 1980 to 2016 in Cochrane Central Register of Controlled Trials" (1,536)

Appendix B: Inclusion/Exclusion Forms

Inclusion/exclusion form

Full Reference:

Inclusion criteria	met?	Comments
Population:		
• Adult male?	Yes/No	
• AND		
• Sexual offender?	Yes/No	
Exposure:		Which tools?
• Risk assessment tool applied	Yes/No	
• Tool specifically designed for sexual offenders	Yes/No	
Outcome:		How measured?
• Reconviction?		
• Re-arrest?		
• Self report?		
• Other?		
Study type:		Which type?
• Case control		
• Cohort		
Exclusion:		
• Not an opinion paper	Yes/No	
Conclusion:	Included/Excluded	

Appendix C: Excluded studies

Study	Reason for Exclusion
Allen and Pflugradt (2014)	Outcome
Babchishin, Hanson and Blais (2016)	Outcome
Babchishin, Hanson and Herman (2011)	Meta-analysis
Babchishin, Nunes, Hermann and Malcom (2015)	Exposure
Baltieri and Boer (2015)	Outcome
Banse, Koppehele-Gossel, Kistemaker, Werner and Schmidt (2013)	Exposure
Bench and Allen (2013)	Population
Blais and Bonta (2014)	Outcome
Blasko, Jeglic and Mercado (2011)	Outcome
Boccaccini, Murrie, Mercado, Quesada, Hawes, Rice and Jeglic (2012)	Exposure
Boccaccini, Turner, Murrie, Henderson and Chevalier (2013)	Population
Briken and Műller (2014)	Non-English language not accessible
Brouillette-Alarie, Babchishin, Hanson and Helmus (2016)	Outcome
Brouillette-Alarie, Hanson, Babchishin and Benbouriche (2014)	Non-English language not accessible
Brown, Harkins and Beech (2012)	Outcome
Buttars, Huss and Brack (2015)	Outcome
Camilleri and Quinsey (2011)	Review
Caudy, Durso and Taxman (2013)	Population
Chevalier, Boccaccini, Murrie and Varela (2015)	Population
Coid, Yang, Ullrich, Zhang, Sizmur, Farrington and Rogers (2011)	Population
Corovic, Christianson and Bergman (2012)	Exposure
Coyle (2011)	Opinion paper
Craissati and Blundell (2013)	Population

Dahle, Biedermann, Lehmann and Gallasch-Nemitz (2014)	Population
de Vries Robbé, de Vogel and Douglas (2013)	Exposure
de Vries Robbé, de Vogel, Wever, Douglas and Nijman (2016)	Population
Dickson, Polaschek and Casey (2013)	Population
Doyle, Carter, Shaw and Dolan (2012)	Population
Doyle, Ogloff and Thomas (2011)	Outcome
Ducro, Pham, Saloppé, Chudzicck and Réveillére (2012)	Non-English language not accessible
Duwe (2012)	Population
Duwe (2013)	Population
Duwe and Freske (2012)	Tool development
Edens, Cox, Smith, DeMatteo and Sorman (2015)	Exposure
Eher, Schilling, Hansmann, Pumberger, Nitschke, Habermeyer and Mokros (2016)	Literature review
Faust, Bickart, Renaud and Camp (2015)	Exposure
Fujita, Watanabe, Yokota, Suzuki, Wachi, Otsuka and Kuraishi (2016)	Population
Gannon and O'Connor (2011)	Exposure
Gardner, Boccaccini, Bitting and Edens (2015)	Meta-analysis
Garombo, Salvadori, Contarino, Castellino, Molinaro, Garofano, Molo, Veglia and Rosso (2016)	Outcome
Giguére and Lussier (2016)	Exposure
Gray, Abel, Jordan, Garby, Wiegel and Harlow (2015)	Exposure
Grubin (2011)	Outcome
Guay (2016)	Non-English language not accessible
Gutierrez, Wilson, Rugge and Bonta (2013)	Meta-analysis
Hamilton (2015)	Critique

Hannah-Moffat (2016)	Opinion paper
Hanson, Babchishin, Helmus and Thornton (2013)	Outcome
Harkins, Howard, Barnett, Wakeling and Miles (2015)	Outcome
Harris and Rice (2015)	Critique
Hart and Cooke (2013)	Opinion paper
Hawes, Boccaccini and Murrie (2013)	Meta-analysis
Hecker (2014)	Opinion paper
Helmus, Hanson, Babchishin, and Mann (2013)	Meta-analysis
Helmus, Hanson, Thornton, Babchishin and Harris (2012)	Meta-analysis
Helmus and Thornton (2015)	Meta-analysis
Hempel, Buck, Goethals and van Marle (2013)	Exposure
Hendry, Douglas, Winter and Edens (2013)	Exposure
Higgins (2012)	Population
Hockenhull, Whittington, Leitner, Barr, McGuire, Cherry, Flentje, Quinn, Dundar and Dickson (2013)	Population
Howard, Barnett and Wakeling (2015)	Population
Janka, Gallasch-Nemitz, Biedermann, and Dahle (2012)	Exposure
Jung, Daniels, Friesen and Ledi (2012)	Population
Jung, Pham and Ennis (2013)	Outcome
Kiland (2016)	Outcome
Kingree and Thompson (2015)	Population
Kingston, Olver, Harris, Wong and Bradford (2015)	Outcome
Kingston, Yates and Firestone (2012)	Exposure
Kingston, Yates and Olver (2014)	Outcome
Kleban, Chesin, Jeglic and Mercado (2013)	Outcome
Krauss and Scurich ((2013)	Critique
Larcombe (2012)	Critique

Lasher, McGrath and Cumming (2015)	Outcome
Lee, Li, Lamade, Schuler and Prentky (2012)	Exposure
Lehmann, Goodwill, Gallasch-Nemitz, Biedermann and Dahle (2013)	Exposure
Lehmann, Goodwill, Hanson and Dahle (2016)	Exposure
Lewis, Olver and Wong (2013)	Population
Lofthouse, Lindsay, Totsika, Hastings, Boer and Haaven (2013)	Population
McDougall, Pearson, Willoughby and Bowles (2013)	Exposure
McNally and Fremouw (2014)	Critique
McPhail, Hermann and Fernandez (2014)	Outcome
Miller, Kimonis, Otto, Kline and Wasserman (2012)	Exposure
Mokros, Gebhard, Heinz, Marschall, Nitschke, Glasgow, Gress and Laws (2013)	Outcome
Mooney and Daffern (2013)	Population
Neller and Frederick (2013)	Outcome
Neller and Petris (2013)	Opinion paper
Nilsson, Wallinius, Gustavson, Anckarsater and Kerekes (2011)	Population
Nunes and Babchishin (2012)	Outcome
Olver, Kingston, Nicholaichuk and Wong (2014)	Exposure
Olver, Nicholaichuk, Gu and Wong (2013)	Population
Olver, M. E., Wong, S. (2011a)	Exposure
Olver and Wong (2013)	Exposure
Osetermann and Salerno ((2016)	Exposure
Pettersen, Nunes and Cortoni (2016)	Exposure
Quesada, Calkins and Jeglic (2014)	Outcome
Quinn, Miles and Kinane (2013)	Population
Ragusa-Salerno, Ostermann and Thomas (2013)	Outcome
Rettenberger, and Eher (2013)	Outcome

Rice and Harris (2014)	Outcome
Rice, Harris and Lang (2013)	Outcome
Rocque and Plummer-Beale (2014)	Outcome
Romine, Miner, Poulin, Dwyer and Berg (2012)	Population
Rossegger, Gerth, Seewald, Urbaniok, Singh and Endrass (2013)	Systematic review
Ryan, Wilson, Kilgour and Reynolds (2014)	Exposure
Sandler, Freeman, Farrelland Seto (2013)	Outcome
Schmidt, Gykiere, Vanhoeck, Mann and Banse (2014)	Outcome
Schmucker and Losel (2015)	Exposure
Seto and Fernandez (2011).	Outcome
Seto, Hanson and Babchishin (2011)	Outcome
Singh, Grann and Fazel (2011)	Meta-analysis
Singh, Grann, Lichtenstein, Langstrom and Fazel (2012)	Population
Smid, Kamphuis, Wever and Van Beek (2013)	Outcome
Smid, Kamphuis, Wever and Verbruggen (2015)	Outcome
Stinson (2016)	Outcome
Storey, Watt, Jackson and Hart (2012)	Outcome
Trinh (2011)	Population
Tully and Browne (2014)	Critique
Turner, Rettenberger, Lohmann, Eher and Briken (2014)	Outcome
van Leeuwen, van Baaren, Chakhssi, Loonen, Lippman and Dijksterhuis (2013)	Exposure
Varela, Boccaccini, Cuervo, Murrie and Clark (2014)	Population
Vitacco, Erickson, Kurus and Apple (2012)	Case law survey
Wakeling, Beech and Freemantle (2013)	Population
Wakeling, Mann and Milner (2011)	Outcome
Walters (2016)	Meta-analysis

Walters, Deming and Casbon (2015)	Outcome
Wilson, Abracen, Looman, Picheca and Ferguson (2011)	Outcome
Woessner and Schwedler (2014)	Exposure
Wormith, Hogg and Guzzo (2012)	Population

Appendix D: Quality assessment form: Cohort

Source Database:

Full Reference:

Question	Score	Comments
	Y(2) P(1) N(0) U	

Were the study objectives clear?
Will a cohort study address the objectives?

Selection bias
Was the cohort recruited in an acceptable way?
Was the cohort representative?
(or special in some way)

Measurement bias
Was exposure accurately measured?
(was exposure uniform to all)
Was risk tool described clearly?
Was the outcome measure clearly stated?
(recidivism data source and definitions clearly stated)
Was the outcome assessed uniformly across the sample (same procedure)?
Was the risk assessor blind to outcome (recidivism)?
Were assessors trained/experienced enough to be competent in applying the tool?
Was the tool applied using enough information? (multiple sources – list them)
Was inter-rater reliability assessed?
Was inter-rater reliability 0.8 or above?
Was the follow-up time long enough? (min 2 years)
Was missing information dealt with appropriately

Attrition
Was drop-out/non-completion rate recorded?
Was drop-out/non-completion stage discussed?

Results
Are the results reported? What are they – ROC AUC or specificity/sensitivity etc reported?

Was predictive validity clearly stated?
Was concurrent validity discussed/addressed? (test correlates well with previously validated tests)
Are the results reliable?
Do results fit with other available evidence?
Can the results be generalised (or are participants different enough from adult male sex offenders to cause concern)?
Were confounding factors discussed/taken into account?

Quality score = _____

No. Unclear = _____

Quality assessment forms adapted from the Critical Appraisal Skills Programme (CASP, 2004). Studies were scored as follows in relation to each question:
0=condition not met.
1=condition partially met.
2=condition fully met.
U=unclear/insufficient information provided.
Scores were summed in order to obtain an overall quality rating, with higher scores indicating better quality studies. Lucidity of reporting was assessed by summing the number of items rated 'U', with a higher score indicating less accurate reporting.

Appendix E: Quality assessment form: Case control

Source Database:

Full Reference:

Question	Score	Comments
	Y(2) P(1) N(0) U	

Were the study objectives clear?
Will a cohort study address the objectives?

Selection bias
Was the cohort recruited in an acceptable way?
Was the cohort representative?
(or special in some way)

Measurement bias
Was exposure accurately measured?
(was exposure uniform to all)
Was risk tool described clearly?
Was the outcome measure clearly stated?
(recidivism data source and definitions clearly stated)
Was the outcome assessed uniformly across the sample (same procedure)?
Was the risk assessor blind to outcome (recidivism)?
Were assessors trained/experienced enough to be competent in applying the tool?
Was the tool applied using enough information? (multiple sources – list them)
Was inter-rater reliability assessed?
Was inter-rater reliability 0.8 or above?
Was the follow-up time long enough? (min 2 years)
Was missing information dealt with appropriately

Attrition
Was drop-out/non-completion rate recorded?
Was drop-out/non-completion stage discussed?

Results
Are the results reported? What are they – ROC AUC or specificity/sensitivity etc reported?

Was predictive validity clearly stated?
Was concurrent validity discussed/addressed? (test correlates well with previously validated tests)
Are the results reliable?
Do results fit with other available evidence?
Can the results be generalised (or are participants different enough from adult male sex offenders to cause concern)?
Were confounding factors discussed/taken into account?

Quality score = _____

No. Unclear = _____

Quality assessment forms adapted from the Critical Appraisal Skills Programme (CASP, 2004). Studies were scored as follows in relation to each question:
0=condition not met.
1=condition partially met.
2=condition fully met.
U=unclear/insufficient information provided.
Scores were summed in order to obtain an overall quality rating, with higher scores indicating better quality studies. Lucidity of reporting was assessed by summing the number of items rated 'U', with a higher score indicating less accurate reporting.

Appendix F: Data extraction form

Source Database:

Full Reference:

Study type?
Location(s) of study sample?
Name of tool(s) used?
Total sample size?
Sample status e.g. sub groups of sample (treated/untreated, rapist/child offender etc.)
Risk group if applicable? (e.g. numbers of low/high)
Outcome measure(s)?
Length of follow-up?
Re-offence rate?
    (No % re-offend?)
    (No % not re-offend?)
Statistical analysis of results for each tool:
    AUC ROC?
    Sensitivity?
    Specificity?
    Likelihood ratio?
Inter-rater reliability?
    (% agreement or kappa)
Quality assessment score
Clarity score

Appendix G: Data from cohort studies

Reference	Sample	Risk Tools	Average Follow-Up (Location)	Outcome	Inter-rater Reliability	Sexual Recidivism Rate	Predictive Accuracy (AUC values)	Quality Score
Barnett, Wakeling, Mandeville-Norden, & Rakestrow (2012)	3402 treated sex offenders	RM2000/S	3 years, SD = (England & Wales)	Proven re-offence	Not assessed	4.9-11%	0.60, [.55, .64]	34
Beggs & Grace (2011)	218 treated child sex offenders	VRS: SO SGAS	12.24 years (New Zealand)	Identifiable victim	VRS: SO, Pre .90, Post .92 SGAS, .88	13.3%	VRS: SO, 0.70 SGAS, 0.66	38
Brouillette-Alarie & Proulx (2013)	711 Max security psychiatric patients	STATIC-99R	Fixed 5 years (Canada)	Charge or conviction	Not assessed	12.8% whole sample (ws) 11% offenders against women (ow) 12.6% offenders against children (oc)	WS 0.73, p<.001, [.66, .81] OW 0.73, p<.01, [.61, .86] OC 0.77, p<.001, [.68, .86]	31
de Vries Robbé, de Vogel, Koster, & Bogaerts (2015)	83 discharged forensic psychiatric patients	SVR-20 SAPROF	1 year fixed 3 years fixed 15.1 years (SD = 5.3, range 3-24) (Netherlands)	Conviction	SVR-20 .85 SAPROF .84	2% 1 year 7% 3 year 19% 15 year	SVR-20 0.58 [.42, .74] SAPROF 0.71 [.56, .86]	40
Eher, Matthes, Schilling, Haubner-MacLean, & Rettenberger (2012)	263 released offenders STATIC-99 (N=257) SORAG (N=250) STABLE-2000 (N=262) STABLE-2007 (N=256)	STATIC-99 SORAG STABLE-2000 STABLE-2007	6.4 years (Austria)	Conviction	STATIC-99 .98 SORAG .93 STABLE-2000 .89 STABLE-2007 .90	10.3%	STATIC-99 0.75 [.66, .84] SORAG 0.72 [.64, .80] STABLE-2000 0.62 [.51, .73] STABLE-2007 0.71 [.61, .81]]	34
Eher, Olver, Heurix, Schilling, & Rettenberger (2015)	261 child sex offenders with diagnosis of paedophilia	SSPI STATIC-99R STABLE-2007 VRS: SO	6.28 years (SD = 2.19 years, range 2.06-10.70 years) (Austria)	Not specified	SSPI .95, p<.00,1 95% CI: [.89, .97] STATIC-99 .75, 95% CI: [.66, .84] STATIC-99R not specified STABLE-2007 .89-.92	general sexual reoffense 8.5% sexual contact reoffense 7.9% sexual noncontact reoffense 1.6%	SSPI 0.71, p<.01, [.57, .86] STATIC-99R 0.67, p<.05, [.47, .82] STABLE-2007 0.60, ns, [.44, .76] VRS: SO 0.76, p<.001, [.66, .86]	36

					VRS: SO .93, p<.0001, 95% CI: [.81, .97]			
Ennis, Buro, & Jung (2016)	345 psychiatric inpatient and outpatient offenders	STATIC-2002R	Minimum 2 years (Canada)	Charge and/or conviction	STATIC-2002R .52 – 1.00 SSPI Not specified	Cluster 1 8.7% Cluster 2 10.7% Cluster 3 16%	0.62-0.69	35
Fedoroff, Richards, Ranger, & Curry (2016)	112 Intellectually Disabled offenders	SORAG	31.88 months (SD = 11.95, range 5-49) (Canada)	Re-offense or PSB (problematic sexual behaviour)	.946, 95% CI: [.817, .992]	Not specified	0.70, [.53, .88]	33
Hanson, Harris, Helmus, & Thornton (2014)	7740 Various 21 samples	STATIC-99R	8.2 years (SD = 5.2, range 0.01-31.5) (Various)	Various 21 samples	.89	11.9%	Not specified	34
Hanson, Helmus, & Harris (2015)	768 STATIC-99R (N=764) STATIC-2002R (N=710) STABLE-2000 (N=616) STABLE-2007 (N=615) Offenders or probation/parole	STATIC-99R STATIC-2002R STABLE-2000 STABLE-2007	7.4 years (SD = 2.2, range 0.2-10.1 years) (Canada)	Various 21 samples	STATIC-99R .91 STATIC-2002R Not specified STABLE-2000 .89 STABLE-2007 Not specified	10.8% new sexual crimes 12.9% sexual crimes/sexually motivated breaches	STATIC-99R 0.788 [.714, .863] STATIC-200R 0.798 [.726, .870] STABLE-2000 0.747 [.668, .825] STABLE-2007 0.784 [.710, .859]	40
Helmus, Babchishin, & Blais (2012)	597 offenders on probation/parole	STABLE-2007	3.4 years (SD = 1.0, range 1.0-5.4) (Canada)	Offences	Not specified	11.4% aboriginal offenders 7.3% non-aboriginal offenders	aboriginal 0.529 [.331, .726] non-aboriginal 0.701, [.611, .791]	34
Helmus, Thornton, Hanson, & Babchishin (2012)	8390 STATIC-99 (N=8390) STATIC-2002 (N=2609)	STATIC-99 STATIC-99R STATIC-2002 STATIC-2002R	Fixed 5 years (N=5937) Fixed 10 years (N=2479) (Various, 24 samples)	Various 24 samples	Not specified	12.4% 11.1% 5 years 16.6% 10 years	STATIC-99 0.713 5 years, 0.706 10 years STATIC-99R 0.720 5 years, 0.710 10 years STATIC-2000 0.709 5 years, 0.689 10 years STATIC-2000R 0.713 5 years,.690 10 years	36

Helmus, Ciardha & Seto (2015)	410 offenders on probation/parole	SSPI STATIC-99R STATIC-2002R STABLE-2000 STABLE-2007	7.5 years (SD = 2.0, range 0.20-10.0) (Canada)	Sexually motivated offence including breaches	Not specified	8.8%	SSPI 0.641 [.557, .726] STATIC-99R 0.771 [.700, .843] STATIC-2002R 0.780 [.715, .846] STABLE-2000 Not specified STABLE-2007 0.709 [.622, .796]	32
Helmus, Hanson, Babchishin, & Thornton (2015)	710 offenders on probation/parole RM2000/S (N=710) STABLE-2007 (N=570)	RM2000/S STABLE-2007	7.7 years (SD = 1.7, range 1.6-10.1 years) (Canada)	Sexually motivated offence including breaches	Not specified	13.7%	RM2000/S 0.685 [.625, .744] RM200/S + STABLE-2007 0.709 [.642, .776]	32
Hill, Rettenberger, Habermann, Berner, Eher, & Briken (2012)	90 released sexual homicide offenders	SVR-20 STATIC-99	10.22 years (Germany)	Not specified	0.77-0.87	15.6%	SVR-20 0.52 [.36, .68] STATIC-99 0.56 [.41, .71]	36
Lee & Hanson (2016)	947 offenders on probation/parole	STATIC-99R STATIC-2002R STABLE-2000 STABLE-2007 BARR-2002R SSPI	7.4 years (SD = 2.2, range 0.2-10.1) (Canada & USA)	Sexually motivated offence including breaches	STATIC-99 0.91 STABLE-2000 0.89	1.80 hazard ratios	STATIC-99R 0.577 [.519, .635]** STATIC-2002R 0.588 [.520,.656]** STABLE-2000 0.612 [.547, .677]** STABLE-2007 0.649 [.585, .714]** BARR-2002R 0.557 [.490, .624]** SSPI 0.553 [.470, .637]**	36
Lehmann, Hanson, Babchishin, Gallasch-Nemitz, Biedermann, & Dahle (2013)	940 offenders reported to police	RRASOR STATIC-99R STATIC-2002R	9 years (Germany)	Not specified	Not specified	7.53%	RRASOR 0.58-0.60 STATIC-99R 0..68-0.69 STATIC-2002R 0.67-0.69	27
Looman, Morphett, & Abracen (2012)	272 assessed and/or treated offenders	STATIC-99R	6.7 years (SD = 3.4, range 0.4-16.4 years) (Canada)	Conviction	0.84	15.4%	0.70	34
McGrath, Lasher, & Cumming (2012)	759 offenders under supervision	VASOR-2 SOTIPS	5 years (Canada)	Charges including	0.89	4.6%	VASOR-2 0.74 SOTIPS 0.61-0.72	40

				breaches				
McGrath, Lasher, Cumming, Langton, & Hoke (2014)	1581 community sample	VASOR-2	5 years (Canada & USA)	Charge and/or conviction including breaches	0.88	8.6%	0.74	37
Montana, Thompson, Ellsworth, Lagan, Helmus, & Rhoades (2012)	337 treated clergy child sex offenders	STATIC-99	16.05 years (SD = 5.12 (UK)	Sexual contact, use of child pornography or behaviour leading to sexual contact (interrupted)	Not specified	6.2%	0.672	26
Nicholaichuk, Olver, Gu, & Wong (2014)	2158	BARS	12 years (SD = 1.7) (Canada)	Not specified	Not specified	12.6%	0.67-0.73 older offenders 0.65-0.66 younger offenders	30
Nunes, Hermann, Malcom, & Lavoie (2013)	462 supervised offenders	STATIC-99R SSPI	6909.33 days (SD = 772.199 days) (Canada)	Conviction	Not specified	23.2%	STATIC-99R 0.07 [-0.19, 0.34] SSPI 0.71 [0.44, 0.98]	30
Nunes, Pettersen, Hermann, Looman, & Spape (2016)	146 treated offenders	MOLEST & RAPE SCALE	7.59 years MOLEST 7.53 years RAPE (Canada)	Conviction	Not specified	19.7 MOLEST 18.8% RAPE	MOLEST Scale Pre-TX 0.53 [.41, .64] Post-TX 0.50 [.39, .62] RAPE Scale Pre-TX 0.53 [.41, .64] Post-TX 0.53 [.41, .64]	24
Olver & Wong (2011)	321 treated offenders	STATIC-99 VRS: SO	10 years (Canada)	Charge and/or conviction	STATIC-99 0.82 VRS: SO Pre-treatment 0.74 Post-treatment 0.79	Charges 22% LRLC 24%LRHC 43% hrlc 27% hrhc Convictions 14% 16% 36% 24%	STATIC -99 0.66-0.67 low risk 0.64-0.65 moderate risk 0.57-0.66 moderate high 0.55-0.56 high risk VRS: SO 0.69-0.73 low risk 0.66-0.70 moderate risk 0.64-0.65 moderate high 0.60-0.65 high risk	36
Olver, Nicholaichuk, Kingston, & Wong (2014)	676 treated offenders	STATIC-99 VRS: SO	6.31 years (SD = 2.32) (Canada)	Charge and/or conviction	Not specified	6.2%	STATIC-99 0.71-0.78 VRS: SO 0.66-0.69	34
Olver, Nicholaichuk, & Wong (2014)	267 treated offenders	STATIC-99 VRS: SO	18.2 years (SD = 4.7) (Canada)	Conviction	STATIC-99 0.82 VRS: SO Pre-treatment	27.3%	0.53-0.54**	40

					0.74 Post-treatment 0.79			
Olver, Christofferson, Grace, & Wong (2014)	539 treated offenders	STATIC-99 VRS: SO	15.5 years (SD = 4.4) (Canada & New Zealand)	Conviction	VRS: SO	22.4%	STATIC-99 0.71 [.66, .76] VRS: SO 0.73 [.69, .78]	42
Olver, Klepfisz, Stockdale, Kingston, Nicholaichuk, & Wong (2016)	668 incarcerated offenders	VRS: SO	10.2 years (Canada)	Charge and/or conviction	Not specified	10.4%	0.68-0.74	34
Parent, Guay, & Knight (2012)	414 sexually dangerous persons	STATIC-99 RM2000/S STATIC-2002 MnSOST-R SORAG	5 years (USA)	Charges	STATIC-99 RM2000/S STATIC-2002 MnSOST-R 0.78 SORAG 0.91	4.9%-29.1%	STATIC-99 0.70 RM2000/S 0.65 STATIC-2002 0.68 MnSOST-R 0.69 SORAG 0.67	36
Rettenberger, Haubner-Maclean, & Eher (2013)	1077	STATIC-99 STATIC-99R	6.35 years (Austria)	Conviction	STATIC-99 0.98 STATIC-99R	6.6%	STATIC-99 0.73 STATIC-99R 0.71	40
Seto & Eke (2015)	266 child pornography offenders	CPORT	5 years (Canada)	Contact and non-contact sexual offences	0.70	11%	0.74 [.63, .84] 0.63 [.41, .86] for pornography only offenders	37
Smallbone & Rallings (2013)	339 released offenders	STATIC-99	29 months (range 15-53 months) (Australia)	Arrest	Not specified	4.8%	0.81 [.72, .90] Indigenous 0.76 [.61, .91] Non-indiginous 0.82 [.68, .91]	29
Smid, Kamphuis, Wever, & Van Beek (2014)	397 treated and untreated sex offenders	STATIC-99 STATIC-99R RRASOR RM2000/S STATIC-2002 STATIC-2002R SACJ-Min SORAG SVR-20	145 months (SD = 30, range 51-201 months) (Netherlands)	Charge and/or conviction	0.87	14.1% overall 10.1% 5 years 14.6% 10 years	STATIC-99 0.72 5 year 0.73 10 year STATIC-99R 0.74 5 year 0.74 10 year RRASOR 0.68 5 year 0.69 10 year RM2000/S 0.72 5 year 0.71 10 year STATIC-2002 0.76 5 year	42

							0.75 10 year STATIC-2002R 0.77 5 year 0.75 10 year SACJ-Min 0.69 5 year 0.71 10 year SORAG 0.63 5 year 0.64 10 year SVR-20 0.53 5 year 0.48 10 year	
Smid, Kamphuis, Wever, & Van Beek l (2016)	266 untreated community and inpatient offenders	STATIC-99R	148 months (SD = 29.6, range 51-201 months) (Netherlands)	Charge and/or conviction	Not specified	15%	0.78 overall 0.83 untreated 0.66 inpatient	35
Tamatea (2014)	245 assessed offenders	STABLE-2007 ASRS	6.4 years (New Zealand)	Recidivism (less breaches) and re-imprisonment	Not specified	0.8% sexual 44.5% any offence	STABLE-2007 0.66 ASRS 0.65	33
Thornton & Knight (2015)	566 480 @ 5years 391 @ 10years	STATIC-99R RM2000/S SRA-FV	Fixed 5 & 10 years (USA)	Charges	STATIC-99R 0.94 RM2000/S 0.90 SRA-FV 0.78	19.2% @ 5years 23.3% @ 10years	STATIC-99R 0.686 [.627, .745] RM2000/S 0.665 [.603, .727] SRA-FV 0.727 [.673, .782]	41
Tully, Browne & Craig (2015)	496 treated community sample	SARN-TMA	2 & 4 years (UK)	Caution or conviction	Not conducted	5.6% 2 years 16.8% 4 years	0.59 2 years 0.57 4 years	37
Turner, Rettenberger, Yoon, Klein, Eher, & Briken (2016)	277 incarcerated offenders	STATIC-99 SORAG SVR-20 SAPROF	5.55 CSA-W 5.65 CSA-E 5.79 CSA-I (Austria)	Conviction	0.85	13.5% CSA-W 25.8% CSA-E 2.4% CSA-I	STATIC-99 CSA 0.83 CSA-W 0.78 CSA-E 0.79 CSA-I ANY RECIDIVISM 0.65 SORAG CSA 0.77 CSA-W 0.76 CSA-E 0.74 CSA-I ANY RECIDIVISM 0.66	34

							SVR-20 CSA 0.75 CSA-W 0.77 CSA-E 0.73 CSA-I ANY RECIDIVISM 0.73 SAPROF CSA 0.52 CSA-W 0.53 CSA-E 0.58 CSA-I ANY RECIDIVISM SAPROF 0.64	
Varela, Boccaccini, Murrie, Caperton, & Gonzalez (2013)	1911 sexually violent predators	STATIC-99 STATIC-99R	4.85 years white (SD = 1.50) 4.89 black (SD = 1.57) 4.58 latino (SD = 1.49) (USA)	Arrest	0.79	STATIC-99 White 12.4% Black 9.3% Latino 13.2% STATIC-99R White 13.1% Black 10.1% Latino 15.0%	STATIC-99 White 0.57 [.45, .70] Black 0.58 [.43, .73] Latino 0.59 [.45, .73] STATIC-99R White 0.59 [.45, .72] Black 0.65 [.51, .78] Latino 0.57 [.41, .73]	34
Woodrow & Bright (2011)	117 treated and released offenders	STATIC-99	45 months (range 5-87 months) (Australia)	Conviction	Not specified	8.5%	0.679-0.718**	29

Appendix H: Likert scale questionnaire: Phase one

EVALUATION OF THE ACTIVE RISK MANAGEMENT SYSTEM (ARMS) PILOT STUDY

THANK YOU FOR AGREEING TO TAKE PART IN THIS STUDY. PLEASE READ THE FOLLOWING STATEMENTS AND INDICATE HOW MUCH YOU AGREE WITH THEM ON A SCALE OF 1 TO 5

1 = STRONGLY DISAGREE
2 = DISAGREE
3 = NEITHER AGREE NOR DISAGREE
4 = AGREE
5 = STRONGLY AGREE

	Please Circle				
The training I received in ARMS will be sufficient for me to use it in my daily duties	1	2	3	4	5
There are aspects of the training that I would change	1	2	3	4	5
I am confident I will be able to complete ARMS	1	2	3	4	5
I anticipate I will use ARMS with all of the offenders I manage	1	2	3	4	5
ARMS can be completed in a reasonable time	1	2	3	4	5
ARMS is user friendly	1	2	3	4	5
The level of paperwork required for ARMS is manageable in the course of my daily duties	1	2	3	4	5
I would recommend ARMS to my colleagues	1	2	3	4	5
The training has improved my confidence in assessing sex offenders	1	2	3	4	5
I would like to use ARMS during the course of my daily duties	1	2	3	4	5
I am confident that ARMS will reduce the resources required to manage offenders	1	2	3	4	5
I believe offender managers will welcome training in ARMS	1	2	3	4	5
I believe offender managers will benefit from training in ARMS	1	2	3	4	5
ARMS will reduce the workload of offender managers	1	2	3	4	5
My knowledge of psychosexual development has improved through this training event	1	2	3	4	5
I feel more able to address sexual issues with offenders	1	2	3	4	5
The training will enable me to feel less anxious in addressing sexual matters with offenders	1	2	3	4	5
The training I have received will enable me to be more confident in discussing sexual issues with the offenders	1	2	3	4	5
I believe I will be able to transfer what I have learned to my daily practice	1	2	3	4	5
Some offender managers struggle to discuss sexual issues with offenders they manage	1	2	3	4	5
Offender managers require training to address sensitive sexual issues with offenders	1	2	3	4	5

I believe day 3 of the training enhanced the overall training experience	1	2	3	4	5
Day 3 should be incorporated within any future training events	1	2	3	4	5
I would like to see ARMS rolled out nationally	1	2	3	4	5

THANK YOU FOR YOUR TIME

Appendix I: Likert scale questionnaire: Phase two

EVALUATION OF THE ACTIVE RISK MANAGEMENT SYSTEM (ARMS) PILOT STUDY

THANK YOU FOR AGREEING TO TAKE PART IN THIS STUDY. PLEASE READ THE FOLLOWING STATEMENTS AND INDICATE HOW MUCH YOU AGREE WITH THEM ON A SCALE OF 1 TO 5

1 = STRONGLY DISAGREE
2 = DISAGREE
3 = NEITHER AGREE NOR DISAGREE
4 = AGREE
5 = STRONGLY AGREE

	Please Circle				
The training I received in ARMS was sufficient for me to train others	1	2	3	4	5
There are aspects of the training that I would change	1	2	3	4	5
I am confident in training others in the use of ARMS	1	2	3	4	5
I will train all offender managers in the use of ARMS	1	2	3	4	5
I am confident that ARMS will reduce the resources required to manage offenders	1	2	3	4	5
I believe offender managers will welcome training in ARMS	1	2	3	4	5
I believe offender managers will benefit from training in ARMS	1	2	3	4	5
ARMS will reduce the workload of offender managers	1	2	3	4	5
The training I have received will enable me to be more confident in discussing sexual issues with the offenders I manage	1	2	3	4	5
I have learned through this training event to address sexual issues with offenders in different ways	1	2	3	4	5
Offender managers require training to address sensitive sexual issues with offenders	1	2	3	4	5
Some offender managers struggle to discuss sexual issues with offenders they manage	1	2	3	4	5
I have learned how to address sexual issues with offenders that may have caused me discomfort previously	1	2	3	4	5
I have learned through this training event how to empower offender managers in addressing sexual issues with offenders	1	2	3	4	5

THANK YOU FOR YOUR TIME

Appendix J: Focus group interview schedule

Evaluation of the Active Risk Management System (ARMS) Pilot Study
Interview Schedule

What were your general thoughts
What did you learn
What areas need to be addressed further
Will it impact upon your practice
What did you feel about it

Reflections on the Training

- Introduction to training
- Risk Factors
- Desistance and protective factors
- Case study evaluations
- Case study assessments
- Case formulation
- Practitioner assessment
- Interview techniques and strategy
- Asking the questions/challenging
- Worksheets/handouts helpful
- Language and how we use it

Introduction

- What were your initial thoughts about doing the training
- Did you have a choice in whether to attend the training
- How long have you worked with sexual offenders
- What are your initial thoughts on using ARMS within your daily work

General Delivery

- In what ways was the training helpful
- In what ways was the training unhelpful
- What would you change about the training
- How well informed were the trainers
- Would you recommend the training to your colleagues

ARMS Tool

- Relevant factors
- Experience of collating information
- Helpfulness of rating guide
- Repetition in factors
- Salient factors
- Overall helpfulness
- Changes

Staff Outcomes

- Confident in ARMS

- Impact on Risk Management
- Strengths – weaknesses
- Helpful/unhelpful factors
- Personal impact

Offender Outcomes

- Engagement/collaboration
- Awareness/ownership of risk
- Motivation
- Anything else

Partnership Working

- Agency sharing
- Method of sharing
- Impact of sharing
- Benefits and challenges

Resources

- Replaced/added to other tools/practices
- Level of training
- Requirements to use to full potential

General

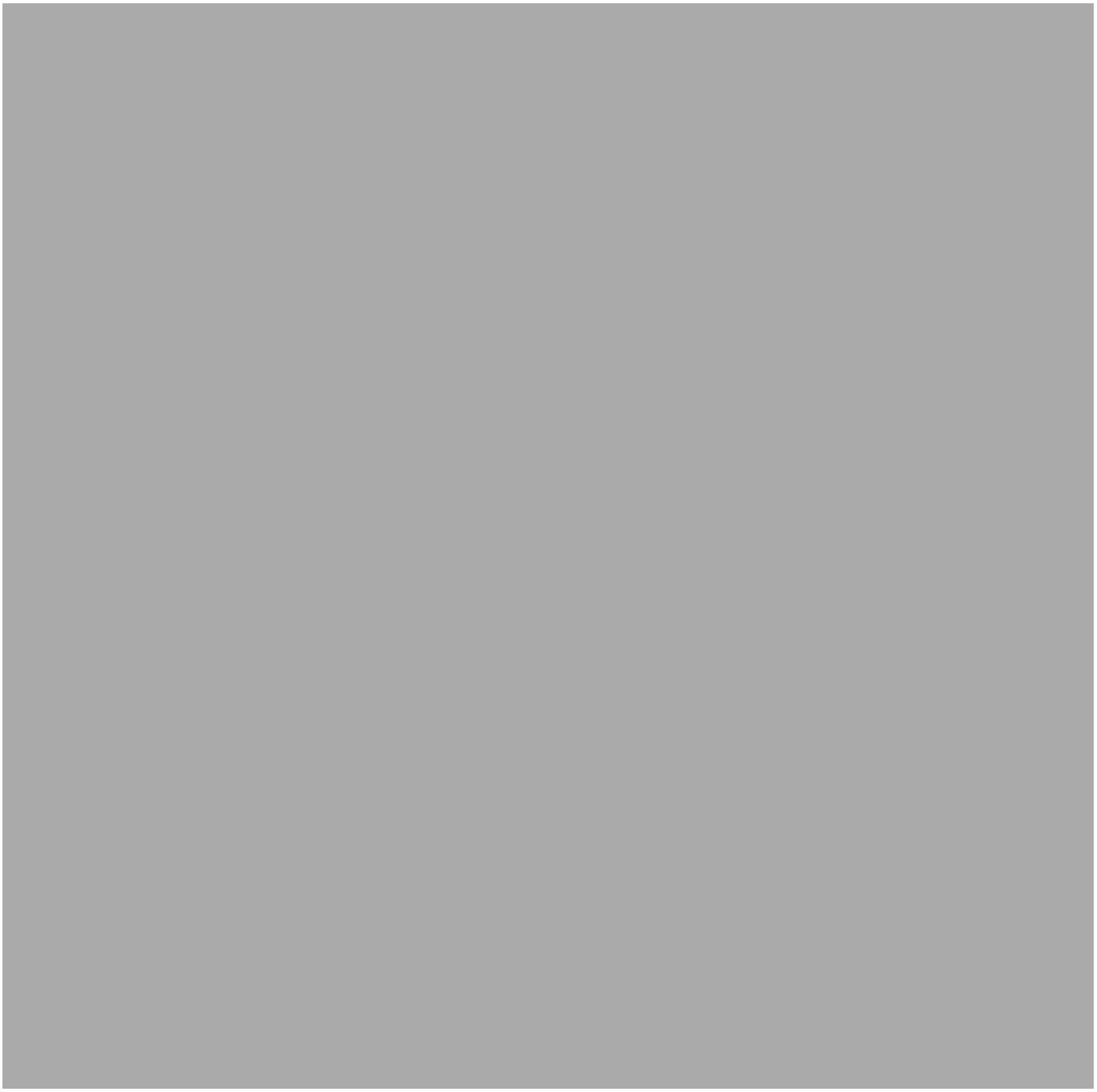
- Confidence to roll out
- Changes
- Messages for developer

Appendix K: Participant information leaflet

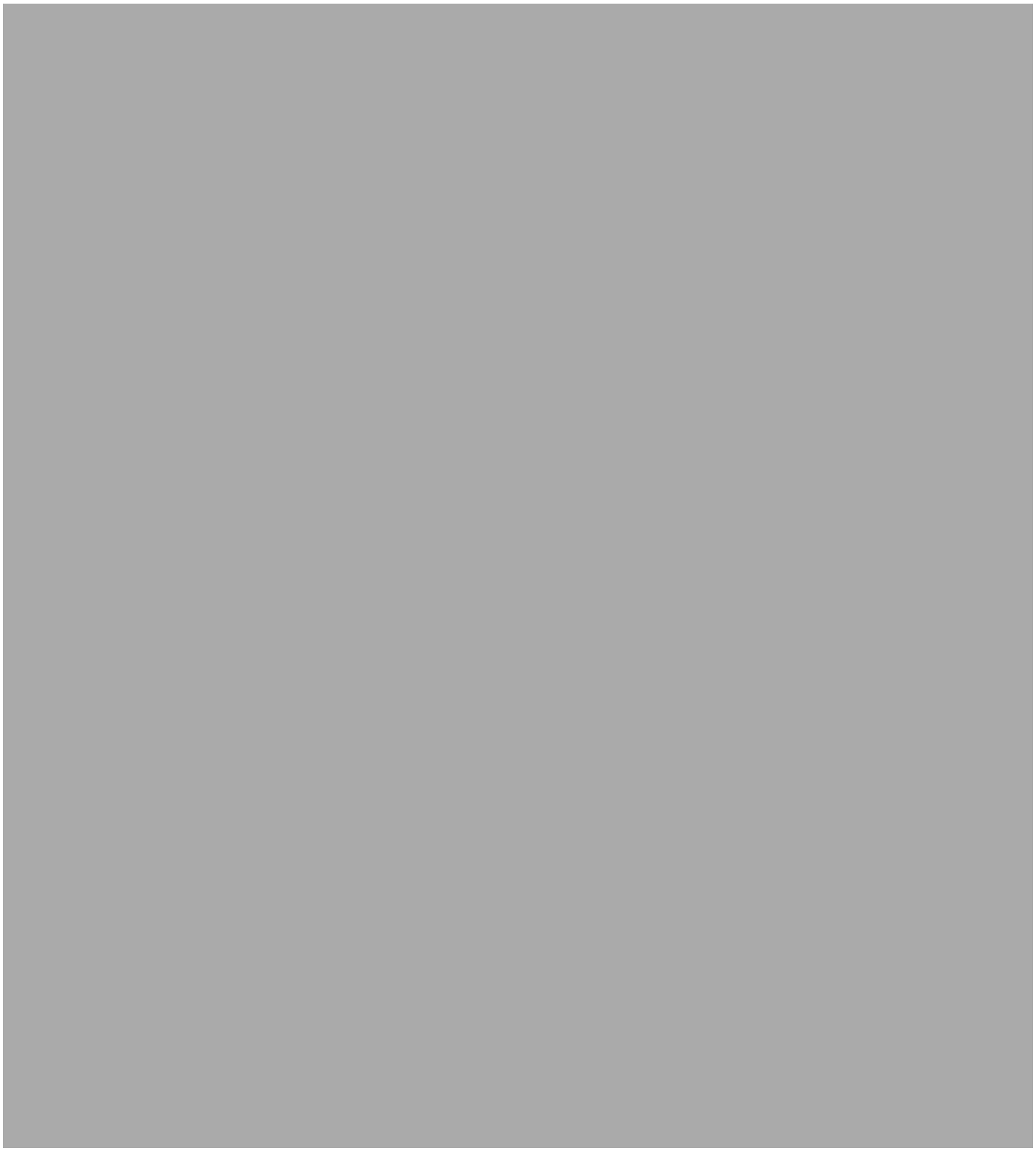
Information Leaflet for Prospective Participants

Title of the Proposed Study
Evaluation of the Active Risk Management System (ARMS) Pilot Study

Description of the Proposed Study
Over the past several months, you have been employing the Active Risk Management System (ARMS) in the course of your work managing registered sex offenders within the community. The current study aims to evaluate the utility of this tool by employing a two staged approach and providing an evaluative report incorporating both quantitative and qualitative analyses.



Appendix L: Participant consent form



Appendix M: SPSS output: Phase three

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Age	434	18	84	39.58	14.669
Nonsexual	434	0	54	1.70	4.895
Sexual	434	1	49	4.30	5.682
Valid N (listwise)	434				

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.900
Bartlett's Test of Sphericity	Approx. Chi-Square	1735.116
	df	55
	Sig.	.000

Correlations

		RM2000S	ARMS
RM2000S	Pearson Correlation	1	.291**
	Sig. (2-tailed)		.000
	N	434	434
ARMS	Pearson Correlation	.291**	1
	Sig. (2-tailed)	.000	
	N	434	434

**. Correlation is significant at the 0.01 level (2-tailed).

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	145.320	11	13.211	70.515	.000[b]
	Residual	72.504	387	.187		
	Total	217.824	398			

a. Dependent Variable: ARMS

b. Predictors: (Constant), Socialinvestment, Relationship, Congruence, Hostile, Selfmanagement, Preoccupation, Socialinfluences, Opportunity, Desistance, Employment, Interests

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.157	.092		1.697	.091
	Opportunity	.242	.042	.240	5.772	.000
	Preoccupation	.110	.042	.108	2.615	.009
	Interests	.113	.045	.114	2.535	.012
	Congruence	.035	.045	.029	.785	.433
	Hostile	.025	.038	.024	.658	.511
	Selfmanagement	.096	.043	.090	2.247	.025
	Socialinfluences	.185	.045	.168	4.163	.000
	Desistance	.122	.042	.122	2.922	.004
	Relationship	-.078	.045	-.054	-1.733	.084
	Employment	.120	.043	.117	2.761	.006
	Socialinvestment	.085	.041	.084	2.041	.042

a. Dependent Variable: ARMS

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	98.003	1	98.003	324.711	.000[b]
	Residual	119.821	397	.302		
	Total	217.824	398			
2	Regression	122.970	2	61.485	256.689	.000[c]
	Residual	94.854	396	.240		
	Total	217.824	398			
3	Regression	131.309	3	43.770	199.839	.000[d]
	Residual	86.515	395	.219		
	Total	217.824	398			
4	Regression	137.615	4	34.404	168.998	.000[e]
	Residual	80.209	394	.204		
	Total	217.824	398			
5	Regression	140.945	5	28.189	144.100	.000[f]
	Residual	76.879	393	.196		
	Total	217.824	398			
6	Regression	142.725	6	23.788	124.166	.000[g]
	Residual	75.099	392	.192		
	Total	217.824	398			
7	Regression	143.690	7	20.527	108.266	.000[h]

	Residual	74.134	391	.190		
	Total	217.824	398			
8	Regression	144.589	8	18.074	96.248	.000[i]
	Residual	73.235	390	.188		
	Total	217.824	398			

a. Dependent Variable: ARMS

b. Predictors: (Constant), Opportunity

c. Predictors: (Constant), Opportunity, Socialinfluences

d. Predictors: (Constant), Opportunity, Socialinfluences, Employment

e. Predictors: (Constant), Opportunity, Socialinfluences, Employment, Preoccupation

f. Predictors: (Constant), Opportunity, Socialinfluences, Employment, Preoccupation, Desistance

g. Predictors: (Constant), Opportunity, Socialinfluences, Employment, Preoccupation, Desistance, Interests

h. Predictors: (Constant), Opportunity, Socialinfluences, Employment, Preoccupation, Desistance, Interests, Socialinvestment

i. Predictors: (Constant), Opportunity, Socialinfluences, Employment, Preoccupation, Desistance, Interests, Socialinvestment, Selfmanagement

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.824	.078		10.595	.000
	Opportunity	.679	.038	.671	18.020	.000
2	(Constant)	.461	.078		5.911	.000
	Opportunity	.487	.038	.481	12.663	.000
	Socialinfluences	.428	.042	.388	10.209	.000
3	(Constant)	.300	.079		3.794	.000
	Opportunity	.437	.038	.432	11.621	.000
	Socialinfluences	.322	.044	.292	7.376	.000
	Employment	.239	.039	.234	6.170	.000
4	(Constant)	.201	.078		2.565	.011
	Opportunity	.337	.040	.333	8.328	.000
	Socialinfluences	.294	.042	.266	6.935	.000
	Employment	.236	.037	.231	6.333	.000
	Preoccupation	.208	.037	.205	5.566	.000
5	(Constant)	.157	.077		2.026	.043
	Opportunity	.292	.041	.289	7.110	.000
	Socialinfluences	.242	.043	.220	5.597	.000
	Employment	.205	.037	.200	5.469	.000

	Preoccupation	.189	.037	.187	5.141	.000
	Desistance	.164	.040	.165	4.126	.000
6	(Constant)	.140	.077		1.825	.069
	Opportunity	.259	.042	.256	6.155	.000
	Socialinfluences	.232	.043	.210	5.388	.000
	Employment	.198	.037	.194	5.353	.000
	Preoccupation	.126	.042	.125	3.008	.003
	Desistance	.163	.039	.163	4.126	.000
	Interests	.129	.042	.130	3.048	.002
7	(Constant)	.113	.077		1.457	.146
	Opportunity	.253	.042	.250	6.025	.000
	Socialinfluences	.217	.043	.197	5.006	.000
	Employment	.163	.040	.159	4.048	.000
	Preoccupation	.123	.042	.121	2.942	.003
	Desistance	.148	.040	.148	3.715	.000
	Interests	.120	.042	.121	2.835	.005
	Socialinvestment	.093	.041	.092	2.256	.025
8	(Constant)	.096	.077		1.239	.216
	Opportunity	.245	.042	.242	5.824	.000
	Socialinfluences	.200	.044	.182	4.579	.000
	Employment	.126	.043	.123	2.893	.004
	Preoccupation	.114	.042	.112	2.724	.007
	Desistance	.138	.040	.138	3.455	.001
	Interests	.126	.042	.127	2.985	.003
	Socialinvestment	.091	.041	.090	2.230	.026
	Selfmanagement	.093	.042	.087	2.188	.029

a. Dependent Variable: ARMS

Excluded Variables[a]

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics Tolerance
1	Preoccupation	.259[b]	6.071	.000	.292	.700
	Interests	.283[b]	6.521	.000	.311	.666
	Congruence	.192[b]	4.827	.000	.236	.833
	Hostile	.216[b]	5.547	.000	.269	.849
	Selfmanagement	.315[b]	8.231	.000	.382	.809
	Socialinfluences	.388[b]	10.209	.000	.456	.761
	Desistance	.355[b]	8.729	.000	.402	.704

	Relationship	-.159[b]	-4.290	.000	-.211	.964
	Employment	.344[b]	9.276	.000	.423	.831
	Socialinvestment	.339[b]	8.818	.000	.405	.787
2	Preoccupation	.208[c]	5.382	.000	.261	.688
	Interests	.218[c]	5.481	.000	.266	.646
	Congruence	.147[c]	4.079	.000	.201	.820
	Hostile	.124[c]	3.357	.001	.167	.784
	Selfmanagement	.201[c]	5.196	.000	.253	.691
	Desistance	.233[c]	5.652	.000	.274	.598
	Relationship	-.077[c]	-2.229	.026	-.111	.903
	Employment	.234[c]	6.170	.000	.297	.702
	Socialinvestment	.226[c]	5.832	.000	.282	.674
3	Preoccupation	.205[d]	5.566	.000	.270	.687
	Interests	.205[d]	5.373	.000	.261	.644
	Congruence	.128[d]	3.710	.000	.184	.813
	Hostile	.085[d]	2.359	.019	.118	.757
	Selfmanagement	.120[d]	2.842	.005	.142	.555
	Desistance	.189[d]	4.631	.000	.227	.572
	Relationship	-.066[d]	-1.974	.049	-.099	.900
	Socialinvestment	.152[d]	3.623	.000	.180	.551
4	Interests	.132[e]	3.045	.002	.152	.484
	Congruence	.077[e]	2.181	.030	.109	.735
	Hostile	.072[e]	2.051	.041	.103	.753
	Selfmanagement	.099[e]	2.428	.016	.122	.550
	Desistance	.165[e]	4.126	.000	.204	.564
	Relationship	-.063[e]	-1.974	.049	-.099	.900
	Socialinvestment	.128[e]	3.124	.002	.156	.544
5	Interests	.130[f]	3.048	.002	.152	.484
	Congruence	.064[f]	1.833	.068	.092	.728
	Hostile	.034[f]	.956	.340	.048	.693
	Selfmanagement	.081[f]	1.993	.047	.100	.542
	Relationship	-.052[f]	-1.634	.103	-.082	.892
	Socialinvestment	.103[f]	2.516	.012	.126	.529
6	Congruence	.031[g]	.848	.397	.043	.642
	Hostile	.033[g]	.925	.355	.047	.692
	Selfmanagement	.089[g]	2.215	.027	.111	.540
	Relationship	-.051[g]	-1.615	.107	-.081	.892
	Socialinvestment	.092[g]	2.256	.025	.113	.524
7	Congruence	.026[h]	.714	.475	.036	.640
	Hostile	.023[h]	.650	.516	.033	.681

	Selfmanagement	.087[h]	2.188	.029	.110	.540
	Relationship	-.049[h]	-1.578	.115	-.080	.891
8	Congruence	.029[i]	.793	.429	.040	.639
	Hostile	.017[i]	.490	.624	.025	.678
	Relationship	-.053[i]	-1.701	.090	-.086	.889

a. Dependent Variable: ARMS

b. Predictors in the Model: (Constant), Opportunity

c. Predictors in the Model: (Constant), Opportunity, Socialinfluences

d. Predictors in the Model: (Constant), Opportunity, Socialinfluences, Employment

e. Predictors in the Model: (Constant), Opportunity, Socialinfluences, Employment, Preoccupation

f. Predictors in the Model: (Constant), Opportunity, Socialinfluences, Employment, Preoccupation, Desistance

g. Predictors in the Model: (Constant), Opportunity, Socialinfluences, Employment, Preoccupation, Desistance, Interests

h. Predictors in the Model: (Constant), Opportunity, Socialinfluences, Employment, Preoccupation, Desistance, Interests, Socialinvestment

i. Predictors in the Model: (Constant), Opportunity, Socialinfluences, Employment, Preoccupation, Desistance, Interests, Socialinvestment, Selfmanagement

Correlations

		ARMS	Opportunity	Preoccupation	Interests	Congruence	Hostile	Selfmanagement	Socialinfluences	Desistance	Relationship	Employment	Socialinvestment
Pearson Correlation	ARMS	1.000	.671	.548	.576	.434	.444	.548	.623	.615	-.281	.561	.576
	Opportunity	.671	1.000	.548	.578	.408	.388	.437	.488	.544	-.190	.411	.461
	Preoccupation	.548	.548	1.000	.664	.469	.294	.339	.366	.411	-.143	.274	.364
	Interests	.576	.578	.664	1.000	.573	.302	.314	.405	.411	-.161	.326	.413
	Congruence	.434	.408	.469	.573	1.000	.223	.242	.301	.360	-.141	.277	.346
	Hostile	.444	.388	.294	.302	.223	1.000	.391	.412	.511	-.104	.391	.436
	Selfmanagement	.548	.437	.339	.314	.242	.391	1.000	.513	.481	-.161	.612	.469
	Socialinfluences	.623	.488	.366	.405	.301	.412	.513	1.000	.550	-.308	.515	.519
	Desistance	.615	.544	.411	.411	.360	.511	.481	.550	1.000	-.260	.475	.517
	Relationship	-.281	-.190	-.143	-.161	-.141	-.104	-.161	-.308	-.260	1.000	-.213	-.217
	Employment	.561	.411	.274	.326	.277	.391	.612	.515	.475	-.213	1.000	.604
	Socialinvestment	.576	.461	.364	.413	.346	.436	.469	.519	.517	-.217	.604	1.000
Sig. (1-tailed)	ARMS	.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Opportunity	.000	.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Preoccupation	.000	.000	.	.000	.000	.000	.000	.000	.000	.002	.000	.000
	Interests	.000	.000	.000	.	.000	.000	.000	.000	.000	.000	.000	.000
	Congruence	.000	.000	.000	.000	.	.000	.000	.000	.000	.002	.000	.000
	Hostile	.000	.000	.000	.000	.000	.	.000	.000	.000	.015	.000	.000
	Selfmanagement	.000	.000	.000	.000	.000	.000	.	.000	.000	.000	.000	.000
	Socialinfluences	.000	.000	.000	.000	.000	.000	.000	.	.000	.000	.000	.000
	Desistance	.000	.000	.000	.000	.000	.000	.000	.000	.	.000	.000	.000
	Relationship	.000	.000	.002	.000	.002	.015	.000	.000	.000	.	.000	.000

	Employment	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.	.000
	Socialinvestment	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.
N	ARMS	434	433	425	425	428	433	428	423	421	434	425	407
	Opportunity	433	433	425	425	427	432	427	422	420	433	425	406
	Preoccupation	425	425	425	418	419	424	419	414	413	425	418	399
	Interests	425	425	418	425	420	424	420	416	414	425	417	399
	Congruence	428	427	419	420	428	427	423	418	416	428	419	403
	Hostile	433	432	424	424	427	433	427	422	421	433	424	407
	Selfmanagement	428	427	419	420	423	427	428	421	417	428	423	405
	Socialinfluences	423	422	414	416	418	422	421	423	412	423	417	401
	Desistance	421	420	413	414	416	421	417	412	421	421	413	399
	Relationship	434	433	425	425	428	433	428	423	421	434	425	407
	Employment	425	425	418	417	419	424	423	417	413	425	425	402
	Socialinvestment	407	406	399	399	403	407	405	401	399	407	402	407

Model Summary[i]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.671[a]	.450	.449	.549	.450	324.711	1	397	.000
2	.751[b]	.565	.562	.489	.115	104.232	1	396	.000
3	.776[c]	.603	.600	.468	.038	38.075	1	395	.000
4	.795[d]	.632	.628	.451	.029	30.977	1	394	.000
5	.804[e]	.647	.643	.442	.015	17.020	1	393	.000
6	.809[f]	.655	.650	.438	.008	9.293	1	392	.002

7	.812[g]	.660	.654	.435	.004	5.091	1	391	.025
8	.815[h]	.664	.657	.433	.004	4.786	1	390	.029

a. Predictors: (Constant), Opportunity

b. Predictors: (Constant), Opportunity, Socialinfluences

c. Predictors: (Constant), Opportunity, Socialinfluences, Employment

d. Predictors: (Constant), Opportunity, Socialinfluences, Employment, Preoccupation

e. Predictors: (Constant), Opportunity, Socialinfluences, Employment, Preoccupation, Desistance

f. Predictors: (Constant), Opportunity, Socialinfluences, Employment, Preoccupation, Desistance, Interests

g. Predictors: (Constant), Opportunity, Socialinfluences, Employment, Preoccupation, Desistance, Interests, Socialinvestment

h. Predictors: (Constant), Opportunity, Socialinfluences, Employment, Preoccupation, Desistance, Interests, Socialinvestment, Selfmanagement

i. Dependent Variable: ARMS

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	98.003	1	98.003	324.711	.000[b]
	Residual	119.821	397	.302		
	Total	217.824	398			
2	Regression	122.970	2	61.485	256.689	.000[c]
	Residual	94.854	396	.240		
	Total	217.824	398			
3	Regression	131.309	3	43.770	199.839	.000[d]
	Residual	86.515	395	.219		
	Total	217.824	398			
4	Regression	137.615	4	34.404	168.998	.000[e]
	Residual	80.209	394	.204		
	Total	217.824	398			
5	Regression	140.945	5	28.189	144.100	.000[f]
	Residual	76.879	393	.196		
	Total	217.824	398			
6	Regression	142.725	6	23.788	124.166	.000[g]
	Residual	75.099	392	.192		
	Total	217.824	398			
7	Regression	143.690	7	20.527	108.266	.000[h]
	Residual	74.134	391	.190		
	Total	217.824	398			
8	Regression	144.589	8	18.074	96.248	.000[i]
	Residual	73.235	390	.188		
	Total	217.824	398			

a. Dependent Variable: ARMS

b. Predictors: (Constant), Opportunity

c. Predictors: (Constant), Opportunity, Socialinfluences

d. Predictors: (Constant), Opportunity, Socialinfluences, Employment

e. Predictors: (Constant), Opportunity, Socialinfluences, Employment, Preoccupation

f. Predictors: (Constant), Opportunity, Socialinfluences, Employment, Preoccupation, Desistance

g. Predictors: (Constant), Opportunity, Socialinfluences, Employment, Preoccupation, Desistance, Interests

h. Predictors: (Constant), Opportunity, Socialinfluences, Employment, Preoccupation, Desistance, Interests, Socialinvestment

i. Predictors: (Constant), Opportunity, Socialinfluences, Employment, Preoccupation, Desistance, Interests, Socialinvestment, Selfmanagement

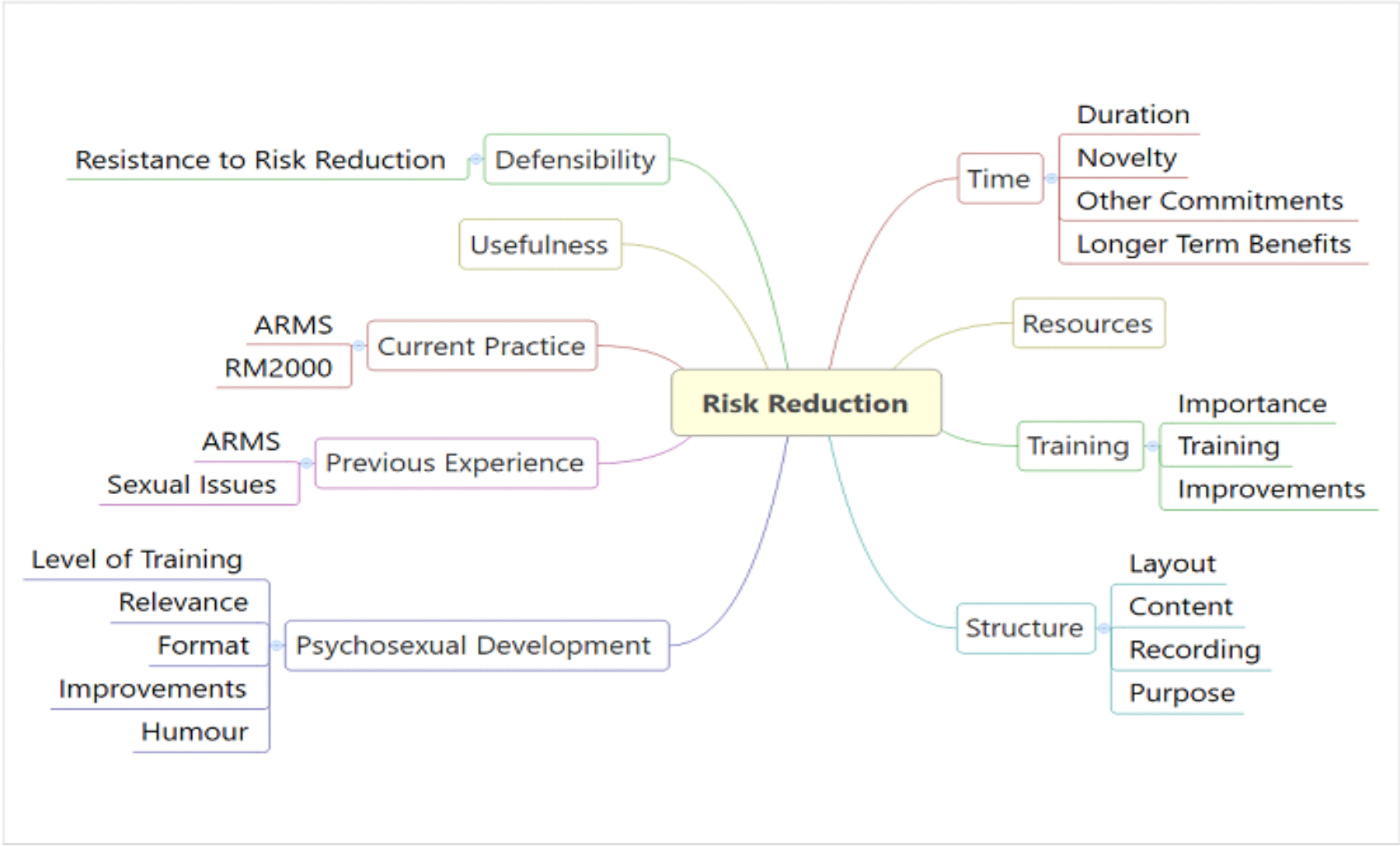
Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	.824	.078		10.595	.000					
	Opportunity	.679	.038	.671	18.020	.000	.671	.671	.671	1.000	1.000
2	(Constant)	.461	.078		5.911	.000					
	Opportunity	.487	.038	.481	12.663	.000	.671	.537	.420	.761	1.313
	Socialinfluences	.428	.042	.388	10.209	.000	.623	.456	.339	.761	1.313
3	(Constant)	.300	.079		3.794	.000					
	Opportunity	.437	.038	.432	11.621	.000	.671	.505	.369	.727	1.376
	Socialinfluences	.322	.044	.292	7.376	.000	.623	.348	.234	.643	1.555
	Employment	.239	.039	.234	6.170	.000	.561	.297	.196	.702	1.425
4	(Constant)	.201	.078		2.565	.011					
	Opportunity	.337	.040	.333	8.328	.000	.671	.387	.255	.583	1.714
	Socialinfluences	.294	.042	.266	6.935	.000	.623	.330	.212	.634	1.578
	Employment	.236	.037	.231	6.333	.000	.561	.304	.194	.702	1.425
	Preoccupation	.208	.037	.205	5.566	.000	.548	.270	.170	.687	1.455
5	(Constant)	.157	.077		2.026	.043					
	Opportunity	.292	.041	.289	7.110	.000	.671	.338	.213	.543	1.842
	Socialinfluences	.242	.043	.220	5.597	.000	.623	.272	.168	.582	1.719
	Employment	.205	.037	.200	5.469	.000	.561	.266	.164	.672	1.489
	Preoccupation	.189	.037	.187	5.141	.000	.548	.251	.154	.678	1.476
	Desistance	.164	.040	.165	4.126	.000	.615	.204	.124	.564	1.773
6	(Constant)	.140	.077		1.825	.069					
	Opportunity	.259	.042	.256	6.155	.000	.671	.297	.183	.507	1.973

	Socialinfluences	.232	.043	.210	5.388	.000	.623	.263	.160	.578	1.730
	Employment	.198	.037	.194	5.353	.000	.561	.261	.159	.670	1.493
	Preoccupation	.126	.042	.125	3.008	.003	.548	.150	.089	.512	1.954
	Desistance	.163	.039	.163	4.126	.000	.615	.204	.122	.564	1.773
	Interests	.129	.042	.130	3.048	.002	.576	.152	.090	.484	2.065
7	(Constant)	.113	.077		1.457	.146					
	Opportunity	.253	.042	.250	6.025	.000	.671	.291	.178	.505	1.982
	Socialinfluences	.217	.043	.197	5.006	.000	.623	.245	.148	.564	1.772
	Employment	.163	.040	.159	4.048	.000	.561	.201	.119	.565	1.770
	Preoccupation	.123	.042	.121	2.942	.003	.548	.147	.087	.511	1.956
	Desistance	.148	.040	.148	3.715	.000	.615	.185	.110	.548	1.824
	Interests	.120	.042	.121	2.835	.005	.576	.142	.084	.480	2.084
	Socialinvestment	.093	.041	.092	2.256	.025	.576	.113	.067	.524	1.909
8	(Constant)	.096	.077		1.239	.216					
	Opportunity	.245	.042	.242	5.824	.000	.671	.283	.171	.500	1.999
	Socialinfluences	.200	.044	.182	4.579	.000	.623	.226	.134	.548	1.826
	Employment	.126	.043	.123	2.893	.004	.561	.145	.085	.479	2.088
	Preoccupation	.114	.042	.112	2.724	.007	.548	.137	.080	.506	1.976
	Desistance	.138	.040	.138	3.455	.001	.615	.172	.101	.541	1.848
	Interests	.126	.042	.127	2.985	.003	.576	.149	.088	.478	2.093
	Socialinvestment	.091	.041	.090	2.230	.026	.576	.112	.065	.524	1.909
	Selfmanagement	.093	.042	.087	2.188	.029	.548	.110	.064	.540	1.853

a. Dependent Variable: ARMS

Appendix N: Final thematic map

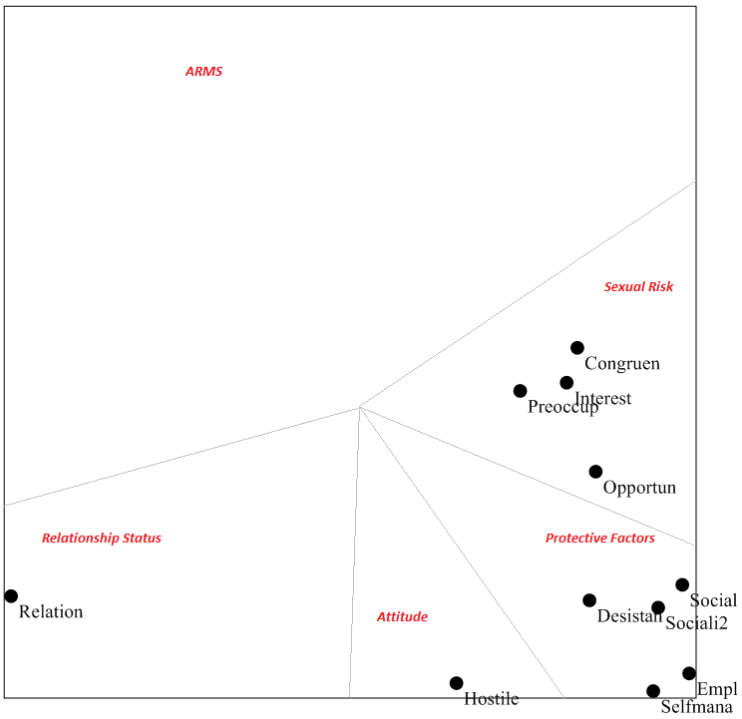


Appendix O: Pattern matrix and factor analysis

	Component		
	Protective factors	**Risk factors**	**Other**
Employment	.846		
Self-management	.817		
Hostile orientation	.720		
Social investment	.700		
Social influences	.641		
Desistance	.626		
Interests		.880	
Preoccupation		.838	
Congruence		.800	
Opportunity	.344	.552	
Relationship			.972

Protective Factors		**Risk Factors**	
Item	**Result**	**Item**	**Result**
Employment	.846	Sexual Interests	.880
Self-Management	.817	Preoccupation	.838
Social Investment	.700	Emotional Congruence	.800
Social Influences	.641		
Desistance	.626		

Appendix P: Multidimensional scaling



Appendix Q: Allocated visits by police officers

	Allocated Visits based on RM2000	Allocated Visits based on ARMS
Low Risk	74	87
Medium Risk	348	414
High Risk	524	536
Very High Risk	660	72
Cumulative Total	1606	1109