AN EXAMINATION OF CURRENT PRACTICE AND NEW DEVELOPMENTS IN THE FORENSIC ASSESSMENT OF OFFENDERS

by

CHLOE LOUISE WHATSON

A thesis submitted to the University of Birmingham for the degree of Doctor in Applied Forensic Psychology Practice (Foren.Psy.D)

Centre for Forensic and Criminological Psychology

School of Psychology

College of Life and Environmental Science

University of Birmingham

B15 2TT

September 2015

Abstract

Exploring and understanding the practice of accurately assessing offenders is an important area of research for forensic practice and the risk management of offenders. Examining the validity, reliability and predictive accuracy of tools used to assess risk of recidivism in forensic mental health settings is important, in order to ensure more accurate risk assessment and management. Furthermore, the inclusion of additional information or approaches in offender assessment such as Index Offence Work (IOW) or Index Offence Analysis (IOA), have been indicated to enhance practitioners' assessments of offenders (West & Greenhall, 2011). They have also been evidenced to enhance the predictive accuracy of existing tools (Lehman, Goodwill, Gallasch-Nemitz, Biedermann & Dahle, 2013), however at present it appears that they are not commonly used within forensic practice.

This thesis aims to explore these different approaches to offender assessment, specifically their clinical utility in forensic mental health settings. The first chapter provides an introduction to the importance of accurate offender assessment and presents current models proposed within the literature to direct practitioners in their work. The second chapter then provides a systematic review of historically used approaches in the assessment of risk (clinical judgement and actuarial assessment), and compares their predictive accuracy and clinical utility, in relation to a mentally disordered offending population. The third chapter, critically appraises a widely used assessment tool to assess risk within forensic mental health settings, the HCR-20 (Version 3.0; Douglas, Hart, Webster & Belfrage, 2013), which utilises a structured professional judgement (SPJ) approach to assessment. The fourth chapter moves away from specific tools used in current practice, instead it explores clinicians' current understanding and use of newer concepts in offender assessment (IOW/IOA), through a qualitative research study.

Finally, the fifth chapter presents a thorough discussion of the overall content, findings, and conclusions of this thesis. This includes implications for forensic research and practice.

This thesis provides support for offender assessment tools currently used by practitioners in forensic mental health settings, although suggests that some have more clinical utility than others. It does however, evidence gaps and inconsistencies in practitioners' understanding and use of IOW/IOA. Findings have important implications for the practice of assessing and managing offenders effectively.

Dedication

This thesis is dedicated to my wonderful grandfather Ron Whatson. Thank you for giving me the courage and confidence to do this. I am so grateful and wish you were here to see me graduate!

Acknowledgements

I would like to express my gratitude and thanks to my academic supervisors, Dr Darren Bishopp and Dr Jessica Woodhams for their knowledge, expertise and support throughout this process. For providing constructive criticism, as well as positive encouragement and enabling me to achieve the final submission of this work. I would also like to thank Sue Hanson, the other staff and my course colleagues for answering questions when needed and always providing encouragement and support.

I would also like to make special thanks to my placement supervisor Sarah Pancholi and colleague Rachel Pike, who throughout this process have provided me with support and have given me invaluable advice on my journey to becoming a qualified psychologist. A thankyou as well to all my friends who have shared the highs and lows with me throughout this process. Special thanks goes to those who have completed the journey with me (Danielle, Lizzi & Nic) – we did it girls! Also Danielle, for getting me through coding and just for generally always being there!

A huge and very special thanks to all of my wonderful family for their love and support but in particular to my mum, dad, brother, sister and Nan who have supported me in every way possible. Thank you for all your guidance, love and unswerving confidence in my ability to achieve this. I am so grateful to all of you for everything. Finally, I would like to acknowledge my partner and best friend, Ashley. Thank you for your constant love and support. For looking after me, putting up with my crazy mood swings and being there no matter what! I could not have achieved this without any of you!

Contents **Page Number**

List of Appendices

List of Tables

List of Figures

CHAPTER ONE

Introduction

Introduction

"The effective assessment and treatment of dangerous offenders has important implications both for society in general and for the offenders themselves" (Harkins, Ware & Mann, 2012, p. 350). It has become a critical aspect of offender treatment both in terms of accurately assessing an individual's risk and appropriately directing their treatment pathway. The assessment of offenders is a core skill used by forensic and clinical psychologists; it requires them to systematically collate a wide variety of information relating to the characteristics of an individual and their offences. As Wood, Garb, Lilienfeld and Nezworski (2002) highlight, assessments carried out by psychologists are increasingly used to inform parole decisions and criminal appeals. Furthermore, Heilbrun (2001) emphasises that clinicians' conclusions will be scrutinized by the legal system and therefore it is their responsibility to provide accurate information. Assessments are therefore of upmost importance and need to be as accurate as possible. Despite this Gacono (2002) highlights that, a growing number of psychologists' are limited in their ability to conduct in-depth assessments of offenders. It is possible that this could be for a variety of factors including; the limited availability of accurate assessment tools, limited access to offender information and a lack of clear guidelines pertaining to the assessment process.

The Ministry of Justice's (MOJ, 2015) most recently published proven reoffending rates, indicate that between October 2012 and September 2013, 60,000 adult offenders were released from custody. Statistics highlight that 27,000 of these individuals' reoffended (45.4%) within one year. Recidivism rates, further emphasise the need for such assessments to be accurate and robust, in order to ensure that offenders address their risk factors through appropriate intervention and therefore reduce their risk, ensuring public protection. This suggests that assessment is more than just part of a practitioner's role, but a big responsibility

which evidently impacts upon wider society. Research has driven the development of approaches to assessment, and within this, psychology has developed a variety of tools and models to assist clinicians in understanding and assessing offenders complex difficulties. Understanding what led to an individual's offence, requires practitioners to understand the person, their life, attitudes and beliefs, coping strategies and core beliefs. Approaches for achieving this have been developed such as formulation or functional analysis (Hart, Sturmey, Logan & McMurran, 2011), as well as more formal models developed for offender assessment and interventions, such as the Risk-Need-Responsivity model (RNR; Andrews, Bonta & Wormith, 2006) and the Good Lives Model (GLM; Ward & Stewart, 2003). Such models were developed in order to assist practitioners in taking a broader and more holistic view of offenders, including their strengths and weaknesses. Despite this, Bonta (2002) highlights that the use of the best and most current offender assessment instruments is not widespread.

Current models of offender assessment

RNR model (Andrews et al., 2006). McGuire (2012, p.316) defines the model as a "risk management rehabilitation model that seeks to reduce offenders' predisposition to reoffend by eradicating, reducing or controlling personality and/or situational variables". The overall aim of the model is to target dynamic risk factors through treatment in order to reduce recidivism rates and it does this through three key principles (see Table 1.). The framework has been instrumental in directing the development and implementation of a number of assessment tools utilised within the criminal justice system (Bonta & Andrews, 2010). Examples of these include: the Offender Group Reconviction Scale (OGRS; Howard, Francis, Soothill & Humphreys, 2009); the Level of Service Inventory-Revised (LSI-R; Andrews & Bonta, 1995); the Offender Assessment System (OASys); Historical, Clinical, Risk-20 (HCR-20; Webster, Douglas, Eaves & Hart, 1997; Douglas, Hart, Webster & Belfrage, 2013);

Structured Assessment of Risk and Need (SARN, Thornton 2002); Risk Matrix 2000 (RM2000; Thornton et al., 2003) and for mentally disordered offenders, the Psychopathy Checklist Revised (PCL-R; Hare, 1991).

Table 1.

Summary of the Key Principles of the RNR Model

Risk Principle	Need Principle	Responsivity Principle
• The intervention should be equivalent and appropriate to meet the offenders level of risk • Offenders who pose a greater risk should receive higher levels of intervention and resources • Accurate assessments of offenders' risks are requires in order for practitioners to appropriately allocate them to treatment	• Focus on treatment targets • Interventions should target needs/risk factors (criminogenic needs)which are subject to change • Treatment should target factors that are relevant to each individual offender, e.g. violent supportive attitudes and anti-social peers	• Interventions which meet an offenders preferred learning style and method of delivery will be most effective • To result in the greatest level of therapeutic change, it should take into account: cognitive ability; motivation; maturity; personal/inter-personal characteristics • Interventions therefore must be responsive to offenders needs

As Table 1. indicates, the risk principle highlights the importance of the accurate assessment of offenders' risk, in order to appropriately design interventions to address these. Whilst the efficacy of risk assessment relies on the ability of assessments to accurately predict future behaviour (Hatcher, 2012), conducting accurate and robust assessments is also the

responsibility of clinicians working within the field and is dependent on the information that is available to conduct such assessments. As Hatcher asserts, the consequence of incorrect prediction can be problematic. For example, someone who is assessed as unlikely to recidivate may be released from custody. If this prediction is incorrect and the offender goes on to reoffend, this has a real and detrimental impact upon society and the offender as well.

The framework suggests that interventions that focus on targeting non-criminogenic needs, for example self-esteem or communication skills, are non-essential and as such should be secondary to criminogenic needs. Given the often complex nature of offenders, particularly those in forensic mental health settings who present with an array of difficulties, focusing only on criminogenic needs both during assessment and intervention would seem inappropriate. The focus on risk and criminogenic needs as opposed to broader needs is likely to demotivate the offender. Whilst the RNR model does provide guidance on dynamic risk factors and therefore indicates what clinicians should be targeting via assessment and intervention, it does not provide practitioners with information regarding how to incorporate such factors into clinical formulations, treatment plans and intervention designs (Polascheck, 2012). Such factors indicate and support the ideas that whilst there are models of offender assessment and treatment, more needs to be known about the tools and processes utilised, in order to ensure offenders are assessed and treated according to their needs, and therefore their risk is reduced.

The Good Lives model (GLM; Ward & Stewart, 2003). The GLM, whilst not strictly a model of offender assessment, informs the assessment and treatment of offenders effectively and overcomes some of the shortcomings of the RNR framework. The approach is defined as a strengths based approach and whilst highlighting the importance of practitioners deriving and addressing criminogenic needs in assessment and treatment, it also places an emphasis on paying attention to offenders' non-criminogenic needs, in order to address and manage their

risk effectively (Ward, Mann & Gannon, 2007). Despite this, in the context of offender rehabilitation the majority of offenders' needs may be viewed as being criminogenic.This again highlights the importance of accurate assessment, in order to delineate those which directly impact upon risk and more general non-criminogenic needs.

The model proposes that all human beings inherently strive to achieve a number of 'primary goods' including: life, knowledge, excellence in work, play and agency, inner peace, relatedness, spirituality, happiness and creativity (Ward & Brown 2004). As such, offending occurs when individuals directly, or indirectly, implement problematic strategies in order to achieve such goals. In contrast to the RNR model, Ward, Yates and Willis (2012) suggest that it provides practitioners with guidance in how to engage and motivate offenders. Furthermore, utilising its strength-based approach, it allows practitioners to identify treatment strategies whilst in secure settings and upon release. This then enables offenders to access interventions and services which more readily address the breadth of their complex needs. The model promotes collaborative-assessment, with offenders taking an active role in identifying their primary goods, alongside identifying how they met these through offending previously. Subsequently, the model then lends itself to interventions which allow offenders to develop skills that will enable them to meet such goods in more pro-social ways, promoting and encouraging desistance from offending (Ward et al., 2012). Ward et al. (2012) argue that this rehabilitation framework guides practitioners in their work with offenders. Overall, it again highlights the importance of practitioners taking an all-encompassing approach to assessment, in order to appropriately meet offenders' needs and target risks effectively.

What does this mean for the practice of offender assessment?

As highlighted by Borum (1996), there remains a long-standing controversy about the ability of mental health professionals to assess and particularly predict risk.. Whilst there are models which indicate principles of offender assessment and intervention (i.e. RNR and GLM), these tell us little about the appropriate selection and utility of assessment methods or tools utilised by practitioners in their everyday roles. There remains an ethical and legal obligation to assess and manage offenders accurately and effectively. In order to do this, a greater understanding of assessment methods utilised is required. In addition to this, some highlight that defined clinical guidelines and additional training for professionals may also be required (Borum, 1996). Standardised assessment tools are important in clinicians' assessments of risk in terms of improving the reliability and validity of risk judgements, although as already mentioned, the predictive accuracy and validity of tools do have limitations. What is clear however, is that additional research is required regarding historical and more recent developments in assessment methods, in order to ensure that practitioners are fulfilling their role and contribution to offender management effectively.

Thesis aims:

The overall aim of this thesis is to examine the pertinent issue of the assessment of offenders, with a specific focus on the clinical utility of different assessment methods. It attempts to achieve this aim through several objectives:

1. To assess the predictive accuracy of historical approaches to risk assessment i.e clinical and actuarial and their clinical utility within secure forensic mental health settings. Chapter two presents a review which examines clinical judgement and actuarial approaches to risk assessment in a population of mentally disordered offenders.

Consideration is given to which assessment tool has greater predictive accuracy and clinical utility for this population.

2. To carry out a critique of a widely utilised tool within clinical practice to assess risk in offenders. To achieve this, Chapter three comprises a critical review of the Historical, Clinical, Risk-20: Assessing Risk for Violence, Version 3 (HCR-20 V3; Douglas et al., 2013). The critique reviews the literature in relation to the tool's development, and assesses its reliability and validity by appraising the findings within the literature where others have used it.

3. To explore newer concepts within the literature suggested to improve the assessment of offenders. Chapter four addresses this by presenting a thematic analysis of discussions amongst clinical and forensic psychologists working within secure forensic hospitals, regarding their understanding and use of index offence work/analysis (IOW/IOA). The data extracted from these accounts provides a perspective on the utility of IOW/IOA when assessing offenders.

As such this thesis includes: a systematic literature review, exploring the predictive accuracy and clinical utility of actuarial and clinical judgement approaches to risk assessment, in a mentally disordered offending population; a critique of an SPJ tool, widely used to assess risk of long-term violent recidivism, the HCR-20 V3 (Douglas et al., 2013); and a qualitative exploration of practitioners' understanding, and use of IOW/IOA, in the assessment of offenders within secure forensic mental health settings. The final chapter provides an overall discussion of the work within this thesis and reflects upon what this means for forensic clinicians whose responsibility it is to assess offenders within their everyday role.

CHAPTER TWO

The Accuracy of Clinical Judgements vs. Actuarial Assessments in Predicting Recidivism in a Population of Mentally Disordered Offenders

A Systematic Literature Review

Abstract

Aim

Using a systematic method the review examined the literature base relating to the accuracy of clinical judgement and actuarial assessment, in predicting recidivism in a population of mentally disordered offenders.

Method

Existing reviews were identified in order to establish the requirements for a review in this area. Five electronic databases were searched and all studies were assessed. Specific inclusion and exclusion criteria were applied. Data were extracted from included studies and those meeting the appropriate quality level were reviewed and synthesised. Results were reported and discussed.

Results

13 out of 15 included studies followed a cohort study design, with two taking a case-control approach. There was variance amongst the studies with regard to whether clinical judgement or actuarial assessment had a higher predictive accuracy and therefore the better clinical utility with the population studied.

Conclusions

The fact that all studies included do not directly compare clinical judgement and actuarial assessment limits the conclusions that can be drawn from this review, however it does highlight the need for further research in this area. Due to the large variance in findings, definite conclusions in regard to the predictive accuracy of assessment methods are difficult to make. Conclusions and limitations of the review are identified and discussed.

Introduction

As discussed in Chapter 1, there is an increasing amount of pressure put on clinicians to make decisions regarding offenders' level of risk within forensic settings. Philipse, Koeter, van den Brink, & Van Der Staak, (2004) highlight that "the assessment of reoffending in patients is a daily routine in most branches of forensic mental health care" (p.264). Decision making within psychology as a whole is an uncertain process and within forensic psychology it surrounds identifying, assessing and quantifying the risk of an individual. Risk assessment in this setting occurs on a daily basis and the quality of patient care is often determined by the accuracy of clinical decision making during this process (Dawes, Faust & Meehl, 1989). It is therefore important that the most appropriate and most accurate risk assessment processes are utilised for working with a specific population of offenders.

Risk has been suggested to be a multidimensional concept that looks at an undesirable outcome and the probability of that outcome occurring (Hurst, 2011). It is a complex entity and can be even more complex for clinicians to make predictions about. Risk assessment is an inexact science and therefore ultimately decisions about levels of risk are made based on an individual's clinical judgment. It has been acknowledged within the literature that the accuracy and adequacy of risk predictions, specifically with populations of mentally disordered offenders, has been questionable (Reed, 1997). Ennis & Emery (1979) argued that mental health professionals' predictions of dangerous behaviour, were incorrect 95% of the time. This therefore indicates that the accuracy of judgements relating to risk varies and may be dependent on the professional's discipline and experience.

There are two major approaches to risk assessment that have been widely discussed within the psychological literature, clinical judgement and actuarial assessment. The most

common approach historically used by clinicians is unstructured, clinical or professional judgement. As defined by Aeigisdottir et al. (2006, p.342), clinical judgement or prediction refers to "any judgement using informal or intuitive processes to combine or integrate client data". This process relies on expertise in gathering, interpreting and assimilating large amounts of information regarding a patient or client. Alternatively actuarial assessment, has been suggested to be a method which strives to achieve accurate predictions with validated instruments and algorithms (Falzer, 2013).

Clinical Judgement

The use of the clinical judgement approach allows the professional to have complete control over which information is considered to inform their judgement of an individual's risk. Hart (1998) proposed that one advantage of using this method for risk assessment is that it is flexible and allows a case-specific approach. In contrast to this however, Hart also highlights that the approach has low interrater reliability and that decisions made by clinicians who fail to justify these, are difficult to question. Further criticism of the approach has been made by Grove & Meehl, (1996) who criticised the approach for being unstructured, subjective and suggestive. Kemshall (1996) adds to this by suggesting that this inherent bias is due to the fact that information is based upon interviewing, observation and self-report. Buchanan (1999) in the context of the prediction of violence risk, alternatively argues that using clinical judgement to predict risk is an advantage, as it focuses on the mechanisms by which violence occurs and thus enhances the validity of risk assessment. Although it could be argued that assessment and prediction may be different things and as such require different methods.

Some research into clinical judgement in the past twenty years has been much more optimistic in its ability to accurately predict recidivism (Lidz, Mulvey & Gardener, 1993;

Borum, 1996; Monahan, 1996). Lidz et al. found that those who were judged by clinicians as low risk, showed fewer violent incidents in the community, the opposite was found for those judged as high risk. In a review Mossman (1994) found whilst actuarial assessment performed better than clinical judgement in long term follow-up, the average accuracy during short-term follow-ups, were comparable to the average for clinical judgements. It is likely that this is as a result of clinicians being able to judge information accurately in the here and now, when working closely with offenders. Predicting an individual's behaviour in the future however, becomes more problematic, as clinicians may be unable to predict the change of factors/circumstances which may impact upon that individual and their risk, reducing the accuracy of clinical judgement predictions in the longer term. Mills (2005) suggests that in the past twenty years of research, one of the key lessons learned is that clinical judgement is a poor and inconsistent method, by which to make estimates regarding violent recidivism. One reason for this may be as a result of the cognitive biases that occur when humans make such judgements. Tversky and Kahneman (1973) highlighted that the true probabilities or likelihood of events, in this case recidivism, are elusive. Due to the fact that these "cannot be assessed objectively" (p.231). They argue that the only way to understand the role of such cognitive biases and more about why human judgements are too high or low, would be to analyse the heuristics a person uses to judge the probability of an event. Their research looks at availability as one of these heuristics and this may indicate that judgements made are dependent on the information that is available to the assessor, inherently introducing bias into this process.

Actuarial Assessment

Actuarial assessment methods allow clinicians to make decisions based on data which can be coded in a predetermined manner. They predict risk based on the relationship between

specific cues or risk factors and the occurrence of the behaviour, for example violence (Convit, Jaeger, Lin, Meisner & Volavka, 1988). Dolan & Doyle (2000) suggest that decisions regarding risk are determined according to rules and that this approach undoubtedly improves the consistency of risk assessments. One criticism, as proposed by Hart (1998), is that actuarial approaches ignore individual variations or differences in risk, instead focusing on static variables. They therefore fail to prioritise clinically relevant variables and thus cause passive predictions of risk. Additionally, they also have been suggested to limit and undermine the role of the clinician and their experience, as well as the fact that data collection is not standardised and different clinicians will go about the assessment in different ways (Lennings, 2005).

In terms of predictive accuracy for recidivism, Quinsey and colleagues have promoted the validity of actuarial assessments extensively, developing the Violence Risk Appraisal Guide (VRAG; Quinsey, Harris, Rice & Cormier, 1998) and the Sex Offender Risk Appraisal Guide (SORAG; Quinsey et al., 1998). Harris et al. (2003) compared a variety of actuarial assessment tools in the prediction of recidivism in sexual offenders. They found that all four instruments significantly predicted recidivism at a greater accuracy than chance. Barbaree, Seto, Langton & Peacock (2001) also found that the VRAG, SORAG, RRASOR and Static-99 successfully predicted general recidivism.

Clinical Judgment vs. Actuarial Assessment

Clearly there is conflicting evidence within the literature as to the relative strengths and limitations of clinical judgement or actuarial assessment approaches to risk assessment. The research however has gone further than just identifying limitations and looking at predictive accuracy for the individual methods. It has also compared the predictive accuracy and therefore the relative utility of both clinical judgement and actuarial assessment.

Monahan (1984), reviewed 'first generation' research in the clinical vs. actuarial debate and concluded that clinicians were accurate, in no more than one out of three predictions, in relation to their predictions of violent recidivism. In support of this a recent review by Hilton, Harris & Rice (2006) reported that actuarial assessments had an effect size 88% larger than did clinical judgement in predicting sexually violent recidivism. They suggest that this is due to the fact that clinical judgement is less tied to empiricism and that clinical experience adds little to the accuracy of clinical judgement. Although, Gardener, Lidz, Mulvey & Shaw (1996) stated that while actuarial measures performed better than clinical ratings in predicting violent recidivism in mentally ill patients, clinical ratings were still better than chance. In contrast, Fazel, Singh, Doll and Grann (2012) suggest that the predictive validity of actuarial assessments is not high enough to justify their sole use in the risk assessment process.

There is however much empirical evidence to show that clinical judgment is inferior to such formal assessments (Grove & Meehl, 1996; Grove, Zald, Lebow, Snitz, & Nelson, 2000; Hilton, Harris, & Rice, 2006). However, as outlined by Borum, Fein, Vossekuil, & Berglund, (1999), even if this method is consistently superior in terms of predictive accuracy, actuarial methods can only be applied when appropriate measures exist and have been adequately validated on the population in question. Many actuarial tools are developed on specific populations and therefore if they are not relevant to the individual undergoing the risk assessment, then they are unable to be used effectively. In practice this becomes more complex. Clinicians make decisions daily which impact upon whether an individual will be given the opportunity to offend, whereas risk assessment only focuses on the actual recidivism rates of released offenders. As such the accuracy of such tools is influenced somewhat by the accuracy of the clinical judgements which led to an offender's release, therefore highlighting the importance of both approaches.

The requirement for clinicians working with offenders to provide accountable, accurate and transparent assessments of risk, requires greater research into specific populations of offenders. At present there is conflicting evidence within the literature and whilst it generally tends to highlight the superior accuracy of actuarial assessments, this may not be applicable with certain populations such as mentally disordered offenders (i.e. those with a diagnosis of mental illness and/or personality disorder). This suggests that further investigation into this area is appropriate and required.

The Current Review

Scoping exercises which involved a preliminary search prior to commencing the review, were conducted in order to see the potential size of the literature and to identify existing and ongoing reviews in order to avoid duplication. A total of four meta-analyses, one meta-review, one systematic review and meta-regression analysis, and one systematic review were found in this area. Of these, none looked specifically at the comparative predictive accuracy of clinical judgement and actuarial assessment for recidivism, in a mentally disordered offending population. Therefore a review of the literature specifically focusing on this population was deemed to be a valuable addition to the research area.

Aim of the Current Review

This review identified and examined the literature relating to the clinical judgement and actuarial assessment conflict, specifically, in relation to their predictive accuracy in identifying the recidivism of mentally disordered offenders. This review had the following objectives:

- To determine whether clinical judgement or actuarial assessment has better predictive accuracy in identifying recidivism within a mentally disordered offending population

- To determine which assessment method is more applicable and beneficial for use with a mentally disordered offending population

Method

Sources of Information

A scoping exercise was carried out. Searches were conducted using a variety of databases on the 7th May 2015: The Cochrane Database of Systematic Reviews (CDSR); The Campbell Collaboration Library of Systematic Reviews and PsychINFO. In addition, a search was also conducted using Google Scholar search engine. Four meta-analyses, one meta-review, one systematic review and meta-regression analysis, and one systematic review were found during the scoping exercise. With regard to the four meta-analyses found, none looked specifically at the predictive accuracy of clinical judgement or actuarial assessment in recidivism, in regard to a mentally disordered offending population. Aeigisdottir et al. (2006) in their meta-analysis, looked at clinical vs. 'statistical' methods of prediction in general within the psychological literature. Whilst touching on the development of 'statistical' or actuarial measures in forensic settings, authors did not look at recidivism as an outcome. Additionally, a lot of the literature is dated pre-1987, thus is not part of the 'modern' research base, instead focusing on when decisions were made about 'dangerousness' as opposed to risk. Finally, although effect sizes within the literature reviewed, indicated a 13% increase in accuracy when using statistical compared to clinical methods, some of the instruments reviewed were not specific to forensic risk assessment. Therefore this review was not similar in terms of research question or outcomes to the current review. Meta-analyses by Grove, Zald, Lebow, Snitz, & Nelson, (2000) and Spengler et al. (2009) looked at clinical judgement and actuarial assessment, however this was not in relation to a forensic population. Instead, they examined human health

behaviours and mental health and psychosocial issues, respectively. Finally, in another meta-analysis by Hanson and Morton-Bourgon (2009), the predictive accuracy of various approaches to the prediction of recidivism was analysed. They concluded that, empirically derived actuarial measures were more accurate than unstructured professional judgement, however they looked specifically at recidivism amongst sexual offenders.

Singh, Grann and Fazel (2011) in their systematic review and meta-regression analysis, and Fazel et al. (2012) in their systematic review, investigated the predictive validity of specific tools in general and specifically for predicting violent or sexual recidivism. Neither of these reviews however, compare the accuracy of such tools with clinical judgement approaches and only look at a general offending population. Finally, the meta-review by Singh and Fazel (2010), was the most relevant, synthesising nine systematic reviews and 31 meta-analyses between 1995 and 2009 within the forensic risk assessment literature. They were concerned with calculating the probability that anti-social behaviour or criminal, violent or sexual offending will occur. Examining a variety of themes including, validity of actuarial tools compared with unstructured and structured clinical judgement; a comparison of risk assessment tools and the predictive validity of these tools for ages and gender. Authors found mixed evidence regarding the predictive accuracy of clinical judgement and actuarial assessment on recidivism in the reviews and meta-analyses. Whilst authors examined the predictive validity of tools for different genders and ethnicities, they did not specifically examine the effectiveness of actuarial assessment or clinical judgement in mentally disordered offenders. Furthermore, the mixed research questions which were addressed resulted in mixed findings and as such indicate a need for further investigation of specific topic areas.

Based on the reviews found during the scoping exercise, there is a need for a more specific systematic review which analyses specifically the predictive accuracy of clinical

judgement and actuarial assessment on recidivism in a mentally disordered population. In addition there is a need for the more recent literature to be reviewed as often reviews have looked at more historical research.

Search Strategy

On 8th May 2015 an electronic search was conducted. Five databases were searched: Psych INFO; Medline; EMBASE; Web of Science and ProQuest. Additionally, Google Scholar search engine was also utilised. The closest date parameters from 1987 to the present (May Week 1 2015) were set where possible: Psych INFO (1967 to May Week 1 2015); Medline (1946 to May Week 1 2015); EMBASE (1974 to 2015 May 07); Web of Science (After 1987) and ProQuest (After 1987). On Google Scholar this was not possible so papers from all dates were retrieved. Searches were also restricted by language to include English papers only and document type (grey literature was excluded due to the large volume of studies in this area).

Additionally bibliographies of retrieved papers and reviews were hand searched for relevant studies based on PICO, (see below). Several key authors in this area were also contacted, with regard to obtaining papers as well as with regard to unpublished studies (papers in preparation) or information about other pertinent studies in the area that they would consider important for review. Three authors responded with papers to be included in the review.

Search Terms

When conducting the search, keyword and other search terms associated with risk assessment, clinical judgement, offenders and mentally disordered offenders were used. Where possible, mapping to subject headings was utilised in the searches to maximise the inclusivity of available literature. In addition, key word searching was also used in order to account for

variation in coding across the different databases. Wildcards were applied to searches, where appropriate, to additionally maximise the amount of articles that were sourced.

(Risk assessment*) OR (actuarial assessment*) OR (assessment* tool*)

AND

(Structured clinical judgement*) OR (unstructured clinical judgement*) OR (clinical judgement*) OR (professional judgement*) OR (decision making) OR ("Clinical Judgment (Not Diagnosis)") OR (psychological assessment) OR (forensic evaluation) OR (clinical decision making)

AND

(predict*) adj2 (recid*) OR (predict*) NEAR/2 (recid*)

AND

(offend*) OR (criminal*) OR (convict*) OR (delinquent*)

AND

(patient*) OR (mental* ill*) OR (mental* disorder*) OR (inpatient*) OR (psychiatric patient*) OR (mentally ill offender) OR (forensic mental health) OR (mentally ill persons) OR (mental patient)

The search syntax for each database can be found in Appendix A.

Study Selection

Following the searches having been conducted using electronic databases, Google Scholar, hand searching and contacting experts, a total of 3,524 citations were found. After accounting for duplicates (*n*=689), title and abstracts were reviewed and obviously irrelevant papers excluded (*n*=2,747). Finally the inclusion and exclusion criteria and PICO (Table 2.) were applied to remaining papers, 71 papers were excluded at this point (please see Appendix B.), full text articles were reviewed where there was insufficient information provided, leaving a total of 17 papers for review. Figure 1. provides an overview of the process of study selection.

Table 2.

Inclusion/Exclusion (PICO) Criteria

PICO	Inclusion	Exclusion
Population	• Male and/or female adult offenders • Mentally disordered offenders • Offenders residing in or being discharged from forensic secure mental health settings • Include individuals with criminal convictions with and without charges	• Juvenile offenders (below 18 years) • Older adults (aged 75 plus) • Offenders residing in or being discharged from prison settings • Offenders who are in community settings • Other non-forensic settings
Intervention	• The practical use of clinical judgement and/or actuarial assessment tools in risk assessment	• Papers that compare clinical judgement and/or actuarial assessment to the structured professional judgement approach
Comparator	N/A	N/A
Outcome	• The efficacy of clinical judgement in predicting: reoffending; reconviction; recidivism or repeat violent or sexual behaviour AND/OR • The efficacy of actuarial tools in predicting; reoffending, reconviction, recidivism, repeat violent or sexual behaviour	
Study Type	• Study type: Any, prospective and retrospective designs to be included • Papers from 1987 onwards • Published, peer reviewed papers	• Papers prior to 1987 • Opinion papers, commentaries, reviews, unpublished dissertations, books (Grey literature)

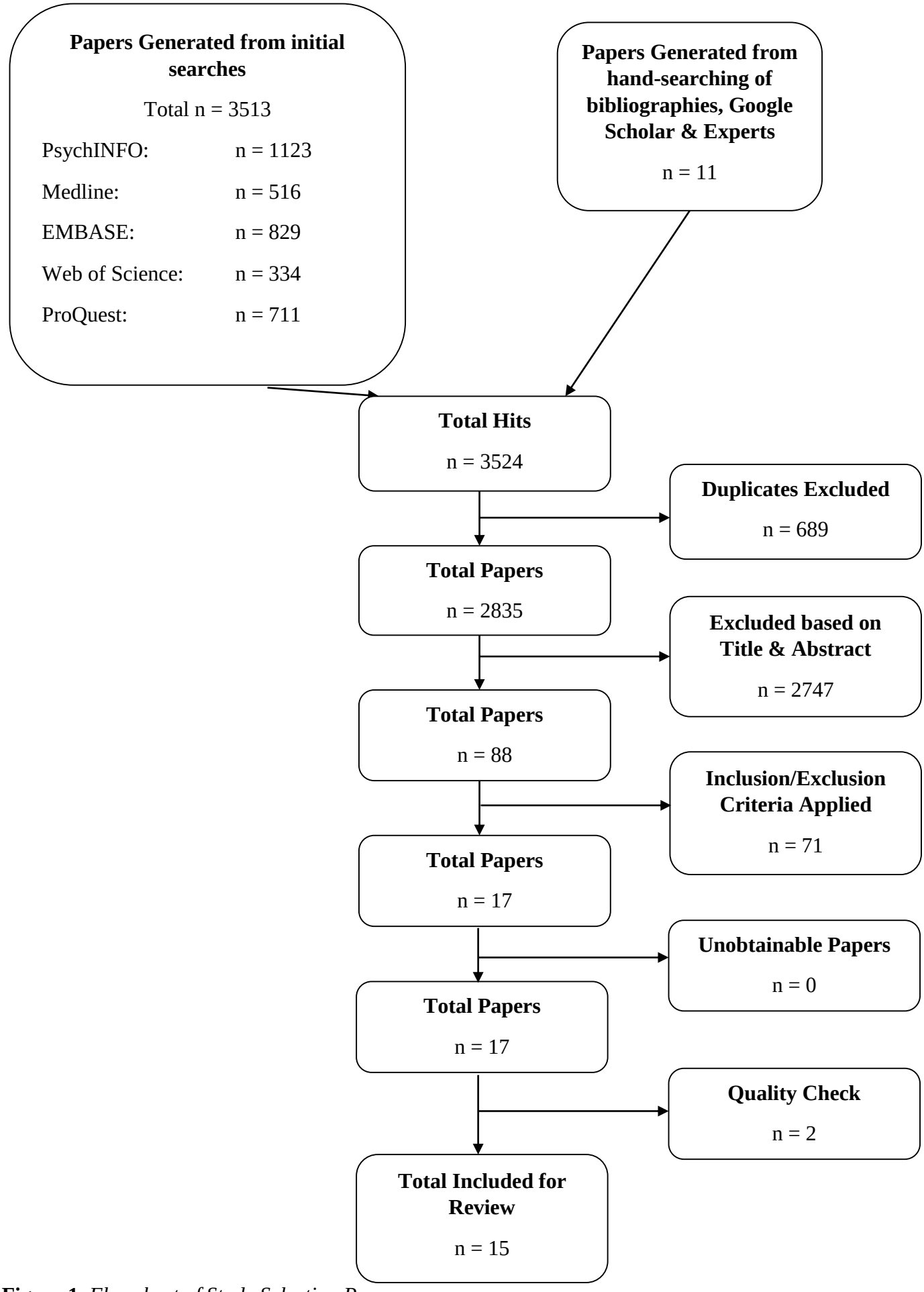


Figure 1. *Flowchart of Study Selection Process.*

Quality Assessment

Following the process of inclusion and exclusion of studies, the studies that met criteria were assessed for methodological quality. The criteria for quality assessment was based upon the Critical Appraisal Skills Programme (CASP) website. Two quality assessments were developed based on checklists for cohort and case-control studies and were adapted in order to make it applicable to studies being reviewed (see Appendix C.). Quality assessments contained screening questions in order to verify that papers selected met inclusion criteria. The cohort quality assessment contained a total of 13 questions and assessed if the cohort was appropriately recruited, whether recidivism was appropriately measured, whether authors accounted for confounding variables and whether the cohort were followed-up for a sufficient time-period. Case-control quality assessments looked at whether biases were addressed during the selection process, the measurement of recidivism and the appropriate use of statistics. This quality assessment also had a total of 13 questions, therefore a maximum score of 26 could be achieved on either of the assessments utilised. On both assessments criteria was scored as 2, 1 or 0 where:

2 = Yes, fully meets criteria

1 = Unclear/ Insufficient information provided

0 = No, does not meet criteria

For each score that was achieved this was converted into a percentage in order to determine its inclusion in the final data synthesis stages. A cut-off score of 60% was selected for inclusion in the final review. It was determined that due to the limited number of studies specifically looking at this area with this population, as well as the lack of randomised control trials, that this would be an appropriate cut-off. Two papers were excluded due to not meeting the cut-off for quality, instead achieving 54% (Fuller & Cowan, 1999; Hilton & Simmons,

2001). Following quality assessment a total of 15 papers were included for review. Quality assessment scores (QAS) and strengths and limitations of included studies can be found in Table 6.

In order to ensure the reliability of the quality assessment, 50% of the quality-checked studies were assessed by a second reviewer, a colleague of the current author. An interrater reliability analysis using the Kappa statistic was run in order to determine the level of agreement between the raters. The results were, Kappa = 0.76, which is described by Vierra and Garrett (2005) as representing *substantial agreement*.

Data Extraction

Following the quality assessment of all included studies, data was extracted from all articles included for review using a data extraction form (Appendix D.), this was developed in order to enhance the information gained through the quality assessment and was carried out for each study. The data extraction form obtained further information into the study's methodology. All information extracted from studies included can be found in Tables 3, 4 and 5.

Results

Following the process of study selection, a total of 15 studies were included for review (Appendix E.). Tables 3, 4 and 5 summarise the key characteristics from each included study, as identified during the data extraction process. The QAS for each of the included studies can be found in Table 6. The QAS scores for studies ranged from 62% (Monahan et al., 2000) to 92% (Bengtson & Langstrom, 2007).

Table 3.

Participant Characteristics of Included Studies

Authors / Year	Sample Size	Gender	Age	Ethnicity	Mental Illness	Offence
Bengtson (2008)	304	Male	Mean at time of release = 32.7 (SD 10.5, range 18-67)	93% Danish	15% medico-legal insanity declaration 96% of medico-legal subjects diagnosed as having a non-psychotic psychiatric disorder (e.g. personality disorder, developmentally disabled, mildly retarded etc.)	160 Rapists 144 Child Molesters 23% sentenced previously for a sexual offence 19% sentenced previously for non-sexual violence 49% sentenced previously for any crime 60 (50%) Rapists
Bengtson & Langström (2007)	121	Male	Not stated	Not stated	Not stated	51 (42%) Extra-Familial Child Molesters

Authors / Year	Sample Size	Gender	Age	Ethnicity	Mental Illness	Offence
						8 (7%) Intra-Familial Child Molesters 2 (2%) Exhibitionists
Brouillette-Alarie & Proulx (2013)	711	Male	18-77 (Mean = 40.85, SD = 12.08) at release.	Not stated.	Not stated.	352 sexual aggressors of children. 251 sexual aggressors of women. 90 mixed offenders. 18 offenders with unknown victims. All offenders had at least one hands on sexual offence.
Ferguson, Ogloff & Thomson (2009)	208	Male 157 (75.5%) Female 51 (24.5%)	17-64 (Mean = 30.77, SD = 9.87)	Caucasian 155 (74.5%) Asian 16 (7.7%) Aboriginal 14 (6.7%) Other 23 (11.1%)	Schizophrenia or psychotic 148 (66.8%) Other 50 (24.0%)	Not stated.

Authors / Year	Sample Size	Gender	Age	Ethnicity	Mental Illness	Offence
Grann, Belfrage & Tengstrom (2000)	560 PD Cohort = 358 Scizophrenia cohort = 202	PD Cohort: Male 322 (90%) Female 36 (10%) Schizophrenia cohort: Male	PD Cohort mean age = 32 (SD=) Schizophrenia cohort = 33 (SD=9.1)	Not stated.	PD Cohort: 62% Concomitant abuse/dependency on alcohol (51%) and/or drugs (27%). Schizophrenia cohort: 50% Concomitant abuse/dependency on alcohol (37%) and /or drugs (33%).	Violent offenders.
Hanson, Helmus & Thornton (2010)	Total = 3,304 Psychiatric =311	Male	Total mean = 39 (SD=12) Psychiatric mean = 33 (SD = 10)	Not stated.	Not stated.	Total: Rapists (39%), Child molesters (53%) Psychiatric: Rapists (50%), Child molesters (49%)

Authors / Year	Sample Size	Gender	Age	Ethnicity	Mental Illness	Offence
Hanson & Thornton (2000)	Total = 1,210 Secure psychiatric = 486	Male	Pinel Secure Psychiatric mean age at release = 36.2 (SD= 10.9) Oak Ridge Secure Psychiatric mean age at release = 30.4 (SD= 9.5)	Not stated.	Not stated.	Pinel: Child molesters 70.4%; Oak Ridge: Child molesters 49.3%
Harris, Rice & Cormier (2002)	406	Male (mostly) Female	Not stated.	Not stated.	Violent cohort: 72% Psychotic diagnosis; 25% Personality Disorder; 85% not guilty by reason of insanity. Non-violent cohort: 86% Psychotic diagnosis, 8% Personality Disorder; 73% Not guilty by reason of insanity.	Not stated.
Huss & Zeiss (2004)	Cases = 26 violent patients	Male		Caucasian 62%	Schizophrenia 57.6%	Not stated.

Authors / Year	Sample Size	Gender	Age	Ethnicity	Mental Illness	Offence
	Matched controls=26 non-violent patients		Mean = 35.4 years (SD = 9.4)	African American 26.6% Hispanic 5.4% Asian-American 3.7% Pacific islander 2%	Schizoaffective disorder 15.3% Bipolar 11% Schizotypal 3.7% Substance use 6.5%, where primary diagnosis, 19.2% had a secondary diagnosis pf substance use disorder Cognitive disorders 4.2% Paraphilia 1.7%	
Kroner, Stadtland, Eidt & Nedopil (2007)	113	Male = 93 Female = 20	Mean age at accusation = 33.7 years (SD= 11.0)	Not stated.	63 people diagnosed with a mental illness according to the ICD-10.	58 offences were violent & 55 non-violent. Biggest offence groups were murder (22), theft (18) and causing bodily harm (17).

Authors / Year	Sample Size	Gender	Age	Ethnicity	Mental Illness	Offence
McNiel, Sandberg & Binder (1998)	317	Male 164 (52%) Female 153 (48%)	Mean = 42.2 (SD = 16.2)	White 217 (69%) African American 52 (16%) Asian American 31 (10%) Other 17 (5%)	Schizophrenia 72 (23%) Manic Disorders 70 (22%) Major Depressive Disorder 44 (14%) Unspecified Psychotic Conditions 31 (10%) Adjustment disorders 36 (11%) Organic psychotic conditions 31 (10%) Other 27 (9%)	Not Stated
Monahan et al (2000)	1136	Male 57.3% Female 42.7%	Mean age = 29.9 (SD=6.2 years)	White 68.7% African American 29.1% Hispanic 2.2%	Depression 41.9% Alcohol/drug abuse/dependence 21.8% Schizophrenia 17%	Not stated.

Authors / Year	Sample Size	Gender	Age	Ethnicity	Mental Illness	Offence
					Bipolar 14.1% Personality Disorder only 2.1% Other psychotic disorder 3.1% Major mental disorders with a co-occurring diagnosis of substance abuse/dependence: Depression 49.6% Schizophrenia 41% Bipolar disorder 37.7% Other psychotic disorder 45%	
Odeh, Zeiss & Huss (2006)	52	Male	Mean = 35.4	Caucasian 62%	Schizophrenia 57.6%	Not Stated

Authors / Year	Sample Size	Gender	Age	Ethnicity	Mental Illness	Offence
				African American 26.6% Hispanic 5.4% Asian American 3.7% Pacific Islanders 2%	Schizo-Affective Disorder 15.3% Bipolar Disorder 11% Schizotypal Personality Disorder 3.7% Substance Abuse 6.5% Organic Disorders 4.2% Paraphilias 1.7%	
Snowden, Gray, Taylor & MacCulloch (2007)	996	Male	Mean age at discharge = 37.7 (SD=9.2, range 16.9-71.2)	Caucasian 69.2% Black Caribbean/African 21.6% Asian 2.4% Mixed 1.5% Unknown 5.2%	Affective disorder 9.9% Personality Disorder 9% Schizophrenia/Psychotic disorder 56.2% Drug-induced psychosis 4.7% Mental Retardation 8.5% 'Other' diagnoses 8.4% Unknown 3.2%	Not stated.

Authors / Year	Sample Size	Gender	Age	Ethnicity	Mental Illness	Offence
Tengstrom (2001)	106	Male	Mean = 33.07 (SD = 10.57)	Not stated.	Schizophrenia 100%	Attempted/completed homicide 4%Assault 53% Unlawful threats 16% Sex Crimes 9% Armed Robbery 10% Kidnapping 3%

Table 4.

Data Extraction of Included Studies

Author / Year	Sample Structure	Assessment Method	Standardised Measures Used	Variables	Statistical Analyses
Bengstson (2008)	304 Sexual offenders underwent a formal psychiatric evaluation between Jan 1978 and Dec 1992	Actuarial Assessment	Static-99 Static-2002 RM-2000	Socio-demographics: age, ethnicity, marital status, education level Forensic history: previous offences, substance misuse Clinical: Time at risk, length of follow-up	ROC Analysis Sensitivity, Specificty, NPV and PPV
Bengtson & Langström (2007)	121 Sexual Offenders subjected to pre-trial forensic psychiatric assessment	Actuarial Assessment vs. Clinical Judgement	Static -99 Static-2002	Socio-demographics: age, ethnicity, marital/cohabitation, education level, employment. Forensic history: previous sexual offences Clinical: length of follow-up	X^2 and Mann-Whitney U Test Spearman's rank ROC analysis Sensitivity, Specificity, NPV and PPV

Author / Year	Sample Structure	Assessment Method	Standardised Measures Used	Variables	Statistical Analyses
Brouillette-Alarie & Proux (2013)	Database of 711 adult male sexual offenders. Drawn from two institutions: Maximum Security Psychiatric Facility and Maximum security penitentiary.	Actuarial assessment	Static-99R	Socio-demographics: age, cohabitation Forensic history: Static-99R items Clinical: length of follow-up	Factor Analysis Cox regression analysis ROC analysis
Ferguson, Ogloff & Thomson (2009)	208 participants admitted to a secure mental health facility in Victoria, Australia.	Actuarial assessment	LSI-R:SV	Socio-demographics: gender, age, ethnicity, education, employment, source of income, marital status Forensic history: LSI-R:SV items, substance abuse, legal status	ANOVA Chi-Square ROC analysis
Grann, Belfrage & Tengstrom (2000)	560 offenders convicted of violent crimes and subjected to court	Actuarial assessment	VRAG H-10 (Historical part of HCR-20)	Socio-demographics: age, gender	T-test ROC analysis

Author / Year	Sample Structure	Assessment Method	Standardised Measures Used	Variables	Statistical Analyses
	ordered forensic psychiatric evaluations in Sweden during 1988-1993.			Forensic history: VRAG and H-10 items Clinical: diagnosis, length of follow-up	Sensitivity, specificity, PPV, NPV.
Hanson, Helmus & Thornton (2010)	3034 raw datasets from nine samples obtained representing all known static-2002 replications as of December 2006. 311 psychiatric patients who had pre-trial forensic psychiatric evaluations between 1978 and 1992 at two settings in Denmark, offenders deemed high risk by courts, suspected of mental disorder and accused of serious offences.	Actuarial assessment	Static-2002 Static-99	Socio-demographics: age Forensic history: previous offences Clinical: length of follow-up	ROC analysis Logistic Regression
Hanson & Thornton (2000)	Institute Phillipe Pinel: 344 psychiatric	Actuarial assessment	RRASOR SACT-MIN	Socio demographics: Age at release.	ROC analysis

Author / Year	Sample Structure	Assessment Method	Standardised Measures Used	Variables	Statistical Analyses
	patients, sexual offenders treated between 1978 &1993. Oak Ridge: 142 psychiatric patients, sexual offenders referred for treatment or assessment between 1972 &1993.		STATIC-99	Forensic history: Prior offences Clinical variables: Averages years of follow-up	Pearson's correlation coefficient Survival analysis
Harris, Rice & Cormier (2002)	467 forensic patients occupying beds in secure psychiatric units in Ontario in June 1990. 406 of these were deemed to have the opportunity to recidivate.	Actuarial assessment vs. Clinical judgment.	VRAG	Socio-demographics: childhood history, adult adjustment Forensic history: Offence history, offence characteristics Clinical variables: Length of follow-up, scores on clinical scales.	Cox regression Pearson's correlation co-efficient ROC analysis
Huss & Zeiss (2004)		Clinical Judgement	None.	Socio-demographics: age, gender, marital status.	Chi-square

Author / Year	Sample Structure	Assessment Method	Standardised Measures Used	Variables	Statistical Analyses
	Cases: 26 vignettes from patients who had been assaultive Matched control: 26 vignettes from patients who were on the same inpatient unit at about the same time as the cases.			Clinical variables: Length of follow-up, clinicians decision vs. aggregate decision	T-test Cox Regression (survival analysis)
Kroner, Stadtland, Eidt & Nedopil (2007)	113 patients accused of crimes and under psychiatric evaluation from criminal responsibility in the department of forensic psychiatry at the University of Munich between 1994 and 1995.	Actuarial assessment	VRAG	Socio-demographics: Age, Gender, VRAG items. Forensic history: Offence history, current offence, VRAG items. Clinical variables: Length of follow-up.	Kaplan-Meier analysis ROC analysis
McNiel, Sandberg & Binder (1998)	317 patients resident in short-term inpatient psychiatric unit	Clinical Judgement	Overt Aggression Scale	Clinical: Clinician confidence in judgements	Logistic Regression analysis ROC analysis

Author / Year	Sample Structure	Assessment Method	Standardised Measures Used	Variables	Statistical Analyses
					Chi-Square
Monahan et al (2000)	1136 admissions sampled from acute psychiatric facility at three sites in America.	Actuarial Assessment	ICT	Socio-demographics: age, social networks Clinical: ICT items, follow-up	ROC analysis
Odeh, Zeiss & Huss (2006)	26 violent & 26 non-violent patients in admission evaluations during the first 24 hours of hospital stay	Clinical Judgement	None	Socio-demographics: age, ethnicity Forensic history: previous assaults Clinical: Clinician occupation	Logistic Regression Analysis Linear Regression Analysis
Snowden, Gray, Taylor & MacCulloch (2007)	996 patients discharged from four independent medium secure units in the UK.	Actuarial assessment	VRAG OGRS	Socio-demographics: age, ethnicity Forensic history: VRAG and OGRS items Clinical: Time to offence	ROC analysis

Author / Year	Sample Structure	Assessment Method	Standardised Measures Used	Variables	Statistical Analyses
Tengstrom (2001)	106 offenders referred for the first time to a court-ordered pre-trial forensic psychiatric assessment between 1988-1993.	Actuarial assessment	VRAG H-10 (HCR-20)	Socio-demographics: age, VRAG and H-10 items Forensic history: offence type, VRAG and H-10 items Clinical: Time of follow-up	ROC analysis

Table 5.

Design Characteristics of Included Studies

Authors / Year	Study Title	Design	Research Question	Outcome Indicators	Results/Main Findings
Bengtson (2008)	Is never better? A cross-validation of the Static-2002 in a Danish sample of sexual offenders	Retrospective Cohort	To assess the ability of three actuarial risk assessment tools in predicting any sexual, non-sexual violent and any violent recidivism	Accuracy in predicting recidivism during follow-up (Mean = 16. 2 years, SD = 4.3, range = 1.90-24.2). Recidivism defined as: 1) A sexual offence 2) A non-sexual violent offence 3) Any violent offence	All three instruments successfully predicted any sexual, non-sexual violent and violent recidivism with moderate accuracy for child molesters, lower accuracy for the total sample and the rapist sub-group. For the total cohort and child molesters, any sexual recidivism was predicted with lower accuracy than for non-sexual violent and violent recidivism. The Static-2002 and RM2000 were almost consistently slightly higher than the Static-99 across the three main outcomes.

Authors / Year	Study Title	Design	Research Question	Outcome Indicators	Results/Main Findings
					Existing actuarial assessments may be limited in their practical value.
Bengtson & Langström (2007)	Unguided clinical and actuarial assessment of re-offending risk: A direct comparison with sex offenders in Denmark.	Retrospective Cohort	To simultaneously test the accuracy of unstructured clinical judgement-based risk assessments by psychiatrists and actuarial risk-scale based assessments in the same sample of sexual offenders.	Accuracy in predicting sexual and/or violent recidivism during follow-up (Mean = 16.3 years). Recidivism defined as: Any sexual reconviction, any violent reconviction, severe sexual reconviction.	Actuarial instruments identified an increased risk of sexual recidivism during follow-up, association between clinical judgements was less convincing. Actuarial assessments significantly more accurate in predicting any sexual reconviction and severe sexual reconviction. No significant difference between clinical judgement and actuarial assessment in predicting violent reconviction. Both types of assessment performed poorly for short-term predictions, actuarial approach predictions grew

Authors / Year	Study Title	Design	Research Question	Outcome Indicators	Results/Main Findings
					stronger as time at risk increased. Socio-demographic variables, forensic history and clinical variables not associated with predictive accuracy.
Brouillette-Alarie & Proux (2013)	Predictive validity of the Static-99R and its dimensions.	Retrospective cohort	To identify the dimensions of the Static-99R, the most commonly used actuarial risk assessment and to test their predictive validity.	Accuracy in predicting sexual, non-sexual violent and non-sexual non-violent recidivism. Recidivism defined as; a new charge or conviction subsequent to assessment with the Static-99R.	Psychiatric offenders had higher Static-99R scores than prison offenders. Psychiatric offenders had a higher sexual recidivism rate than non-psychiatric offenders. The Static-99 and the Static-99R significantly predicted non-sexual violent recidivism in the psychiatric offenders. The Static-99R was a good predictor of sexual recidivism for the whole sample as well as with sexual aggressors of

Authors / Year	Study Title	Design	Research Question	Outcome Indicators	Results/Main Findings
					women and sexual aggressors of children. This was better than the Static-99 which was only moderate.
Ferguson, Ogloff & Thomson (2009)	Predicting recidivism by mentally disordered offenders using the LSI-R:SV.	Retrospective cohort	To assess the ability of the LSI-R:SV in predicting recidivism in an Australian forensic psychiatric population of both dually diagnosed and non-substance abusers.	Accuracy in predicting violent and non-violent recidivism for offenders with a mental disorder.	Results indicate that for the sample as a whole the LSI-R:SV significantly predicts recidivism for any new offence and for violent new offences. The LSI-R:SV significantly predicts any new offence and for violent new offences in non-substance abusers, but not in substance abusers. More than half of patients reoffended and two thirds of this is attributed to patients with a dual-diagnosis. Across all new offences, this group was more likely to reoffend than the mentally disordered only group.

Authors / Year	Study Title	Design	Research Question	Outcome Indicators	Results/Main Findings
					For all mentally ill offenders the LSI-R:SV predicts recidivism at a moderate level of accuracy.
Grann, Belfrage & Tengstrom (2000)	Actuarial assessment of risk for violence. Predictive validity of the VRAG and the historical part of the HCR-20	Retrospective Cohort	To test the empirical validity of the VRAG and the historical part of the HCR-20 in a Swedish setting and to further clarify their relative importance in two offender sub-groups: offenders with personality disorder; and offenders with Schizophrenia.	Accuracy in predicting violent reconviction. Violent reconviction = homicide, assault, robbery or rape Dichotomous criterion variable = a violent crime within two years from release or discharge leading to reconviction.	Amongst 404 mentally disordered offenders both the VRAG and H-10 predicted violent reconviction within 2 years from release or discharge significantly better than chance. Despite the VRAG being mathematically optimised to predict violent recidivism it is not any better suited than the non-weighted H-10. In the PD cohort the H-10 predicted violent reconviction better than the VRAG. In the Schizophrenia cohort the H-10 also performed better

Authors / Year	Study Title	Design	Research Question	Outcome Indicators	Results/Main Findings
					than the VRAG in predicting violent reconviction.
Hanson, Helmus & Thornton (2010)	Predicting recidivism amongst sexual offenders: A multi-site study of Static-2002.	Retrospective Cohort	To analyse and assess the predictive accuracy of the Static-2002 on a dataset created from all known Static-2002 studies.	Accuracy in predicting sexual, violent and any recidivism.	For the whole sample the Static-2002 was more accurate than the Static-99 for the prediction of sexual, violent and general recidivism. In the psychiatric part of the sample the Static-2002 had higher predictive accuracy for sexual, violent and any recidivism that the Static-99. Recidivism rates for the psychiatric population who were pre-selected to be high risk were significantly higher for sexual recidivism and violent recidivism.
Hanson & Thornton (2000)	Improving risk assessments for sex offenders: A	Retrospective Cohort	To compare the predictive accuracy of three sex offender's risk-assessment	Accuracy in predicting sexual or any violent	The RRASOR and the SACT-Min showed roughly equivalent predictive accuracy

Authors / Year	Study Title	Design	Research Question	Outcome Indicators	Results/Main Findings
	comparison of three actuarial scales.		measures: the RRASOR; Thornton's SACJ-Min and the Static-99.	recidivism during follow-up. Average follow-up: Pinel = 4 years; Oak Ridge = 10 years.	and the combination of the two scales (Static-99) was more accurate than either original scale. The predictive accuracy of the scales was relatively consistent across samples. For the prediction of sex offence recidivism the Static-99 was more accurate than the RRASOR or the SACJ-Min. For any violent recidivism the Static-99 was more accurate than either of the other two tools. Recidivism rates were similar between the Pinel psychiatric sample and the two prison samples.

Authors / Year	Study Title	Design	Research Question	Outcome Indicators	Results/Main Findings
Harris, Rice & Cormier (2002)	Prospective replication of the violence risk appraisal guide in predicting violent recidivism among forensic patients.	Retrospective Cohort	A prospective replication of the predictive accuracy of the VRAG and a test of the predictive accuracy of clinical assessments of risk. To examine how well clinicians ratings predicted violent recidivism and made any incremental contributions to the validity of VRAG scores.	Accuracy in predicting violent recidivism during follow-up (8 years). Violent recidivism defined as: any criminal charge for a violent offence against persons, which occurred subsequent to the index offence. Also included any actions that resulted in patients being returned to maximum security.	Overall patients who met the criteria for violent recidivism had higher VRAG scores that patients who had the opportunity but did not violently recidivate, The accuracy of VRAG scores in predicting recidivism was related to gender. For male patients the accuracy of the VRAG constituted a large effect size, it was also significantly correlated with the speed of violent recidivism. Composite clinical judgement was related to violent recidivism overall, especially among male patients. However, this was significantly less so than actuarial scores.

Authors / Year	Study Title	Design	Research Question	Outcome Indicators	Results/Main Findings
					The accuracy of the VRAG in predicting violent recidivism was high in a constant five-year follow-up.
Huss & Zeiss (2004)	Clinical assessment of violence from inpatient records: A comparison of individual and aggregate decision making across risk strategies.	Case-control	Assessed mental health professional's ability to make clinical decisions.	Asked clinicians to make predictions for a 2-year follow-up, also looked at probability estimates and predictions for severity. Compared individual decisions to group decisions about violence. Violence= any physical contact between a patient and a victim that "placed a victim in serious jeopardy". Criminal charges and convictions deemed to meet threshold for violence, verbal aggression did not.	There were no significant differences in predictive accuracy of the four types of clinicians (psychologists, psychiatrists, nurses, social workers) across the three methods or across severity ratings. Results suggest that differences exist in comparing individual and group decisions as well as the conceptualisation of risk. Individual clinicians unable to predict actual violence and time until violence whether risk described via dichotomous, probabilities or risk categories.

Authors / Year	Study Title	Design	Research Question	Outcome Indicators	Results/Main Findings
					Individual clinicians were somewhat more accurate in predicting the severity of violence.
Kroner, Stadtland, Eidt & Nedopil (2007)	The validity of the violence risk appraisal guide (VRAG) in predicting criminal recidivism.	Retrospective cohort	To test the predictive validity of the VRAG on a German sample.	Accuracy of the VRAG in predicting violent recidivism vs. non-violent recidivism during time at risk. Mean time at risk = 58.06 months (SD 3.39, range = 1-115 months). Recidivism defines as: committing at least one further offence during the observational period.	VRAG total score showed high predictive accuracy for recidivism. 'Elementary maladjustment' showed highest predictive accuracy as a single variable. When the outcome was limited to violent recidivism, the predictive accuracy of the VRAG total score increased. For patients who had a violent index offence VRAG score significantly predicted general recidivism, however this was

Authors / Year	Study Title	Design	Research Question	Outcome Indicators	Results/Main Findings
				Non-violent = driving without a licence, theft, fraud or exhibitionism. Violent = imminent threat of violence such as armed robbery, as well as bodily harm and sexual contact crimes.	not significant for violent recidivism. Patients with higher VRAG scores were more likely to re-offend and after a shorter time period than those with lower VRAG scores.
McNiel, Sandberg & Binder (1998)	The Relationship Between Confidence and Accuracy in Clinical Assessment of Psychiatric Patients Potential for Violence	Retrospective Cohort	To assess the relationship between confidence and accuracy in clinical evaluation of patients risk of violence following admission to a short-term psychiatric in-patient unit.	Accuracy of clinicians estimates of violence risk. The effect of confidence of clinicians in their judgements on accuracy.	Clinician's estimation of violence risk showed a substantial relationship with actual violent episodes. A strong relationship was found between predicted and actual violence where clinicians had high confidence in their judgements, moderate for moderate confidence and practically no relationship for low confidence. As the level of confidence increased the accuracy of

Authors / Year	Study Title	Design	Research Question	Outcome Indicators	Results/Main Findings
					clinical judgements significantly increased. The confidence clinicians have in their evaluations is an important moderator of predictive validity.
Monahan et al (2000)	Developing a clinically useful actuarial tool for assessing violence risk.	Retrospective Cohort.	Applying the ICT methods to a sample of patients assessed in the MacArthur Violence Risk Assessment Study in order to increase the clinical utility of the actuarial method	Accuracy in predicting the prevalence of violence by discharged patients in the community. Follow-up = 20 weeks. Violence= acts of battery, sexual assaults, weapon use and threats made with weapons.	The ICT method significantly predicted violence in all 11 risk groups. The ICT partitioned 72.6% of the sample into one of two categories with regard to risk of violence. Findings demonstrate that the ICT method may be adapted for clinical use.
Odeh, Zeiss & Huss (2006)	Cues They Use: Clinicians' Endorsement of Risk Cues in Predictions of Dangerousness.	Case-control	To identify cues clinicians use in making clinical judgements regarding future violence and the association between	Cues used to predict future violence in patients.	13 risk cues were found to be significantly correlated with clinicians' dichotomous predictions of violence.

Authors / Year	Study Title	Design	Research Question	Outcome Indicators	Results/Main Findings
			clinicians judgements of future violence and actual future violence.	Ability of clinicians to make judgements about future violence risk.	Clinician occupation, previous assaults, hostility, medication non-compliance, paranoid delusions and family problems cues were found to significantly influence probability predictions of violence. Clinician occupation per se was not found to significantly influence ability in predicting future violence, although differences were found in the risk cues used by clinicians. Risk cues used in prediction of violence were not found to be associated with actual violence outcomes. Clinical predictions were not found to be significantly related to actual violence outcomes.

Authors / Year	Study Title	Design	Research Question	Outcome Indicators	Results/Main Findings
Snowden, Gray, Taylor & MacCulloch (2007)	Actuarial prediction of violent recidivism in mentally disordered offenders.	Prospective Cohort	To compare twoinstruments that differ in what they are designed to predict (general versus violent recidivism) and the inclusion of mental health variables.	Reconvictions during follow-up. 641 patients with OGRS scores, mean follow-up = 6.32 years (SD=2.06, range = 2.02-11.96). 421 patients with VRAG scores, mean follow-up = 6.22 years (SD=2.22, range = 2.02-11.39). Violent offences = all offences classified as violence against a person by the home office. Including: kidnap, criminal damage endangering life, robbery, rape and indecent assault. General offences = all offences including those also classified as violence.	Results confirm the utility of the VRAG and OGRS in predicting violent recidivism over a period of six months to five years in mentally disordered offenders. Both do well in predicting violent recidivism but overestimated the absolute probability of violence in a mentally disordered population. Both tools significantly predicted general and violent recidivism above chance levels. VRAG performs better when predicting in short-term. Whereas OGRS slightly better for longer-term prediction for both general and violent recidivism.

Authors / Year	Study Title	Design	Research Question	Outcome Indicators	Results/Main Findings
Tengstrom (2001)	Long-term predictive validity of the historical factors in two risk assessment instruments in a group of violent offenders with schizophrenia	Retrospective cohort	To investigate the long-term predictive validity using the H-10 and the VRAG in a group of violent offenders with schizophrenia.	Accuracy in predicting violent recidivism during follow-up. Mean follow-up = 86.0 months. Violent recidivism = a reconviction of attempted or completed homicide, assault, all sex crimes, armed robbery, and forcible confinement.	Both the VRAG and H-10 significantly predicted violent recidivism in the sample at a moderate level. The H-10 performed better than the VRAG in predicting violent recidivism. The predictive validity of both tools decreased when psychopathy items were removed. An increase of scores on the VRAG were related to higher absolute risk of recidivating.

Descriptive Data Synthesis

The 15 included studies were investigated from a qualitative position, rather than combined statistically and analysed. Combining results statistically and analysing results in a quantitative format may have rendered results meaningless and therefore this was not deemed to be appropriate.

The studies included within this review were conducted in a variety of countries: one in the UK, two in Denmark, two in Canada, one in Germany, one in Australia, one in America and two in Sweden. Two of the studies took samples from mixed settings, however the 'mentally disordered' sample were taken from Denmark (Hanson, Helmus & Thornton, 2010) and Canada (Hanson & Thornton, 2000). Of the studies included, the majority utilised a prospective or retrospective cohort design, although two studies utilised case-control designs (Huss & Zeiss, 2004; Odeh, Zeiss & Huss, 2006). Both prospective and retrospective cohort designs have advantages and disadvantages. Prospective studies are carried out from the present time into the future and thus have the advantage of being tailored to collect the specific data required, potentially making the data more complete. Although the disadvantage of this is that there may be long follow-up periods. Alternatively retrospective studies whilst carried out at the present time, look back at the past to analyse outcomes. A disadvantage of such methods is that the researcher has limited control over the data collection, resulting in data being inaccurate, inconsistent or incomplete however, such methods mean data is available immediately and is less costly and time consuming than prospective cohort studies. All studies sampled a mentally disordered offending population discharged from secure psychiatric institutions. Two of the included studies sampled mixed populations (Hanson & Thornton, 2000; Hanson et al, 2010), however the results for the mentally disordered sample specifically, could be easily identified and separated out, and therefore studies were included. The majority of samples within the

studies consisted of male and female participants. Across the 15 studies a total of 9,596 participants were accounted for; with a total 5,879 of these participants being specifically a mentally disordered offending population.

Studies differed in their reporting of mean age within each sample, with some studies not reporting this at all (Bengtson & Langstrom, 2007; Harris, Rice & Cormier, 2002). Some studies reported the mean age at release for their sample (Bengtson, 2008; Brouillette-Alarie & Proux, 2013; Hanson & Thornton, 2000; Snowden, Gray, Taylor & MacCulloch, 2007). The mean age at release within these studies was 35.57. The majority of studies just reported a mean age for the sample, 33.86 (Ferguson, Ogloff & Thomson, 2009; Grann, Belfrage & Tengstrom, 2000; Hanson et al, 2010; Huss & Zeiss, 2004; McNiel, Sandberg & Binder, 1998; Monahan et al, 2000; Odeh et al, 2006; Tengstrom, 2001). In contrast one study specified mean age at accusation, 33.7 (Kroner, Stadtland, Eidt & Nedopil, 2007). Ethnicity was reported in seven out of the 15 studies that were reviewed. Between the studies that did report ethnicity, recording patterns were similar, looking at between three and five categories (Bengtson, 2008; Ferguson et al, 2009; Huss & Zeiss, 2004; McNiel et al., 1998; Monahan et al., 2000; Odeh et al., 2006; Snowden et al., 2007).

Regarding specifics detailing participant's mental disorders and offence details, there was variability among the studies. Out of the 15 studies, 10 detailed participant's mental disorders (Bengtson, 2008; Ferguson et al., 2009; Grann et al., 2000; Harris et al., 2002; Huss & Zeiss, 2004; McNiel et al., 1998; Monahan et al., 2000; Odeh et al., 2006; Snowden et al., 2007; Tengstrom, 2001). Recording patterns differed among these studies, with the majority only distinguishing between one to three categories and Monahan et al. (2000) distinguishing between 11. Others looked at around seven categories although looked at slightly different conditions (Huss & Zeiss, 2004; McNiel et al., 1998; Odeh et al., 2006; Snowden et al., 2007).

Those that looked at wider categorical division of mental disorder are perhaps more in line with the diversity that is seen within these settings. In terms of offence details, eight of the 15 studies reported the types of offenders within their sample. Of these, five reported specifics regarding numbers of participants with each offence types (Bengtson, 2008; Bengtson & Langstrom, 2007; Hanson et al., 2010; Kroner et al., 2007; Tengstrom, 2001). In regard to the mental illness and offence details, studies reviewed have taken an either/or approach as only three of the studies report both (Bengtson, 2008; Grann et al., 2000; Tengstrom, 2001). This possibly indicates a flaw in other studies, as it limits the generalisability of their findings, as it does not provide transparent information about the samples studied.

Out of the 15 studies reviewed, two considered and compared actuarial assessment and clinical judgement (Bengtson & Langstrom, 2007; Harris et al., 2002), three looked just at clinical judgement (Huss & Zeiss, 2004; McNiel et al., 1998; Odeh et al., 2006) and the other ten looked solely at actuarial assessment (Bengtson, 2008; Brouillette-Alarie & Proux, 2013; Ferguson et al, 2009; Grann et al., 2000; Hanson et al., 2010; Hanson & Thornton, 2000; Kroner et al., 2007; Monahan et al., 2000; Snowden et al., 2007; Tengstrom, 2001). For those studies that incorporated actuarial assessment, several standardised measures were used across studies. Some looked at actuarial assessments aimed at looking at sexual recidivism: Static-99; Static-99R; Static-2002; RM-2000; RRASOR; SACJ-Min (Bengtson, 2008; Bengtson & Langstrom, 2007; Brouillette-Alarie & Proux, 2013; Hanson et al., 2010; Hanson & Thornton, 2000). Others used standardised measures aimed at predicting violent recidivism: Violence Risk Appraisal Guide (VRAG) and Historical-10 from the Historical-Clinical Risk-20 (H-10, HCR-20) (Grann et al., 2000; Kroner et al., 2007; Snowden et al., 2007; Tengstrom, 2001). Other standardised measures which looked at general recidivism which were also utilised, including: Level of Service Inventory-Revised: Screening Version (LSI-R:SV; Ferguson et al., 2009),

OGRS (Snowden et al., 2007). Monahan et al. (2000) also utilised a method called the Iterative Classification Tree (ICT). All studies considered the predictive accuracy of the assessment method used in predicting violent, sexual or general recidivism during a set follow-up period. The studies by McNiel et al. (1998) and Odeh et al. (2006) also reviewed accuracy in assessment method of predicting violent recidivism, but additionally looked at the influence of confidence judgements on this and the cues relating to the prediction of outcomes respectively.

Scrutiny of the articles revealed the research question and outcomes being measured by the study. Papers by Bengtson (2008), Hanson et al. (2010), Kroner et al. (2007) and Snowden et al. (2007), revealed some ambiguity in respect of the research question. Research questions or aims were not clearly set out and it is only upon reading both articles in more detail, that the reader is made aware of the specific outcomes being measured. In addition to this, there were differences amongst studies in what outcomes were specifically identified. Nine of the studies specifically defined violent, sexual or general recidivism as an outcome (Bengtson, 2008; Bengtson & Langstrom, 2008; Brouillette-Alarie & Proux, 2013; Hanson et al., 2010; Hanson & Thornton, 2000; Harris et al., 2002; Kroner et al., 2007; Tengstrom, 2001). Others looked at violent reconvictions (Grann et al., 2000; Snowden et al., 2007) and some looked at future violence potential (Huss & Zeiss, 2004; Odeh et al., 2006). Slight differences in outcomes and the definitions used make comparison between studies more difficult and those that did not specifically use recidivism as an outcome may be more limited.

There are some similarities across the 15 studies in regard to the variables used for statistical analysis. There was some evident consistency in the choice of socio-demographic variables selected, such as gender and age, through to the reporting of ethnicity, marital status, employment status and IQ. One study did not consider socio-demographic variables, considering clinical variables only (McNiel et al., 1998). There was some variation between

the clinical variables considered across the 15 studies, but most considered length of follow-up apart from Ferguson et al. (2009) and McMillan et al (2004), who considered no clinical variables, and McNiel et al. (1998) and Odeh et al (2006), who looked at variables relating to the clinicians. . Others looked at time at risk (Bengtson 2008), individual clinician decisions vs. aggregate decisions (Huss & Zeiss, 2004) and diagnosis (Grann et al., 2000). Where standardised measures were used, variables relating to items on these measures were also included, for example VRAG and H-10 items (Grann et al., 2000; Kroner et al., 2007;Snowden et al., 2007; Tengstrom, 2001) or LSI-R:SV (Ferguson et al., 2009). Only three studies did not consider forensic history variables (Huss & Zeiss, 2004; McNiel et al., 1998 and Monahan et al., 2000). Remaining studies tended to look at previous offences and offence characteristics, as well as previous assaults and items from standardised measures for example, psychopathy (Grann et al., 2000; Tengstrom, 2004). Studies which did not consider all variables could lack internal validity, as confounding variables may have not been addressed and therefore controlled for. It is particularly surprising given the nature of the population that all studies did not specifically look at mental health diagnoses as a variable and this could further result in studies lacking internal validity.

When reporting the key findings of all 15 papers, it is important to consider the range of statistical analyses used in the studies that were reviewed, as this may have a bearing on significance of results found and reflect limitations based on sample size and also the use of selected tests for testing outcomes effectively. Across the 15 studies there was a considerable range in the selection of tests used, from Logistic Regression Analysis, Cox-regression Analysis, Linear Regression Analysis, Chi-square, Mann-Whitney U Test, T-test, ANOVA, Kaplan-Meier analysis and ROC Analysis. All tests were used in order to examine the relationships between variables and outcomes. The majority of studies utilised ROC analysis

in order to assess the predictive accuracy of the measures used, only two studies did not utilise this method of statistical analysis (Huss & Zeiss, 2004 and Odeh et al., 2006) opting for chi-square and a t-test, and logistic and linear regression respectively, in order to assess this.

Considerable variability and contradictions are evident across the studies, largely because they do not all consistently attempt to measure the same outcomes of the predictive accuracy of clinical judgement vs. actuarial assessment. Two studies that considered the predictive accuracy of actuarial assessment compared to clinical judgement, (Bengtson & Langstrom, 2007; Harris et al., 2002) and actuarial assessment alone, as well as looking at socio-demographic variables, found mixed evidence of the effect of these. The first study found that such variables were not found to be significantly associated with the predictive accuracy of assessment methods. Alternatively, Harris et al. found that the accuracy of VRAG scores in predicting recidivism were associated with gender and that the VRAG performed better for males than females. They reported a Cohen's d of 1.6 and an effect size of 0.75 $p<0.001$.

Reviewing the relationship between predictive accuracy of assessment type and clinical variables, a greater number of significant relationships were found. Firstly for the variable time of follow-up, Bengtson & Langstrom (2007) reported that both clinical judgement and actuarial assessment performed poorly, in terms of predictive accuracy during a short-term follow-up period, with no significant differences between the two approaches (x^2= 0.05-0.99, df = 1, p= 0.32-0.83). However, the predictive accuracy of actuarial assessments grew as the time that the patient was at risk increased. Snowden et al. (2007) found that the VRAG performed better in short-term follow-up (1 year), AUC = 0.86, than long-term follow-up AUC=0.75. Whereas in a five year follow-up (longer-term), Harris et al. found the VRAG to have high predictive accuracy r (329) =0.42, $p<0.001$. Grann et al. (2000) found that within a two-year follow-up,

both the VRAG (AUC=0.68, 95% CI=.63-.76), and H-10 (AUC=0.71, 95% CI = .66-.76), predicted violent recidivism at levels significantly better than chance.

McNiel et al. (1998) found that where clinicians had high confidence in their clinical judgements regarding future violence, there was a strong relationship with actual violence (Wald = 18.947, $p = 0.00$). Further, in regard to variables relating to the clinician, Odeh et al. (2006) found that clinical occupation was not a significant factor in any of the prediction models, however, differences in types of risk cues used were found. A significant logistic regression $X^2(13) = 33.15$, $p = 0.002$, showed that nurses and social workers were more likely than psychologists and psychiatrists to base clinical judgements on hostility (Wald = 6.715, $p = 0.010$), delusions (Wald = 4.248, $p = 0.039$), medication compliance (Wald = 7.683, $p = 0.006$), and family problems (Wald = 4.380, $p = 0.036$). Finally, Huss and Zeiss (2004) found that there was no significant difference in clinician profession on predictive accuracy. They also found that aggregate decisions were accurate above chance levels $X^2(1, N=354)= 14.05$, p=0.001. They also found that clinicians were more accurate, although not impressively so, at predicting the severity of future violence, their predictions correlated modestly with actual severity of violence in the sample, r=0.12, p=0.03. It is likely that this is as a result of the level of detailed information that was available to clinicians.

For other clinical variables that were considered such as sample type, Brouillette-Alarie and Proux (2013), found that the mentally disordered part of the sample had higher scores on the Static-99 (mean= 3.89) than the prison based sample (mean= 2.11). This could be as a result of the greater deviance of mentally disordered sex offenders. Also Ferguson et al. (2009), found that the LSI-R:SV significantly predicted any new offences in non-substance abusers (AUC= 0.78, $p<0.01$), but not in substance abusers (AUC=0.51, $p>0.05$).

Where forensic variables were taken into account, Bengtson & Langstrom (2007) found that previous sexual offences were not associated with the predictive accuracy of clinical judgement or actuarial methods, $X^2(3, N=69)$, $p = 0.06$. Bengtson (2008), reported that in all three assessment tools that they looked at, there was lower accuracy in predicting all types of recidivism for rapists than child molesters ($X^2=6.16$, $d.f. = 1$, $p<0.05$). It is possible that this could be as a result of the fact that child molesters have a deviant sexual interest in children and as such are more difficult to treat. Alternatively it could be as a result of the wide variety of motives behind, and explanations there are, for rape. Brouillette-Alarie and Proux (2013), found differences for the predictive accuracy of the Static-99R and the Static-99 for sexual aggressors of women and sexual aggressors of children. They found the Static-99R to have better predictive accuracy for both (women: AUC = 0.73, $p<0.01$; children: AUC = 0.77, $p<0.001$). Other studies also found forensic variables such as offence type (Kroner et al., 2007) or psychopathy (Tengstrom, 2001) to impact upon the predictive accuracy of actuarial tools. Kroner et al. found, for patients who had a violent index offence their VRAG score predicted general recidivism but not violent recidivism significantly (AUC=0.702, $p=0.001$). Tengstrom found that when psychopathy was removed, the predictive accuracy of the VRAG and H-10 decreased, although this was not significantly (AUC = 07.4, $p=0.65$).

In looking at overall findings for the studies, again there is conflicting evidence in regard to the overall predictive accuracy of clinical judgement and actuarial assessment, in predicting recidivism in mentally disordered offenders. This variation in findings may be explained by the fact that not all studies looked at both actuarial and clinical judgement methods, choosing instead to focus on one or the other.

In studies that considered both actuarial assessment and clinical judgement, Bengtson & Langstrom (2007) found that for predictions of any sexual reconviction the Static-2002 had

significantly higher predictive accuracy than clinical judgement (X^2= 4.98, df = 1, p< .05). There was no significant difference in the accuracy of predictions made by the Static-99 and those made using clinical judgement. For predictions of a severe sexual reconviction, the Static-99 had a significantly higher predictive accuracy than clinical judgement (X^2 = 5.11, df = 1, p< .05), whereas the predictive accuracy of the Static-2002 over clinical judgement fell short of statistical significance. In regard to the predictive accuracy of clinical judgement in predicting violent recidivism there were no differences found between the two measures. Overall findings suggest superiority of the actuarial assessment method.

Similarly, Harris et al. (2002) compared clinical judgements with actuarial assessments which assess violent recidivism. They found that composite clinical judgment predictions did relate to violent recidivism overall, r (383)=0.17, p<0.001, especially among male patients, r (329)=0.23, p<0.01. Suggesting that such clinical judgements do have value in assessing risk of violence in a mentally disordered population. However, they also found that these correlations were significantly lower (p<0.05) than achieved by the VRAG on the same patients, r (329) = 0.42, p<0.001. This therefore again highlighted when comparing measures directly, the superiority of the actuarial assessment method.

For those studies that solely looked at clinical judgement, there was again disagreement in regard to its predictive accuracy. McNiel et al. (1998) findings suggest that clinical judgement does accurately predict actual violence outcomes, when clinicians have high confidence (Wald = 18.947, p = 0.000), however, this accuracy does decrease when clinicians only have moderate (Wald = 4.336, p = .0373) and low confidence (Wald = 0.208, p = .6480) in their clinical judgements. In contrast, Odeh et al. (2006) found that clinical judgements of violence were not found to be associated with actual violent outcomes (X^2(14) = 16.803, p = 0.267) and that clinical occupation did not significantly impact upon predictive abilities.

Similarly, Huss and Zeiss (2004) found that experienced clinicians were not able to predict violence at levels above chance (X^2 1(354) = 1.68, p=0.24). In contrast however, they found that group dichotomous predictions were accurate above chance levels (X^2 1(354) = 14.05, p=0.001), although when probability ratings of violence were averaged, aggregated clinicians were unable to predict future violence accurately, t (1,354) = 16.4, p=0.10. Findings suggest that in general clinical judgement lacks predictive accuracy, however it is possible that when clinician's predictions are grouped this accuracy is improved. It remains unclear however, whether this would be more accurate than actuarial predictions of risk, although such approaches are used in practice to enhance the validity of assessments.

Included studies which specifically looked at actuarial assessment methods were split into: sexual offending measures, violent offending measures and general offending measures. Four studies looked at sexual offending measures (Bengtson, 2008; Brouillette-Alarie & Proux, 2013; Hanson et al., 2010; Hanson & Thornton, 2000). Hanson and Thornton (2000) found out of the measures they looked at, the Static-99 was more accurate (AUC=0.71) than the RRASOR (AUC=0.68, p<0.05) or the SACJ-Min (AUC=0.67, p<0.01) in predicting sexual recidivism. The Static -99 was also found to be more accurate in predicting violent recidivism (AUC=0.69) than the RRASOR (AUC= 0.64, p<0.001) or the SACJ-Min (AUC = 0.64, p<0.001). Highlighting the Static-99's superiority in actuarially assessing risk of recidivism. Bengtson, also looked at the Static-99 in comparison to two newer measures, the Static-2002 and the RM-2000. It was reported that all measures predicted any sexual, non-sexual violent and violent recidivism significantly higher than chance levels (p<0.05) in a mentally disordered offending population. Similarly, Hanson et al (2010) found that the Static-2002 predicted all types of recidivism with greater accuracy than Static-99, but differences were not large (sexual AUC=0.68 vs.0.66; violent AUC= 0.70 vs. 0.66; any AUC = 0.71 vs. 0.66). Finally,

Brouillette-Alarie and Proux, found that in the mentally disordered offenders within their sample, the Static-99R (AUC=0.65, $p<0.05$) significantly predicted non-sexual violent recidivism at higher levels than the Static-99 (AUC=0.68, $p<0.01$). However, for non-violent recidivism predictive accuracy was not significant unless looking at the whole sample (which included non-mentally disordered offenders as well). Results suggest that recent actuarial tools developed primarily for predicting sexual recidivism are successful in doing so.

Four included studies also looked at actuarial measures for assessing risk of violence (Grann et al., 2000; Kroner et al., 2007; Snowden et al., 2007;Tengstrom 2001). Grann et al looked at the H-10 and VRAG and found that for the whole sample the VRAG's ability to predict violent recidivism, AUC = 0.68. The H-10's AUC = 0.71, was larger, however this was not significantly so ($X^2(1) = 2.07$, $p=0.1505$). The AUC's found by Kroner et al. for the VRAG were slightly higher (AUC= 0.703, $p=0.000$) and increased further to 0.763 ($p=0.004$) when the outcome was limited to violent recidivism as opposed to general. Similarly, Snowden et al. found that the predictive accuracy for the VRAG increased from AUC=0.743 to 0.776 for general and violent recidivism respectively. In addition they found that the OGRS also had high levels for general (AUC=0.785) and violent (AUC=0.762) recidivism. All of the AUC's were found to be significant above chance levels for both of these measures ($p<0.0001$). Finally, Tengstrom also found the VRAG to have a slightly lower predictive accuracy (AUC=0.68) in comparison to the H-10 (AUC=0.76) in this sample. Findings suggest that all actuarial measures significantly predicted violent and/or general recidivism at levels above chance and therefore emphasise their utility in clinical settings when working with mentally disordered offenders.

Two studies were also included that looked at other actuarial measures. Ferguson et al. (2009) assessed the predictive validity of the LSI-R:SV in predicting recidivism. They found

that it did so at a significant level above chance for any new offence (AUC=0.67, p<0.001), for a non-violent new offence (AUC=0.67, p<0.001) and for a violent new offence (AUC=0.60, p<0.05). Monahan et al. (2000) looked at developing a new actuarial method for use in predicting risk in a mentally disordered population in clinical settings, the ICT. They also found that this did so at a level significantly above chance (AUC=0.80, p<0.001). These findings further support the use of actuarial assessment methods when assessing risk.

As aforementioned, the 15 studies included for review had a range of QAS. They showed similarities in regard to their methodological approaches and outcomes that were being considered. Although not all studies clearly outlined outcomes (Grann et al., 2000; Huss & Zeiss, 2004; McNiel et al., 1998; Odeh et al., 2006), and consequently this was a limitation of these studies. The outcomes considered and methodologies used, are comparable with other studies, which have also considered the predictive accuracy of clinical judgement and actuarial assessment methods in other populations. Much of the literature in this area however, is comprised of descriptive papers, as opposed to research studies, and where observational studies have been conducted this has been on general offending populations. One of the studies' strengths is that they represent a mentally disordered offending population. Sample sizes ranged from very small (Odeh et al., 2006) to very large (Monahan et al., 2000; Snowden et al., 2007). Those with small samples are likely to have less power and predictive weight, resulting in them being less representative of the population than the studies with larger sample sizes. Whilst Hanson et al's. (2010) total sample was 3,304 only 311 pf these participants were mentally disordered, this was similar to that of Hanson and Thornton (2000), whose mentally disordered sample was also smaller than the total sample. This again questions the representativeness of the cohorts in these studies and may account for observed differences in reports of predictive validity, of the different methods across studies.

Differences in quality of studies were evident across the papers that were reviewed. Strengths and limitations of each are summarised in Table 6. One limitation that is evident across papers is that not all papers use appropriate statistical analysis for outcomes being measured. Fawcett (2006) states, that carrying out receiver operating characteristic (ROC) analysis allows an area under the curve (AUC) to be determined by plotting the relationship between specificity and sensitivity. This AUC represents the probability that a test or process will classify a randomly chosen positive finding, higher than a randomly chosen negative finding, regardless of base rates. Therefore to determine the predictive accuracy of a measure (clinical judgement or actuarial assessment) it would suggest this type of analysis would be required, although not all studies used this (Huss & Zeiss, 2004; Odeh et al., 2006). Some researchers also suggest that reporting sensitivity and specificity as well as AUC's is important in providing a full understanding of a test's accuracy (Zwieg & Campbell, 1993; Swets, Dawes & Monahan, 2000). Despite this, four of the included studies reported AUC's but not sensitivity and specificity, suggesting that findings regarding the predictive accuracy of the measures in these studies may be limited (Bengtson & Langstrom, 2007; Hanson et al., 2010; Hanson & Thornton, 2000; Snowden et al., 2007).

Additionally, a further limitation across all studies with the exception of Monahan et al. (2000), is that they relied on case notes of the patients being studied in order to score measures or make clinical judgements. Whilst other factors were relied upon as well, this still poses a variety of problems. For example, authors were unable to control how data was originally collected and there may be missing data which would have impacted upon the final data set. Further to this, this meant that in the case of looking at the accuracy of clinical judgement, clinicians will not have interacted, or observed patients, therefore resulting in them lacking

information on which to make predictions. Finally, often case records and actuarial assessments are completed using patient self-report, leaving the studies open to an inherent bias.

Whilst all studies made some attempts to identify and control for confounding variables, this was limited. Particularly striking, considering the population being considered (mentally disordered offenders), was the number of studies that failed to consider forensic history and the impact that this may have had on the predictive accuracy of assessment methods. In particular those that looked at clinical judgement (Bengtson & Langstrom, 2007; Harris et al., 2002; Huss & Zeiss, 2004; McNiel et al., 1998; Odeh et al., 2006) failed to acknowledge the impact clinicians previous knowledge of a patients' offences (i.e. nature and severity) may have had on their judgements. Additionally, the majority of studies failed to consider more clinical variables such as length of stay. Snowden et al. (2007), did include this variable but others did not consider it at all and this could have been a potentially important confounding variable. If a patient had had a longer stay, and clinicians were aware and knowledgeable about the patient, then this is likely to have influenced their clinical judgements. In addition, if a patient has had a longer stay, it is likely that they may have completed numerous psychometric tests and therefore are able to 'fake good' and provide an inaccurate profile, as many of the measures were scored using file data and retrospectively, the actuarial assessments may have inaccurately identified their risk level. Alternatively the increased length of stay could also impact in that individuals become more honest and open, thus making assessments more accurate.

There was also a poor description of sample characteristics in some cases (Bengtson, 2008;Bengtson & Langstrom, 2007; Brouillette-Alarie & Proux, 2013; Grann et al., 2000; Hanson et al., 2010; Hanson & Thornton, 2000; Harris et al., 2002; Kroner et al., 2007), which made it difficult to understand the choice of variables selected. Studies were penalised accordingly in QA for this, along with the studies lack of consistency in statistical analyses,

bringing the generalizability of studies reviewed into question. Whilst they may be comparable with other studies, which have looked at the predictive accuracy of clinical judgement and actuarial assessment, it is difficult to comment with any great certainty as to whether these findings can be extrapolated to the wider mentally disordered population. Greater specificity in the selection of statistical analysis used, as well as transparency in establishing outcomes and variables to be assessed, may have strengthened the quality of some of the studies.

Table 6.

Strengths, Limitations & Quality Assessment Scores of Included Studies

Author / Year	Strengths	Limitations	Quality Assessment Score
Bengtson (2008)	Outcomes clearly defined Appropriate statistical analysis used Replicates other findings Limitations clearly set out	Longer follow-up than usually applied Highly selected sample, limited external validity, results can only be generalised to more severe offenders Performance statistics were at the lower end of what has already been published, therefore caution is required when interpreting these A large number of offenders were excluded on various grounds for analysis, which may have affected the accuracy of risk scales	81%
Bengtson & Langström (2007)	Directly compares both clinical judgement and actuarial assessment Representative sample Replicates other findings Limitations clearly set out Outcomes clearly defined	Highly selected sample, limited external validity Relatively small sample Limited internal validity, clinicians differing internal scales for risk categories Findings limited to more severe offenders with respect to offence	92 %

Author / Year	Strengths	Limitations	Quality Assessment Score
	Appropriate statistical analysis used	characteristics and psychological impairment	
Brouillette-Alarie & Proux (2013)	Replicates other findings Limitations clearly set out Outcomes clearly defined Appropriate statistical analysis used Large sample, increased external validity	Only part of sample psychiatric and therefore 'mentally disordered', limits external validity Predictive validity differences between non-sexual violent and non-sexual non-violent recidivism may reflect variations between psychiatric patients and correctional offenders Factor analysis of dichotomous variable is not optimal Did not have complete recidivism information for both settings, sample dependent, need bigger and more unified sample Use of sentence length to determine release dates for psychiatric sample was suboptimal and could have led to inaccuracies in cox regression	73%
Ferguson, Ogloff & Thomson (2009)	Representative sample	The full version of LSI-R:SV was not used which may account for	85%

Author / Year	Strengths	Limitations	Quality Assessment Score
	First study to assess the predictive abilities of the LSI-R:SV in this population Replicates some other findings on LSI-R:SV predictive abilities in other populations Limitations clearly outlined Appropriate statistical analysis used	lack of prediction in sub-groups, lack of internal validity Lack of information regarding time at risk, unable to assess predictions of long or short-term risk better Australian sample, restricts external validity of findings Obtained offending history possibly incomplete Retrospective design, inferences made where information not available Direct comparisons with other studies limited due to being first study of its kind	
Grann, Belfrage & Tengstrom (2000)	Large representative sample, high external validity Replicated previous findings Appropriate statistical analysis used Outcomes clearly defined Limitations clearly outlined	Only VRAG and H-10 tested in this study, alternative approaches have been suggested Other factors may have inflated the relative importance of actuarial data in this study Violence inclusion criteria differed between two sub-groups, the two groups might differ in respects other than just diagnosis	73%

Author / Year	Strengths	Limitations	Quality Assessment Score
		Groups may have recidivated at different paces Highly selected sample	
Hanson, Helmus & Thornton (2010)	Large total sample, sufficient statistical power to detect small differences Appropriate statistical analysis used Outcomes clearly defined Raises important questions about the application of such assessment tools	Not a representative sample, relatively small proportion mentally disordered offenders and higher number of high risk offenders that would normally be found Limitations not clearly stated or detailed	69%
Hanson & Thornton (2000)	Large sample Appropriate statistical analysis Findings challenge those who claim that sexual recidivism cannot be predicted with sufficient accuracy Replicates and adds to previous findings/research Outcomes clearly defined	Not clearly mentally disordered sample Limitations not clearly stated or detailed Follow-up times differed Used different recidivism criteria for different samples Lacks internal validity	69%
Harris, Rice & Cormier (2002)	Large representative sample Addresses concerns raised by previous researchers	Limitations not clearly set out or detailed	85%

Author / Year	Strengths	Limitations	Quality Assessment Score
	Directly compares clinical judgement and actuarial assessment Replicated previous findings Outcomes clearly defined No contamination of clinical judgement by actuarial scores and vice versa. Uses appropriate statistical analysis	Clinicians had more information for judgements than was used to score actuarial measure	
Huss & Zeiss (2004)	This study has clinical utility Examined a larger sample of clinicians and patients than previous studies Limitations clearly outlined Uses appropriate statistical analysis	Used vignettes, not clients who clinicians had contact with and had less information available to them to do this Discussion in aggregate decision making not taken into account, not realistic and inherently flawed Results do not have direct application to clinical practice and require replication Risk categories were artificial in their construction Patients used in the study were at a heightened degree of risk because they had previously been committed, therefore base rates	76%

Author / Year	Strengths	Limitations	Quality Assessment Score
		were higher than would normally be expected	
Kroner, Stadtland, Eidt & Nedopil (2007)	Replicates other findings Appropriate statistical analysis used Confirms use of VRAG with a German sample, external validity of findings	Small, selective sample, limits applicability of findings Selection bias, not a discharge sample Limitations not clearly set-out	85%
McNiel, Sandberg & Binder (1998)	Moderately large representative sample Limitations clearly set out Examines confidence as affecting clinical judgement building on previous research Appropriate statistical analysis used Outcomes clearly defined Naturalistic treatment setting, increases external validity (this is however limited)	Prediction estimates could have been influenced by response scale Individual differences between clinicians not considered Information regarding violent patients could have facilitated accuracy of judgements Confounding variables not addressed or controlled for Not representative of community settings	81%
Monahan et al (2000)	Findings show that a clinically useful tool for assessing the risk of violence amongst acute psychiatric patients exists Outcomes clearly defined	Not an epidemiological study of violence among people with mental disorder, limits external validity	62%

Author / Year	Strengths	Limitations	Quality Assessment Score
	Limitations clearly set out Appropriate statistical analysis used Classifies patients realistically	The extent to which the accuracy of this clinical tool can be generalised to other settings is unclear Looks at the development of the tool more than the predictive accuracy of an actuarial tool	
Odeh, Zeiss & Huss (2006)	Replicates findings from previous research Limitations clearly set out Representative sample Appropriate statistical analysis used Outcomes clearly defined	Very small sample, not representative, looks at civilly committed U.S veterans Risk cues are not necessarily indicative of risk factors Precise frequency of risk cues are an artefact of specific protocols used in sample Clinicians only read protocols, did not interact with patients and therefore may have lacked information to make valid predictions Suffers from under-reporting of violent incidents Low agreement of clinicians for each individual cue with respect to patient protocol	77%

Author / Year	Strengths	Limitations	Quality Assessment Score
Snowden, Gray, Taylor & MacCulloch (2007)	Replicates previous findings Large representative sample, high external validity Outcomes clearly defined Limitations clearly outlined Appropriate statistical analysis used	Outcome variable of reconvictions not ideal as the majority of violent acts do not result in formal convictions Lengthy follow-up Both tools could not be used with whole sample, therefore limits generalisability of findings	92%
Tengstrom (2001)	Replicates and builds on previous findings Representative sample Appropriate statistical analysis used	Limitations not clearly set-out Results unclear Limited external validity due to highly selected sample Due to nature of sample the schizophrenia/major mental disorder item in both tool could not contribute to overall predictive validity	69%

Discussion

Main Findings

Fifteen studies were included in this systematic review, which aimed to examine the literature pertaining to the predictive accuracy of clinical judgement and actuarial assessments, in predicting recidivism in a population of mentally disordered offenders. In addition, it attempted to identify the best assessment method to use with this population type and whether there are any other factors which may relate to the predictive accuracy of clinical judgement or actuarial assessment, in accurately predicting recidivism.

The 15 studies included for review all varied in their results. There was some consistency between the studies' findings in general in regard to predictive accuracy of assessment tools, but divergence, in terms of what variables may be associated with this. Studies that considered socio-demographic variables were generally united in their findings that these did not influence predictive accuracy, although one study did find that gender of the offender played a role in the predictive accuracy of the VRAG actuarial assessment tool (Harris et al., 2002). It is possible that gender was not assessed as a variable in other studies as although overall studies looked at a mixed populations, nine out of the 15 studies had male samples only and therefore were unable to make comparisons regarding predictive accuracy and gender. This finding does however indicate a need for further research in relation to potential gender differences on predictive accuracy.

There was disparity between those studies where clinical variables were considered. Bengtson & Langstrom (2007) found no association of clinical variables with accuracy of prediction. Snowden et al. (2007) on the other hand, found that follow-up time did affect the predictive accuracy of measures. For studies that assesses clinician related variables, Odeh et

al. (2006) found that clinician occupation, previous assaults, hostility, medical non-compliance, paranoid delusions and family problem cues, all significantly influence the probability other judgements would be made of violent outcomes. However, none of these findings are linked to accuracy of prediction methods used. Only one study (McNeil et al., 1998) that considered clinician confidence as a clinical variable, found that this significantly influenced the accuracy of clinical judgement. These findings suggest that to an extent, clinical variables are associated with predictive accuracy of assessment methods used and in particular where clinical judgement is used. This is consistent with other research in the literature base. For example, Elbogen, Calkins, Scalora and Tomkins (2002), found that when asking clinicians to rate the relevance of risk factors for influencing their judgements of recidivism, in general, clinicians rated clinically derived behavioural variables as significantly more relevant. This was even when compared to the HCR-20 and VRAG where cues came from historical, contextual and dispositional domains.

One of the main focuses of this review is that the population is a forensic psychiatric one (mentally disordered offenders). It is therefore surprising that none of the studies included really considered diagnoses in relation to the accuracy of clinical judgement and actuarial assessment. One study did consider this (Grann et al., 2000) and whilst 10 of the 15 studies considered it as a variable within their sample descriptions, it was only Grann et al. who made reference to this in their findings, in relation to the difference between the personality disorder and schizophrenia cohorts. Furthermore, only two studies made reference in their findings to more forensic variables (Hanson et al., 2010; Hanson & Thornton, 2000), distinguishing between the predictive accuracy of the tools in reference to patient's offence type. Other studies made no reference to this at all, which shows an inconsistency with other research. For example, Lindsay and Beail (2004) suggested that studies of clinical judgement have generally

found that clinicians tend to focus on clinical history, previous violence, substance abuse history and anger control rather than demographic characteristic such as age. Findings are therefore consistent and relevant in terms of socio-demographic variables but do fall down on clinical and forensic variables in particular.

In attempting to identify what factors may be associated with the predictive accuracy of clinical judgement and actuarial assessment, these fifteen studies present a mixed picture. This suggests that their overall findings in regard to the accuracy of both clinical judgement and actuarial assessment methods may be questionable with this population. It is likely that this reflects the difficulty in predicting risk in mentally disordered offenders, as it can vary with the nature of their disorder and/or stage of illness, making it almost impossible to predict how they will be in the future. All studies that have assessed actuarial assessment methods suggest that they are suitable to use with this population, whereas findings from clinical judgement studies are less convincing, although do suggest that this method has some clinical utility. Due to the variance between study findings it is difficult to conclude with confidence that any factors are consistently associated with predictive accuracy.

Strengths and Limitations of the Review

Fifteen studies is a moderate number of research papers to be included in a review and this is a strength of this review. The fact that only two of these directly compared clinical judgement and actuarial assessment and only a further three studies looked at clinical judgement alone may be a potential limitation. Whilst actuarial assessment vs. clinical judgement has acquired a lot of attention in the literature base, much of this is descriptive in nature and does not test out the topic in research studies, specifically with a population of mentally disordered offenders. Whilst in one sense this is a useful issue, in that it highlights a move within the

literature and clinical practice towards actuarial assessment methods, the limited number of studies presented in this review relating to clinical judgement could undermine conclusions in terms of generalisability. Both in relating to the wider offending population, as well as more specifically the wider secure psychiatric population, and limited the conclusions that could be drawn regarding comparisons between the two assessment methods. There could be several contributing factors to the limited number of clinical judgement studies included for review. One of these could be the difficulty experienced by the researcher in sourcing citations. Whilst a number of electronic databases were searched as well as utilising Google Scholar, the number of citations ranged from tens of thousands to very few numbers of citations registering and there was significant disparity in the subject matter, with many studies looking at general offending populations rather than a mentally disordered population. This meant that at application of inclusion/exclusion criteria many citations were excluded.

Further to this, as the literature reviewed in this study was the 'modern' literature, it is possible that this is the reason for the lack of studies assessing clinical judgement. It may have been more useful to expand the date range of studies included, rather than keeping a focus on just the 'modern' literature (post 1987). This would have been advantageous as, with hindsight, if it has been done, other existing cohort studies by well-known proponents in the field, may have been able to be included. For example, Quinsey and Maguire (1986) considered dangerous men in maximum security and clinicians' ability to predict accurate recidivism during follow-up. Also Menzies, Webster and Sepejak (1985a) evaluated the accuracy of actuarial assessment predictions of violence amongst forensic psychiatric patients. Neither of these studies were included for review because of their date, however these and other papers from the more dated research base may have been useful additions to the findings of the final review, as such studies laid the foundation for risk evaluation studies. Finally, it is unrealistic to propose that clinicians

would predict high risk in released offenders, if offenders were considered high risk it is likely that they would be detained for longer periods under the mental health act. Without the presence of large longitudinal research studies it is therefore difficult to assess how accurate clinical judgements are within this population.

With reference to the initial search syntax devised, any problems with this were identified and rectified by the researcher at the initial searching stages. These difficulties surrounded the need to account for alternative spellings of search terms, as well as making them broad enough to produce citations, but specific enough to produce as relevant information as possible. The use of subject mapping where possible was employed in order to alleviate this issue, as well as making search terms specific enough to minimise overlap with physical health papers. Retrospectively reviewing the search terms and parameters used within this review, suggests that the review protocol and number of citations found may have been strengthened by the inclusion of supplementary search terms associated with forensic psychology and assessment methods used. For example as well as using search terms like 'actuarial assessment' and 'clinical judgement' it may also have been beneficial to utilise other terms used for these concepts such as 'mechanical prediction' or 'statistical prediction' and 'unguided clinical judgement' or 'clinical assessment'. As these were not included this may have limited the extent to which all relevant papers in this research area were found.

Whilst hand-searching of individual references was carried out, due to time constraints this was limited to reviews and meta-analyses found, surrounding the general research area of clinical judgement and actuarial assessment. If the researcher had had a greater amount of time and was able to hand search additional reference lists, this could have potentially resulted in a wider literature base. Additionally, this may have improved the quality of the review and its findings, adding additional papers that PICO criteria could be applied to and excluded

accordingly. Reflecting on the process that this systematic review took, if it was to be repeated it would need to take account of the limitations discussed thus far and amend them accordingly in order to make the best use of the literature available in relation to this topic area. However, this review had the addition that, as part of the research strategy, that the key authors in the field were contacted in regard to unpublished work or directions to additional literature that should have been included in this review. This process only resulted in one additional paper, which hadn't been found through electronic searches or hand sourcing and was ultimately excluded at the PICO stage due to the population that it considered. Therefore this would suggest that relevant research that was available had been sourced successfully by the researcher.

Another limitation of this systematic review is that only one researcher was involved in the process from start to finish. The sole researcher conducted the search, data extraction and quality assessment within this review. It is possible that, had there been another researcher, they may have picked out different information at data extraction stage, resulting in the inclusion of some studies that were in fact excluded by the sole researcher. Whilst this may be a potential limitation to the validity of this review, having a second researcher could also have resulted in additional limitations. For example, having a second assessor may have resulted in studies that were included, being excluded, therefore would have limited the small number of studies included for review following the study selection process. A considerable strength of this review however was that a second assessor did quality rate 50% of included papers and there was substantial interrater agreement, thus suggesting the correct studies were included by the author for final review.

The studies themselves also have limitations which impact upon the conclusions that have been drawn by authors and therefore additionally have an impact upon the strength of this review. Firstly, in several of the studies, there is a lack of clarity in regard to only partially

stating the sample characteristics (Bengtson, 2008;Bengtson & Langstrom, 2007; Brouillette-Alarie & Proux, 2013; Grann et al., 2000; Hanson et al., 2010; Hanson & Thornton, 2000; Harris et al., 2002; Kroner et al., 2007). This therefore makes it difficult to identify how authors have reached outcomes and variables to be considered. More limiting it would seem, is the fact that only one of the studies reviewed (Tengstrom, 2001) considers characteristics of patients' mental disorders as well as their offence details. None of the other studies reviewed, report both participants' mental disorders and offences, and therefore do not consider how these variables could impact upon the predictive accuracy of clinical judgement or actuarial assessment. This brings into question the internal validity of both the studies themselves and the systematic review. Additionally, this does not enable the author of this this review to be certain that the participants within the studies reviewed, provide a representative sample of a mentally disordered offending population. This limitation is particularly important as it lowers this reviews external validity, meaning that findings could not be extrapolated to the wider population of mentally disordered offenders.

The studies included used either a retrospective cohort (Bengtson, 2008; Bengtson & Langstrom, 2007; Brouillette-Alarie & Proux, 2013; Ferguson et al., 2009; Grann et al., 2000; Kroner et al., 2007; Hanson et al., 2010; Hanson & Thornton, 2000; Harris et al., 2002; McNiel et al., 1998; Monahan et al., 2000; Tengstrom, 2001), prospective cohort (Snowden et al., 2007) or case-control study design (Huss & Zeiss, 2004; Odeh et al., 2006). The cohort methodology does appear to be an appropriate one for these studies, as both retrospective and prospective cohort studies allow for the long-term follow-up required in order to accurately assess recidivism. Whilst the case-control methodology, still produced relevant results, it would appear not as appropriate as a cohort study, largely due to its requirement for a matched control group, this limited the sample size, resulting in limitations on the extrapolation of findings for

the study and also limiting the review in this way. Also the small number of participants meant a limited statistical power, which may be the reason for some of the results not reaching statistical significance and again causing questions in its validity. Despite these differences in methodology, all studies also relied on case notes regarding patients, whether retrospective or current, in order to inform clinical judgements in particular. This makes all studies and therefore the review, inherently open to bias, due to the individual differences that would be present in the reporting of case notes, as well as the inevitability of reliance on self-report and the potential for inaccurate accounts or missing information regarding patient's presentation or behaviour. It is possible that, as services move towards the use of electronic recording systems this further limits the availability of information for clinicians, for example the quality and richness of information may be lost due to the use of different recording systems and often it is this information which is used to validate evidence against. This has the potential to limit these kinds of studies in the future and poses significant implications for clinical practice.

There are several possibilities that could be considered in order to strengthen the methodological approaches used. Firstly, as all studies are related to the outcome of predictive accuracy on assessment, an improvement would be that they all directly consider both clinical judgement and actuarial assessment. Secondly, a way of addressing the methodological limitations described would be to use more standardised measures within the studies and not just in the case of when looking at actuarial assessment. For example, those assessing accuracy of clinical judgement could also use standardised measures such as the overt aggression scale as used by McNiel et al., (1998). This would potentially limit the individual differences in reporting style and missing content regarding patients, as there would be a more standardised process. Finally, where the accuracy of clinical judgement is assessed as much of the information taken into account is retrospective and maybe interpreted and reported differently

by different clinicians', interrater reliability would need to be assessed to minimise levels of inaccurate or missing information as well as controlling for individual differences.

A final point for consideration in regard to the fifteen studies reviewed and the impact of these on the strength of the review, is to consider the choice of statistical analyses used within the studies. The studies use a combination of parametric and non-parametric tests. Although what is of more interest is the disparity between those that have used ROC analysis and those studies which haven't (Huss & Zeiss, 2004; Odeh et al., 2006). As aforementioned, it has been suggested that this analysis is beneficial in establishing the predictive accuracy of assessment methods, particularly when sensitivity and specificity is reported. For example, as Doyle and Dolan (2002) highlight, an instrument or clinician that predicted violent from non-violent patients with nearly perfect accuracy would have ROC-AUC curves approaching 1.0. It is possible that, as not all studies utilised such methods, and further to this some of the studies that did, did not report sensitivity and specificity and cut-off scores, this limited the review. This potentially limits generalisation of findings of the review, as because ROC analysis was not carried out in all studies, effect sizes could not be compared, resulting in a percentage not being determined as to which assessment method has a higher predictive accuracy. For example in their meta-analysis Aeigisdottir et al. (2006) found a 13% increase in predictive validity for actuarial assessments over clinical judgements when comparing effect sizes. Furthermore, it raises possible questions regarding the conclusions made by authors who did not utilise these methods.

Conclusions

This systematic review found that there is a conflict between whether clinical judgement or actuarial assessment, is more accurate in predicting recidivism, in a mentally disordered offending population. Although overall, not withstanding the limitations, the findings of the

studies reviewed would suggest that actuarial assessment is superior in terms of predictive accuracy as opposed to clinical judgement. All of the fifteen studies within this review considered different socio-demographic, clinical and forensic variables in relation to their influence on predictive accuracy. The inconsistency between the variables considered and the overall findings suggest strong evidence that there is a necessity for further research in this area, which further considers mentally disordered offending variables.

The majority of studies considered in this review look at violent recidivism although there are those that consider both violent and sexual recidivism (Bengtson & Langstrom, 2007), this discrepancy may account for some of the inconsistency in findings. It may be that clinical judgement and actuarial assessments are affected by offence type, and have differing predictive accuracy for different types of offending behaviour as Bengtson and Langstrom's study would suggest. Further research therefore would benefit from considering such forensic variables as this.

Alternatively, further research could draw more heavily on socio-demographic variables such as gender as such factors are important when considering the tools use in both male and female population groups. Elbogen et al. (2001), investigated the relationships between gender and clinicians' judgements of dangerousness in a civil psychiatric facility. They found that there were higher judgements of dangerousness for males than females and that there was also a significant interaction between clinicians' own gender and patient gender. Whilst not in a forensic population this suggests that gender could play an important role in the judgement of recidivism, and rather than aiding risk assessment in psychiatric populations, gender may contribute to the inaccuracy of risk assessment. It is important to note however, that like risk, dangerousness is ascribed to a person and does not help clarify the immediacy, severity or nature of harm of the individual (Scott, 1977). All of these are also missing from risk

assessments and evidence their limitation in assisting clinicians in making appropriate judgements regarding someone's risk or release.

The studies reviewed here had a mixed population of males and females and despite this, gender was not widely considered as a variable. It is possible that future research that took such variables into account, could add significantly valuable information to what is already known about the predictive validity of assessment methods used by clinicians working with mentally disordered offenders. Further to this it would also be beneficial for researchers to consider male and female populations as separate entities and reporting the differences observed in predictive accuracy. This would have the potential to make findings more externally valid to the wider population of mentally disordered offenders, particularly due to the comparatively smaller percentage of females to males (Ministry of Justice, 2010).

Given the small number of studies analysed in this review, which directly compare the two methods, this clearly emphasises the need for further research in this area. In addition, it is an indicator that not many firm conclusions can be drawn in regard to the predictive accuracy of different assessment types. Research by Monahan et al. (2005) has suggested that a two-stage procedure of risk assessment is employed that begins with actuarial assessment and then moves to a second stage of clinical judgement. A process that keeps the two approaches conceptually distinct but, utilises the clinical advantages of both of them. However, Seto (2005) suggests that combining assessments does not increase the accuracy in prediction and can in fact reduce it. In recent years there has been a move to the development of the structured professional judgement approach (SPJ), with a number of risk assessment tools which utilise this approach being developed. One of these, the HCR-20 Version 3 is discussed in greater detail in the following chapter. Overall, what is clear from this review, is that there is a need for additional research on clinical and actuarial prediction, not just of violence as Litwack

(2001) has highlighted but, for all types of offences. A greater understanding would assist clinicians in having a sufficient understanding of the individuals they work with in order to adequately assess their risk. Furthermore, further research would allow for the development of better tools to aid clinicians. In turn allowing tribunals or other professional bodies to make accurate decisions about patients discharge and risk management.

CHAPTER THREE

The Historical, Clinical, Risk-20 (HCR-20): Assessing Risk for Violence (Version 3.0)

Critique of a Psychometric Assessment

Introduction

Violence and its control is an issue which remains of high importance within today's society. Yang, Wong and Coid (2010) suggest that it is a major public health issue which affects perpetrators, victims and witnesses. As demonstrated in Chapters 1 and 2 it is vital that practitioners are able to predict and manage risk effectively. Douglas, Hart, Webster and Belfrage (2013) highlight that as recently as twenty years ago, best guidance concerning how to practice violence risk assessment and management came from individual research studies, reviews by clinicians, and advice from legal scholars. In contrast the field of risk assessment has now developed, with considerable attention being given to the development of more rigorous and robust tools to assess risk, in order to overcome shortfalls of the clinical judgement approach (Douglas & Reeves, 2010). One such instrument, is the Historical Clinical Risk-20: Assessing Risk for Violence (HCR-20), developed by; Webster, Eaves, Douglas and Wintrup (1995), Webster, Douglas, Eaves and Hart (1997) and Douglas et al. (2013).

This review appraises the HCR-20 Version 3 (Douglas et al., 2013), in addition it aims to explore its psychometric properties as a risk prediction tool and also consider its clinical utility throughout. Previous versions of the HCR-20 have been extensively reviewed. Overall there have been over 200 empirical studies which have evaluated it (Douglas et al., 2013). A recent large scale study by Singh (2013), found that of 2,135 clinicians, across 44 countries, the HCR-20 was the most commonly used violence risk assessment. In terms of studies which look at the HCR-20's scientific validity and reliability there is also a considerable body of evidence; (deVogel, van den Broek & de Vries Robbe, 2014; Douglas 2014; Douglas & Belfrage, 2014; Doyle et al., 2013; Strub, Douglas & Nicholls, 2014).

The HCR-20 is not a formal psychological test, it is a framework widely adopted within forensic psychiatry to assess violence recidivism (Douglas & Reeves, 2010). It is an extensively used clinical tool within forensic settings, and the use of SPJ approaches in risk assessment such as the HCR-20, is formally outlined by the Department of Health (2007) to be important in forensic practice. It has broad clinical utility, and is used in a variety of forensic settings and also in research surrounding violence risk.

Background to the HCR-20 (Version 3)

The assessment of violence risk has historically been a challenging task for practitioners (Douglas, Ogloff & Hart, 2003), some researchers questioned whether it would ever be appropriate to estimate violence risk in a quantitative way (Litwack, 1993). Morris and Miller (1985), argued that robust violence risk assessment and management was impossible and should be discarded. This pessimistic view has now changed and it would seem that the reason for this is an increase in research. Douglas et al. (2013), identify that the evidence base expanded rapidly in the 1990's in three ways: epidemiological research identified that the perpetration of violence was a real and pervasive issue; epidemiological and clinical research established valid risk factors for various types of violence; and research confirmed the predictive validity of multi-factor tools designed to assess violence risk.

As discussed in Chapter 2, historically practitioners would assess violence using unstructured clinical judgement, an approach which has been criticised as being fundamentally flawed (Grove & Meehl, 1996). To overcome the shortcomings of the clinical judgement approach, the actuarial approach was developed, so beginning the clinical vs. actuarial debate which has been discussed widely within the literature (Grove & Meehl, 1996; Litwack, 2001; Webster, Hucker & Bloom, 2002) and is discussed in greater detail in the previous chapter.

In summary, both clinical judgement and actuarial assessment may be useful in some contexts, however as aforementioned they both have their flaws. Douglas et al. (2013) argue that it is as a result of their flaws that the HCR-20 risk assessment tool has developed. The SPJ approach, was developed as a method to assist clinicians in structuring risk assessments. Pedersen, Ramussen and Elsass (2012), outline that this approach, ensures all relevant factors for future violence are included in the clinicians' assessments. SPJ outlines not only what information practitioners should gather and how to gather it, but also how they should use this information, in order to make decisions in relation to the presence and relevance of risk factors. This change in approach is supported by an increasingly growing evidence base, which incorporates empirical research studies on specific risk factors, and also theoretical models of violence. Examples of tools which were developed and use the approach include: the Spousal Assault Risk Assessment guide (SARA; Kropp &Hart, 2000); HCR-20V3 (Douglas et al., 2013); the Sexual Violence Risk-20 (SVR-20; Boer, Hart, Kropp & Webster, 1997) and the Risk for Sexual Violence Protocol (RSVP; Hart et al., 2003).

Overview of the HCR-20 Version 3

The assessment and the manual. The HCR-20 was first published in 1995 (Webster et al.), this was then revised in 1997 (Webster et al.). Version 2 (V2) of the HCR-20 was a major revision, whilst original risk factors were retained, names and definitions were changed in order to improve clinical utility. The HCR-20V2, has been the focus of narrative reviews (Douglas & Reeves, 2010) and meta-analytic reviews (Singh et al., 2011; Yang et al., 2010). Despite having extensive research support authors still felt there was a need for its revision. These revisions relate to the growth in literature relating to violence and the SPJ approach. Authors felt that it was important that the HCR-20 should reflect this and include it in the definitions of risk factors. Additionally, through review and discussions with others it was

decided by authors that there were areas which required clarification, for example the definitions of risk factors. In Version 3 (V3), these include indicators for the risk factor and also clarification regarding administration and scoring.

An overview of the HCR-20V3 items can be found in Table 7. In contrast to V2, a number of the factors in V3 are now divided into sub-items. Douglas et al. (2013), indicate that inclusion of sub-items ensures that the multi-faceted nature of some risk factors is considered. V3 also includes indicators when defining risk factors, the primary goal of this is an attempt to enhance interrater reliability (IRR), it also allows for the evaluator to identify specific ways that the risk factor may manifest for the individual.

The manual details a series of steps that the evaluator should take in order to arrive at a final judgement of risk. All items are coded on a three-point scale (Y = Present, P = Partially/Possibly Present, N = Not Present), evaluators also have the option to omit the item, if there is no reliable information by which to judge the presence of the risk factor. The timeframe for coding historical factors is lifetime and for clinical factors is usually in the past six to twelve months. Risk management factors can be coded for both inside secure settings and also in the community, the evaluator makes judgements on whether the risk factor would be present in the future.

The assessor, is then asked to make a judgment regarding the relevance of each risk factor, to assist with development of risk management strategies. Douglas et al. (2013, p.50) state, “By relevance, we mean the extent to which the factor is critical to the evaluator’s formulation of what caused the evaluee to perpetrate violence and how best to prevent future violence”. Evaluators must code the relevance of the factor on a 3-point scale (High, Moderate or Low). Smith, Kelley, Rulseh, Sorman and Edens (2014), suggest that assessing the presence

and relevance of a risk factor, means that the HCR-20V3 is in line with research which identifies the importance of assessing intra-individual risk factors for violence.

Table 7.

Items and Sub-Items in the HCR-20 Version 3.0.

Historical (10)	***Clinical (5)***	***Risk Management***
H1 – History of Problems with Violence: *a)Child, 12 years and under* *b)An adolescent, 13 to 17 years old* *c)An adult, 18 years and older*	C1 – Recent Problems with Insight: *a)Mental Disorder* *b)Violence Risk* *c)Need for Treatment*	R1 – Future Problems with Professional Services and Plans
H2 – History of Problems with Other Antisocial Behaviour: *a)Child, 12 years and under* *b)An adolescent, 13 to 17 years old* *c)An adult, 18 years and older*	C2 – Recent Problems with Violent Ideation or Intent	R2 – Future Problems with Living Situation
H3- History of Problems with Relationships: *a)Intimate Relationships* *b) Non-intimate Relationships*	C3 – Recent Problems with Symptoms of Major Mental Disorder: *a)Psychotic Disorders* *b)Major Mood Disorders* *c)Other Major Mental Disorders*	R3 - Future Problems with Personal Support
H4 – History of Problems with Employment	C4 – Recent Problems with Instability: *a)Affective* *b)Behavioural* *c)Cognitive*	R4 – Future Problems with Treatment or Supervision Response: *a)Compliance* *b)Responsiveness*

Historical (10)	***Clinical (5)***	***Risk Management***
H5 – History of Problems with Substance Use	C5 – Recent Problems with Treatment or Supervision Response: *a)Compliance* *b)Responsiveness*	R5 – Future Problems with Stress or Coping
H6 – History of Problems with Major Mental Disorder: *a)Psychotic Disorders* *b)Major Mood Disorders* *c)Other Major Mental Disorders*		
H7 -History of Problems with Personality Disorder: *a)Personality Disorders of the Anti-Social type* *b)All other Personality Disorders*		
H8 - History of Problems with Traumatic Experiences: *a)Victimization/Trauma* *b)Adverse Childrearing Experiences*		
H9 – History of Problems with Violent Attitudes		
H-10 History of Problems with Treatment or Supervision Response: *a)Compliance* *b)Responsiveness*		

In the final stages, the assessor is asked to integrate the separate risk factors relevant to that individual into a formulation, in order to explain how an individual came to perpetrate violence. Following this, risk scenarios are then developed, in an attempt to determine what that individual may do in the future and how likely this is. Risk scenario planning is stipulated as necessary in V3, whereas in the past it was not. This then allows for the development of robust risk management strategies, allowing for a structured and comprehensive plan, which assists in reducing the likelihood that violent behaviour will occur. The final step for the assessor is to develop conclusions regarding an individual's risk and provide summary risk ratings (SRR's).

Psychometric Properties of the HCR-20 Version 3

Kline (1986), outlines that there are a variety of factors to consider when developing a 'good' psychological test, including; reliability, validity and appropriate norms. Despite the HCR-20 not being a typical 'psychometric test', using the criteria outlined within literature it is possible to assess its adherence to the properties of a good psychometric test, and therefore draw conclusions regarding its clinical utility.

Reliability. Reliability relates to the consistency of a test (Kline, 1998). This particularly relates to a test's stability over time (test-retest reliability), internal consistency and IRR. Test-retest reliability can be understood by, if a test were administered to one individual, several times without any changes to that individual, then the scores would be consistent. Kline reports, that a correlation of .8 or higher is required in order for the test to be considered as having good test-retest reliability. Internal consistency relates to how much a test measures the construct it is intending to measure, it looks at how highly correlated the items of a measure are, to establish whether they are all measuring the same thing. It is measured using Cronbach's

Alpha, often a figure of .7 is considered to be a minimum value for a good test (Kline, 1998). Finally, IRR explores the variance between one assessor and another when they are administering the same tool. For this a Cohen's kappa of 0.6 would be required in order to deem that the raters have substantial agreement (Viera & Garrett, 2005). The reliability of the HCR-20V3 is discussed below. The lack of literature regarding test-retest reliability could be considered to be an area that needs developing. As Kline (1998, p.29) suggests that this is "an essential attribute for any good measure whether psychometric or not" and therefore this could be considered as a limitation.

Studies have assessed the IRR of the HCR-20V3. For example, de Vogel et al. (2014), coded the HCR-20V3 on a sample of 86 Dutch forensic psychiatric patients. They found IRR to be .72 which indicates good agreement for an intraclass correlation coefficient (ICC). The study also compared this to the sum of numerical ratings of presence made using the HCR-20V2, and found the IRR to be higher at .83. This would indicate that V3 performs equally well or better on IRR compared to V2, suggesting that changes made have improved its consistency between one rater and another. Douglas and Belfrage (2014), evaluated IRR of the HCR-20V3 in 35 forensic psychiatric patients in Sweden. They found that, IRR of the sum of numerical presence ratings were good to excellent. For total and scale scores they averaged .85, and for relevance ratings .70. IRR for summary risk ratings (SRR's) was found to be .81 (institutional violence) and .75 (community violence). Across a total of 138 paired ratings, 86.15% were in perfect agreement. Further in support of good to excellent IRR in the HCR-20V3, Doyle et al. (2013), analysed it in a sample of 20 forensic psychiatric patients in England and Wales. They reported the following values; Total (.92), Historical (.91), Clinical (.90) and Risk management (.93). These findings indicate high IRR for the HCR-20 V3 and would meet the criteria for a good psychometric test, however the sample utilised was small and therefore

this reduces external validity of findings. Interestingly, when comparing reliability scores on the scales, it would appear that the changes made in V3 have increased the IRR of these scales. Douglas and Reeves (2010), using V2, found; the Historical scale to produce a median coefficient of .86, Clinical scale, .74 and Risk scale, .68. This suggests that the coefficients for IRR for V3 are all higher and there is less difference between them. Indicating that, inputting indicators into risk definitions may have increased IRR on all scales. It could be argued however that IRR alone is not an ideal way of assessing the quality of a test, it is possible that assessors could agree on items even if they are not actually assessing anything relevant to the construct, in this case violence, and therefore this presents a limitation.

Other studies (Smith et al., 2014; Kotter et al., 2014) reported coefficients that were lower. Smith et al. reported the following coefficients; Historical (.92), Clinical (.67) and Risk Management (.68 institutional; .88 community). Kotter et al. used five clinicians who previously had no experience of the HCR-20 and trained them in its use. The coefficient reported for SRR's was .86. Average item coefficients were reported as; .65 (Historical), .66 (Clinical) and .73 (Risk management). Kotter et al. found lower coefficients for IRR for all three scales, with only the risk management scale meeting criteria for a good psychometric test. It is unclear whether the historical, clinical and risk scales are reliable on Cronbach's alpha and therefore, it is not clear whether any of the scales are actually measuring anything well.

Evidence would indicate that the HCR-20V3 would meet the criteria in relation to IRR for a good test. In regards to internal reliability relating to internal consistency, at present research evidence relating to the HCR-20V3 is limited. Previous evaluations of the HCR-20 assessments internal consistency, have produced coefficients of .95 for the whole HCR-20 assessment (Belfrage, 1998). Further to this, variations in coefficients have been found across the three scales from .85 to .96 (Dunbar, Quinones & Crevecoeur, 2005). These studies looked

at V1 (Belfrage, 1998) and V2 (Dunbar et al., 2005) and indicate internal consistency across both tools. Ross, Hart and Webster (1998) reported Cronbach's alpha of, .74 (historical scale), and .64 (clinical scale) for V2. There is therefore research evidence, which suggests that both Version 1 and 2 of the HCR-20 have acceptable (.6 -.7) to good (.8 or higher) reliability, although the clinical scale does not look good and may not be acceptable. It is likely therefore that similar findings would be found for V3 and possible that the clinical scale may need further development.

There is a lack of empirical research evidence at present in relation to the internal consistency of the HCR-20V3. Eidhammer, Selmer and Bjorkly (2013) did however conduct such a study. They examined a Norwegian sample of 20 forensic psychiatric mental health patients and compared versions 2 and 3. They reported coefficients of .85, .59 and .81 for the scales. They also found equivalent coefficients of .84 for the total sum of scores in V2 and the presence ratings in V3. This suggests that they found moderate to good estimates of internal consistency between the two versions, although the clinical scale again would not meet criteria for a good test. This indicates that the two versions reflect common underlying dimensions in relation to violence risk. They reported however that there were still differences between ratings of the same patients using different versions. This and the fact that this appears to be the only study examining its internal consistency, suggest that further empirical research evidence is needed.

Validity. There are various ways in which a test can be considered valid, unlike reliability there is no one validity coefficient for a test, as validity is theoretical. Face validity relates to how the test appears to the evaluator. Reviewing items of the HCR-20V3, it would seem that it could be said to have good face validity. Items appear to be relevant and logical in relation to the literature regarding violence risk assessment. The revision of the HCR-20 takes

into account developments in literature. For example, research relating to trauma and abuse has developed. Swanson et al. (2002), found that trauma experienced as an adult does impact on an individual's propensity to commit interpersonal violence. The HCR-20V3 accounts for this in item 'H8' where it looks at victimization and trauma across developmental stages. In addition more recent research evidence (Logan, Nathan & Brown, 2011), suggests that individuals who are not well understood and who are not risk managed with confidence, are more likely to continue their behaviour. Logan (2014) states that HCR-20V3 takes such evidence into account, with its focus on formulation as a way of overcoming these issues.

Content validity, relates to a tool's ability to incorporate all of the content of a particular construct. In relation to the HCR-20, authors could claim excellent content validity. Since it was first developed in 1995, it has evolved through authors examining and reviewing the scientific and professional literature. In addition to this, it takes into account what is needed for practical utility by practitioners. As outlined by Douglas et al. (2013) the HCR-20 has been widely adopted due to its extensive evaluation by independent researchers and also because of its clinical utility. It is however, important to note that further research evidence is required, specifically on the HCR-20V3 in relation to violent behaviour in male and female offenders. Further empirical research evidence which specifically establishes content validity of the measure and reports coefficients for this may be useful in supporting this further.

Criterion validity relates to how useful the measure is in predicting the criteria that is being assessed (Kline, 1998). Therefore, in relation to the HCR-20 this would be the ability of the final risk judgement in predicting violent behaviour or recidivism. There are two forms of criterion validity, predictive and concurrent. Predictive validity relates to the ability of test in predicting the outcome. Concurrent validity relates to the correlation of one test, with another

test taken at the same time (Kline, 1998). An example of this which relates to the HCR-20 is comparing the outcomes of the test with different versions of the test, taken at the same time.

There are several studies which explore the concurrent validity of HCR-20. De Vogel et al. (2014), reported a correlation of .93 between the two versions (2&3) total scores. Strub et al. (2014) similarly found a correlation of .91. The correlation for SRR's in this study was .98 and for the scales it was reported to be; Historical (.89), Clinical (.76) and Risk Management (.81). Douglas and Belfrage (2014), computed Pearson correlation coefficients between the averages of the two versions, for each scale. They found; Historical (.87), Clinical (.76), Risk Management-In (.67) and Risk Management-Out (.82). They suggest that reliability and validity between the two versions are comparable and this is evidenced by the high concurrent validity scores. Douglas et al. (2013, p.18) stipulate that "continuity of concept" was one of the guiding principles in the measures revision and that findings exemplify this. Similarly, Bjorkly, Eidhammer and Selmer (2014), reported the following coefficients; Historical (.85), Clinical (.59), Risk Management (.81) and Total (.84). This further emphasises that the HCR-20V3 has high concurrent validity relating to its predecessors and therefore would meet the criteria for a good test. At present there is no research evidence that establishes the concurrent validity of HCR-20V3 with other risk assessment tools, research which does compare V3 with other measures would be useful in exploring this further.

The predictive validity of the HCR-20 has also been examined, looking at the association between the tool and subsequent violence that is perpetrated. Much of this research, uses receiver operating characteristic (ROC) analysis and reports the area under the curve (AUC) to assess the tools predictive validity. Douglas et al. (2013) highlight that AUC values of .70 may be considered moderate, and .9 and above may be considered large. Doyle et al. (2013) prospectively followed 387 male and female patients and measured violent outcomes at

six and twelve month time increments. They reported the following significant AUC values; Total (.73), Historical (.63), Clinical (.75) and Risk Management (.67). At twelve months AUC values were; Total (.70), Historical (.63), Clinical (.71), Risk Management (.63). In addition logistic regression analyses was carried out on the total scale. Authors found that the total scale indicated it was significantly predictive of both six and 12 month violence even when variables such as age and gender were controlled for. These values only indicate moderate levels of predictive accuracy however and therefore may indicate that V3 is lacking in this area. De Vries Robbe and de Vogel (2010), also found moderate values for the predictive validity of the HCR-20V3, for community violence at one, two and three years (.77,.75 and .67 respectively). Whilst these values indicate moderate predictive accuracy and thus to some extent may suggest the HCR-20V3 meets criteria for a good test, this is not sufficient alone to determine whether a test is a good one. It should be noted that AUC values only predict levels better than chance, meaning there is still considerable error. It is therefore questionable whether this is 'good enough' when it informs decisions which affect someone's life.

Comparative (Normative) Data. Comparative data is not reported in the manual for the HCR-20V3, the absence of this is a potential flaw of the assessment. Comparators provide a basis by which assessors can compare test data and would offer evaluators an insight into the base-rate of violence occurring within a population. Providing comparators would also make it easier to establish potential gender and population differences. Kline (1998), relates such data with providing meaning for that test. Therefore whilst there are empirical research findings which suggest that the HCR-20V3 meets criteria for a good test in terms of reliability and validity, this criteria could be used to draw conclusions which indicate otherwise.

It is surprising that despite all the research conducted in relation to the HCR-20 since its inception, no comparative data has been developed. By presenting such data, this would

more easily allow for comparison research to be conducted and for clinical utility of the measure to improve. Due to the lack of comparators provided, questions arise in relation to how the HCR-20 can claim to be a standardised risk assessment tool. This then questions the degree to which it can discriminate between different groups such as males or female in relation to violence prediction. Growing literature does however provide information in regard to its use as a clinical tool within various population groups; for example North America, Western Europe and the United Kingdom, however at present in relation to the HCR-20V3 this remains limited and further expansion is required.

Limitations of the HCR-20 Version 3

As aforementioned, one obvious limitation of the HCR-20 V3, is the fact that it has a lack of comparators, in addition to this there are a number of other limitations. The HCR-20V3 does meet the criteria in several areas for a good psychometric test, although this is limited in certain areas. Further to this, limitations relate to its clinical utility as opposed to its psychometric properties. Such limitations could however potentially impact upon the overall validity and reliability of the assessment.

De Vogel et al. (2014), comment that their overall conclusions regarding the HCR-20V3 are that it is a strong revision with enhanced clinical utility. Despite this, they highlight that new studies carried out in different institutions and countries are much needed. This would improve information regarding the validity and reliability of the tool and also mean that more would be known about its use in different settings, by different practitioners, and with different populations. This critique presents the empirical research studies which have examined the HCR-20V3. Whilst these do provide information regarding the reliability, validity and clinical utility of the tool, these also emphasise the need for more research. In addition, it is difficult to

be certain regarding the statistics that are reported, as the studies completed analyses on different drafts of V3 and some of the research was not published. This is a limitation and additional empirical research evidence is now needed. Retrospective and prospective research which focuses on IRR, concurrent and predictive validity would provide further understanding regarding the clinical utility of the tool in the assessment and management of violence risk in other countries. It would also allow for more robust conclusions regarding reliability and validity to be drawn.

Furthermore, Logan (2014) emphasises that research relating to formulation, which provides the link with risk management strategies, is still limited. Whilst this link has been implied, there is at present no evidence which clearly links risk formulation and effective risk management plans. There is research which outlines a model for evaluating risk formulations (Hart et al., 2011), that allows for the determination of a formulations acceptability. As yet however, Logan suggests that there is no evidence which establishes whether acceptable risk formulations are linked to the development of effective risk management strategies. This would be useful for the HCR-20V3, particularly as authors place so much emphasis of the importance of formulation in improving the measures of clinical utility.

Unlike other psychometric tests such as those which measure anxiety or anger, the HCR-20V3 does not provide the assessor with a composite score. This may assist with the clinical utility of the tool as it would allow practitioners to infer the level of violence risk and assess how best to address it. However, a composite score would provide no information with regard to what type of violence an individual is likely to commit and the severity of such violence, which would make the development of risk management plans difficult. Previously criticism has been levelled at earlier versions of the HCR-20 (Douglas, 2008), due to the 'over-breadth' of content in each of the items, thus making them ambiguous. This links to arguments regarding

a lack of composite score. Depending on which factors a rater weighs as relevant to that individual, individuals could come out as having similar risk levels, but significantly different risk profiles. The HCR-20V3 however appears to have come some way in overcoming this. The addition of risk indicators into the content of each item allows for the consideration of more dynamic risk profiles. Not having a composite score allows the assessor to think about the individual, particularly in relation to the emphasis on formulation and risk management and this could be considered a significant strength in assisting with the tool's clinical utility. Although this does open the assessment process to greater bias and misinterpretation, for example, if clinicians can interpret items in multiple ways then this could mean the tool is fundamentally flawed.

Finally, another limitation of the tool is that it is costly in terms of time, effort and training. It can only be used by individuals who are trained and have sufficient knowledge of the completion of risk assessments, clinical practice and theory. Harris and Rice (2015), suggest that a significant limitation of V3 is that it reduces a total score to a rating with three values, resulting in considerable information loss. It is possible, however, that if an experienced risk assessor does conduct the risk evaluation it is unlikely that salient information will be lost. Additionally, the inclusion of relevance ratings should ensure that important information that contributed to an individual's risk of violence is carried through to formulation and risk management planning. Douglas et al. (2013) highlight in their manual, that it may be useful to take a multi-disciplinary approach to scoring and de Vogel and de Ruiter (2006), found that AUC's for consensus ratings were stronger ranging between .77 and .86 for the HCR-20V2. This would limit misinterpretation and bias in the assessment procedure as a whole and therefore would also improve clinical utility.

Conclusions

This critique has aimed to review and appraise the literature relating to the reliability, validity, and clinical utility of the HCR-20V3. It has assessed this in relation to previous versions and outlined what is now needed in order to further validate this tool. In addition it has aimed to evidence strengths, limitations and areas of improvement particularly in relation to clinical utility, risk formulation and the development of risk management plans.

The HCR-20 is not what would be considered as a classic psychometric tool, it is instead a framework for use by clinicians to assess risk of violence, and is strongly based within theory. Inevitably however, there has been discussion within the literature in relation to ideas of classical test theory such as reliability, validity and normative data. An appraisal of this research evidence which related to the HCR-20V3 would indicate that it fails to meet all criteria to be considered a good psychometric test. For example its validity is questionable, items ambiguous and it measures a concept which has multiple meanings and interpretations for different professionals. It could also be criticised for using single items to assess complex clinical features. Despite this, it has a number of strengths. Although much research evidence which supports these comes from the author of the tool, highlighting therefore its main limitation that further empirical research evidence is required. The HCR-20 remains one of the most widely used tools for the purpose of assessing violence risk. Going forward, research must now focus on comparing it to other measures of violence risk and explore this in a variety of different populations and settings.

CHAPTER FOUR

An Exploration of Practitioners' Understanding and Use of Index Offence 'Work'

Abstract

Aim

Due to the lack of literature regarding index offence work/ analysis, this study aimed to explore clinical and forensic psychologists' understanding and use of these, in the context of their work with offenders.

Method

A thematic analysis was performed on discussions held between practitioners during four focus groups which were conducted by the author. Discussions surrounded practitioners' use and understanding of such work.

Results

The qualitative methodology used, resulted in the extraction of a number of themes relating to the processes utilised by practitioners in index offence work, these were consistent with models of offender assessment and rehabilitation. A number of themes were also obtained, which highlighted the lack of consistency and understanding amongst professionals and services.

Conclusions

Whilst there is evidence of common areas of good practice across services in relation to index offence work, understanding regarding what the work involves, evidenced inconsistencies. Index offence work within forensic mental health settings, requires an individualised approach and as such specific guidelines whilst useful, may be difficult to implement. Findings indicate that at present IOW is not routinely implemented by practitioners. Areas for future research are discussed.

Introduction

The Assessment of Offenders

As evidenced in previous chapters, the importance of assessing offenders and their risk accurately, is of significant importance. West (2000) suggests, that any assessment which does not take account of offence data is deficient and it is possible that detailed information regarding an individual's index offence, can provide more in depth information regarding personality difficulties. In addition, other researchers (Daffern et al., 2007; Weist, 1981; West & Greenhall, 2011), have suggested, that index offence analysis (IOA) should be a core task of any forensic clinician who is engaged in the assessment of offenders. It is proposed that this should provide as full an account as possible, of an offender's relevant criminal events and should also use available crime scene photographs and associated reports. It is argued that this information can then be used to aid formulation related to the individual's risk and treatment planning.

Completing a full assessment including detailed information in relation to an offender's index offence is inevitably difficult; offenders' accounts of their offence may not always be genuine. For example, Harry (1992) found that offenders, who had committed more violent index offences, are more inclined to deny, minimise the severity of, or blame their crimes on accidents, alcohol, drugs, uncontrolled emotional arousal or situational factors, than offenders who have committed less violent offences. Spence (1989) reports that relying on an offence account from an offender, can result in a 'consolidated narrative' which is constructed by that individual, and in accepting this account, colluded with, by the practitioner. Similarly, Melton, Petrila, Poythress and Slogbin (1997) also highlight that relying on an offender's narrative can result in possible distortions of the truth. It seems that assessments not taking this information into account are inevitably impoverished.

The assessments which clinicians complete feed into important decisions regarding admission, treatment, risk of reoffending and discharge from secure services. It is difficult to establish how such decisions can be made if practitioners fail to have a full awareness and understanding of an individual's index offence. It would seem that this would then result in the practitioner's knowledge of the individual being limited, potentially making their assessment redundant. Information gained from an in depth review of the offence, may enable practitioners to devise hypotheses regarding the aetiology of an individual's offending, by clearly identifying intrapersonal, interpersonal and situational variables, that had an impact upon the individual. However, if clinicians working with and assessing offenders in their everyday practice do not have knowledge of the benefits of index offence work (IOW) or what such work entails, it is possible that they will be ineffective in using it. The implementation therefore of a structured framework to assist practitioners in the completion of IOW/IOA may be useful. Such frameworks have been developed over the last twenty years in relation to the assessment and prediction of risk, for example SPJ tools. As indicated by Douglas and Belfrage (2014), the SPJ approach allows for a structured professional decision making system to help facilitate professional risk assessment and management. This approach is now widely used with positive results within the risk assessment field. This may indicate that a similar structured framework or tool in the area of IOA/IOW may also be useful for practitioners working within forensic settings.

Research relating to Index Offence Analysis/Work

Psychologists working within secure forensic settings often have a number of methods which they utilise in order to assess their clients. For example, the traditional ABC functional analysis model has been used in the assessment of sex offenders (Beech, Fisher & Thornton, 2003), stalkers (Westrup & Fremouw, 1998), fire setters (Murphy & Clare, 1996) and violent

offenders (Ireland, 2009). A functional analysis approach looks at the outward presentation of the individual's behaviour but focuses more on the function of such behaviours. It typically involves the practitioner obtaining detailed information about the antecedents, the behaviours, and the consequences of offending (the ABC model). Beech et al. (2003) stated that it should include the actual behaviours carried out, along with the accompanying thoughts and emotions. Other methods include a SORC analysis, as proposed by Ireland (2009). The SORC (S: setting conditions; O: organism variables; R: response variables; C: consequences) incorporates the developmental history and learning experiences of the individual. Whilst these methods of analysis could be and are applied to understanding an individual's index offence, as highlighted by West and Greenhall (2011), currently there appears to be no such structured tool which incorporates these ideas, and assists practitioners in specifically carrying out the analysis of an individual's index offence. It is suggested that little is known within the research literature regarding what IOA/IOW comprises and there appears to be no formal definition. Although, West and Greenhall (2011) suggest that IOA can be defined as;

> "The formal and structured examination of the events, circumstances, and behaviours that occurred before, during and after the last set of criminal actions that brought an offender into contact with the criminal justice system". (p.144-145)

Knauer and Wilkinson-Tough (in press) suggest that as clinicians, their experience of IOW is a process which assists the multidisciplinary team to understand the factors which brought the individual into secure services. In their chapter they refer to it as a piece of work which is similar to a functional analysis which focuses on the individual's offence. A piece of work such as this is evidently likely to assist practitioners in identifying the factors involved in the offence and enable them to link this to potential future risk. In addition it could aid practitioners in identifying treatment targets for the individual, creating an active account of the

offence and provide an understanding of potential cognitive distortions (i.e. thoughts associated with justification and neutralisation of the individual's offending behaviour).

Within the research literature there appear to have been some attempts to look at the ways in which clinicians may be able to incorporate IOA into the assessment process. Gresswell and Hollin (1992) propose a multiple sequential functional analysis methodology to be applied by clinicians retrospectively to an offender's criminal behaviours. They recommended that this would provide a clear and coherent summary of the individual and enable practitioners to identify situational variables that may promote further offending. Similarly Green (2008), proposed the use of functional analysis on an individual's offending behaviour, with practitioners asking specific questions in relation to the location of the offence, the victim and what the offence entailed. The concept of practitioners using formulation or functional analysis in order to incorporate IOA into their assessments of offenders is not a radical idea. It is already used widely in the general assessment of offenders and the idea itself has been around since Lazarus (1971). It is common practice that clinical and forensic psychologists utilise this model when working with complex individuals (Hanley, Iwata & McCord, 2003). Best practice guidelines for formulation as indicated by the British Psychological Society (BPS), state that formulations of presenting problems or situations, should integrate information from assessments within a coherent framework. This framework should draw upon psychological theory and evidence which incorporates interpersonal, societal, cultural and biological factors (BPS, 2011).

Wood et al. (2002), suggest that case formulation is one of the most difficult tasks faced by practitioners, having found that asking clinicians to accurately identify why individuals behave in a specific way is surprisingly difficult. More recent research (Davies, Jones & Howells, 2010; Jones, Daffern & Shine, 2010), has identified the need for further development

of an approach to the assessment of offenders, which focuses on understanding more about the complexity of violent behaviour, in order to aid decisions in relation to risk. McDougal, Pearson, Willoughby and Bowles (2013) in their research, looked at the contribution of examining offence-related behaviour of offenders in prison, prior to their release, and the contribution of this to risk prediction. They found a strong correlation between observed negative behaviours in prison and their frequency in the community. Whilst this is a correlation and therefore only indicates a relationship between the two variables, findings also showed that the frequency of these behaviours, significantly predicted the offenders who would reoffend or be recalled to prison. This therefore suggests that exploring offence related behaviours in greater detail can have an impact upon the prediction of risk and is something that practitioners could potentially utilise in their assessment of offenders.

McDougal et al. (2013) mention that criminological research indicates that one of the major predictors of recidivism is the type and frequency of previous convictions, with Clark, Fisher and McDougal (1993) highlighting that past behaviour is the best predictor of future behaviour. These findings suggest, that it is possible to accurately predict offence-related behaviour in the prison environment, on the basis of an objective behavioural analysis of the offence. In practice however, this may not be as easy, for example, in cases where the individual has committed murder, the offender may not have a history of previous offences. Dobash, Dobash, Cavanagh, Smith and Medina-Ariza (2007) found that 13% of murderers had no previous convictions; they therefore suggested that further investigation in relation to the assessment of these types of offenders is needed, including a focus on the type of murder and specific situation and contextual factors that may further explain their offending. They indicate that this supports the need for a robust case formulation approach to the assessment of individual cases and offences by practitioners. Herman (1990) suggests that, often in

psychological formulations in regard to the motives of sex offenders, “the sexual offence virtually disappears” (p.182).

Jones (2004), in contrast to others within the literature, suggest that a case formulation approach to assessment such as functional analysis, only focuses on historical discrete episodes of an individual’s offending behaviour. It could be argued that this is as a result of the fact that this is the only information available to practitioners working within certain forensic settings. This does make sense as generally within forensic settings, interventions with offenders involve a systematic exploration of an individual’s offence; for example the ‘decision chain’ used in sex offender treatment programmes within the prison environment. This looks at a sequence of choices leading to an individual’s offence. Beech et al. (2003) highlight that it is characterised by the situation in which it took place, the thoughts that made sense of and responded to the situation, and the emotions and actions that arose from those thoughts. Whilst this may be useful in providing a practitioner with information regarding an individual’s index offence, it relies heavily on an individual’s ability and willingness to self-report these events, which can be problematic in itself. In addition, the utility of such a process is dependent upon what is then done with it, whether it is further analysed to provide information regarding an individual’s problematic personality characteristics relevant to their offence for example, or whether it is just completed as a tick box within a wider intervention. The latter could potentially result in a failure to address some aspects of individual’s offence related thinking, feeling and behaviour. Further to this, the offenders that clinicians are working with often have to go over their offence a number of times, for example in court, parole hearings, assessments and therapeutic settings, and as suggested by Jones (2004), this may result in the account having a lack of emotional impact upon the offender. In addition this type of work can be difficult, with offenders reporting that they have poor recollections of the offence due to psychosis or

substance use (Pyszora, Barker & Kopelman, 2003), or a self-defensive strategy that could be conscious or unconscious. There is therefore a suggestion that a case formulation exploring 'Offence Paralleling Behaviours' (OPB's) by clinicians would be more beneficial in this type of work (West and Greenhall, 2011), however this formulation approach would still assume that there was a knowledge of the index offence.

There is evidence within the OPB research which suggests that situational and contextual factors play a significant role in the understanding and prediction of future behaviour (Daffern et al., 2007; Jones, 2002a). Therefore, it seems logical that practitioners working with complex individuals, should pay more attention to specific aspects of the index offence and use these when developing formulations that will inform treatment and risk related decisions. The tool proposed by West and Greenhall (2011) to be used for IOA, places an emphasis on including empirical evidence such as that relating to typologies of offenders, (sexual murderers, rapists, child molesters and serious violence), by clinicians in their IOA, rather than just the self-report of behaviours from the offender themselves. They also emphasise the need to corroborate and collate this information in relation to other collateral sources such as witness depositions and crime scene evidence. This is likely to provide a more robust assessment of an individual's index offence, as for example looking at typologies alone may be limited due to their lack of empirical support and are in fact relatively descriptive, resulting in the assessment being somewhat impoverished. What is clear from the research literature is that there is evidence to suggest some form of IOW should be conducted when working with offenders, however there is disagreement in regard to what approach would be most beneficial and a lack of information in regard to what such work would involve.

The use of Index Offence Analysis/Work by Clinicians

There is official guidance which identifies that IOW/IOA should be undertaken in forensic and clinical practice, and clinicians are often required to have knowledge of the offenders' index offence and report on it to relevant parties at various stages of their work. For example, those working with offenders who provide the Ministry of Justice with reports in relation to restricted offenders in secure hospitals, have to respond to questions relating to the factors underpinning the index offence. This inevitably is not always an easy task when working with individuals who present with very complex needs. In addition, the Department of Health (2008) outline that doctors assessing offenders for admission to secure hospital should request 'relevant' documentation and information in relation to the offence. However, such guidance poses questions regarding what different professionals may interpret as relevant. The Risk Management Authority (RMA; 2006) stipulates a requirement for clinicians to undertake an analysis of offenders past and current offending, which sets out the specific criminogenic factors relating to the offender. The RMA state that this should include a detailed analysis of patterns of behaviour, motivation, antecedents and diversity of offending. Without this robust understanding of offenders' forensic histories, it is possible that failures within the criminal justice system may occur, such as that indicated by Fallon, Bluglass, Edwards and Daniels (1999) or in other reports (Reed, 1997; Hill, 2009; HM Inspectorate of Probation, 2006). The findings of these investigations highlight that practitioners may have failed to have a robust understanding of their clients' forensic histories and their potential to commit further offences, resulting in significant offender management failures. Inquiring more thoroughly into the criminal history of the offender they are working with, should result in a more accurate appraisal of an individual's risk.

Despite extensive guidance outlining the requirement for such work to be carried out, some master's research found that this still does not happen in practice and that in actual fact there is little understanding in relation to what IOW/IOA is or what it involves. Fallon (2007, as cited in West & Greenhall, 2011), investigated the level of knowledge of the clinical team in relation to patients' index offences in a medium secure unit. Findings indicated that staff who were qualified and more senior had a greater awareness of index offence information, however most staff had not seen witness depositions or crime scene photographs and many staff working with offenders on a daily basis had a limited awareness of their index offence. It is evident that this could be inevitably problematic as a full picture of an individual's offending behaviours and potential OPB's cannot be identified and therefore inform risk management decisions and treatment plans appropriately.

It is likely that there are practitioners who include IOW/IOA within their assessment of offenders. However, as identified by West and Greenhall (2011) due to the absence of a formal protocol or guide, the process that they use is likely to be unstructured which then risks compromising the validity of the assessment. In addition, it means that the use of IOW as part of the assessment process, is likely to be not as widespread or as thorough as it should be. They also provide a proposed guide using the principles of functional analysis which ensures that index offence information is collected, analysed and incorporated, into the assessment process by practitioners. It uses an offender's account of their index offence as a baseline to compare collateral evidence against, and relates this to relevant empirical research evidence for that individual's offence (West & Greenhall, 2011). There is however, no research evidence which indicates that practitioners throughout the United Kingdom are currently using such a framework or whether they have developed their own for use within their service. There is also no further research similar to that of Fallon (2007, as cited in West & Greenhall, 2011), which

indicates the extent to which practitioners have knowledge of, or understand the importance of, including index offence information in assessments and how they may do this.

Without clear research evidence, it is difficult to know how many psychologists review collateral sources of information in their assessment of offenders and if this involves an analysis of the individual's index offence. It is unclear whether clinicians just rely on the most recent data in relation to the offender, in order to avoid a time consuming trawl through original documents, which may provide valuable information in informing an assessment. Doyle and Dolan (2002) suggest that an improved assessment procedure, will result in an improved formulation, which will provide a comprehensive explanation of offending behaviour in each individual case. Furthermore, it is likely that this will further assist in leading practitioners to be more specific in the identification of treatment targets. MacCulloch, Bailey and Robinson (1995) emphasise that a complete assessment is essential for formulation and appropriate treatment, and it is apparent that IOW/IOA is something which is necessary to ensure a complete assessment.

As outlined here, to date, there is limited research in relation to the concept of IOA/IOW, including practitioners' understanding of what it involves, how they carry it out, the existence of standardised protocols, and how these are all incorporated into their assessment of offenders. Some clinicians may undertake other work prior to, or following offence work that allows offenders to engage with the work, process the experience, find a way to live with the emotions resulting from an offence, and commit to a life without offending. This highlights the fact that different practitioners may employ different methods and procedures which comprise IOW. At present there is no research which analyses this and more specifically there has been no qualitative research carried out, which explores clinicians' understanding and use of IOW/IOA.

An approach such as this has the potential to identify indicators and common areas of good practice in relation to the assessment of offenders and their index offence.

Aims

This study aimed to explore clinical and forensic psychologists' understanding and use of IOA/IOW, in the context of their work with offenders in secure hospitals. This was achieved through applying a bottom-up thematic analysis (TA) of practitioners' accounts of the content of their IOW, distinctions between IOW and IOA and personal challenges that they face when conducting such work.

The main research question is therefore:

How do clinicians working within secure forensic hospitals understand and use Index Offence Work and Analysis?

Method

Analytic Approach

This study adopted a qualitative approach to data collection and analysis, some quantitative analysis was also conducted on the questionnaire data that was collected, in order to provide descriptive statistics regarding IOW. The data gathered from the questionnaires was put into a spreadsheet and analysed using SPSS statistics package (version 22) and then was interpreted by the researcher. The quantitative analysis was minimal as the primary analytic approach in this study is qualitative. This was chosen in order to ensure that sufficiently rich data was collected, in order to generate unique and distinctive information relating to practitioners' understanding, and use of, IOW. Within qualitative approaches it is regarded as

important to explain the rationale behind method selection (Guest, Bunce & Johnson, 2006; Guest, McQueen & Namey, 2012). A full explanation of the decision making process and the qualitative methodology utilised by the author can be found in Appendix F.

Participants

Participants in this study were recruited from four secure forensic hospitals who provided their ethical approval for the research to be carried out using their staff. The secure hospitals were located in the Midlands and South West of England and included low and medium security. Three of the hospitals were NHS hospitals and one was a private hospital, the hospitals had a mixture of male and female wards. Following ethical approval being gained, the head of psychology at each site was approached through a formal letter or email, which invited qualified clinical and forensic psychologists to participate in the research. A summary of the research (Appendix G.), was included and then cascaded to potential participants, this was so that they were able to opt in to focus groups, prior to seeing the formal participant information sheet (Appendix K.) and signing a consent form (Appendix L.) at the focus group stage. The researcher then liaised with participants to establish an appropriate time to visit each site and conduct the focus groups.

As recommendations for qualitative studies suggest (Guest et al., 2012), a purposive sampling method was used by the author, which means that the sample were chosen because they fulfil a common criteria, the inclusion criteria (Guest et al., 2006). The inclusion criteria for participants in this study, were that all participants were qualified forensic or clinical psychologists who work with service users, and undertake assessments and treatment with them. This approach was taken as it was deemed that the qualified clinical and forensic

psychologists that would participate in this study, would be able to provide rich descriptions of the specific phenomenon that was being explored.

Following the recruitment process a total sample size of 21 participants was recruited. Four focus groups were run (one at each site) and each focus group lasted between 75 and 90 minutes. Three of the focus groups had five participants and one had six. Braun and Clarke (2013) recommend that when conducting a moderate sized study such as this one, between 3-6 focus groups should be run. The author had hoped that they would be able to run two focus groups at each site, resulting in a larger sample size, however due to staffing issues and availability, as well as the fact that one service was undergoing a significant restructure, this was not possible. Despite this, the sample recruited is still in line with the recommendations for qualitative research (detailed above). Of the 21 participants recruited, 15 were female and 6 were male. A summary of participant details can be found in Table 8.

Table 8.

Summary of Participant Details

Description	Total Number
Gender	
Male	6
Female	15
Job Description	
Clinical Psychologist	12
Forensic Psychologist	5
Clinical and Forensic Psychologist	1
Consultant Forensic Psychologist	1
Consultant Clinical and Forensic Psychologist	1
Lead for Psychology Secure Services	1
Hospital Type	
NHS	15
Private	6

Data Collection

Data were collected via engaging participants in focus groups regarding IOW. The focus groups were conducted by the author, a Trainee Forensic Psychologist. The author liaised with participants in order to establish a suitable time and location to run the focus groups. The rooms were of reasonable size (to fit a maximum of six people), private and had refreshments available for participants. To avoid any major disruptions focus groups were conducted at participants' place of work. Focus groups were carried out at a variety of dates and times to suit the services involved.

The author developed a questionnaire (Appendix H.) which aimed to explore clinicians' understanding of IOW. The questionnaire was designed to provide some descriptive information in relation to practitioners' understanding and use of IOW including: the frequency of this work; the presence of existing policies/protocols within the service; and the number of hours spent completing this work on a weekly basis. A semi-structured interview schedule (Appendix I.) was also developed by the author, and questions were considered to be pertinent to the research question being explored. This was for promoting discussion of participants' understanding of IOW, how they conduct this work, as well as the challenges that this type of work poses. In addition, the schedule contained prompts for the author to use for each question, in order to promote discussion and ensure that participants fully understood what they were being asked to discuss. The prompts also enabled flexibility and enabled the researcher to explore issues which emerged during the discussion. Each focus group was video recorded, as opposed to just audio recording, due to the fact it may have made it difficult to distinguish between participants when it came to transcription.

Participants were instructed to arrive at the room, ten minutes prior to the commencement of the focus group. This was in order for the author to re-administer the participant information sheets, answer questions, obtain informed consent and allow participants to get refreshments and get comfortable. Once informed consent had been gained participants were tasked with completing the questionnaire. Participants were reminded that they were being video recorded and the focus group then began. Following the completion of the focus group, all participants were thanked for their participation and were provided with contact details should they have any questions following the completion of the research.

All focus groups were transcribed by the author from the video recordings, verbatim with all identifying information being redacted during the transcription process, and only

numbers used to identify participants. Any significant non-verbal information was also inserted into the transcripts by the author, including any pauses in the focus groups for comfort breaks or disruptions. Following transcription video recordings were deleted and only the transcriptions were kept for analytic purposes. McLellan, MacQueen and Neidig (2003) highlight that transcripts are useful in qualitative research when it is being carried out at multiple sites. All data were stored securely with only the author having access to it.

Ethical Considerations

Ethical approval for this study was granted by the University of Birmingham's Science, Technology, Engineering and Mathematics ethics committee. It was also granted by the respective R&D departments at each of the proposed sites (see Appendix J.).

Informed consent. Participants were provided with the research summary (Appendix G.) initially so that they were able to opt in to the focus groups. Upon confirmation of a participant's interest in taking part in the research they were provided with a participant information sheet (Appendix K.) and this was again provided prior to gaining informed consent at the focus group stage. This outlined to all potential participants the purpose of the study, as well as information regarding the storage of personal information. The information sheets also provided information regarding withdrawal and the secure disposal of data.

Prior to the focus groups all participants had the opportunity to read participant information sheets again and ask any questions. Each participant was then provided with a consent form (Appendix L.) to confirm that they were willing to take part in the study. All information was stored securely by the author. Participants were reminded at this stage of their right to withdraw prior to the commencement of the focus group. No participants

withdrew their consent to take part. Participants were not deceived in any way during this study, they were provided with details of the study and aims and objectives from the outset.

Confidentiality. All information and data were stored in encrypted files on a password protected computer that only the current author has access to. The video recordings of the focus groups were also kept on a password protected computer and were deleted once transcribed. Transcriptions of the focus groups are also stored in encrypted files.

Personal information and consent forms were stored and anonymised and kept apart from questionnaire information. Participants were notified through the participant information sheet and again when providing informed consent that they would be quoted in transcripts for the purpose of analysis, but would not be identifiable. Upon the completion of this research all personal information was destroyed. Due to data protection, transcripts must be kept for a period of 10 years, the author has saved these to an encrypted memory stick and these will be kept securely for this time period.

In order to ensure the confidentiality of service users which were referred to during the focus groups, all participants were informed, to refrain from using names or other identifying information about service users, where this did occur this was redacted from transcripts. At the start of each focus group participants were reminded to keep discussions confidential.

Data Treatment

The present study followed the comprehensive seven-step guidelines for conducting TA as outlined by Braun and Clarke (2013), a summary of this can be found in Table 9. In order to be consistent with the main principles of this approach, the stages were completed in an iterative manner by the author. This was opposed to retaining all initial codes and themes that were developed as the final product of the analysis. Appendix M. details a step-by-step

explanation of the data treatment process as carried out by the author.

A TA approach was selected in order to analyse data qualitatively and highlight the broader themes surrounding the nature of IOW. TA is a method for identifying, analysing and reporting patterns or themes within a set of data, in the case of this research study this relates to the transcripts produced by the author following conducting focus groups. It organises the data and describes it in rich detail (Braun & Clarke, 2006). TA is not wedded to any pre-existing theoretical framework and therefore can be used within different theoretical frameworks. It was therefore considered by the author to be a good choice for this type of research, where there is little existing information and no theories/models already developed or defined.

Table 9.

Process for Data Analysis based on Braun & Clarke (2013)

Stage	Process
1	Transcription – Transcribing video data from focus groups using orthographic transcription.
2	Reading and familiarisation – Reading and re-reading the data and noting down initial ideas in relation to the research question.
3	Coding (complete), across the entire dataset – Coding interesting features of the data in a systematic fashion and collating data relevant to each code.
4	Searching for themes – amalgamating codes into potential themes and gathering all data into each potential theme
5	Reviewing themes – Checking if the themes in relation to the sections of coded data. Generating a thematic map of the analysis.
6	Defining and naming themes – This is ongoing and involves refining the specifics of each theme and the overall story that the analysis tells about the data. Generates clear definitions and names for each theme.
7	Writing and finalising analysis – The final opportunity to perform analysis on the data, final analysis of selected data extracts, relating back the analysis to the research question and literature, producing a scholarly report of the analysis.

Validity and Quality

Guest et al. (2012) suggest that assessing the quality of qualitative research requires the use of different criteria than those that are used for assessing the validity of research which utilises a quantitative methodology. In TA, they suggest that face validity is an important concept, and that transparency of the process is critical for making a convincing case for research findings and interpretations. As such all stages of the data collection and analysis process of this research have been reported. This should enable those utilising the research to

make informed assessments on the credibility and validity of the research findings. Furthermore, the use of a transcriptions protocol as discussed in the data treatment section, enhances validity of the data by ensuring consistency.

Other ways in which the validity of this research study has been enhanced is through the production of a codebook (Guest et al., 2012). During the development of this codebook, themes, sub-themes and codes were reviewed by an academic peer of the author, who was conducting research utilising the same methodology. This enabled a discussion regarding the names and meaning of themes and enabled the author to review and re-name these. This will have increased the face validity of the themes and sub-themes developed by the author. Furthermore, the author has supported all themes with quotes in the results section, to further increase the validity of findings, as it directly connects interpretations with what participants actually reported within the focus groups.

Interrater reliability (IRR) is important when considering the reliability of qualitative research. IRR signifies the extent to which two or more data analysts code the same qualitative data set in the same way (Guest et al., 2012). The development of a codebook assisted with the author establishing IRR. The academic peer independently applied the codebook to 50% of the transcripts. An agreement matrix was developed to check how consistently the author and peer had assigned codes and themes to the transcripts. The independent rater reviewed the data, so provided checks on individual biases and this also accounted for the variance in interpretation of code definitions. An IRR analysis using the Kappa statistic was used to determine consistency between the coders. The Kappa statistic was found to be 0.51, indicating a moderate level of agreement between coders (Vierra & Garett, 2005). Cohen's kappa is however a conservative estimate of agreement (Fleiss, Cohen & Everitt, 1969), which is influenced by the level of chance agreements present which were high in this case, thus

impacting and reducing the Kappa statistic. Overall coders agreed on 87% of themes and sub-themes within the data which was coded, indicating that overall agreement was high.

Reflexivity. In order to maintain transparency and aid the reader in understanding potential biases and perspectives in relation to the research area, it was deemed important to include a section on reflexivity. This assists in providing recognition from the researcher that they are involved in the study and as a result, potentially impact upon its outcomes (Willig, 2001). In this study the researcher a Trainee Forensic Psychologist had undergone similar doctoral training as some of the participants. As such the researcher had similar experiences to working with offenders and conducting IOW within secure forensic settings, therefore the researcher had prior knowledge and understanding of processes and concepts within this area, and this may have impacted on how data was interpreted and coded during the analysis stage. Furthermore at the time of the analysis the researcher was working within a prison environment as opposed to a secure hospital and the potential differences in the ways IOW is conducted in different settings could also have impacted upon data interpretation.

Results

Questionnaire Data

Descriptive analysis of the questionnaire data was performed using SPSS statistics package (version 22) and produced descriptive statistics, a summary is provided in Table 10. Findings showed that on average staff identified that there are policies in place which require them to carry out IOW, and that such work is carried out both on a one to one basis, and in group format with service users. Staff identified that IOW/IOA takes up a moderate proportion of the work carried out with service users and on average this is a total of 5.71 hours per week. Total hours spent carrying out such work ranged between two and ten hours. The majority of

participants indicated that there are policies and processes in place which staff adhere to when conducting such work and also that they feel well equipped to conduct such work, although a proportion of participants indicated feeling unsure about both of these areas.

Clinical psychologists' reported on average, a lesser number of hours spent conducting IOW compared to forensic psychologists, although several staff members indicated that this time varies. In addition a greater proportion of clinical psychologists' reported being unsure or not well equipped to complete IOW in comparison to forensic psychologists. This may indicate a difference in training between the two disciplines and suggests that more information is needed in regard to how to conduct IOW.

Table 10.

Summary of Descriptive Statistics for Questionnaire Data

Question	Yes (%)	No (%)	Unsure (%)
1. *Working as a psychologist, within your service are there policies in place that require you to carry out specific index offence work?*	43	29	29
2. *Do you carry out some form of index offence work/analysis when working with service users either on a one to one basis or as part of a group?*	100	0	0
3. *Within your service is there a pre-prescribed process that everyone adhered to for completing such work?*	67	24	0
4. *Does index offence work/analysis take up a large proportion of the work that you carry out with service users?*	67	24	0
5. *Do you feel that index offence work/analysis is important as part of service users' treatment pathway?*	100	0	0
6. *Do you feel well equipped to complete index offence work/analysis with service users?*	71	5	24

Focus Group Data

Analysis of the data identified a total of three overarching themes relating to IOW/IOA. The author asked participants a series of questions in order to ascertain their knowledge, understanding and use of IOW (see semi-structured interview schedule, Appendix I.). Overarching themes and sub-themes that were derived from the data, map onto the questions which were asked. All themes relating to each question and the overall research question are summarised below and are presented in a hierarchical manner. This section examines all data

extracted regarding IOW/IOA. Due to the necessity to be succinct, only a few illustrative items and codes are presented here in order to provide context and support for the extracted themes.

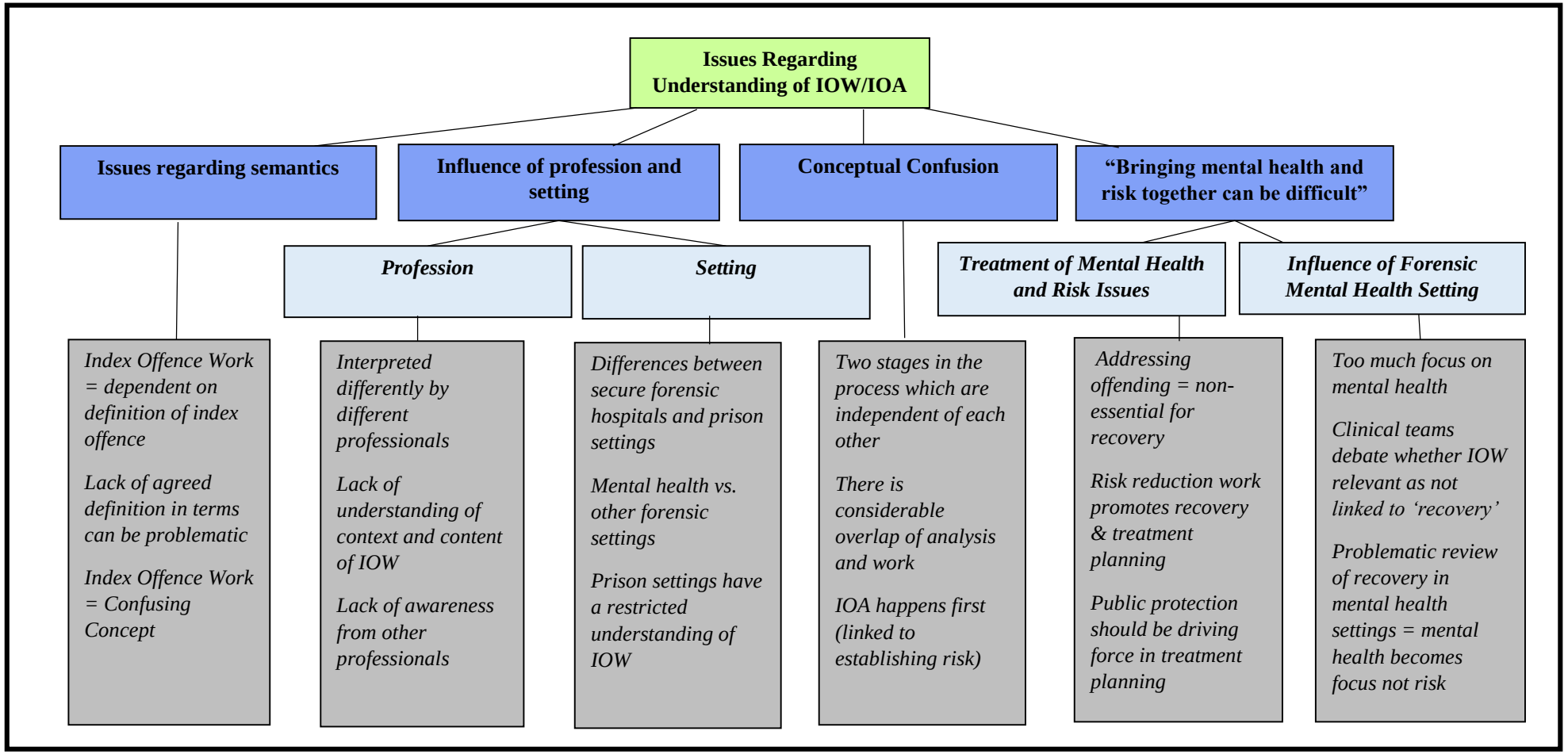


Figure 2. *Thematic Hierarchy of Issues Regarding Understanding of IOW/IOA*

Issues regarding understanding of IOW/IOA. A total of four broad subthemes were extracted from the data relating to issues regarding understanding of IOW/IOA. Figure 2. Presents a thematic map of the hierarchy of factors organised into sub-themes. Where relevant subthemes were further delineated into sub-subthemes, in order to ensure distinct ancillary items were incorporated. The opinions of participants in this area related to the difficulties in understanding relating to IOW/IOA. Salient subthemes related to issues around conducting specific IOW, such as semantic differences, confusion between the concepts of IOW/IOA and the influence of the context in which the work is being conducted.

Issues regarding semantics. In the context of what IOW involved for practitioners, it became evident through participants' discussions that there were a number of issues relating to the definition of IOW and an 'index offence' (*n=18).* Statements made which endorsed this sub-theme related to participants' feelings that labelling it IOW makes it confusing, because the work which is done is broader than this, and doesn't just focus on an individual's index offence. Participants spoke about 'hating the term' and the idea that it has different meanings to different people. This links to ideas endorsed by other participants (*n=2*), that IOW can incorporate a variety of models and therefore cannot be simplistically defined as one thing. Two participants also endorsed the idea that using the term 'IOW' can be problematic in relation to what the client internalises and understands in relation to IOW:

> *"Calling it the index offence is the bit that that makes it confusing and a bit like it's very specific, when actually (Participant 21 – "It's suggests to me the people at the DOH have no idea what it is they're describing"), yeah and the index offence work will never just be the index offence you would never just focus on that."- Participant 20*

"Yeah it's inappropriate, problematic behaviour… I think the focus is just when you talk about index offence, it's just sometimes assumed that you're gonna focus on you know one offence that somebody did and actually you're not often looking at that, you're looking at things way back in the past that led to you know people having the feelings they did about the world and the strategies they then learnt for dealing with their difficulties and that you address those …"– Participant 16

Influence of profession and setting. Across participants, a sub-theme of differences in understanding relating to different professionals, was prevalent from the discussions had (*n=12*). However, this was particularly relevant when participants were asked to discuss what IOW involves for them. Participants' spoke about how the background of the professional can impact upon the understanding of IOW and also how the setting in which it is being conducted can influence this. This resulted in two sub-subthemes being derived *profession* and *setting*.

Profession. A number of participants (*n=5*) across the focus groups presented the view that commissioners directing that practitioners should conduct IOW, have a lack of understanding in relation to IOW and what it really involves. In addition participants also spoke about how this can be prevalent amongst colleagues within their own wider teams, in understanding what the full breadth of IOW is (*n=5*). During discussions, opinions were voiced regarding potential differences between the training received by forensic and clinical psychologists. This theme was only prevalent in focus group 2 and not in any of the other focus groups, although the evident lack of understanding across focus groups may indicate training needs:

"…but I think commissioners don't understand that and unless you're actually talking about the offence itself…that actually talking about the offence itself is not gonna do anything it's all about the precursors to it…" – Participant 16

"I actually think that people are here because of their risk, because if it wasn't about their risk they'd be in adult services predominantly, so and I think maybe that's where the department of health and our perhaps psychiatric colleagues particularly do get very focused on the index offence or the incident of risk behaviour that's typically led to their admission, and I think that's their understanding of index offence, whereas we tend to go okay index offence and then all of the rest of this too, erm whereas I think sometimes there's a sense within the clinical team perhaps and nursing staff just deal with the index offence and that means all of their risk goes away and so there's a complete misunderstanding of risk reduction work." – Participant 21

Setting. A number of differences were observed in relation to participant's views on the influence of setting, this largely surrounded the difference between forensic mental health settings and 'other' forensic settings, namely prisons. Comments from participants may also indicate a consensus that secure forensic hospitals are better able to cater to individual needs, rather than generic programmes for all types of offending. Although one participant who emphasised that he is from a prison service background, highlighted that group offending interventions has been shown to be 'what works' within the literature. Some participants (*n=2*) also discussed how the difference in relation to more formal settings such as tribunals compared with CPA's can impact upon IOW and the understanding of the patient:

"....and that works well I think versus a more formal setting of like a tribunal, where you know the index offence can be very central to panels or decisions you're making about where patients are at with their understanding of that and their risk factors, the CPA process is a more friendly patient friendly forum for that I think." – Participant 15

"Ye I hadn't thought about that and I think there's more of a focus on that in prisons isn't there? In the offending behaviour programmes." – Participant 8

Conceptual confusion. In relation to the differences between IOA and IOW it was apparent that amongst participants there was considerable confusion (*n=15*). In relation to the processes used within IOW and IOA, a number of participants indicated that IOW and IOA are distinct processes, indicating that one stage follows on from another and this usually begins with IOA, with IOW being conducted dependent on factors such as treatment readiness. Other opinions (*n=5*), related to there being an overlap between the analysis and work, and this may relate back to opinions voiced within the focus group of difficulties in the definitions which are used. The idea of IOA and IOW were viewed by some participants (*n=2*) to be a newer and developing concept in forensic mental health services, which again may be the reason for training needs which were identified within focus groups:

"Ye so it's more about getting someone to identify themselves, what do you think you're current, what are your risks, what are we talking about, what's likely to potentially happen again? Ye I don't know about anyone else but I don't, you might have heard more about this participant 6 erm what's it called erm index analysis

work ,index offence analysis that's not something I have trained in or do…" – Participant 7

"I think there's different things isn't there, I think (Participant 7 – "I don't really know what that is") I think the index offence analysis work, sorry I'm answering (Participant 6 – "No go on",) it's just all that came to mind when you said that is functional analysis I think, do you think that's what it is?" – Participant 8

"Ye, just formulation like formulating the index offence and what factors impacted on it." – Participant 6

"So perhaps, this is the interesting thing isn't it about being clinical and not a forensic psychologist, cause I don't know that and I always think there's this mystery thing that I don't know about but actually maybe do we do it?– Participant 7

"There's two parts aren't there, the analysis and then the work and I suppose from my point of view the analysis which you would do with everybody is to establish risk and where the risks are and what the risk factors are and which ones are key and what might need working on…those are usually multiple, but I think with some of our service users you don't then go on to do the second part, because they may be not at a point where they're ready, they can't engage, or they've done some work in the past and don't want to do any more, so the very basic thing you would get from it is information that helps with risk… - Participant 3

"I mean obviously the analysis is key before and that's the formulation really before the work starts." – Participant 15

"Bringing mental health and risk together can be difficult". Participants focused on the idea of mental health issues and risk issues within secure forensic mental health services. This subtheme pertained to the concept that the bringing together of mental health and risk can be problematic (*n=10*). This was highlighted by participant three who explained the idea of two separate models attempting to come together ineffectively. This theme further separated out into two sub-subthemes relating to *Treatment of Mental Health and Risk issues* and *Influence of Forensic Mental Health Setting*.

Treatment of mental health and risk issues. In relation to the opinion provided by participant 3, some other participants spoke about the idea of how IOW and recovery from mental disorder promote risk reduction (*n=3*). Some participants indicated that a focus on either risk or mental health is not required for the other, however some participants highlighted that risk should be the focus:

> *"… it's interesting this idea of offence work and contributing to people's recovery, because in some ways it's like you've got two kind of models they're trying to meld together really in some kind of language and I think offence work, does it help somebody's recovery from mental illness? It might do but that's not the main aim of it, so erm I think it's role is I mean if you take for example we've got this thing 'my shared pathway' which is sort of linked to trying to link somehow recovery and erm kind of risk issues in forensic service users, I'm not sure it actually does that…"-Participant 3*

> *"I'm thinking that maybe index offence work might help recovery in the sense that it's helping the service user think about that aspect of themselves which often they cut off from, they don't often like to think of themselves as a perpetrator or what*

they've done and so it might help people to look at the good and bad aspects of themselves and behaviour I guess that might help with recovery in some way." – Participant 4

Influence of forensic mental health setting. The idea of the influence of the setting impacting on the view of risk and recovery was prevalent across participants. In particular, issues regarding specific views held within the setting were evident, as well as the idea of risk and mental health issues being problematic within the wider team. Participants spoke about the confusion between risk and mental health issues, with one participant highlighting that within secure forensic mental health, mental health "trumps" risk most of the time. Two participants in contrast to this, highlighted that this goes against what the literature indicates in relation to mental health and risk, with another two participants emphasising that the view that is generally held and the focus on mental health factors results in important social factors being missed. Following on from this, a few participants highlighted how such views can be problematic within the wider team, thus indicating the challenges of IOW within forensic mental health settings:

"I've heard something that's even worse than that…deal with the mental health, risk goes away and that was the original battle when we were trying to get a risk programme even established, erm there was significant debates about whether or not that was even required, because it's not relevant to somebody's recovery…erm which is was interesting and at times quite laughable really…"- Participant 21

"Well it goes against all the literature that says basically people with mental health problems are no more prone to committing offences than anybody else so how does that work? If that is the sole cause…" – Participant 17

"Partly I think the thing I struggle with sometimes sitting on a clinical team is when erm…if the team just purely focused on someone's symptoms for example and their mental illness and someone might be moving through in my opinion the system far too quickly and haven't actually done any index offence work or thought about it, they might have just in quite a superficial way done some early warning signs work with somebody and then just because they're stable, all of a sudden someone's on a rehab ward, they're looking for discharge…." – Participant 7

"I think what happens is it's easier to conceptualise it as mental health has had an overbearing factor because then in the back of their minds I imagine the rest of the clinical team will think other people with those social environmental factors might not have had the same kind of offence pathway, therefore the only obvious difference is mental health, however I totally agree with your point that people with mental health problems…aren't more likely to commit offences, however…there is a reality that people do act on their command hallucinations and whether it's just based on that or whether it's because they've had experiences of offences beforehand, I think that's the tension that we end up holding this idea that the dominant narrative is or was at some point mental health causes offences…I think trying to hold both in mind is really difficult and I find myself swaying from and then correcting myself and saying lets really think about the offence and the mental health both as they interplay and individually."- Participant 19

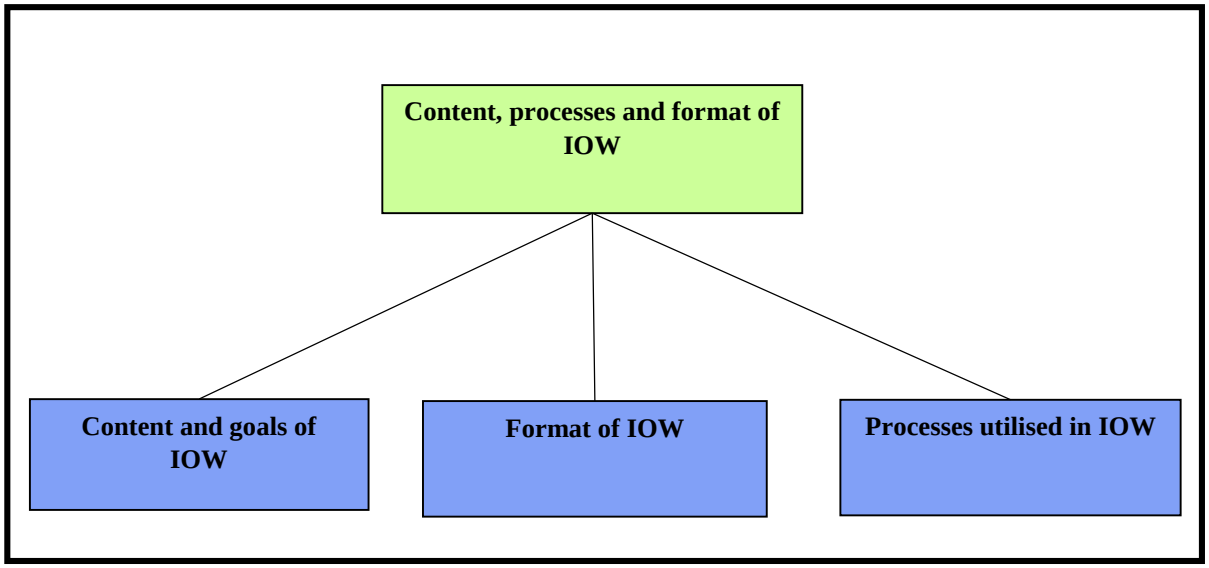


Figure 3. *Thematic Hierarchy of Content, Processes and Format of IOW.*

Content, processes and format of IOW. Three broad subthemes were extracted from the data relating to the overarching theme of Content, Processes and Format of IOW. Figure 3. presents the thematic hierarchy of subthemes. Due to the breadth of this overarching theme and its subthemes, subthemes were further separated (see Figure, 4 and 5). This theme dominated the majority of discussions held within focus groups. Overall in relation to this theme participants debated the content and goals of IOW, what factors impact upon the format on which it takes and key processes which are utilised when conducting IOW.

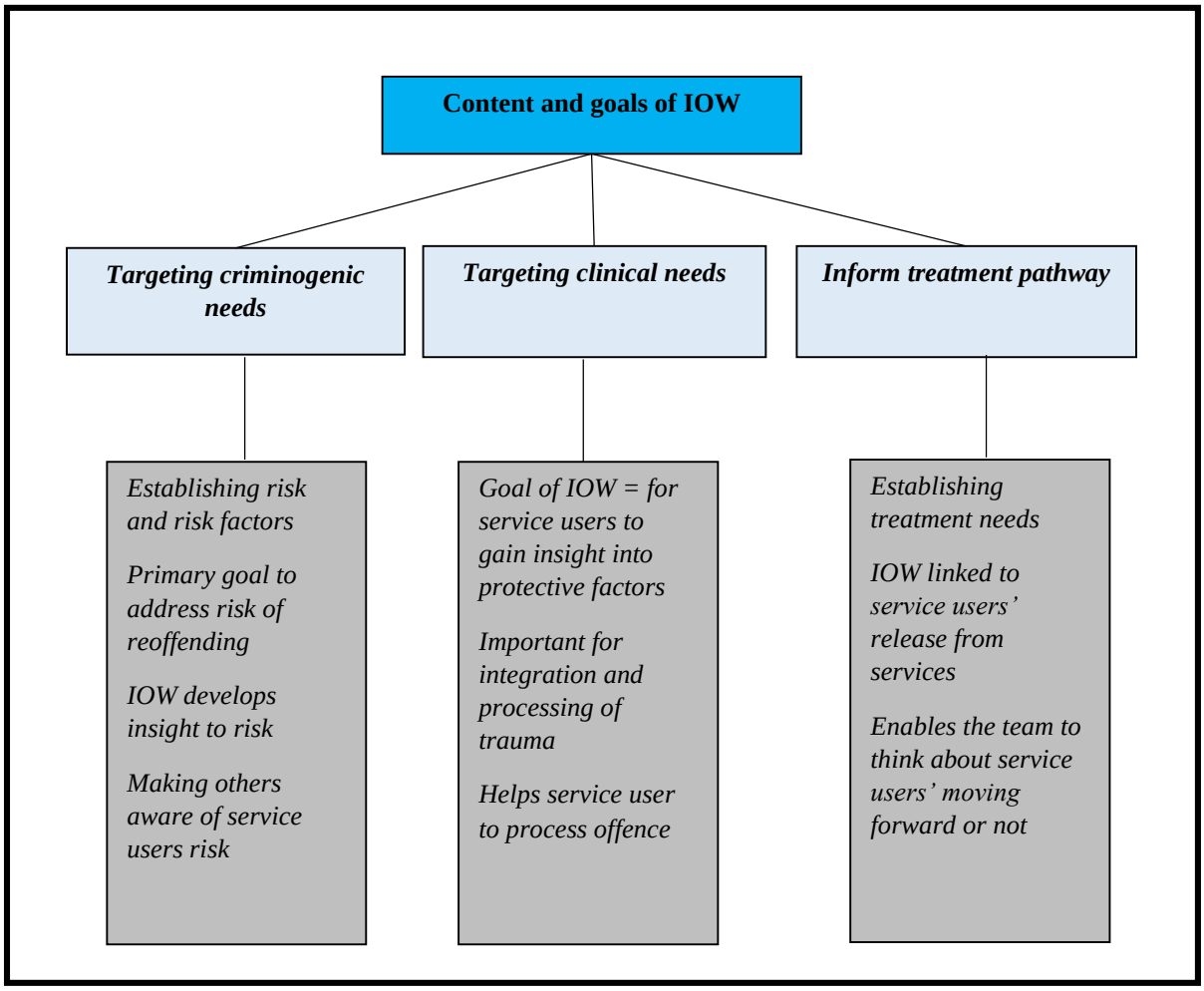


Figure 4. *Thematic Hierarchy of content and goals of IOW subtheme.*

Content and goals of IOW. A pervasive sub-theme relating to the goals of IOW which was expunged from the data was content and goals of IOW. Practitioners (*n=17*) spoke about developing their own, the service users, and the wider teams insight, into factors relating to individual's clinical and risk needs. This was further separated into three sub-subthemes *Targeting Criminogenic Needs*, *Targeting Clinical Needs* and *Inform Treatment Pathway*.

Targeting criminogenic needs. A predominant sub-subtheme within participants' dialogue, was targeting service users' criminogenic needs through IOW (*n=17*). Participants inferred that addressing risk of reoffending is the primary goal of such work

and it allows practitioners to clearly identify risks and risk factors for each individual. Therefore establishing the causal factors and functions of offending behaviour and addressing these appropriately. Two participants' also recognised that IOW has another goal within secure settings itself, in helping to identify Offending Behaviour Programme's (OPB's). No other participants supported this view point and this may be due to the fact that these participants' work on a personality disorder ward where such behaviours may be more prevalent. Participants (*n=6*) did however identify another goal relating to criminogenic factors, developing the wider teams insight into an individual's risk and making the team aware of relevant behaviours. Indicating IOW plays a role in the development of risk management strategies:

> *"I think it's erm, maybe the primary aim is to reduce risks of reoffending."* - *Participant 4*

> *"And I always think that this is kind of the or what I view as the idealised outcome of index offence work, which I was thinking about offence paralleling behaviour and if someone's index offence is relational in nature, if they could get to a point where they can start to notice where they might be enacting something, so I was thinking about some of the men on _____ and someone with sexual offences, and he engages in lots of offence paralleling behaviour, if you can get to a point where he can actually notice that he might be using someone for example in a sadistic way and gain some pleasure from talking to them about their index offence for example, that would be like the ideal outcome is someone gains that level of awareness…I don't know if we really get that far…" – Participant 7*

Targeting clinical needs. A number of participants identified other indirect goals to IOW (*n=14*). These surrounded addressing any problematic behaviour that an individual presents with, having a global understanding of the individual, as well as enhancing the service users understanding of their risk and protective factors, and their insight into managing these. Five focus group members also objectified that a key goal of IOW is to assist offenders in processing the offence and thus indirectly develop insight. In addition one participant identified how specifically working on an individual's mental health issues assisted with indirectly reducing risk:

> *"And I've worked with people where actually the mental health awareness has been the key thing in reducing their risk, cause that is the key factor in reducing their risk…and so kind of a big piece of work on mental health awareness has been the thing to help them obviously for reducing risk and public protection etc…" – Participant 10*

> *"I think it's all of those things really I think it's erm, maybe the primary aim is to reduce risks of reoffending but erm you've also got the goal alongside of erm, I guess helping the service user process erm the offence and that might indirectly reduce risks of reoffending erm, but also to erm, for their own well-being as well I think to process what's happened and to adjust to what's happened and what life might be like after the offence, that's my understanding of it." – Participant 4*

Inform treatment pathway. Thirteen participants endorsed the sub-subtheme of a goal of IOW being associated with informing service users' treatment pathways. Participants (*n=9*) discussed how assessments which form a significant part of the IOW process provide important

information regarding what interventions should target and thus inform individual's treatment pathways:

> *"Well, the treatments informed by the formulation cause it tells you what areas you need to work on, so I'd say that formulation underlies everything really." – Participant 16*

> *"And often, there is I suppose another phase isn't there? Which is assessment, have we got a clear picture of what we need to work on, there's ideal kind of therapy work which addresses risks…" - Participant 2*

A smaller number of participants (*n=7*) also spoke about one of the goals of IOW is to direct service users' release from forensic services. Both in terms of the practitioner and the individual service user thinking about working towards their release, as well as the wider team and external influences, such as tribunals and the decisions that they make. Highlighting the wider impact of IOW:

> *"So often the work might be indirect mightn't it, so we might do a risk assessment, care plan someone's needs for the longer term, but not actually meet with the person at all and you might share that with the team and think about what the risks are but not actually do that work if they're not ready." – Participant 7*

> *"…if they haven't done a good offence formulation in assessment, you can't make that explicit, so it might be more difficult for people if they're not able to*

engage in that part of the work, to argue at a tribunal or in a managers hearing or whatever, that anything's changed." - *Participant 2*

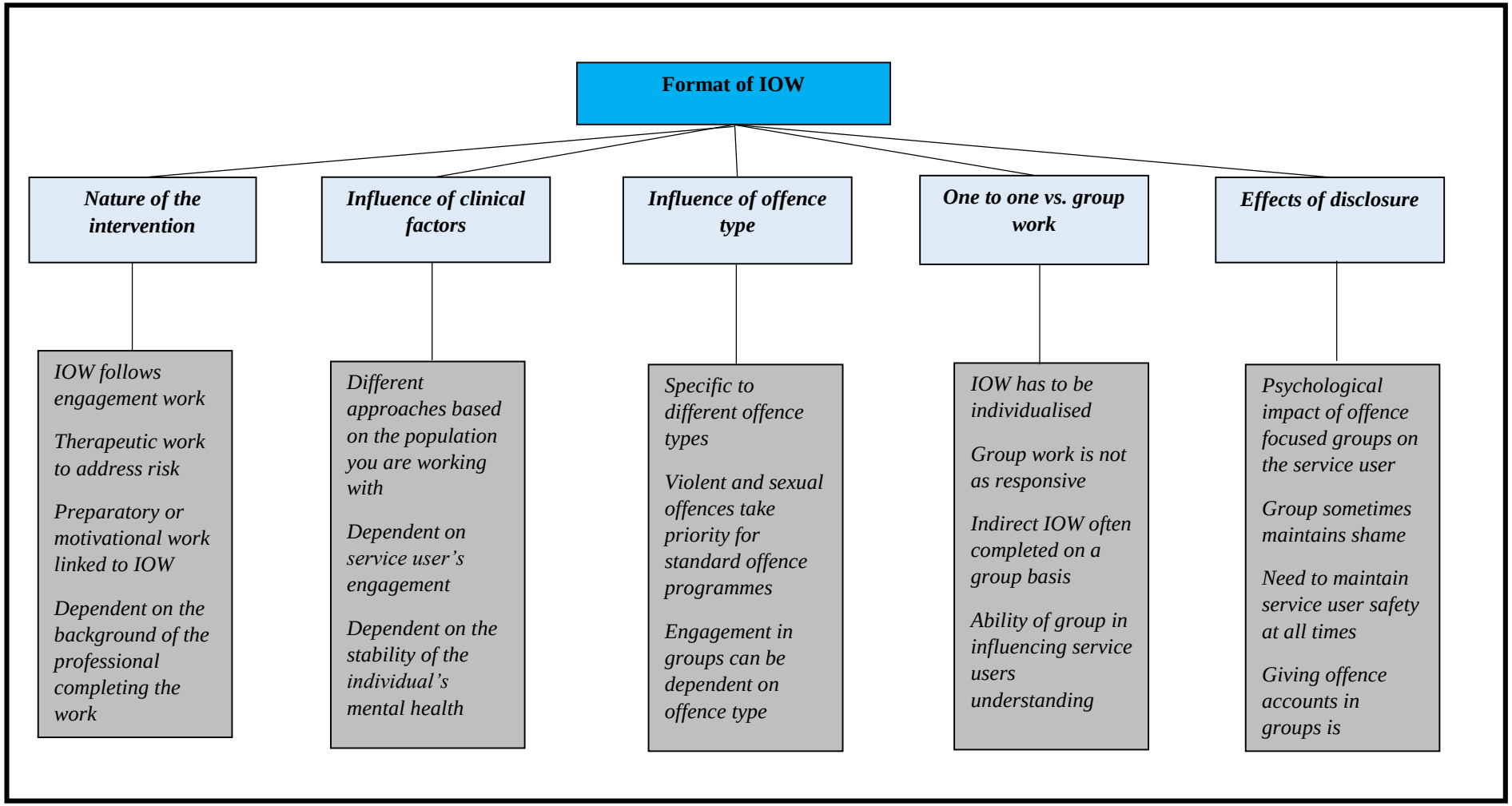


Figure 4. *Thematic Hierarchy of the Format of IOW sub-theme.*

Format of IOW. This sub-theme related to factors which impact or influence the format IOW takes within forensic mental health settings. The topic promoted a lot of discussion which resulted in this theme being endorsed by all focus group participants *(n=21)*. A variety of factors relating to IOW's format were discussed and therefore there was a need to further divide this sub-theme into a number of sub sub-themes. This was done in order to demonstrate the discrete factors which may impact upon this aspect of conducting IOW.

Nature of the intervention. Some participants focused on discussions surrounding the nature of the intervention itself (*n=14*). This dialogue corresponded to the intervention regularly involving preparatory or psychoeducational work. More specifically, motivational work in order to get service users engaged or to provide them with the necessary coping strategies to tolerate risk reduction work. One participant highlighted that the intervention often focused around further exploratory work following a risk assessment, in order to really understand where the risk for that individual lies. Another highlighted that, as part of psychoeducational work, external resources to the service will often be utilised. Generally the consensus was that such work involved educating service users how to stay well if they are mentally ill and as a result this impacted upon their ability to manage their risks and engage in IOW effectively:

> *"…there might be some kind of motivational work in between which is a negotiation of are you gonna do the work (nods and hums of agreement), which could be pretty protracted for some people, is there something we can do that makes it more likely that this person is going to engage in work that they need to do in order to reduce their risks, and therefore get out of hospital and be safe in the community… might be about getting that person engaged in if they can work collaboratively with other people…" – Participant 2*

"I'm working with somebody at the moment and I'm doing exploratory work… because I need to understand and formulate where that individuals risks lie, because I'm not convinced that it's specifically lying purely in his mental illness for example…So I guess my view is almost exploring it with the individual, to say well you tell me where your risks sit and what's going to make them worse or better…" – Participant 17

As part of this sub-subtheme, participants discussed how their own background influences the work that they do (*n=3*). In addition the expectations of other professionals and settings, such as psychiatrists, tribunals or commissioners as to the works format (*n=4*) and how this can impact upon the nature of the intervention were also described. Interestingly participants who spoke about their own background influencing the work were generally service leads/senior psychologists (*n=2*) and all had a background working in the prison service (*n=3*). This may indicate that a wider range of experience enables practitioners to be better able to influence/ understand how IOW should be approached:

"I was doing an assessment in consultation and they were talking about this particular person in the group and they were just saying he's just not taking anything in and erm and there saying he had a horrendous past… and that he's obviously got no feelings for anybody else, least of all himself and I said well has any work been done on his own victimisation and they said well no, I said I think that's the starting point really…so how can you expect him to worry about anyone else, if he doesn't worry about himself, or think he's worthwhile, what's you know, what's gonna motivate him to change his life if you know he just has a real downer on himself?" – Participant 16

"And certainly for people here perhaps who might be going back into the prison service as opposed to being discharged out into the community, the prison service and parole boards are very clear that if you've got somebody who came into prison for a violent offence but they've got a history of sexual offending, before they can be paroled they will have to address their SOTP work, even if that was not their index offence and I think my sense is that actually is that it's not about the specific offence and when it was committed in terms of a pattern of offending, it's about what will help this person to lead a more law abiding and less risky, safer and more fulfilling life…I don't just leave it at the I've done the index offence my job is done." – Participant 21

Influence of clinical factors. A highly prevalent sub-subtheme endorsed by participants (*n=17*) was the role of clinical factors relating to the service user. In particular this related to the treatment readiness or engagement of the client. Resistance and trauma were also spoken about in affecting how the work is conducted, or on occasion preventing practitioners from doing so. Two participants who work on ASD wards, also spoke about how working with client groups who present with significant cognitive or social functioning difficulties, can force the practitioner to change the way in which they conduct it. Another emphasised this through stating that service users are often unwell and unable to encode information; all of these factors highlight the complex nature of the population within forensic mental health settings and how this inevitably affects IOW:

"Not being able to, they might be very unwell at the time and therefore unable to encode information that well and retain it, there's a high proportion of head injury even without as p1 was talking about our specialist unit, just in general

forensic populations prison or hospital, a high risk of head injury and erm there's a lot of people whose best account might be quite a poor one erm…" – Participant 2

"And I think particularly with our ward as well, being ASD we've got to be more careful because of their literal interpretation of things at times and their already deficit in their social awareness and skills so doing group work particularly around kind of well sexual offending could prove detrimental rather than…helpful" – Participant 12

Eleven participants also spoke about the fact that often groups cannot be run as a result of individual clinical factors relating to not having enough service users at the same level of treatment readiness to run groups and the inability of service users to engage:

"And the problem is that's sometimes in a group that people who are very quiet cannot contribute very much and then they slip through the net…I remember seeing one man who had borderline learning disabilities and he's been in a group and they'd been talking about the four steps to offending and erm I was asking him about this and what the four steps were and he looked at me and said erm but I was in a field and there weren't any steps…" – Participant 16

Influence of offence type. A smaller but relevant sub-subtheme which participants spoke about was the influence of service users' offence types on IOW (*n=11*). Consistently, participants spoke about how often work has to be individualised as a result of this. In addition

participants held views which emphasised that violent and sexual offences took priority for IOW and therefore often people who don't have these types of offences miss out on carrying out any specific work:

"Well there was a man that I worked with, he erm was involved in lots and lots of sort of burglaries would then go to prison and then relapse and he's become really unwell and become very violent in prison, but actually when you did the HCR-20 and you looked at his history his actual index offence wasn't violence related…yeah so I think it's difficult because I think that was overlooked in that sense, he didn't do any particular work, we did a lot of work around sort of gambling and what increased his risk of committing an offence but I don't think we didn't do specific work around burglaries." – Participant 9

"I think that's the difficult thing about the population, even though were quite a big unit it's surprising when you actually try to get together a group of people with a similar offence, at a similar stage, it's quite hard to do…I think VOTP will be different because I think that kind of extends beyond just well lots of people have got violence in their history haven't they?"– Participant 3

Participants also discussed the fact that for some offence types such as sex offending, group work can put them at risk from others, making service users not want to engage:

"It's balancing those things isn't it…but we have, we have tended to do the sex offender work individually, because you know we've been doing sort of a needs analysis over a period of time and we just don't have that group of people at the same level that would benefit from a group intervention…" – Participant 3

One to one vs. group work. The topic of whether to conduct IOW on a one to one or group basis promoted considerable discussion across all four focus groups. Whilst participants spoke considerably more about group work, one to one work was also reviewed by practitioners (*n=16*). This was mostly regarding the need for IOW to be individualised, as views were held that often group work risks not meeting service users' needs effectively. Whilst participants acknowledged the benefits of group programmes, they evidenced that these were often difficult in practice, and as such much of their offence focused work or work relating to IOW, was carried out on a one to one basis:

> *"So I think the actual work focusing on the index offence in this service, at the moment as things are, would mainly be one to one…" – Participant 6*

> *"Yeah and it's the factors that lead you to offend in the first place are the factors that need addressing… I mean just as you compared two different fire setters, you've got somebody who has no interest in fire whatsoever but are desperate for help because something else has gone wrong in their life and fire was I guess maybe the impulsive act, never been interested in it whatsoever so putting them on a six week programme for fire setting for that particular individual…it's not going to be very helpful, or an effective use of time to be honest and that's where I think CTM's struggle…it's trying to convince people that manualised programmes have their place, but I think they're very difficult to work into practice in these type of environments…– Participant 17*

IOW in a group format was also spoken about by the majority of participants (*n=18*), indicating it was an area that is central to IOW and it's execution in secure settings. Focus

group members spoke about both advantages and disadvantages of groups. Some participants expressed views which identified advantages of group IOW (*n=9*). Essentially, dialogue focused on the ability of the group process to enhance service users' understanding and how often challenges are more powerful when coming from peers. In contrast one participant identified how sometimes 'group think' (Janis, 1982) can be problematic for practitioners to manage. Further to this a few participants identified that groups can be useful for conducting indirect IOW, such as substance related groups, and also that groups can provide important information for the assessment of offenders. For example a better understanding of individual risk factors, indicating that groups do have a role in IOW:

> *"I suppose what it can do group work like that, you do see people suddenly jumping on an idea so readily, it makes you realise that just how close to the surface those desires are and with just minor amount of prompting from somebody, you know perhaps it's about them wanting to hear that, and somebody only has to say it and then boom they're there, erm but at least that you're aware of it so that you can deal with it." – Participant 14*

> *"...we've had a group that did really well, quite a mixed ability group actually who did erm... work together really well and challenged each other and pushed each other and seemed to get a lot out of that from the different challenges that they had, but we even saw people who weren't very confident saying 'oh that doesn't make sense', 'you're saying that's gonna help you avoid this but it's not you're just gonna end up here' and they were quite accepting of that..." – Participant 11*

Ten participants also discussed responsivity issues relating to groups and one to one work and the importance of considering these. They stipulated that such considerations result in service users obtaining more benefit from groups, responsivity across IOW was evidently viewed as a key issue:

> *"There's also lots of anxiety isn't there about the service users sharing their index offence within a group setting, erm which is something I guess it's come from I don't know perhaps it's more from the psychiatry erm within the service and that that would make someone more vulnerable potentially…." – Participant 9*

> *"But for me also it's about understanding what a manualised approach is, that actually a manualised approach is not a you will simply say this by wrote, that actually the manuals are written in a way that supports the strategies and processes and techniques to get the best out of the group and they're flexible enough to consider the individual formulation, so you may do the same task with each person but actually the way in which it's delivered in terms of the skill of the practitioner, is that you would hold that formulation and work that process through with that individual, thereby individualising the technique or the skill or the focus of that discussion to that individual, which would be different to the person next to you…"- Participant 21*

Effects of disclosure. Offence disclosure was a smaller part of the overall sub-theme relating to the format of IOW and was spoken about by a few participants as being problematic when thinking about running IOW in group formats (*n=7*). Inevitably it is difficult to ask service users not to disclose, as this may be problematic in itself, as highlighted by participant

one. However, through disclosing offences this brings into question service users' safety and well-being and it was suggested may have a potential negative psychological impact. Participants emphasised that this was particularly relevant in secure hospitals due to the small environments, but this again poses questions about differences between forensic mental health and other forensic services:

> *"I imagine if you were to have a group that was specifically about index offences, erm then you would end up with interesting questions about what it would mean to disclose or not to disclose…what is spoken about what is not spoken about who knows what, and I could imagine that being a tricky thing to think about, in that for people's safety and well-being, in many ways it's the obvious thing not to ask not to disclose their index offence but then I also wonder what the psychological meaning of let's have a group about this that we then want you to disclose, talk about something but who do you and who knows you and who are you so I think there are there are lots of err…potential challenges to doing it thinking specifically about index offence and although I don't you'll know this p3, in prison settings with sex offender treatment programme, and isn't one of part of the group process about each individual relating or retelling their offence?" – Participant 1*

> *"The thing I wanted to say about the sex offender groups and index offence work is that it used to be the case that people would be expected to go into the actual offence in gory detail, which is completely and utterly inappropriate and also very boring for a lot of people and titillating for others, which is you know both cases are quite detrimentally harmful, I'm pleased to say that people don't do that*

anymore erm I've never done it I've always thought it was completely inappropriate…" – Participant 16

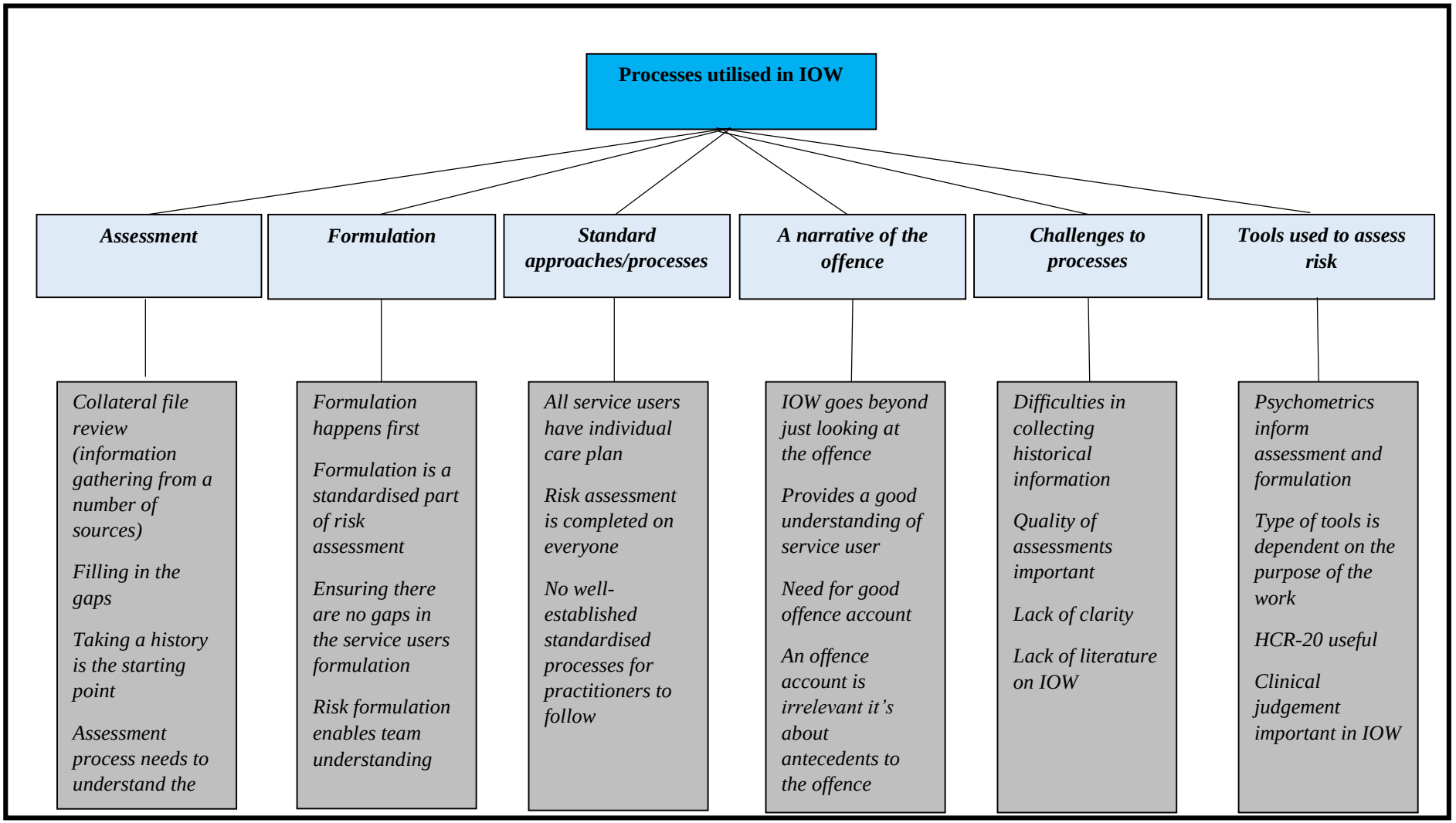


Figure 5. *Thematic Hierarchy of Processes Utilised in IOW sub-theme.*

Processes utilised in IOW. Processes utilised by practitioners was a central aspect across the discussions had in focus groups. This sub-theme embodies the processes involved and used by practitioners when conducting such work as part of their job role and the challenges faced which impact upon this process. Due to the expansive nature of the discussions relating to this area, the sub-theme was further separated into sub-subthemes in order to demonstrate the variety of factors relating to the process utilised within IOW.

Assessment. A key component of IOW that was singled-out by participants is the assessment process (*n=16*). Many participants (*n=9*) highlighted the assessment process in enabling them to gain an understanding of service users risks and needs. In particular, they spoke about taking a history as being the starting point to the assessment in order to establish where a service user is at. It was evident that key components in the context of a broader psychological assessment involves taking a history and completing a file review. A view was also held that such assessments should be completed on an individual basis:

> *"I always tend to start with talking about the antecedents to the offence cause it can be helpful for the person to tell a story about what was going on for them before erm…so…I think I nearly always start with that, often doing a timeline erm but ye then … very much make a choice about what they can tolerate."- Participant 5*

> *"So……from my perspective I will do index offence work with everybody, but it has the potential to be different depending on who I'm working with so…the first part for me is always the formulation and understanding the risk issues erm…that are much more global than just the index offence erm… and trying to think about how we cover all of the risk behaviours, because we'll get people who have index*

offences that are perhaps regarded as less serious than perhaps some of their previous offending, so I suppose coming from a prison background my erm……my experience teaches me that just dealing with an index offence is not always enough…"- Participant 21

Formulation. A number of participants spoke about formulation happening first in the assessment process and that this underlies everything that follows. Such processes were explained as allowing practitioners to identify the risk and need of the service user and make recommendations for to target risk factors/ needs. Additionally, through observing participants' discussions it was established that formulation is now a formal part of the risk assessment process, and allows practitioners to formulate and understand an individual's offending behaviour. It seemed that this was a key factor in assisting practitioners' execution of IOW. One participant highlighted how the assessment process allows for everyone involved in an individual's care pathway to have an understanding of what can realistically be achieved with that service user. This highlights the importance of the process of formultion in IOW and specifically within the wider context of a secure forensic hospital:

"Well, the treatments informed by the formulation, cause it tells you what areas you need to work on so I'd say that formulation underlies everything really…" – Participant 16

"You start with assessment really to try and identify what someone's err clinical needs are in terms of their yeah what their risks are, so what they are linked to particularly…obviously thinking about risk of violence to others and whether it's

more about someone's sexual offending or whether it's something about you know you kind of identify what the risks are that you need to address that's where you would start…" - Participant 7

Participants' also voiced opinions which tended to focus on the purpose of IOA (*n=10*). The majority of participants who spoke about IOA, endorsed the idea that it is the formulation aspect of work with a service user (*n=6*). Only one participant evidenced that IOA was a specific formulation of an individual's offence, highlighting a lack of understanding amongst practitioners generally:

"It's just semantic in a way, because analysis is about you know doing a formulation and what's it all about and the work is actually addressing the deficits and helping people find other coping strategies, or ways of meeting their needs" – Participant 16

"Yeah I think it's all the same, I think there's a difference between index offence work and index offence analysis, cause I think that index offence work is the formulation, but I think the analysis is when you really go into that specific incident and your yes it might fit into the wider formulation, in terms of coping skills, and alcohol, and or substance misuse, but I think your kind of looking into it in a way of you know all that kind of like seemingly irrelevant decisions and how did you get there and why, I think that's more analysis whereas having a formulation that looks at their patterns, is different" – Participant 20

Nine participants also postulated that formulation allows for the whole team to have an understanding, in assisting with managing and containing service users, again highlighting its importance in the context of secure forensic settings:

> *"And I think of in terms of how it impacts into our work, I think that's how it impacts but at every stage, so I'll talk about it right from the care plan that we produce as the team we'll be introducing concepts to the nursing team, we'll be thinking about the service users themselves…I think people almost rely on a timeline, but the timeline has and whatever you call a timeline, a trajectory whatever it is I think it gives a space for both the team, us as professionals, us as psychologists to actually understand where it stemmed…" – Participant 19*

Standard approaches/processes. A number of participants spoke about standard processes (*n=15*) that are in place within their services which are linked to IOW. Participants identified care plans, risk assessment, and formulation as standard processes that they use when conducting IOW. This theme was not prevalent in focus group one, where participants instead focused on the lack of standard processes, which may indicate a specific view. All participants (*n=21*) identified the need for IOW to be individual in process and two participants identified that there are no guidelines regarding the process or format that IOW should take. Overall, standardised processes were viewed as not useful in this setting, due to the influence of service user and organisational factors:

> *"I suppose when it really comes down to it, it is the processes that are standardised in people's treatment like CPA's and ward rounds and that sort of thing, assessment tools and everything else is individual." – Participant 14*

"But the initial question was, did we have or would we say we have a standardised approach and I said I answered no to that and part of that is purely I don't think we do have that approach and ____ was very clear when I first started, actually there is no guideline that you follow and you know, you base it on your formulation and you do that work." – Participant 20

Some participants also spoke about the lack of well-established standardised guidelines for IOW within forensic mental health settings and how having to have such guidelines would have implications for practice (*n=9*). Participants in focus group four focused on the lack of guidelines and the lack formal protocols for working with risk in general (*n=4*). This may indicate that participants feel ill-equipped when conducting such work. In general however, participants identified how a standardised approach would risk not meeting service users' needs, and thus it needing to be individualised and not prescriptive. One participant suggested that where standardised approaches are used this may be as a result of a lack of resources and another indicated that such 'manualised' approaches were used within the prison service in conducting IOW. This may indicate further differences between secure forensic hospitals and other secure forensic services:

"I think that it's about whether we become so prescriptive that all these things have to be covered and then you get into that awful position of I can't do this, or this person can't do this, what if I haven't done it, what does that mean what are the consequences of not doing that? So we sort of say, most of the risk literature in terms of intervention will say an assessment, a formulation, an intervention which usually covers a focus on a behaviour, a focus on a trajectory, a focus on victims, a focus on staying away from further offending, to me that's as tight as I'd want to see

it…probably rebel against that despite the fact that that's exactly what I used to do whilst working in the prison service and I just think if most people know that typical risk work will include certain blocks of work then actually how you do them is about being responsive to the client…"- Participant 21

"You'd be doing the patients a disservice, so people use set protocols when they haven't got resources to do, to offer more, I mean obviously we're lucky because we can work with people individually, we don't need a sort of one size fits all type approach which often ends up then not meeting many people's needs at all." – Participant 16

A narrative of the offence. A pertinent sub sub-theme that was extracted from the data was the centrality of an offence narrative to the process of conducting IOW. Participants spoke about this in relation to the practical aspect of conducting IOW (*n=15*) and also the necessity of its interpretation (*n=14*). Differences between focus groups were observed, with participants in focus group 2 particularly concentrating on the relevance to the assessment process, and participants in focus group 4 highlighting the need for IOW and work around an individual's offence, to be guided by the literature. This may indicate differences in priority given to components in the process of conducting IOW within different services. Overall participants' spoke about the relevance of having a narrative regarding an offence and how this helped them in their role as psychologists, as well as helping the service user themselves.

Through participants' accounts in relation to the narrative around service users' offences, it was evident that its interpretation is deemed important in providing information about the service user and their risk (*n=14*). Some participants also highlighted that having a

clear narrative of the offence can also be therapeutic for the service users. There were however differences observed, in that some participants (*n=2*) endorsed the need for a good offence account, otherwise the work that is conducted can't be made explicit, whereas others (*n=2*) suggested that this was not the case. This may indicate alternative ways of achieving the same thing, an understanding of the individual and the antecedents to their offence:

> *"I think that usually index offence assessment and analysis is what they are talking about there so, at the most basic level that would be something about do we have a good story about what happened around the offence…" – Participant 2*

> *"I see so many patients who've had to trot out the details of the actual offence to umpteen people again and again and its sort of this rehearsed story and it's just irrelevant really, because you know as we've said it's all the stuff that led to it in the first place that's the important thing…"- Participant 16*

Challenges to process. Participants also highlighted a number of challenges in relation to the process surrounding obtaining a narrative of the offence (*n=10*). This ranged from the difficulties met in regard to actually completing the process, such as collecting historical information. This may indicate a lack of organisational policy in relation to accessing patient files, as in practice it shouldn't be as difficult to get access to these. Additionally, the introduction of electronic records in many services, could also have impacted upon this. Challenges were implied as being as a result of a lack of clarity regarding what IOW entails. This included one participant identifying the difficulty in relation to a lack of literature on IOW in general, but specifically for working with service users with learning disabilities, or Autistic

Spectrum Disorder (ASD), this emphasises the complex and diverse nature of service users that practitioners conduct IOW with and may indicate a feeling of being ill-equipped to do so:

> *"I think one of the things I am learning about, is I erm work in a ward for people with learning disabilities and ASD, is that there seems to be even less written about how to approach a piece of work with that population… There isn't a great deal written about index offence work per se but that particular erm area is more bereft of research and guidance" – Participant 1*

> *"Yeah, because either the gaps in existing formulations or gaps in the assessment…even though there could be volumes of notes it doesn't necessarily translate into volumes of useful information (p3 "Yeah yeah")"- Participant 2*

Tools used to assess risk. A comparatively small number of participants spoke about the role of assessment tools in the process of conducting IOW (*n=9*) in comparison to other sub-subthemes. Typically participants spoke about the selection of tools being dependent on the service user that they are working with. Within this factors determining tool selection such as practitioner preference and offence type, highlighted the idiographic nature of IOW. Strengths and limitations of actuarial and clinical judgement approaches to assessment in IOW were also discussed, as well as the role of assessment tools in the assessment of risk. .

A more prevalent discussion across focus groups (*n=11*) was unsurprisingly the role of tools in assessing risk. Common themes surrounded the use of psychometric and other actuarial tools to inform assessment as well as distinguishing this from conducting specific risk assessments. Participants (*n=6*) described the use of psychometric tests to inform IOW, during

the assessment and formulation stages and also in therapy in order to measure progress. Some explicitly spoke about not having a standard battery of psychometrics that they would use in IOW and this linked to views presented by other participants, indicating selection is based on the individual service user, offence type, and also the preference of the practitioner:

> *"Yeah, and then I suppose in terms of assessments, there are certain psychometrics you might use during your assessment and formulation period and then also during therapy, you may erm use measures to look at progress and how the patients getting on and all that." – Participant 15*

> *"I mean we would probably if we were going to do a personality assessment would mostly do the PAI, or if you wanted something that was kind of slightly more diagnostic it might be the MCMI, but I think we would probably use that less… so we might have ones that we favour I think within the department, but we still have access to a whole range of things, I think if you're dealing with a sex offender we would use some of the standardised measures…" – Participant 3*

When discussing the use of tools in IOW one area which was spoken about was risk assessment and the fact that this is a standard process within IOW. In particular a number of practitioners (*n=6*), indicated completing a HCR-20, or alternative risk assessment, are standard within IOW. One participant highlighted the fact that such risk assessments combine the actuarial and clinical judgement approaches and as such are more useful:

"And its part I mean now its part of standardising for the HCR-20, so every patient will have that a formulation for the most prevalent violent behaviour."- Participant 14

"I think there's enough evidence in the literature to be fairly flexible, it depends on kind of the individuals experience and skills, we do have an initial guide which says use the HCR-20, the SVR-20…"- Participant 18

A number of limitations of actuarial approaches were highlighted by participants during discussions. These surrounded the lack of applicability of actuarial assessment methods to the client group within secure forensic mental health services and therefore IOW. It was acknowledged that clinical judgement approaches are less accurate than actuarial assessment tools:

"It's finding a comparison data group for a guy with frontal lobe injury, who raped his wife and has substance misuse problems which never happens, so you might be able to be quite precise, but whether or not it relates towards the person you're working with is something that I don't find easy…" – Participant 2

"It also gives you a number, but it doesn't give you the why's or wherefores as to what you should do to try and help minimise the risk…"- Participant 1

A unique discussion to focus groups one and two, pertained to the view that practitioners work within secure settings goes beyond the actuarial (*n=8*). Highlighting therefore, the superior role of clinical judgement and the need for dynamic risk assessment in IOW:

"It depends what you mean by that really you know, it depends what you're if you're writing a report and what purposes you're writing it for, so I suppose in our general clinical work with this population it goes a bit beyond the actuarial doesn't it and you've got bits of that incorporated in the HCR-20 which you may not agree is the best methodology, but if somebody was to come to us and say look for this tribunal report we need to have an actuarial judgement and ok to do this, or thi,s is very basic and there are usually four points to the question and that's it, but for us clinically because most of our job is about the dynamic assessment…" – Participant 3

"And then depending on as we have just discussed in terms of what makes index offence work, what that means, what that looks like, it can be about anything so we use an awful lot of psychometrics that assess all sorts of different things,, which form part of the index offence work, then there's loads of it isn't there? Of course there's a lot of clinical judgement that's about you're assessment and you're formulation and you're conversations with the service user erm so yeah." – Participant 8

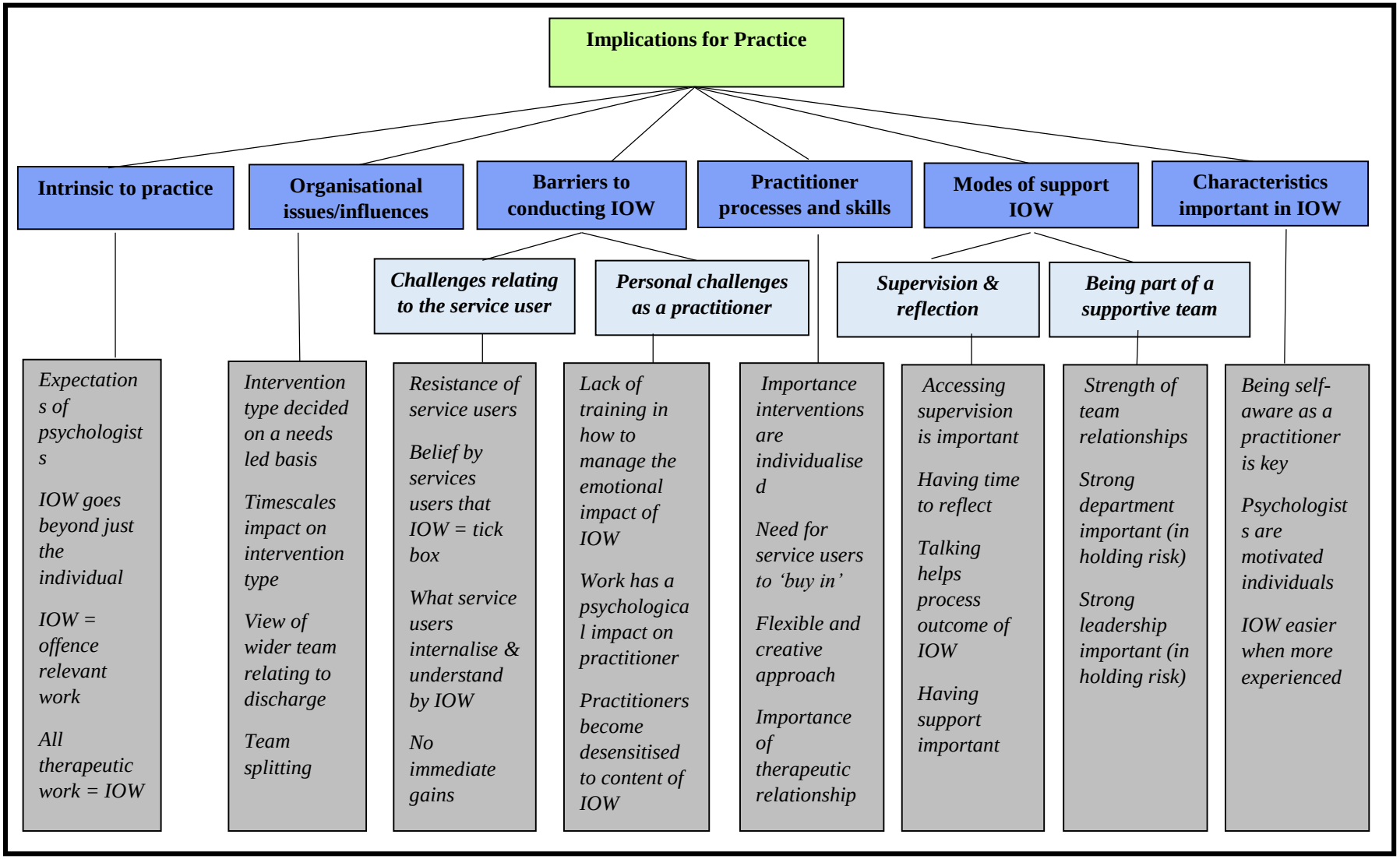


Figure 6. *Thematic hierarchy of Implications for Practice.*

Implications for practice. Across focus groups practitioners' discussions often surrounded the implications of IOW in terms of it practical implementation. These viewpoints ranged from the importance and centrality of IOW to practitioners roles, to the characteristics considered as important for practitioners to possess, in order to effectively conduct IOW. Due to the range of areas discussed in relation to this overarching theme, it was further delineated into a total of six sub-themes, as can be seen in Figure 6.

Intrinsic to practice. A number of participants (*n=15*) described IOW in the context of their role as psychologists and the embodiment of work that they do in their everyday job role. This ranged from participants endorsing that IOW is intrinsic to being a psychologist, to going beyond the individual and relating to their role in the context of the wider team. This is exemplified by participant 21 who stated *"…its risk that's the remit of the psychologist, mental health is the remit of the psychologist, everything else is the remit of the psychologist…".* This indicates that there may be a common view held within secure services, that the psychologist deals with a wide range of factors relating to the individual, and that it is expected that they will do this. This may further indicate a lack of understanding of other professionals within secure services as to what such work involves.

Specifically characterising what IOW involves as a practitioner working within a secure setting appeared difficult for the majority of participants. They endorsed the idea that it is difficult to view IOW as a standalone piece of work and the majority of participants provided information which indicates that IOW is a broad area. Such comments from participants would suggest that IOW is difficult to define and comprises a number of different components. It also indicates a view that it is difficult to separate out IOW from somebody's entire treatment pathway, instead running intrinsically through everything that both the psychologists and the service users do:

"…It's a regular topic on the agenda for ward rounds every week and it's always there isn't it? Within the treatment planning, within the team..."- Participant 13

"I find it yeah, I find it really difficult to think that the work that we do isn't index offence work, because whatever decisions sort of happen within that are all based on whatever's happened before, like you know if it's regarding coping skills well erm in some cases it may be that coping skills fit in with the commission of that offence, if it's about relationships, if it's about attachment, so index offence work to me is just the work cause that's how I always view it…" – Participant 20

Organisational issues/influences. Practitioners (*n=14*) spoke about the influence of the organisation on decisions regarding intervention type (group or one to one) when conducting IOW. It was highlighted that decisions are made on a needs led basis and group work is often not suitable for all service users, but timescales also influence practitioner decisions. One participant also accentuated the financial impact and how this influences intervention type, as one to one work is often more costly, with another illustrating the influence of the wider organisation such as other professionals and anxieties held regarding offence focused groups. Discussions on this topic were conducted mostly in focus groups one and two, this may indicate that offence focused groups are newer concepts within these services. An interesting idea proposed by participant 19, suggested that the concept of whether groups or one to one work is most effective is always changing, and this may imply a periodic changing in the wider organisations preferences for the format of offence focused work:

"But groups aren't suitable for everyone (Participant 6 – They're not for everyone no") or it might not at that particular time, so I agree its erm needs led

but groups can be really powerful…It's cost effective as well to do group work as opposed to individual interventions." – Participant 10

"I think it's both and I think from my understanding and you guys can correct me, my understanding is we move often through cycles of going towards more manualised approaches to then rebelling against it and thinking after you know that didn't work so let's go back towards the individual stuff, go back towards group stuff…" – Participant 19

A discrete aspect of this sub-theme also extracted from participants' accounts (*n=11*), related to the difficulties relating to the wider organisation and this being a barrier to IOW. Participants alluded to the lack of understanding of the wider team, often resulting in splits within teams. There was a particular consensus amongst participants, of the regular debates regarding the completion of IOW and service users' release from services, within the wider clinical team and disagreement over service users' treatment pathways:

"…they very much appreciate our work but they don't necessarily relate to what is happening and what it is that we're holding…" – Participant 3

"You can also get splitting in the team then as well, when you get people thinking why are they still here, they don't need to be here, well they haven't done the work and their offence was really violent…" – Participant 5

Barriers to conducting IOW. When reflecting on personal challenges to conducting IOW, the majority of participants (n=18) identified barriers to conducting it. They

expressed opinions which indicated challenges as a result of the service user's presentation, as well challenges personal challenges. Sub sub-themes which comprised the barriers to conducting IOW, included Challenges relating to the service user and Personal challenges as a practitioner.

Challenges relating to the service user. A moderately sized sub sub-theme was extracted from the data relating to barriers when conducting IOW, with 11 out of 21 participants expressing views which pertained to this. The participants described a level of resistance from the service users, who are often unwilling to think about their risk and regularly view IOW as no more than a tick box exercise. This led on to other participants who identified other challenges as being related to high levels of shame, a lack of emotional readiness and the trauma relating to the index offence:

> *"I guess a lot of other people in the system who that isn't their primary task or first thing on their agenda and there might be lots of reasons not to think about someone's risk when they really don't want to and it makes them very upset and it might make everyone's job harder at that time…" – Participant 2*

> *"Erm it can be draining work to do, I mean it's challenging for you as a psychologist, I think, I think it's yeah it depends what level, how motivated someone is to do the work because I've certainly worked with someone before who felt they needed to sort of tick a box to get through and do the index offence work and it was a real it was a real struggle cause they felt sort of forced into doing it really…" – Participant 9*

Personal challenges as a practitioner. This was a larger sub-subtheme which was endorsed by a greater number of participants (*n=16*), participants spoke about different individuals impacting upon them in different ways, and the challenge of needing to maintain rapport with the service user despite this. For example, the impact of transference and countertransference and the emotional challenges as a result of this. Some participants highlighted the fact that practitioners become desensitised to the content of IOW. Interestingly one participant highlighted that practitioners receive a lack of training in regard to how to manage aspects of IOW. This was the only participant who highlighted this factor as being a personal challenge in conducting IOW, but it evidences possible further training needs relating to IOW:

> *"That's really interesting, I've gone totally the other end I actually don't give a damn now, I'm like, if I'm gonna get attacked I'm gonna get attacked, I might as well go out he's gonna blindside me anyway,, so I'll carry on doing whatever I'm doing." – Participant 19*

> *"I think it's such a difficult thing to do, you know for lots of different reasons because it's also I think your often aware of other people's embarrassment and their view of themselves, erm and it's also trying to continue to maintain the rapport with somebody under those circumstances and also knowing when it's the right time to approach certain issues with patients, cause you know you have to pick your moment as well, sometimes you don't really get the opportunity to pick the moment but you know if you can you should…" – Participant 14*

Practitioner processes and skills. All participants (*n=21*) attested this sub-theme and discussed topics relevant to the skills and processes which they utilise in their practice, relating to IOW. Within this theme the need for involvement of the service user themselves in conducting IOW was identified as being of paramount importance.

Focus group members articulated the need for IOW to have a focus on the individual (*n=17*). This included the importance of interventions being individualised, in order to meet service users' needs, and be meaningful to them. Simultaneously, participants emphasised the need to be responsive to the individual, working with whatever the presenting difficulties are for that individual, irrelevant of the approach used. One participant, a consultant clinical and forensic psychologist, articulated that this should be done regardless of there being the presence of an offence. This may indicate a more clinical as opposed to forensic take on IOW, as it suggests a broader focus than just risk, a prevalent view of the forensic mental health settings across focus groups. It was also emphasised that acknowledging the impact of IOW on service users, and taking a flexible and creative approach to the work is necessary, particularly when working with client groups which present with complex needs:

> *"I think it's being creative, but it's also thinking about what skills does someone need to have to be able to provide an account of what are your expectations, so for example things like sequencing is a big problem for erm people with learning disabilities and ASD, so my last index offence is on lots of little bits of paper with very bad stick figure drawings so that we can move them around…" – Participant 1*
>
> *"It's very individualised it has to be." – Participant 16*

The salience of the therapeutic engagement of service users, was also a prevalent aspect of this sub-theme (*n=14*). Participants articulated the importance of collaborative working in encouraging this and enabling service users to have an understanding of the need to complete the work. The importance of the therapeutic relationship in conducting such work and promoting this was also made clear:

> *"I guess, erm in an ideal scenario what one would hope but not always achieved is if you're able to share your formulation of the index offence and have that be at least partially collaborative where possible, then there can be something therapeutic in itself in terms of people having some understanding of they ended up in that situation…"– Participant 1*

> *"…I think you know some service users may not see the sort of relevance of doing that particular piece of work and it's our job to sort of be able to engage them, to sort of do that piece of work meaningfully." – Participant 9*

Modes of support. Participants (*n=14*) reported accessing different modes of support in order to manage the personal challenges relating to IOW. Sub-subthemes focused on different modes of support were categorised into *Supervision and reflection* and *Being part of a supportive team*.

Supervision and reflection. A highly prevalent part of the overall sub-theme (*n=16*), within participants' responses, emphasised the need to access supervision and utilise reflective practice to manage IOW. Perceptions which characterised this included the need to be reflective when carrying out this work in order to manage what you experience when working with clients with complex presentations:

"We have clinical supervision which is one of the ways you can reflect on your work but a lot of the time it is like informal peer support as well isn't it?" – Participant 7

"It's like an informal de-brief isn't it? You come back from you're session and you're like whoa and we do have those conversations in the office quite a lot don't we? (Participant 7 – "Yeah we do"), because it is about kind of supporting one another and just giving time to one another, time to be able to think and reflect" – Participant 8

"Absolutely, but I don't know, I think and I suppose for us that's what we get in certain elements of our supervision sessions, if that's what we want to use it for, as well as doing absolutely mandated this is protected time specifically for this…" – Participant 21

Being part of a supportive team. A small number of participants (*n=4*) described being part of a supportive team as being important in managing the impact of IOW. These participants referred to the role of a strong team, with strong leadership, in helping practitioners to hold the emotions that IOW often provokes. One participant highlighted the necessity of team resilience and this may have been highlighted particularly due to the fact, that she is the service lead for a service who has recently gone through restructuring. This may indicate that psychology teams may not only need to be resilient in relation to IOW, but also to other issues within the wider organisations in which they work:

"…in terms of coping I think it helps when you are part of a team and can talk to people, I think that really is critical and just acknowledging that what we do isn't

normal it might be normal for us, but it you know it's not the sort of thing that most people go to work and do." – Participant 14

"…I think the strong department, I think we're a strong department, I would say that wouldn't I? Is really important and although this may be going off track a little bit, I think what I've rediscovered which has been really pleasing to me, is that I think that strong leadership, colleagues might disagree with me, is actually quite important as well for how you manage the job… we've had some real challenges about I think how we're viewed or how our role is viewed and I think the resilience of the team is not a surprise to me but it is a surprise to me in some ways..." – Participant 3

Characteristics important in IOW. Clinicians (*n=13*) identified a number of important characteristics of practitioners conducting IOW. They spoke about the need to be self-aware, objective and motivated when carrying out such work, as well as the need to have robust coping strategies, such as planning activities in order to be able to leave work at work. In particular, recognising your own difficulties appeared to be an important characteristic and was endorsed by a number of participants (*n=10*):

"There's some cases where I've found it really hard to switch off and I've been for weekends away and I've seen that person on the Friday afternoon, sillily or stupidly should I say, and it's ruined my weekend away because that's all I've thought about…so you have to be careful that it doesn't." - Participant 6

"…I think here for me as a supervisor hearing some of this really difficult stuff in supervision at a time when everything else is really difficult really makes you realise what how important this work is and erm how psychologists, the reason why psychologists do it, because their training is very intensive, you can't get to be a psychologist unless you're massively motivated because it's so hard to get there in the first place and that actually that training and the thing that comes from people innately, that makes them want to become a clinical psychologist is key to how you manage the work that you do…" – Participant 3

Discussion

The current study aimed to explore qualified clinical and forensic psychologists' understanding and use of IOW/IOA. This was examined by the author, through conducting a TA on the accounts and discussions of practitioners during focus groups. The content and processes of IOW, as understood by practitioners, were identified by the author using a bottom-up atheoretical analysis of practitioners' descriptions of the practical aspects of their role and how this related to IOW within their services. Qualitative data analysis resulted in the extraction of themes relating to the format and processes used by practitioners in conducting IOW, as well as issues relating to this. The ways in which practitioners identified that they conducted such work, were found to be consistent with the 'what works literature', in relation to working with offenders. For example the RNR model (Bonta & Andrews, 2007).

The assessment of offenders impacts upon important decisions relating to an offender (Wood et al., 2002) and therefore it is vital that they are accurate and that their quality is consistent. Findings indicate that in general IOA is not utilised in practice and there are inconsistencies in the understanding of what IOW really is. These results support the findings

of previous research, for example Fallon (2007) who found that in general within a medium secure unit there was a lack of understanding of service user's index offences and that it was often difficult to conduct such work.

Issues Regarding Understanding of IOW/IOA and Implications for Practice

The salient finding of this research across themes was a clear lack of consistency in understanding relating to IOW in general and more specifically of IOA. Findings highlighted the importance of the assessment process in IOW. The centrality of IOW to their general practice was acknowledged by practitioners, and this is consistent with the literature, which identifies that work focusing on a service user's offence is a core task of forensic clinicians (Daffern et al., 2007; West & Greenhall, 2011). Across focus groups however, inconsistencies between what constitutes IOW were observed and as such findings indicate some lack of knowledge, or understanding amongst professionals, which may indicate training needs. Different services focused on different factors as priority components of the work and as such this may indicate a lack of consistency across forensic practice. However characteristics required and requirements for support when completing such work, were consistent.

A widely endorsed theme derived from the data and seen across focus groups related to issues regarding the terms and definitions used. Participants appeared to struggle in providing any clear definitions in regard to IOA or IOW. They conveyed difficulty in defining the work they complete with service users as IOW, due to the fact, the work that they do is much broader than just focusing on the index offence. Further to this, a lack of or different understanding of the concept and content of IOW from other professionals, was identified by participants as inherently problematic. Particularly when making decisions as part of the wider team regarding service users' risks or treatment pathways. Findings therefore indicate that within forensic

mental health services, there remains inconsistency in the definitions of IOW used, and whilst practitioners complete IOW as part of their practice, what it involves is down to the individual's understanding of what it should. As a result of this, it is evident that IOW has the potential to look different and be conducted differently depending on what terminology the professional is using, making it a subjective concept and resulting in work with offenders being inconsistent.

Some researchers highlight that forensic practitioners analyse the patterns of crime for risk factors, in order to predict recidivism, as part of the assessments that they conduct with their clients (Moore & Drennan, 2013). Others have suggested that this is in a similar way to investigative psychologists, who analyse crimes to predict offender characteristics in order to prioritise suspects (Alison, Goodwill, Almond, van den Heuvel & Winter, 2010). Within the field of forensic psychology itself, researchers are increasingly emphasising the importance of crime scene analysis for risk assessment. For example, Beech et al. (2003) highlight that functional analysis including the identification of modus operandi, is one of the four major components of risk assessment. Further to this, Lehman et al. (2013) emphasise that one of the main reasons for low levels of predictive accuracy in offender recidivism, is the lack of examination of the index offence. As previously noticed, instead of focusing on the index offence practitioners within the current study view IOW as broader than this. The lack of evidence from findings which indicates that practitioners make clear links between offence information and the accuracy of their risk assessments, further emphasises gaps in knowledge and understanding relating to IOW.

This is further exemplified by the fact that whilst IOW in general was spoken about by all participants, IOA was only spoken about by a proportion of these, indicating a more limited understanding. Whilst these participants did advocate that IOA relates to formulation or functional analysis, only one participant evidenced that IOA is a specific formulation of an

individual's offence. Thus emphasising the limited knowledge of practitioners in general in regard to IOA. This may indicate a significant lack of training or understanding in this area and may suggest that there are gaps in the assessments which practitioners in forensic mental health settings currently conduct or that they have a lack of experience in doing so. The BPS (2011), emphasise that the quality of a practitioners' formulations are dependent on the quality of the assessment and information derived from it. Therefore if practitioners are failing to utilise important information regarding a service user's index offence in the assessments which they conduct, it is likely that there may be gaps in both their formulations and assessments.

Further to this, in one of the focus groups IOA was highlighted by a group of clinical psychologists as something which is better understood by forensic psychologists. Whilst these practitioners did agree that IOA is functional analysis, they did not clearly demonstrate that they would utilise such skills in understanding an individual's offence. This may suggest that there are differences in training across the different disciplines. It could also be as a result of the higher proportion of clinical than forensic psychologists in this sample. Despite this, the sample is however reflective of clinicians working within forensic mental health settings and as such demonstrates limited understanding of specific offence analysis and its incorporation into wider assessment. Arguably, this could indicate that practitioners working within these settings, are ill-equipped to carry out such work, and findings from questionnaire data supported this. Future research examining and comparing forensic and clinical psychologists' use and understanding of IOW/IOA would be useful in exploring this further.

A less prevalent but important finding and one which may provide further insight into the differences in understanding and use of IOA/IOW was the concept that bringing risk and mental health together is problematic. Participants spoke about it being two separate models trying to fit together coherently, which is not always possible, and as a result this means that in

forensic mental health settings, mental health becomes the focus as opposed to risk. This causes problems both within wider teams, regarding disagreements as to what takes priority, but could also have wider and more problematic implications. If there is a lack of clear understanding of IOW within secure forensic mental health settings and work surrounding mental health is viewed as meeting this need, this may result in other risk factors remaining "untreated" and offenders' risk needs failing to be met. The differences of opinion as to what should be the focus of IOW further exemplifies the need for clearer guidelines and training in order to ensure that the work is consistent across services. There are also issues for service users on 37/41 sentences who go between hospital and prison settings. If IOW hasn't really addressed their needs, but it is stipulated that they have already completed offence focused work, they may miss out on opportunities to complete offence focused groups or other risk focused work. As such they may be re-released into the community with outstanding treatment needs, thus making it more likely that they will offend again.

This finding is consistent with the wider literature in that often practitioners working within forensic mental health settings are tasked with delivering services to two clients, both the service user themselves, and also the organisation and the wider general public (Halleck, 1987; Howells, Day, & Thomas-Peter, 2004; Ward & Salmon, 2009). Maden (2008), highlighted the potential struggle of working within health systems or criminal justice systems, in seeing the perspective of the other. For example, failing to see risk as the primary issue or failing to see the broader issues such as mental health and instead just focusing on risk. Further to this, Robertson, Barnao and Ward (2011) argue that the challenges relating to this are greatest in forensic mental health settings, where treatment and risk perspectives attempt to converge into one coherent system. This therefore results in considerable debates within teams, as found within this study, regarding the purpose of "treatment" and therefore may impact on assessment

procedures which are utilised. Consistent with the findings of the current study, Duggan (2008) identifies, that questions have been raised as to whether the focus of interventions should be to reduce the risk of reoffending or to address symptomatic distress and an individual's other psychological needs. Such debates are likely to cause tensions within the wider team (Davies, Heyman, Godin, Shaw & Reynolds, 2006) and it is therefore evident that differences of opinion are likely to impact upon the format and focus that IOW takes.

A lack of understanding and differences in opinion regarding the focus and processes used in such work may cause significant splits within teams and could be problematic for the smooth running of organisations. Sturmey and McMurran (2011), suggest that sharing with the wider team, between all professionals including relevant non-clinical team members is critical to safe offender management. It is evidently important to have agreement amongst the wider team. If a definition could be agreed upon with clear guidelines for all professions to follow in relation to the conduct of IOW then this may be beneficial. Findings also indicated that such differences in understanding may vary across different services. There was a view held across participants that other forensic settings such as prisons, hold a restricted view in relation to what IOW is, with practitioners identifying that within forensic mental health they are better able to cater for individuals' needs. There are evidently different processes used in regard to conducting such work across settings, for example there is much more offence focused group work conducted within prison settings. Such a finding would indicate that conducting similar research across settings, would provide interesting findings about the consistency in process or differences across forensic settings.

In summary, the findings of this research support what the literature base already indicates, that there is some lack of understanding amongst professionals working within forensic settings regarding IOA/IOW. It is possible that this issue could be rectified through

additional training, however more needs to be known about the differences in understanding from specific professionals, in order to identify what and where training needs should be targeted. It is also possible that it could be the term itself which is unhelpful, as findings from this research indicate a good understanding of practitioner roles.

Content, Processes and Format of IOW and Implications for Practice

Despite prevalent themes relating to a mixed understanding of the necessity and use of IOW and IOA, data treatment resulted in the extraction of several pertinent sub-themes relating to the content, processes, and format that practitioners do currently use as part of their understanding of IOW, within forensic mental health settings. Regardless of issues pertaining to consistencies in understanding, through practitioners' discussions it was evident that work surrounding a service user's index offence, is viewed as core part of their role. These findings are consistent with other research which suggests that IOW should be a core task of forensic practitioners who are engaged in working with and assessing offenders (Daffern et al., 2007; Weist, 1981; West, 2000; West & Greenhall, 2011).

Content and goals of IOW. The most prevalent theme to emerge from the data regarding what IOW involves for practitioners working within forensic mental health settings, related to skills and processes utilised by them. Participants' discussions surrounded the need for participants to focus on the individual in order to fully meet their needs and for the work to be meaningful to them. In particular the need for the client to 'buy in' to the work and to work collaboratively with the practitioner and thus develop a strong therapeutic alliance. Glass and Arnholt (2008) suggest that collaborative working is key when working with service users on offence focused work. They indicate that it is highly valued by service users, due to the fact it

enables more transparency in process and allows for agreed goals and methods between practitioner and service user, thus enhancing the therapeutic alliance.

Other themes which were widely endorsed further evidenced the need to work in a responsive way when conducting IOW. Factors were largely related to the service user themselves and included: treatment readiness and resistance, engagement, cognitive and social functioning, mental state and the impact of trauma. It is likely that such factors impact on what must take priority in terms of the focus of the work and imply that conducting IOW with complex individuals can be difficult. Poor cognitive functioning as a result of psychosis, substance misuse, or other mental health issues, have previously been indicated to impact upon service users' engagement in such work (Pyszora et al., 2003). Factors evidenced by practitioners are likely to impact on service users' engagement in such work and therefore what dictates the content and focus of the work. This would indicate that developing standardised ways of conducting IOW is difficult. In addition to service user centred factors, a key theme was the influence of the wider organisation on such work. For example, the backgrounds of practitioners' themselves and their preferences for the ways in which they conduct such work. This further links to a smaller theme surrounding the influence of the team and the idea that IOW should promote team working, in order to have a wide understanding of the individual. Such ways of working are inevitably likely to increase understanding and thus allow the team to better manage and contain service users (Sturmey & McMurran, 2011).

There is a wealth of research which considers the need for practitioner's work with offenders, to be as much offender focused, as offence focused (Andrews, Bonta & Wormith, 2011; Blackburn, 2004). The complex psychological and social problems of the population within forensic mental health settings mean that individualised assessment and treatment formulations are necessary. Treatment programmes need to have multiple components in order

to address such problems and the findings of this research exemplify this. Themes were drawn from the data which related to IOW being conducted on a one to one and group basis by practitioners in these settings. Whilst practitioners acknowledged the important role of group work for enhancing service user understanding, they highlighted reasons for this being problematic in practice. Reasons included the negative impact of offence disclosure and responsivity issues. It was emphasised that the majority of work regarding service users' index offences is at present carried out on a one to one basis, as it was highlighted that group work often fails to address the real causes of an individual's offence, and therefore risks not meeting their needs effectively. Due to the diverse population, it is often difficult to get a group of individuals at the same level to run a group such as this. Such issues relating to IOW, may be specific to the diverse population in forensic mental health settings. Blackburn (2004) highlights that offence focused programmes for mentally disordered offenders, are seen as an adjunct to working with an individual offender and their needs. This is further supported by the findings of the current research, which identified that groups which focus on more peripheral factors relating to offending, can be a useful aside to one to one offence focused work.

Whilst such ways of working are in line with models for offender rehabilitation and have good validity (Polascheck, 2012), for example the RNR model (Bonta & Andrews, 2007), it is possible that within these settings the index offence becomes more peripheral, resulting in the offence virtually disappearing (Herman, 1990). Some researchers have identified, that failing to identify specific information regarding service users' risks in assessment and intervention and instead attributing behaviour to internal, cognitive deficits can be problematic (Hayles, 2006). Whilst the current study's findings are in line with previous research and argue against a 'one size fits all approach' to offence focused work, they do raise questions regarding

the consistency of IOW conducted across forensic settings. Questions regarding the sufficient flexibility of the manualised and closely audited programmes within the criminal justice system (Hart et al., 2011) are raised. It is possible that what participants exemplified within this study in relation to the way they conduct work with offenders, is a more complex and adapted forensic RNR model, as suggested by Robertson et al. (2011). Such a model however is theoretically underdeveloped at present and lacks empirical support. Therefore whilst forensic mental health practitioners may be effectively meeting the responsivity principle, how can they be meeting the risk principle if they are not effectively assessing and targeting this through specific offence focused intervention? Although, it could be argued that the RNR model fails to account for the unique set of variables that culminate in a mentally disordered individual's offending.

In general participants evidenced that they take an individualised and all-inclusive approach to the IOW conducted with offenders. There was no clear evidence provided that a coherent and detailed exploration of a service user's offence is conducted and by not considering the more exact details of an offence it is possible that the offence gets lost completely. However, this is dependent on the nature of the offence and the driving factors within this, for example substance abuse work alone could potentially reduce an individual's risk. Whilst practitioners through their interventions may conduct a more systematic exploration of an offence, this will rely heavily on a service user's ability to self-report and this is problematic within itself (Harry, 1992; Melton et al., 1997; Spence, 1989). It is possible that in practice the offence itself just becomes a tick box within the wider intervention and as such aspects of it which may provide information regarding an individual's thinking, feeling and behaviour are lost. Given the importance of such assessments this is inherently problematic.

Processes utilised in IOW. A number of themes surrounded the idea of having a standardised process for conducting IOW. Participants indicated that care plans, risk

assessments and formulation are all standardised processes which make up part of conducting IOW, however a lack of guidelines or formal protocols for specifically conducting IOW was evident, although given the broad nature of IOW it is difficult to determine what this would look like. In line with participants' views of the need to conduct work in a responsive way and to best meet all service user's needs, a standardised approach for such work was deemed as being unable to do this, given the complex nature of the population within forensic mental health services. Some participants evidenced that standardised processes were utilised only when services had a lack of resources and this again poses questions regarding the differences between offence focused work conducted across settings, which require further exploration. There was however, some standardisation in processes relating to IOW, in that practitioners stated that they will utilise risk assessments and other tools, in order to inform their assessments and formulations for this work. Such ways of working are again in line with what the research exemplifies. For example, Rice and Harris (1997) outline the characteristics of an ideal treatment programme. They assert that such a programme would: appraise risk of recidivism using actuarial devices, deliver services at an intensity which matches a service user's level of risk, and be focused on criminogenic needs as treatment targets. At present, there is no such device which provides reliable risk estimates, or which actually predicts completely accurately, whether an individual will reoffend and this therefore presents an idealised concept. In addition to this they suggest the use of psychotropic medications and CBT therapies in order to teach offenders other skills would be another important component. Whilst practitioners spoke about broadly looking at antecedents to a service user's offence, they focused much more widely on other factors which if targeted through intervention will indirectly reduce risk. A lack of specific focus on the index offence may result in important risk factors which relate to the service user being missed, although more specifically, this may ultimately come down to the

quality of the clinicians assessment or report. Howells et al. (2004) suggest that a formal risk assessment should be conducted by practitioners in conjunction with a functional assessment approach to understand behaviour in mentally disordered offenders. Conducting IOA as suggested by West & Greenhall (2011) would allow for such an approach, but at present the findings from such research indicate that practitioners do not do this.

Gaining a narrative of the offence was a key theme extrapolated from the data, occurring in the context of a broader psychological assessment, the importance of formulation as a key process within IOW was also evidenced. The purpose of the assessment is to make sense of the individual and their treatment needs and this includes offence behaviours. Practitioners stated that gleaning a clear picture of an individual's offence can often be difficult as there can be gaps in historical information. Gaps in collateral information are likely to result in practitioners relying on the offender in order to gain a narrative of the offence. Aside from the issues this highlights, this also indicates that there may be issues with the information that practitioners review as part of the assessment process. This again questions the robustness of such assessments. It was evident that different services give a different priority to the necessity to gain such an account, with some individuals suggesting that an offence account is irrelevant and IOW should focus in greater deal, on targeting the antecedents and other issues that an individual presents with. As aforementioned, due to the complex nature of the client group it may be that there is a need to take a broader approach and as Robertson et al. (2011) suggest, intervention arguably comprises the bulk of the clinical work undertaken with this population. However, this contradicts participants' views that the primary goal of such work is to target risk and may demonstrate issues pertaining to organisational culture.

The collection of third-party information such as medical, criminal, educational and employment records, and statements from witnesses and victims, have been stipulated to be a

central characteristic of forensic assessment, distinguishing it from other assessment types (Nicholson & Norwood, 2000; Heilbrun, Rosenfeld, Warren & Collins, 1994). Information like this is essential in order to corroborate information taken from the accounts of offenders, who have a tendency to distort responses (Melton et al., 1997). Moore and Drennan (2013) highlight that in forensic clinical practice, there is often missing data, with the multiple contexts and perspectives regarding what happened in an individual's index offence. Within the current research participants identified that collecting historical information can be difficult and this highlights that in practice there could be gaps in assessments. Heilbrun et al. (1994) suggested that gaps in formulations and assessments may be as a result of the 'accessibility effect'. This stipulates that the more readily available the information is, the more likely that it will be incorporated by practitioners into their forensic assessments of service users. Within forensic services there is often information which fails to be includes in clinical records or databases, as such, this may make the assessments conducted inherently bereft. This could be an explanation as to why practitioners currently fail to complete IOA and a formal analysis of an individual's offence, as crime scene data and/or depositions, may be unavailable to them. Police have access to crime scene data and this may highlight another issue regarding a lack of sharing between the police, prisons and health services. This could indicate wider organisational issues, in that such information needs to be made more readily available, in order to ensure that practitioners' assessments of offenders are robust and ensure that all risk factors are identified.

Formulation was accentuated to be key within IOW and the first component in the process, directing everything which follows. Formulation was spoken about as being important for the practitioner, service user and the wider team in allowing everyone to have an understanding of their difficulties and what can be realistically achieved. Participants evidenced that such information is vital in directing the service user's treatment pathway both

in terms of intervention and directing service user's release. Thus enabling practitioners and service users, as well as the wider team to think about working towards release as suggested by West (2000). These findings highlight the importance of effective formulation and assessment in IOW and this is consistent with previous research (Gresswell & Hollin, 1992; Green, 2008; Hanley et al., 2003; Lazarus, 1971). Findings of the current research indicate that practitioners working in forensic mental health settings do conduct formulation as part of IOW, however they indicate a specific formulation of the offence is not conducted. Eells (2007) proposes that case formulation should organise practitioners' hypotheses regarding the causes, precipitants and maintaining factors relevant to an individual's difficulties. The way in which practitioners explained their use of formulation is consistent with this approach, whilst they may use models such as functional analysis to assist in formulation, participants did not in generally speak about this in developing a formulation of the offence. Moore and Drennan (2013) highlight that case formulation goes beyond the application of a single model, and this is consistent with practitioners view in relation to this. They identify that a case formulation allows for the inclusion of a risk assessment, but does not wholly rely upon it. Only one participant spoke about the inclusion of a specific offence formulation which contributes to the broader formulation which may also include a risk assessment.

This further highlights that there may be components missing from practitioners' assessments and the development of formal guidelines may assist in ensuring consistency and the inclusion of all important components as suggested by West & Greenhall (2011). However, this relies on good research to assist practitioners in understanding the key components. Whilst there is consensus generally that mental health professionals should conduct formulation as a core competence of their practice, as indicated by Hart et al. (2011), there is no agreement concerning how practitioners should conduct or evaluate it. Logan and Johnstone (2010)

suggest that formulation might be particularly important when working with individuals who fail to respond to medication, who present with complex problems, or who may pose a risk of harm towards themselves or others. This makes sense, given the population that practitioners who took part in focus groups work with. However a specific analysis of offence behaviours may assist in telling practitioners more about these individuals, and thus allow them to assess and treatment plan more effectively.

As evidenced here, findings indicate that practitioners do complete IOW within forensic mental health settings and that this has a number of key components. What is clear however, is that the work is very broad in nature and requires an individualised approach. As such this would make specific criteria and guidelines difficult. Despite this, the lack of inclusion of specific offence analysis was evident and this may mean that the assessments conducted with offenders have significant gaps and fail to identify all relevant treatment needs. Guidelines therefore relating to the inclusion of a specific offence formulation and the collateral information which should be reviewed as part of this may be more useful.

Limitations

The current research is not without its limitations. Firstly, it is possible that the present study suffers from being unable to generalise findings due to the relatively small sample size used. The sample size was deemed suitable for use with a qualitative methodology (Braun & Clarke, 2013). Due to the broad nature of the research question however, it is conceivable that other factors relating to practitioner's use and understanding of IOW/ IOA, remain unexplored. This study may also be limited due to its reliance on purposive sampling methods and strict inclusion criteria. The inclusion criteria was qualified psychologists working within secure forensic mental health settings. This may have resulted in the population being studied not

being completely representative of all psychologists working with offenders, across secure settings. During analysis differences were observed between the opinions of forensic and clinical psychologists, as well as indications that practitioners viewed IOW being conducted in a different manner in other forensic settings, for example prisons. In addition the sample comprised of qualified psychologists only, the majority of whom were clinically and not forensically trained. As such it may not be appropriate to draw conclusions regarding IOW and IOA practices and processes across psychologists' work with offenders in different services, as problems may be specific to those working in secure forensic mental health services. Despite these limitations, findings have been useful in beginning to highlight different ways of working and discrepancies in training and understanding. They provide a clear focus for future research where practitioners' understanding and use of IOW and IOA could be explored across disciplines and services. In addition to this the inclusion of trainee psychologists could provide important information regarding further differences in training.

It should be noted that only the current author conducted focus groups, which should have ensured that the quality of data generated was consistent. In endeavouring to make certain of this, a standardised interview schedule was also utilised. Despite this, within the focus groups, there were some confounding variables which may have impacted upon the data. For example, during two of the focus groups, participants had to leave prior to the end of the focus groups due to work commitments. Therefore, whether these focus groups truly captured all participants' views, in all areas in relation to IOW remains unknown. Following the first stages of analysis of the data, it did however appear that data saturation was reached and that no new codes/themes would have arisen with the addition of these participants' contributions. Another factor was that within three of the four focus groups that were conducted, the senior psychologist/service lead for that service was present. As well as it being noted that these

individuals often dominated much of the focus group discussion, it is possible that other clinicians may have felt unable to voice their opinions due to a fear of exposing their own vulnerabilities, or fears of looking incapable of carrying out their role effectively. Both of these factors were out of the author's control, but may have resulted in some aspects of the data being biased. The author attempted to minimise this by continuing to ask the questions in an open and inclusive way, in order to attempt to encourage all participants to provide their opinions and be transparent about their practice.

Finally, some questions may arise in regard to the treatment of data in this study. TA is content driven and was chosen by the author due to the fact that it is theoretically neutral. Despite this, there remains a risk that when applying codes to the data and when deriving the thematic hierarchy, the author could have imposed some subjective bias. It is possible that the themes derived from the data were biased by what the coder (the current author, a trainee forensic psychologist) already knew about the theoretical concept of conducting IOW/IOA, and the practical application of it in the role of a forensic psychologist. Such potential issues were however compensated for, through the development and use of a novel codebook, and also through completing IRR checks. Further to this guidelines for conducting TA were also strictly adhered to (Braun & Clarke, 2013).

Conclusions

Within forensic practice, there has been the development of empirically validated tools and the publication of specialised ethical guidelines (Committee on ethical guidelines for forensic psychologists, 1991; Committee on Professional Practice and Standards, 1994). Despite this, many researchers suggest that there remains considerable inconsistency in the quality of assessment practice (Borum & Grisso, 1996; Heilbrun & Collins, 1995; West &

Greenhall, 2011). Findings from the current study would support this. There is limited substantiated or regulatory guidance for many forensic professional activities, one of which would appear to be IOW. Whilst there are guidelines (RMA, 2006; DOH, 2008) which indicate the need for practitioners to conduct such work, there remains inconsistencies in practitioners' understanding and execution of such work.

The results of this study indicate that practitioners within forensic mental health settings do carry out some form of IOW, however this has the propensity to be different, as a result of a variety of factors. Blackburn (2004), emphasises that practitioners are ethically obliged to provide treatment of distress or disability, whether or not this is the cause of offending. It is likely that this blurs priorities within forensic mental health settings, as to what the primary focus of the work should be. As such, practitioners' understanding of what constitutes IOW is much broader than just an offence. It is evident from findings of this research that there is a lack of understanding across teams about the scope and purpose of IOW as well as what it constitutes. The processes utilised by practitioners as identified by this study, would suggest that practitioners possess good assessment and intervention skills, which are in line with models for working with offenders. However, the breadth and differing opinions as to what constitutes IOW would suggest that individuals and services do things differently. Whilst it may be difficult to define, IOW does need to be defined, otherwise this may result in failures of offender management due to important information regarding offences being lost. The necessity to complete individualised assessment, formulation and intervention was evident, thus making the development of specific guidelines for IOW difficult, however through not including specific formulations of an individual's offence, practitioners risk their assessments being fundamentally flawed. Therefore guidelines which assist in the formulation of this, which could contribute to the wider assessment, may be useful.

A key focus for future research should aim to explore IOW and IOA across other forensic settings, in order to provide further evidence of the similarities or differences in the assessment of offenders. Additionally, research which tests out the use of formal guidelines in relation to offence analysis, would provide evidence of its practical utility and specific information as to whether it would make the practice of assessing offenders more robust. There remains no formal definition in regard to IOW/ IOA and there remains a lack of research evidence regarding what it specifically involves. This research has been successful in broadening the research base and thus achieved some of the aims, however improvements in this area can still be made.

CHAPTER FIVE

Discussion and Conclusions

Discussion

This thesis aimed to examine different approaches to assessment utilised by clinicians in forensic practice. This was achieved through exploring and evaluating historical and widely used assessment methods, clinical judgement, actuarial assessment and SPJ approaches. In addition newer ideas for inclusion in the assessment of offenders (IOW/IOA) were also examined through an original research study.

Within the literature the importance of offender assessment and its impact on important decisions has been widely discussed (Andrews et al., 2006; Harkins et al., 2012). Clinicians working with offenders clearly have a legal and ethical responsibility to identify factors relating to offenders' criminal behaviour, that place them at future risk of reoffending (Sreenivasan, Kirkish, Garrick, Weinberger & Phenix, 2000). Risk assessment tools have been widely reviewed within the literature (Fazel et al., 2012; Singh et al., 2011; Yang et al., 2010). Despite this, questions remain regarding which approaches are most appropriate to use with different offender populations in forensic practice. Therefore, the current thesis intended to explore this in more detail, in order to understand which existing approaches and new concepts are most useful in increasing the accuracy and utility of offender assessment, allowing clinicians to carry out more robust assessments.

In order to consider the accuracy of clinical judgement and actuarial assessment approaches, which are widely used by practitioners to predict risk of recidivism within mentally disordered offending populations, a systematic literature review was conducted. Findings indicated that generally actuarial tools performed better than clinical judgement when predicting risk of recidivism and as such suggest that they are therefore most useful for clinicians when assessing risk. Despite this however, literature which assessed clinical

judgement still found it to be useful. Few of the studies found, directly compared clinical and actuarial approaches and therefore it is difficult to categorically conclude which approach is most useful, although overall findings would indicate that the integration of both approaches may be most successful in accurately assessing offenders. Practice reflects how ideas about risk and its assessment have changed and whilst there are now risk assessments which assist practitioners, it would seem that these have failed to have a significant impact on recidivism rates. Furthermore, predictive estimates (AUC's) only provide us with estimates above chance levels for entire samples. This then limits the utility of such actuarial measures, meaning that clinicians are still dependent on their clinical judgement, again highlighting how approaches in practice cannot be distinct.

As indicated by Chapter 2, using both clinical judgement and actuarial approaches in the assessment of risk, may be most useful. The SPJ (Hart et al., 2011) approach combines actuarial scales with structured professional judgement by clinicians. An example of a tool using this approach and widely adopted within forensic settings is the HCR-20 (Douglas & Reeves, 2010). As such it is evidently important to understand its reliability and validity as this impacts upon the clinical utility of the measure. Chapter 3, presented an overview of its most recent revision (version 3.0) and considered literature relating to its validity and reliability, in relation to its ability to predict long-term violence.

The newest version of the HCR-20 framework does meet criteria for having good psychometric properties to an extent. However, its validity and reliability is lacking in a number of areas, for example concurrent validity. At present much of the literature exploring this concept is conducted by the authors of the framework, and fails to compare it to other structured tools used to assess violence risk, such as the VRAG. It is therefore evident that further empirical research evidence is required, this should compare the newest version to other

measures of violence risk, and also validate it on other populations and forensic settings. As with Chapter 2, it may also be particularly useful to assess the tool's predictive accuracy with mentally disordered offenders, in order to assess its use specifically with this population. Furthermore, current research evidence indicates that the clinical scale is not that reliable, this is possibly due to the fact it fails to take into account the individual nature of disorders and their impact upon risk. The clinical utility of such a tool within forensic mental health settings is therefore questionable, given the complex nature of clients and therefore the individualised approach to assessment that is required, this does not lend itself to the use of a structured tool. Also, research indicates that the tool is often used improperly (Webster, Muller-Isberner & Fransson, 2002) and this could mean that it ignores risks related to other offence types or specific populations. Without further research however, it will continue to be used widely by practitioners, due to the fact that there is arguably nothing better at present.

Chapters 2 and 3 explored existing tools utilised by practitioners in their assessment of offenders. Both highlighted the limitations and strengths of such tools and also the role they play in assisting clinicians in accurately assessing offenders. It was considered important to build on this further and explore additional approaches proposed within the literature to be useful in enhancing clinicians' assessments of offenders. Some literature (Daffern et al., 2007; West & Greenhall, 201) indicates that IOA/IOW would enhance clinicians' assessment of offenders. Existing research also indicated that at present this is often not well understood or incorporated by the clinical team (Fallon, 2007, as cited in West & Greenhall, 2011). As something which has been evidenced to increase clinicians' understanding of an individual, their complex needs and also their risk and therefore have potentially important implications for forensic practice, it was deemed appropriate for further investigation through a research study.

A qualitative research study was carried out to explore clinicians' understanding and use of IOW/IOA in their clinical practice. A thematic analysis of practitioners' discussions was conducted and a variety of themes were extracted from the data. Findings were consistent with previous research with regard to the processes used by practitioners, for example, collaborative working (Glass & Arnholt, 2008) and responsivity (Andrews et al., 2011), and highlighted the centrality of IOW within the practice of assessment. The importance of assessment and formulation in particular, was also highlighted as a key process, something which has been widely discussed within the literature (Eells, 2007; Hanley et al., 2003).

Interestingly however, it was evident that a specific formulation of the index offence is not necessarily included in practitioners' assessments. As research in this area develops further, it may be useful for professional bodies to develop guidelines pertaining to this, which would assist its implementation by practitioners. Findings were consistent with research literature pertaining to 'what works' in offender assessment (GLM and RNR). The results argue against a 'one size fits all' approach to offender assessment, but did raise questions, regarding the consistency of processes utilised as part of IOW across different services in particular. Whilst practitioners do conduct some form of IOW within their practice, this did not appear to consistently include an analysis of an offender's index offence.

A salient finding was a lack of consistency in understanding of IOW and more specifically IOA. Whilst it was found that such work was viewed as central to practitioners' roles and to the assessment process, terminology and definitions used, were viewed as problematic. Finally, another key finding which possibly has broader implications for practice, was difficulty in forensic mental health settings of bringing together risk and mental health. This finding again supports the wider literature (Madden, 2008; Robertson et al., 2011), which suggests that the complex needs of a mentally disordered offender group mean that an

individualised approach to assessment and treatment is necessary, again in line with the 'what works' literature. It became evident however, that on occasion this could result in the index offence being lost completely and therefore presents wider issues in terms of risk management failures.

Implications for Practice

Based on findings from Chapters 2 and 3, this thesis supports the idea that tools utilised by practitioners are valid, reliable and have good predictive accuracy to a limited extent. Given these flaws, findings suggest that such tools should be used as part of a broader psychological assessment, which utilises a number of tools and/or psychometrics to inform a formulation and arrive at a judgement of risk. This may enable practitioners to make the most accurate and robust assessment of offenders, although further research is needed in order to further validate these conclusions. Previous research has suggested that low levels of predictive accuracy in offender recidivism is as a result of a lack of examination of the index offence (Lehman et al., 2013). One way of achieving such an examination would be for practitioners to include an analysis of the individual's offence within their assessments or as part of their IOW. Future research could look at the types of assessments and reports carried out by clinicians and the quality of these in order to explore this further.

Whilst formulation is acknowledged as being an important aspect of offender assessment (Hart et al., 2011), there remains at present no agreement on how it should be conducted. This coupled with inconsistencies in the understanding and interpretation of IOW/IOA across disciplines, professions, and settings, could result in practitioners' assessments having key components missing. It is possible that a lack of understanding could be as a result of a number of factors. Firstly, it may indicate that practitioners working within

forensic mental health settings are ill-equipped to conduct such work, this may be due to a lack of understanding and overemphasis on mental health needs. Training aimed at the wider multidisciplinary team may increase awareness of the importance of targeting risk, as well as mental health needs, during assessment and intervention. It is also possible that clinicians may not have direct access to the index offence information needed for such an analysis, or that there are gaps in the information that they have access to. With records held on offenders moving increasingly towards an electronic format, it is possible that this may be the case.

Findings from this thesis would support an argument that providing guidelines to assist clinicians in assessing offenders may be useful in assisting their practice. In particular in how to include IOW/IOA within this. What is evident from the findings of this thesis however, is that this would be difficult due to the complex nature of offenders and the need to be responsive to each individual's needs. This includes, the tools selected by practitioners to assess risk being appropriate to the client group. It is important that particularly in forensic mental health settings that mental health as opposed to risk does not become the focus of assessments.

In summary, without further research relating to the assessment processes used with offenders, in particular the use of tools for specific offender subgroups and the inclusion of IOW/IOA, this could result in the accuracy of clinicians assessments not being as robust and comprehensive as they could be. In the case of mental health settings, if priority is given to mental health over risk issues this could result in outstanding treatment needs. In the general assessment of offenders, it could also result in risks and needs not being accurately assessed and understood, resulting in offenders not undergoing appropriate treatment, therefore resulting in risk management failures. Offender assessment remains an important issue in forensic practice, as such more empirical research evidence is required, to further validate tools and assessment approaches and help further clinicians' understanding and use of them.

Limitations and Future Directions

There are some limitations from the outcomes of this thesis which should be considered. The first relates to the sample used within the research study and pertains to its small and highly selective nature. It is possible that as a result of it being relatively small, some factors relating to IOW/IOA remain unexplored. In addition the sample was highly selective and utilised strict inclusion criteria resulting in the inclusion of qualified clinical and forensic psychologists working within forensic mental health services only. It is possible therefore that findings cannot be generalised to psychologists working within other forensic settings. Therefore conclusions drawn may be specific to mental health settings and therefore suggests, future research relating to IOW/IOA, would benefit from comparing practitioners' understanding and use of IOW/IOA across settings. This would provide a more in-depth understanding of processes used and enable a clearer set of guidelines, terminology and definitions regarding the inclusion of such work in the assessment process.

The fact that Chapters 2 and 4 only consider forensic mental health settings and therefore the assessment of mentally disordered offenders however, is an overall limitation of the current thesis. As a result this limits the ability of findings and conclusions to be generalised to the assessment of offenders in general. Previous research has indicated (Lord & Perkins, 2014; Robertson et al., 2011), that mentally disordered offenders present with more complex problems and as such require different approaches to assessment and treatment. As such this could mean that findings from the research study in particular, would not be applicable to the assessment of offenders in other settings. Overall this thesis has evidenced that greater research is required, both regarding the efficacy of specific assessment tools to aid clinicians' in understanding what it most appropriate and accurate with mentally disordered offenders, and also regarding the offender assessment process, including the use of IOW/IOA in a variety of

other forensic settings. Such research would inevitably assist practitioners in conducting more robust assessments across offending populations, which would in turn result in more consistency across settings and ensure more effective offender management.

Conclusions

This thesis aimed to provide a greater insight and understanding regarding the clinical utility of historically used approaches and tools in the assessment of offenders and newer concepts which have been suggested within the literature. Through the completion of this work it was highlighted that there are mixed findings regarding appropriate assessment approaches in clinical practice. Despite this, findings evidenced that assessment tools currently utilised within forensic practice do have limitations, however at present they are all we have and as such, further empirical research evidence which assists in the development of better and more accurate tools is required. Outcomes of the systematic review indicate that overall, actuarial assessment approaches perform better in accurately predicting risk of recidivism as opposed to clinical judgement, however they also indicated that clinical judgement has an important role to play. Therefore SPJ tools such as that discussed in Chapter 3 may have the most clinical utility for offender assessment at present. It was also illustrated through the research study that clinicians take a responsive and individualised approach to offender assessment, however at present within forensic mental health settings, this does not always include a specific analysis of offence information.

Findings from this thesis may be extended in several ways through future research: firstly, research which directly compares assessment approaches and clinical tools used in forensic practice on a variety of offending populations, to enhance their clinical utility further; and secondly by carrying out further research into practitioners' understanding and use of

IOW/IOA across professional disciplines and settings. Research such as this would assist in generating a more robust and reliable understanding of offender assessment, it would allow for terminology to be clearly defined, and in the case of IOW/IOA in particular could assist in working towards clearer guidelines for the conduct of such work, making assessments more consistent. This would then further assist with the standardisation of processes across settings and help ensure more effective offender assessment and management.

References

Aeigisdottir, S., White, M.J., Spengler, P. M., Maugherman, A.S., Anderson, L.A., Cook, R.S. et al. (2006). The Meta-Analysis of Clinical Judgement Progect: Fifty-Six Years of Accumulated Research on Clinical Versus Statistical Prediction. *The Counselling Psychologist, 34,* 341-382. doi: 10.1177/0011000006295149

Alison, L.J., Goodwill, A.M., Almond, L., van den Heuvel, C., & Winter, J. (2010). Pragmatic solutions to offender profiling and behavioural investigative advice. *Legal and Criminological Psychology, 15,* 115–132. doi: 10.1348/135532509X46334

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

Andrews, D.A., Bonta., J., & Wormith, J.S. (2006). The recent past and near future of risk and/or need assessment. *Crime and Delinquency, 52(1),* 7-27. doi: 10.1177/0011128705281756

Andrews, D.A., Bonta, J., & Wormith, J.S. (2011). The risk-need-responsivity (RNR) model: Does adding the good lives model contribute to effective crime prevention? *Criminal Justice and Behaviour, 38, 7,* 735-755. doi: 10.1177/0093854811406356

Barbaree, H.E., Seto, M.C., Langton, C.M., & Peacock, E.J. (2001). Evaluating the predictive accuracy of six risk assessment instruments for adult sex offenders. *Criminal Justice and Behaviour, 28(4),* 490-521. doi: 10.1177/009385480102800406

Beech, A.R., Fisher, D.D., & Thornton, D. (2003). Risk assessment of sex offenders. *Professional Psychology: Research and Practice, 34, 4,* 339-352. doi: 10.1037/0735-7028.34.4.339

Belfrage, H. (1998). Implementing the HCR-20 scheme for risk assessment in a forensic psychiatric hospital: Integrating research and clinical practice. *The Journal of Forensic Psychiatry, 9, 2,* 328-338. doi: 10.1080/09585189808402200

Bengtson, S. (2008). Is newer better? A cross-validation of the Static-2002 and the Risk Matrix 2000 in a Danish sample of sexual offenders. *Psychology, Crime & Law, 14(2),* 85-106. doi: 10.1080/10683160701483104

Bengtson, S. & Langstrom, N. (2007). Unguided clinical and actuarial assessment of re-offending risk: A direct comparison with sex offenders in Denmark. *Sexual Abuse: Journal of Research and Treatment,19,* 135-153. doi: 10.1007/s11194-007-9044-5

Bird, C.M. (2005). How I stopped dreading and learned to love transcription. *Qualitative Research, 1,* 226-248. doi: 10.1177/1077800404273413

Bjorkly, S., Eidhammer, G., & Selmer, L.E. (2014). Concurrent validity and clinical utility of the HCR-20V3 compared with the HCR-20 in forensic mental health nursing: Similar tools but improved method. *Journal of Forensic Nursing, 10, 4,* 234-242. doi: 10.1097/JFN.0000000000000047

Blackburn, R. (2004). "What works" with mentally disordered offenders. *Psychology, Crime & Law, 10, (3),* 297-308. doi: 10.1080/10683160410001662780

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

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

Bonta, J., & Andrews, D.A. (2007). *Risk-need-responsivity model for offender assessment and rehabilitation.* Ottawa, Ontario: Public Safety Canada.

Bonta, J. & Andrews, D. (2010). Viewing offender assessment and rehabilitation through the lens of the risk-need-responsivity model. In McNiel, F., Raynor, P., & Trotter, C. (Eds.) *Offender supervision: New directions in theory, research and practice* (2nd ed.) (pp. 19-64). New York: Routledge.

Borum, R. (1996). Improving the clinical practice of violence risk assessment: Technology, guidelines, and training. *American Psychologist,* 945-956. doi: 10.1037/0003-066x.51.9.945

Borum, R., Fein, R., Vossekuil, B., & Berglund, J. (1999). Threat Assessment: Defining an Approach for Evaluating Risk of Targeted Violence. *Behavioural Sciences and the Law,* 17, 323-377. doi : 10.1002/1099-0798(199907/09)

Borum, R., & Grisso, T. (1996). Establishing standards for criminal forensic reports: An empirical analysis. *Bulletin of the American Academy of Psychiatry and Law, 24,* 297-317.

Braun, V. & Clarke, V. (2006). Using Thematic Analysis in Psychology. *Qualitative Research in Psychology, 2,* 77-101. doi: 10.1191/1478088706qp063oa

Braun, V. & Clarke, V. (2013). *Successful qualitative research a practical guide for beginners.* London: SAGE Publications, Ltd.

British Psychological Society. (December, 2011). *Good practice guidelines on the use of psychological formulation.* Retrieved from https://www.canterbury.ac.uk/.../DCP-Guidelines-for-Formulation.pdf

Brouillette-Alarie, S., & Proux, 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

Buchanan, A. (1999). Risk and dangerousness. *Psychological Medicine, 29(2),* 465-473. doi: 10.1017/s0033291798008101

Clark, D.A., Fisher, M.J. & McDougal, C. (1993). A New Methodology for Assessing the Level of Risk in Incarcerated Offenders. *British Journal of Criminology, 33,* 436-448. Retrieved from

http://bjc.oxfordjournals.org/content/33/3/436.full.pdf+html?hwshib2=authn%3A1441707505%3A20150907%253Ad414fa66-ad7b-47a1-ab93-20fe8a671127%3A0%3A0%3A0%3AwseUX8Noa4hv3c0xAiNxdw%3D%3D

Committee on Ethical Guidelines for Forensic Psychologists. (1991). Specialty guidelines for forensic psychologists. *Law and Human Behaviour, 15,* 655-665. doi: 10.1007/bf01065858

Committee on Professional Practice and Standards. (1994). Guidelines for child custody evaluations in divorce proceedings. *American Psychologist, 49,* 677-680. doi: 10.1037/e305272003-001

Convit, A., Jaeger, J., Lin, S.P., Meisner, M., & Volavka, J. (1988). Predicting assaultiveness in psychiatric inpatients: A pilot study. *Hospital and Community Psychiatry, 39,* 429-434. doi: 10.1176/ps.39.4.429

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

Davies, J.P., Heyman, B., Godin, P.M., Shaw, M.P., & Reynolds, L. (2006). The problems of offenders with mental disorders: A plurality of perspectives within a single mental health care organisation. *Social Science & Medicine, 63,* 1097−1108. doi: 10.1016/j.socscimed.2006.03.002

Davies, J., Jones, L. & Howells, K. (2010). Evaluating individual change. In M. Daffern, L. Jones, & J. Shine (Eds.), *Offence paralleling behaviour: A case formulation approach to offender assessment and intervention* (pp. 287–302). Chichester, UK: Wiley-Blackwell.

Dawes, R.M., Faust, D., & Meehl, P.E. (1989). Clinical versus actuarial judgment. *Science, 243,* 1668-1674. doi: 10.1126/science.2648573

Department of Health (2007). *Best practice in managing risk: Principles and evidence for best practice in the assessment and management of risk to self and others in mental health services.* London: Department of Health National Risk Management Programme. Retrieved from www.dh.gov.uk

Department of Health (2008). *Code of Practice: Mental Health Act 1983.* London: The Stationary Office. Retrieved from www.dh.gov.uk

de Vogel, V., van den Broek, E., & de Vries Robbe, M. (2014). The use of the HCR-20 V3 in Dutch forensic psychiatric practice. *International Journal of Forensic Mental Health, 13,* 109-121. doi: 10.1080/14999013.2014.906518

de Vogel, V., & de Ruiter. (2006). Structured professional judgement of violence risk in forensic clinical practice: A prospective study into the predictive validity of the Dutch HCR-20. *Psychology, Crime and Law, 12, 3,* 321-336. doi: 10.1080/10683160600569029

de Vries Robbe, M., & de Vogel, V. (2010). Pilot study of the HCR-20 (version 3). In K.S. Douglas, S.D. Hart, C.D. Webster, & H. Belfrage (Eds.). *HCR-20 V3 assessing risk for violence-use guide.* Burnaby, Canada: Mental Health, Law and Policy Institute, Simon Fraser University

Dobash, R.P., Dobash, R.E., Cavanagh, K., Smith, D. & Medina-Ariza, J. (2007). Onset of offending and life course among men convicted of murder. *Homicide Studies, 11,* 243–271. doi: 10.1177/1088767907306850

Dolan, M., & Doyle, M. (2000). Violence risk prediction: Clinical and actuarial measures and the role of the psychopathy checklist. *British Journal of Psychiatry, 177,* 303-311. doi: 10.1192/bjp.177.4.303

Douglas, K.S. (2008). *The HCR-20 violence risk assessment scheme.* In B. Cutler (Ed.), Encyclopaedia of psychology and law (pp.353-354). Thousand Oaks, CA: Sage.

Douglas, K.S. (2014). Version 3 of the historical-clinical-risk management-20 (HCR-20 V3): Relevance to risk assessment and management in forensic conditional release contexts. *Behavioural Sciences and the Law, 32,* 557-576. doi: 10.1002/bsl.2134

Douglas, K.S., & Belfrage, H. (2014). Interrater reliability and concurrent validity of the HCR-20 version 3. *International Journal of Forensic Mental Health, 13,* 130-139. doi: 10.1080/14999013.2014.908429

Douglas, K.S., Hart, S.D., Webster, C.D., & Belfrage, H. (2013). *HCR-20 V3: Assessing risk for violence-user guide*. Burnaby, Canada: Mental Health, Law and Policy Institute, Simon Fraser University.

Douglas, K.S., Hart, S.D, Webster, C.D., Belfrage, H., Guy, L. S., & Wilson, C.M. (2014). Historical-clinical-risk management-20, version 3 (HCR-20 V3): Development and overview. *International Journal of Forensic Mental Health, 13,* 93-108. doi: 10.1080/14999013.2014.906519

Douglas, K.S., Ogloff, J.R.P., & Hart, S.D. (2003). Evaluation of a model of violence risk assessment amongst forensic psychiatric patients. *Psychiatric Services, 54, 10,* 1372-1379. doi: 10.1176/appi.ps.54.10.1372

Douglas, K.S., & Reeves, K. (2010). The HCR-20 violence risk assessment scheme: Overview and review of the research. In R.K.Otto & K.S. Douglas (Eds.), *Handbook of violence risk assessment* (pp.147-185). New York: Routledge/Taylor & Francis Group.

Doyle, M., & Dolan, M. (2002). Violence risk assessment: Combining actuarial and clinical information to structure clinical judgements for the formulation and management of risk. *Journal of Psychiatric and Mental Health Nursing, 9,* 649–657. Retrieved from http://onlinelibrary.wiley.com/doi/10.1046/j.1365-2850.2002.00535.x/pdf

Doyle, M., Power, L.A., Coid, J., Kallis, C., Ullrich, S., Shaw, J. (2013). Predicting post-discharge community violence in England and Wales using the HCR-20 V3. *International Journal of Forensic Mental Health, 13,* 140-147. doi: 10.1080/14999013.2014.906517

Duggan, C. (2008). Focusing on treatment: The main interventions and their implications. In K. Soothill, P. Rogers, & M. Dolan (Eds.), *Handbook of forensic mental health* (pp. 64−84). UK: William Publishing.

Dunabar, E., Quinones, J., & Crevecoeur, D.A. (2005). Assessment of hate crime offenders: The role of bias intent in examining violence risk. *Journal of Forensic Psychology Practice, 5(1),* 1-19. doi: 10.1300/J158v05n01_01

Eells TD (2007) History and current status of psychotherapy case formulation. In Eells TD (ed.) *Handbook of Psychotherapy Case Formulation*. New York: Guilford.

Eidhammer, G., Selmer, L.E., & Bjorkly, S., and the SAFE Pilot Project Group. (2013). *A brief report of a comparison of the HCR-20 (version 2) and (HCR-20 (version 3) concerning internal consistency and clinical utility.* Centre for Research and Education in Forensic Psychiatry, Oslo University Hospital, Norway.

Elbogen E.B., Calkins, M.C., Scalora, M.J., & Tomkins, A.J. (2002). Perceived relevance of factors for violence risk assessment: a survey of clinicians. *International Journal of Forensic Mental Health Services, 1(1),* 37–46. doi: 10.1080/14999013.2002.10471159

Ennis B.J. & Emery R.D. (1978) *The Rights of Mental Patients.* Avon Books, New York.

Fallon, P., Bluglass, R., Edwards, B. & Daniels, G. (1999). *Report of the Committee of Inquiry into the Personality Disorder Unit, Ashworth Special Hospital* (Vol. 1). London: The Stationary Office. Retrieved from www.doh.gov.uk

Falzer, P.R. (2013). Valuing structured professional judgment: Predictive validity, decision-making, and the clinical-actuarial conflict. *Behavioral Sciences & the Law, 31(1),* 40-54. doi: 10.1002/bsl.2043

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters,* 27, 861–874. doi:10.1016/j.patrec.2005.10.010

Fazel, S., Singh, J.P., Doll, H., & Grann, M. (2012). Use of risk assessment instruments to predict violence and antisocial behaviour in 73 samples involving 24,827 people: Systematic review and meta-analysis. *BMJ: British Medical Journal, 345(7868),* 1–12. doi: 10.1136/bmj.e632

Ferguson,A.M., Ogloff, J.R.P., & Thomson, L. (2009). Predicting recidivism by mentally disordered offenders using the LSI-R:SV. *Criminal Justice and Behaviour, 36(1),* 5-20. doi: 10.1177/0093854808326525

Fleiss, J.L., Cohen, J., & Everitt, B.S. (1969). Large sample standard errors of kappa and weighted kappa. *Psychological Bulletin, 72(5),* 323-327. doi: 10.1037/h0028106

Fuller, J., & Cowan, J. (1999). Risk assessment in a multi-disciplinary forensic setting: Clinical judgement revisited. *The Journal of Forensic Psychiatry, 10(2),* 276-289. doi: 10.1080/09585189908403681

Gacono, C.B. (2002). Why there is a need for the personality assessment of offenders. *International Journal of Offender Therapy and Comparative Criminology, 46(3),* 271-273. doi: 10.1177/0306624X02463001

Gardner, W., Lidz, C.W., Mulvey, E.P. & Shae, E.C. (1996). A Comparison of Actuarial Methods for Identifying Repetitively Violent Patients with Mental Illnesses. *Law and Human Behavior,* 20, 35-48. DOI: 10.1007/BF01499131

Glass, C.R., & Arnholt, D.B. (2000). Consumer perspectives on helpful and hindering factors in mental health treatment. *Journal of Clinical Psychology, 56,* 1467–1480. doi: 10.1002/1097-4679(200011)56:11<1467::AID-JCLP9>3.0.CO;2-K

Grann, M., Belfrage, H., & Tengstrom, A. (2000). Predictive validity of the VRAG and the historical part of the HCR-20. *Criminal Justice and Behaviour, 27(1),* 97-114. doi: 10.1177/0093854800027001006

Green, T. (2008). Clinical assessment and formulation. In J.Houston & S. Galloway (Eds.), *Sexual offending and mental health: Multidisciplinary management in the community* (pp. 152-173. London: Jessica Kingsley Publishers.

Gresswell, D.M. & Hollin, C.R. (1992). Towards a new methodology for making sense of case material: An illustrative case involving attempted multiple murder. *Criminal Behaviour and Mental Health, 2,* 329-341. doi: 10.1007/1993-33899-001

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

Grove, W.M., Zald, D.H., Lebow, B.S., Snitz, B.E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analyses. *Psychological Assessment, 12, 1,* 19-30. doi: 10.1037//1040-3590.12.1.19

Guest, G., Bunce, A. & Johnson, L. (2006). How Many Interviews Are Enough?: An Experiment with Data Saturation and Variability. *Field Methods, 18,* 59-82. doi: 10.1177/1525822X05279903

Guest, G., MacQueen, K.M. & Namey, E.E. (2012). *Applied Thematic Analysis.* USA: SAGE Publications, Inc.

Halleck, S.L. (1987). *The Mentally Disordered Offender.* Washington, DC: American Psychiatric Press.

Hanley, G.P., Iwata, B.A., & McCord, B.E. (2003). Functional analysis of problem behaviour: A review. *Journal of Applied Behaviour Analysis, 36, 2,* 147-185. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1284431/pdf/12858983.pdf

Hanson, R.K., Helmus, L., & Thornton, D. (2010). Predicting recidivism amongst sexual offenders: A multi-site study of the Static-2002. *Law and Human Behaviour, 34,* 198-211. doi: 10.1007/s10 979-009-918 0-1

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. (1999). *Static-99: Improving actuarial risk assessments for sex offenders (User Report No.1999–02).* Ottawa: Department of the Solicitor General of Canada.

Hanson, R.K., & Thornton, D. (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

Hare, R.D. (1991). *The Hare Psychopathy Checklist - Revised (PCL-R)* (2nd Ed.). Toronto, Ontario, Canada: Multi-Health Systems.

Harkins, L., Ware, J., & Mann, R. (2012). Crime and Punishment: 'What Works'?. In Davies, G., & Beech, A.R. (Eds.). *Forensic Psychology, second edition* (pp.307-327). Chichester: Wiley.

Harry, B. (1992). Criminals' explanations of their criminal behaviour, part 1: The contribution of criminological variables. *Journal of Forensic Sciences, 37,* 1327-1333. doi: 10.1007/978-1-4614-5690-2_89

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

Harris, G.T., Rice, M.E., & Cormier, C.A. (2002). Prospective replication of the violence risk appraisal guide in predicting violent recidivism among forensic patients. *Law and Human Behaviour, 26(4),* 377-394. doi: 10.1023/a:1016347320889

Harris, G. T., Rice, M.E., Quinsey, V. L., Lalumiere, M.L., Boer, D., & Lang, C. (2003). A multi-site comparison of actuarial risk instruments for sex offenders. *Psychological Assessment, 15,* 413–425. doi: 10.1037/1040-3590.15.3.413

Hart, S.D. (1998). The role of psychopathy in assessing risk for violence: conceptual and methodological issues. *Legal and Criminological Psychology, 3,* 121–137. doi: 10.1111/j.2044-8333.1998.tb00354.x

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 the risk of sexual violence.* Burnaby, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University.

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

Hatcher, R. (2012). Risk assessment and offender programmes. In Davies, G., & Beech, A.R. (Eds.). *Forensic Psychology, second edition* (pp.307-327). Chichester: Wiley.

Hayles, M. (2006). Constructing safety: A collaborative approach to managing risk and responsibility. In Gorman K, Gregory M, Hayle M, Parton N (eds) *Constructive Work with Offenders* (pp 67-85). London: Jessica Kingsley Publishers.

Heilbrun, K. (2001). *Principles of forensic mental health assessment*. New York: Kluwer Academic/Plenum Publishers.

Heilbrun, K., & Collins, S. (1995). Evaluations of trial competency and mental state at the time of the offense: Report characteristics. *Professional Psychology: Research and Practice, 26,* 61-67**.** doi: 10.1037//0735-7028.26.1.61

Heilbrun, K., Rosenfeld, B., Warren, J., & Collins, S. (1994). The use of third-party information in forensic assessments: A two-state comparison. *Bulletin of American Academic Psychiatry Law, 22, 3,* 399-406. doi: 10.1002/0471264385.wei1105

Herman, J. L. (1990). Sex offenders: A feminist perspective. In W. L. Marshal, D. R. Laws & H. E. Barbaree (Eds.), *Handbook of sexual assault: Issues, theories, and treatment of the offender* (pp. 177–193). New York: Plenum Press.

Hill, L. (2009). *Investigation into the issues arising from the Serious Further Offence Review: Dano Sonnex [DS] (d.o.b. 07.07.85, age 23)*. London: National Offender Management Service. Retrieved from www.justice.gov.uk

Hilton, N.Z., Harris, G.T., & Rice, M.E. (2006). Sixty-Six Years of Research on the Clinical Versus Actuarial Prediction of Violence. *The Counselling Psychologist,* 34, 400-409. doi: 10.1177/0011000005285877

Hilton, N.Z., & Simmons, J. (2001). The influence of actuarial assessment in clinical judgements and tribunal decisions about mentally disordered offenders in maximum security. *Law and Human Behaviour, 25(4),* 393-408. doi: 10.1023/a:1010607719239

HM Inspectorate of Probation (2006). *An Independent Review of a Serious Further Offence case: Anthony Rice*. London: HM Inspectorate of Probation. Retrieved from www.inspectorates.homeoffice.gov.uk/hmiprobation

Howard, P., Francis, B., Soothill, K., & Humphreys, L. (2009). OGRS 3: The revised Offender Group Reconviction Scale (Ministry of Justice Research Summary 7/09). London, England: Ministry of Justice. Retrieved from http://www.justice.gov.uk/publications/docs/oasysresearch-summary-07-09-ii.pdf

Howells, K., Day, A., & Thomas-Peter, B. (2004). Changing violent behaviour: Forensic mental health and criminological models compared. *Journal of Forensic Psychiatry & Psychology, 15, 3,* 391-406. doi: 10.1080/14789940410001655907

Hurst, M. (2011). Professional judgement in the assessment of risk: Is there a role for systemic practice? *Journal of Family Therapy, 33(2),* 168-180. doi: 10.1111/j.1467-6427.2011.00533.x

Huss, M.T., & Zeiss, R.A. (2004). Clinical assessment of violence from inpatient records: A comparison of individual and aggregate decision making across risk strategies. *International Journal of Forensic Mental Health, 3(2),* 139-147. doi: 10.1080/14999013.2004.10471203

Ireland, J.L. (2009). 'Conducting individualised theory-driven assessments of violent offenders'. In J.L. Ireland., C.A. Ireland., & P. Birch (Eds.), *Violent and Sexual Offenders: Assessment, treatment and management.* Devon: Willan Publishing.

Ireland, J.L., Ireland, C.A., & Birch, P. (2009). *Violent and Sexual Offenders: Assessment, treatment and management.* Devon: Willan Publishing.

Janis, I. L. (1982). Groupthink: Psychological studies of policy decisions and fiascoes. Boston: Houghton Mifflin.

Jones, L.F. (2002a). An Individual Case Formulation approach to the assessment of motivation. In M.McMurran. (Ed.), *Motivating Offenders to Change*. Chichester: Wiley.

Jones, L. F. (2004). Offence paralleling behaviour (OPB) as a framework for assessment and interventions with offenders. In A. Needs & G. Towl (Eds.), *Applying psychology to forensic practice* (pp. 34–63). Oxford: Blackwell/British Psychological Society.

Jones, L., Daffern, M., & Shine, J. (2010). Summary and future directions. In M. Daffern, L. Jones, & J. Shine (Eds.), *Offence paralleling behaviour: A case formulation approach to offender assessment and intervention* (pp. 317–328). Chichester, UK: Wiley Blackwell.

Kemshall H. (1996) *Reviewing Risk: A Review of Research on the Assessment and Management of Risk and Dangerousness: Implications for Policy and Practice in the Probation Service*. Home Office, London.

Kline, P. (1986). *A handbook of test construction: introduction to psychometric design*. New York, NY, US: Methuen.

Kline, P. (1998). *The new psychometrics: science, psychology and measurement.* Florence, Kentucky, US: Taylor & Frances/Routledge.

Knauer, V. & Wilkinson-Tough, M. (In press). Submitted chapter to *Seminars in Practical Forensic Psychiatry* (2nd Edition).

Kotter, S., von Franque, F., Bolzmacher, M., Eucker, S., Holzinger, B., & Muller-Isberner, R. (2014). The HCR-20 V3 in Germany. *International Journal of Forensic Mental Health, 13,* 122-129. doi: 10.1080/14999013.2014.911784

Kroner, C., Stadtland, C., Eidt, M., & Nedopil, N. (2007). The validity of the violence risk appraisal guide (VRAG) in predicting criminal recidivism. *Criminal Behaviour and Mental Health, 17,* 89-100. doi: 10.1002/cbm.644

Kropp, P.R., & Hart, S.D. (2000). The spousal assault risk assessment (SARA) guide: Reliability and validity in adult male offenders. *Law and Human Behaviour, 24,* 101-118. doi: 10.1023/a:1005430904495

Lazarus, A.A. (1971). *Behaviour Therapy and Beyond.* McGraw-Hill, New York.

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

Lennings, C. (2005). Risk assessment in care and protection: The case for actuarial approaches. *AeJAMH (Australian e-Journal for the Advancement of Mental Health), 4(1),* 27-36. doi: 10.5172/jamh.4.1.27

Lidz, C., Mulvey, E., & Gardner, W. (1993). The accuracy of predictions of violence to others. *Journal of the American Medical Association*, 269, 1007-1011. doi: 10.1016/0736-4679(93)90331-z

Lindsay, W.R. & Beail, N. (2004). Risk assessment: Actuarial prediction and clinical judgment of offending incidents and behaviour for intellectual disability services. *Journal of Applied Research in Intellectual Disabilities*, 229-234. doi: 10.1111/j.1468-3148.2004.00212.x

Litwack, T.R. (1993). Comments on the ethics of dangerousness assessments. *Law and Human Behaviour, 17, 4,* 479-482. doi: 10.1007/bf01044380

Litwack, T.R., (2001). Actuarial versus clinical assessments of dangerousness. *Psychology, Public Policy and The Law*, 7, 409-433. doi: 10.1037/1076-8971.7.2.409

Logan, C. (2014). The HCR-20 version 3: A case study in risk formulation. *International Journal of Forensic Mental Health, 13,* 172-180. doi: 10.1080/14999013.2014.906516

Logan, C., & Johnstone, L. (2010). Personality disorder and violence: Making the link through risk formulation. *Journal of Personality Disorders, 24(5),* 610-633. doi: 10.1521/pedi.2010.24.5.610

Logan, C., Nathan, R., & Brown, A. (2011). Formulation in clinical risk assessment and management. In R.W. Whittington & C. Logan (Eds.), *Self-harm and Violence: Towards best practice in managing risk* (pp.187–204). Chichester, UK: Wiley-Blackwell. doi: 10.1002/9781119991175.ch10

Lord, A., & Perkins, D. (2014). Assessing and treating sexual offenders with mental disorders. *Journal of Forensic Practice, 16(2),* 94-109. doi: 10.1108/JFP-02-2013-0012

MacCulloch, M.J., Bailey, J., & Robinson, C. (1995). Mentally disordered attackers and killers: Towards a taxonomy. *Journal of Forensic Psychiatry, 6(1),* 41–61. doi: 10.1080/09585189508409875

Maden, T. (2008). The process and systems for adults. In K. Soothill, P. Rogers, & M. Dolan (Eds.), *Handbook of forensic mental health* (pp. 124−143). London, UK: Willan Publishing.

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

McGuire, J. (2012). Crime and Punishment: 'What Works'?. In Davies, G., & Beech, A.R. (Eds.). *Forensic Psychology, second edition* (pp.307-327). Chichester: Wiley.

McLellan, E., MacQueen, K.M. & Neidig, J.L. (2003). Beyond the Qualitative Interview: Data Preparation and Transcription. *Field Methods, 15,* 63-84.

McNiel, D.E., Sandberg, D.A., & Binder, R.L. (1998). The relationship between confidence and accuracy in clinical assessment of psychiatric patients' potential for violence. *Law and Human Behavior, 22,* 1998.

Meehl, P.E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. Minnepolis, US: Minnesota University Press.

Melton, G.B., Petrila, J., Poythress, N.G., & Slobogin, C. (1997). *Psychological evaluations for the courts: A handbook for mental health professionals and lawyers* (2nd ed.). New York: The Guilford Press.

Menzies, R.J., Webster, C.D., & Sepejak, D.S. (1985a). The dimensions of dangerousness: evaluating the accuracy of psychometric predictions of violence among forensic patients. *Law and Human Behavior,9,* 35-56. doi: 10.1007/bf01044289

Mills, J.F. (2005). Advances in the assessment and prediction of interpersonal violence. *Journal of Interpersonal Violence,* 20, 236-241. doi: 10.1177/0886260504267745

Ministry of Justice (2010). *Statistics of mentally disordered offenders 2008 England and Wales.* London: Ministry of Justice. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/218001/mentally-disordered-offenders-2008.pdf

Ministry of Justice (2013). *Proven re-offending statistics quarterly bulletin: October 2012 to September 2013, England and Wales*. Retrieved from https://www.gov.uk/government/collections/reoffending-statistics

Moore, E., & Drennan, G. (2013). Complex forensic case formulation in recovery-oriented services: Some implications for routine practice. *Criminal Behaviour and Mental Health, 23,* 230-240. doi: 10.1002/cbm.1885.

Monahan, J. (1984). The prediction of violent behaviour: toward a second generation of theory and policy. *American Journal of Psychiatry, 141,* 10–15. doi: 10.1176/ajp.141.1.10

Monahan, J. (1996). Violence prediction: The last 20 years and the next 20 years. *Criminal Justice and Behaviour*, 23, 107-120. doi: 10.1177/0093854896023001008

Monahan, J., Steadman, H.J., Applebaum, P.S., Robbins, P.C., Mulvey, E.P., Silver, E., Roth, L.H., & Grisso, T. (2000). Developing a clinically useful actuarial tool for assessing violence risk. *British Journal of Psychiatry, 176,* 312-319. doi: 10.1192/bjp.176.4.312

Morris, N., & Miller, M. (1985). Predictions of dangerousness. *Crime and Justice, 6,* 1-50. doi: 10.1086/449103

Mossman, D. (1994). Assessing predictions of violence: Being accurate about accuracy. *Journal of Consulting and Clinical Psychology*, 62, 783-792. doi: 10.1037//0022-006x.62.4.783

Murphy, G.H., & Clare, I.C.H. (1996). Analysis of motivation in people with mild learning disabilities (mental handicap) who set fires. *Psychology, Crime & Law, 2(3),* 153-164. doi: 10.1080/10683169608409774

Nicholson, R.A., & Norwood, S. The quality of forensic psychological assessments, reports and testimony: Acknowledging the gap between promise and practice. *Law and Human Behaviour, 24, 1,* 9-44. doi: 10.1023/A:1005422702678

Odeh, M.S., Zeiss, R.A., & Huss, M.T. (2006). Cues they use: Clinicians' endorsement of risk cues in predictions of dangerousness. *Behavioral Sciences & the Law,* 147-156. doi: 10.1002/bsl.672

Pedersen, L., Ramussen, K., & Elsass, P. (2012). HCR-20 violence risk assessments as a guide for treating and managing violence risk in a forensic psychiatric setting. Psychology, Crime and Law, *18, 8,* 733-743. doi: 10.1080/1068316X.2010.548814

Philipse, M.W.G., Koeter, M.W.J., van den Brink, W., & Van Der Staak, C.P.F. (2004). The structural coherence of clinically derived dynamic indicators of reoffending risk. *Criminal Behaviour and Mental Health, 14(4),* 263-279. doi: 10.1002/cbm.593

Polascheck, D.L.L. (2012). An appraisal of the risk-need-responsivity (RNR) model of offender rehabilitation and its application in correctional treatment. *Legal and Criminological Psychology, 17,* 1-17. doi: 10.1111/j.2044-8333.2011.02038

Pyszora, N.M., Barker, A.F., & Kopelman, M.D. (2003). Amnesia for Criminal Offences : A Study of Life Sentence Prisoners. *Journal of Forensic Psychiatry and Psychology, 14,* 475-490.

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

Quinsey, V.L., & Maguire, A. (1986). Maximum security psychiatric patients: Actuarial and clinical predictions of dangerousness. *Journal of Interpersonal Violence, 1,* 143-171. doi: 10.1177/088626086001002002

Reed, J. (1997). Risk assessment and clinical risk management: The lessons from recent inquiries. *British Journal of Psychiatry, 170,* 4-7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9328515

Rice, M.E., & Harris, G.T. (1997). The treatment of mentally disordered offenders. *Psychology, Public Policy and Law, 3(1),* 126-183. doi: 10.1037/1076-8971.3.1.126

Risk Management Authority (2006). *Standards and Guidelines for Risk Assessment, Version 1*. Retrieved from www.rmascotland.gov.uk

Robertson, P., Barnao, M., & Ward, T. (2011). Rehabilitation frameworks in forensic mental health. *Aggression and Violent Behaviour, 16,* 472-484. doi: 10.1016/j.avb.2011.03.003

Ross, D.J., Hart, S.D., & Webster, C.D. (1998). *Aggression in psychiatric patients: Using the HCR-20 to assess risk for violence in hospital and in the community.* Burnaby, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University.

Scott, P.D. (1977). Assessing dangerousness in criminals. *The British Journal of Psychiatry, 131(2),* 127-142. doi: 10.1192/bjp.131.2.127

Seto, M.C. (2005). Is more better? Combining actuarial risk scales to predict recidivism among adult sex offenders. *Psychological Assessment, 17,* 156-167. doi: 10.1037/1040-3590.17.2.156

Singh, J.P. (2013). Predictive validity performance indicators in violence risk assessment: A methodological primer. *Behavioural Sciences and the Law, 31,* 8-22. doi: 10.1002/bsl.2052

Singh, J.P. & Fazel, S. (2010). Forensic risk assessment: A metareview.*Criminal Justice and Behavior, 37(9),* 965-988. doi: 10.1177/0093854810374274

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,* 499-513. doi: doi:10.1016/j.cpr.2010.11.009

Smith, S.T., Kelley, S.E., Rulseh, A., Sorman, K., & Edens, J.F. (2014). Adapting the HCR-20 V3 for pre-trial settings. *International Journal of Forensic Mental Health, 13,* 160-171. doi: 10.1080/14999013.2014.906520

Smith, J.A. & Osborn, M. (2003). Interpretive Phenomenological Analysis. In J.A. Smith (Ed.), *Qualitative psychology: a practical guide to methods.* London: SAGE Publications, Ltd.

Snowden, R.J., Gray, N.S., Taylor, J., & MacCulloch, M.J. (2007). Actuarial prediction of violent recidivism in mentally disordered offenders. *Psychological Medicine, 37,* 1539-1549. doi: 10.1017/S0033291707000876

Spence, D.P. (1989). *Narrative truth and historical truth: Meaning and interpretation in psychoanalysis.* New York: Norton.

Spengler, P.M., White, M.J., Eigisdottir, S., Maugherman, A.S., Anderson, L.A., Cook, R.S. et al. (2009). The meta-analysis of clinical judgement project - Effect of judgement on accuracy. *The Counselling Psychologist, 37,* 350-399. doi: 10.1177/0011000006295149

Sreenivasan, S., Kirkish, P., Garrick, T., Weinberger, L.E., & Phenix, A. (2000). Actuarial risk asessment models: A review of critical issues related to violence and sex offender recidivism assessments. *The Journal of the American Academy of Psychiatry and the Law, 28,* 438-448. Retrieved from http://www.researchgate.net/profile/Linda_Weinberger/publication/12135329_Actuarial_risk_assessment_models_a_review_of_critical_issues_related_to_violence_and_sex-offender_recidivism_assessments/links/00b4953209fa6eae99000000.pdf

Strub, D.S., Douglas, K.S., & Nicholls, T.L. (2014). The validity of the version 3 of the HCR-20 violence risk assessment scheme amongst offenders and civil psychiatric patients. *International Journal of Forensic Mental Health, 13,* 148-159. doi: 10.1080/14999013.2014.911785

Sturmey, P., & McMurran, M. (Eds.). (2011). *Forensic case formulation.* Chichester, UK: Wiley-Blackwell.

Swanson, J.W., Swartz, M.S., Esscock, S.M., Osher, F.C., Wagner, H.R., Goodman, L.A., Rosenburg, S.D., & Meador, K.G. (2002). The social-environmental context of violent behaviour in persons treated for severe mental illness. *American Journal of Public Health, 92, 9,* 1523-1531. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1447272/

Swets, J.A., Dawes, R.M., & Monahan, J. (2000). Better decisions through science. *Scientific American, 283(4),* 82-87. doi:10.1038/scientificamerican1000-82

Tengstrom, A. (2001). Long-term predictive validity of the historical factors in two risk assessment instruments in a group of violent offenders with schizophrenia. *Nordic Journal of Psychiatry, 55(4),* 243-250. doi: 10.1080/080394801681019093

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

Thornton, D., Mann, R., Webster, S., Blud, L., Travers, R., Friendship, C., & Erikson, M. (2003). *Distinguishing and combining risks for sexual and violent recidivism*. In R. Prentky, E. Janus, M. Seto, & A. W. Burgess (Eds.), Annals of the New York Academy of Sciences: Vol. 989. Sexually coercive behavior: Understanding and management (pp. 225– 235). New York: New York Academy of Sciences.

Tversky, A., & Kahneman, D. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5,* 207-232. doi: 10.1037/e301722005-001

Vierra, A.J., & Garrett, J.M. (2005). Understanding interobserver agreement: The kappa statistic. *Family medicine, 37,* 360-363. Retrieved from: http://www.stfm.org/fmhub/fm2005/May/Anthony360.pdf

Ward, T., & Brown, M. (2004). The good lives model and conceptual issues in offender rehabilitation. *Psychology, Crime & Law, 10(3),* 243-257. doi: 10.1080/10683160410001662744

Ward, T., Mann, R.E., & Gannon, T.A. (2007). The good lives model of offender rehabilitation: Clinical implications. *Aggression and Violent Behaviour, 12,* 87-107. doi: 10.1016/j.avb.2006.03.004

Ward, T., & Salmon, K. (2009). The ethics of punishment: Correctional practice implications. *Aggression and Violent Behaviour, 14,* 239-247. doi: 10.1016/j.avb.2009.03.009

Ward, T., & Stewart, C. (2003). Criminogenic needs and human needs: A theoretical model. *Psychology, Crime & Law, 9,* 125-143. doi: 10.1080/1068316031000116247

Ward, T., Yates, P.M., & Willis, G.M. (2012). The good lives model and the risk need responsivity model a critical response to Andrews, Bonta and Wormith (2011). *Criminal Justice and Behaviour, 39(1),* 94-110. doi: 10.1177/0093854811426085

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

Webster, C.D., Eaves, D., Douglas, K.S., & Wintrup, A. (1995). *The HCR-20 scheme: The assessment of dangerousness and risk.* Burnaby, Canada: Simon Fraser University and Forensic Psychiatric Services Commission of British Columbia.

Webster, C.D., Hucker, S.J., & Bloom, H. (2002). Transcending the actuarial versus clinical polemic in assessing risk for violence. *Criminal Justice and Behaviour, 29,* 659-665. doi: 10.1177/009385402236736

Webster, C.D., Muller-Isberner, R., & Fransson, G. (2002). Violence risk assessment: Using structured clinical guidelines professionally. *International Journal of Forensic Mental Health, 1(2),* 185-193. doi: 10.1080/14999013.2002.10471173

West, A.G., & Greenhall, P.V. (2011). Incorporating index offence analysis into forensic clinical assessment. *Legal and Criminological Psychology, 16,* 144-159. doi: 10.1348/135532510X495124

West, A. (2000). Clinical assessment of homicide offenders: The significance of crime scene in offense and offender analysis. *Homicide Studies, 4,* 219-234. doi: 10.1177/1088767900004003002

Westrup, D., & Fremouw, W.J. (1998). Stalking behaviour: A literature review and suggested functional analytic assessment technology. *Aggression and Violent Behaviour, 3 (3),* 255-274. doi: 10.1016/S1359-1789(97)00023-2

Weist, J. (1981). Treatment of violent offenders. *Clinical Social Work Journal, 9(4),* 271-281. doi: 10.1007/BF00758209

Willig, C. (2001). *Introducing qualitative research in psychology: Adventures in theory and method*. Maidenhead: Mc Graw-Hill.

Wood, J.M., Garb, H.N., Lilienfield, S.O. & Nezworski, M.T. (2002). Clinical assessment. *Annual Review Psychology, 53,* 519-543. doi: 10.1146/annurev.psych.53.100901.135136

Yang, M., Wong, S.C.P., & Coid, J. (2010). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin, 136, 5,* 740-767. doi: 10.1037/a0020473

Zweig, M.H., & Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. *Clinical Chemistry, 39(4),* 561-577. Retrieved from http://www.clinchem.org/content/39/4/561.full.pdf+html

APPENDIX A. Search Syntax for Each Database.

Psych INFO (1967 to May Week 1 2015)

☐	# ▲	Searches	Results	Search Type	Actions
☐	1	exp risk assessment/ or exp evaluation/ or exp decision making/	158460	Advanced	Display More »
☐	2	actuarial assessment.mp.	69	Advanced	Display More »
☐	3	1 or 2	158484	Advanced	Display More »
☐	4	exp "clinical judgment (not diagnosis)"/ or exp psychological assessment/	39347	Advanced	Display More »
☐	5	(structured clinical judgement* or unstructured clinical judgement* or professional judgement* or clinical judgement*).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	458	Advanced	Display Delete More »
☐	6	(predict* adj2 recid*).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	814	Advanced	Display More »
☐	7	4 or 5 or 6	40487	Advanced	Display More »
☐	8	exp psychiatric patients/ or exp mental disorders/	488773	Advanced	Display More »
☐	9	exp mentally ill offenders/ or exp criminals/ or exp forensic evaluation/	19273	Advanced	Display More »
☐	10	forensic mental health.mp.	914	Advanced	Display More »
☐	11	(offend* or crim* or delinquen* or convict*).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	92091	Advanced	Display More »
☐	12	8 or 9 or 10 or 11	570744	Advanced	Display More »
☐	13	3 and 7 and 12	1198	Advanced	Display More »
☐	14	limit 13 to english	1140	Advanced	Display More »

Medline (1946 to May Week 1 2015)

#	Searches	Results	Search Type	Actions
1	exp risk assessment/ or exp evaluation/ or exp decision making/	158460	Advanced	Display More »
2	actuarial assessment.mp.	69	Advanced	Display More »
3	1 or 2	158484	Advanced	Display More »
4	exp "clinical judgment (not diagnosis)"/ or exp psychological assessment/	39347	Advanced	Display More »
5	(structured clinical judgement* or unstructured clinical judgement* or professional judgement* or clinical judgement*).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	458	Advanced	Display Delete More »
6	(predict* adj2 recid*).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	814	Advanced	Display More »
7	4 or 5 or 6	40487	Advanced	Display More »
8	exp psychiatric patients/ or exp mental disorders/	488773	Advanced	Display More »
9	exp mentally ill offenders/ or exp criminals/ or exp forensic evaluation/	19273	Advanced	Display More »
10	forensic mental health.mp.	914	Advanced	Display More »
11	(offend* or crim* or delinquen* or convict*).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	92091	Advanced	Display More »
12	8 or 9 or 10 or 11	570744	Advanced	Display More »
13	3 and 7 and 12	1198	Advanced	Display More »
14	limit 13 to english	1140	Advanced	Display More »

☐	15	exp Risk Assessment/	10357	Advanced	Display More »
☐	16	(actuarial assessment* or assessment* tool*).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	8415	Advanced	Display More »
☐	17	15 or 16	18286	Advanced	Display More »
☐	18	exp Decision Making/	72743	Advanced	Display More »
☐	19	(structured clinical judgement* or unstructured clinical judgement* or professional judgement* or clinical judgement*).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	458	Advanced	Display Delete More »
☐	20	(predict* adj2 recid*).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	814	Advanced	Display More »
☐	21	18 or 19 or 20	73935	Advanced	Display More »
☐	22	exp Mentally Ill Persons/	0	Advanced	Delete More »
☐	23	exp Mental Disorders/	469311	Advanced	Display More »
☐	24	exp Criminals/	16593	Advanced	Display More »
☐	25	(psychiatric patient* or mental* ill* offend* or forensic mental health or offend* or crim* or delinquen* or convict*).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	125259	Advanced	Display Delete More »
☐	26	22 or 23 or 24 or 25	572502	Advanced	Display More »
☐	27	17 and 21 and 26	356	Advanced	Display More »
☐	28	limit 27 to english	339	Advanced	Display More »

EMBASE (1974 to 2015 May 07)

# ▲	Searches	Results	Search Type	Actions
1	(offend* or crim* or delinquen* or convict*).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	92091	Advanced	Display More »
2	exp mental patient/	0	Advanced	Display More »
3	forensic mental health.mp.	914	Advanced	Display More »
4	mental* ill*.mp.	42892	Advanced	Display More »
5	1 or 2 or 3 or 4	129457	Advanced	Display More »
6	exp risk assessment/	10357	Advanced	Display More »
7	actuarial assessment*.mp.	83	Advanced	Display More »
8	6 or 7	10388	Advanced	Display More »
9	(structured clinical judgement* or unstructured clinical judgement* or professional judgement* or clinical judgement*).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	458	Advanced	Display Delete More »
10	exp decision making/	72743	Advanced	Display More »
11	9 or 10	73139	Advanced	Display More »
12	5 and 8 and 11	89	Advanced	Display More »
13	limit 12 to english	85	Advanced	Display More »

Web of Science (After 1987)

#4	**TOPIC:** (("risk assessment*" OR "actuarial assessment*" OR "assessment* tool*")) *AND* **TOPIC:** ((("structured clinical judgement*" OR "unstructured clinical judgement*" OR "professional judgement*" OR "clinical judgement*") OR (("predict*") NEAR/2 ("recid*")))) *AND* **TOPIC:** (("patient*" OR "mental* ill* offend*" OR "mental* ill*" OR "mental* disord*" OR "inpatient*" OR "forensic mental health" OR "offend*" OR "convict*" OR "crim*" OR "delinquen*")) Refined by: **LANGUAGES:** (ENGLISH) AND **DOCUMENT TYPES:** (ARTICLE OR REVIEW) *DocType=All document types; Language=All languages;*
#3	**TOPIC:** (("risk assessment*" OR "actuarial assessment*" OR "assessment* tool*")) *AND* **TOPIC:** ((("structured clinical judgement*" OR "unstructured clinical judgement*" OR "professional judgement*" OR "clinical judgement*") OR (("predict*") NEAR/2 ("recid*")))) *AND* **TOPIC:** (("patient*" OR "mental* ill* offend*" OR "mental* ill*" OR "mental* disord*" OR "inpatient*" OR "forensic mental health" OR "offend*" OR "convict*" OR "crim*" OR "delinquen*")) Refined by: **LANGUAGES:** (ENGLISH) *DocType=All document types; Language=All languages;*
#2	**TOPIC:** (("risk assessment*" OR "actuarial assessment*" OR "assessment* tool*")) *AND* **TOPIC:** ((("structured clinical judgement*" OR "unstructured clinical judgement*" OR "professional judgement*" OR "clinical judgement*") OR (("predict*") NEAR/2 ("recid*")))) *AND* **TOPIC:** (("patient*" OR "mental* ill* offend*" OR "mental* ill*" OR "mental* disord*" OR "inpatient*" OR "forensic mental health" OR "offend*" OR "convict*" OR "crim*" OR "delinquen*")) *DocType=All document types; Language=All languages;*
#1	**TOPIC:** (("risk assessment*" OR "actuarial assessment*" OR "assessment* tool*")) *DocType=All document types; Language=All languages;*

ProQuest (After 1987)

all((risk assessment* OR actuarial assessment* OR assessment* tool*)) AND all(structured clinical judgement* OR unstructured clinical judgement* OR professional judgement* OR clinical judgement* OR ((predict*) NEAR/2 (recid*))) AND all(patient* OR mental* ill* OR mental* disord* OR inpatient* OR psychiatric patient* OR mental* ill* offend* OR forensic mental health OR offend* OR convict* OR crim* OR delinquen*)

APPENDIX B. Table of Excluded Papers.

Ref ID	Study	P	I	C	O	Comments
6414	Allan, Dawson & Allan (2006): *Prediction of the risk of male sexual reoffending in Australia,*	N	Y	N/A	Y	
6490	Barbaree, Seto, Langton & Peacock (2001): *Evaluating the predictive accuracy of six risk assessment instruments for adult sexual offenders*	N	Y	N/A	Y	
3616	Barbaree, Langton & Peacock (2006): *The factor structure of static actuarial items: In relation to prediction*	N	Y	N/A	Y	
3619	Barabaree, Langton & Peacock (2006): *Different actuarial risk measures produce different risk ratings for sexual offenders*	N	Y	N/A	Y	

Ref ID	Study	P	I	C	O	Comments
3606	Bartosh, Garby, Lewis & Gray (2003): *Difference in the predictive validity of actuarial risk assessment in relation to sex offender type*	N	Y	N/A	N	
01	Beauregard & Mieczkowski (2009): *Testing the predictive utility of the STATIC-99: A Bayes analysis*	Y	Y	N/A	Y	Mixed sample, results for mentally disordered population not-distinguishable. Excluded during data extraction stage.
3405	Belfrage (1998): *Making risk predictions without an instrument*	Y	Y	N/A	N	

3658	Blasko, Jeglic & Mercado (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*	N	Y	N/A	Y	
5683	Canales, Campbell, Wei & Totten (2014): *Prediction of general and violent recidivism among mentally disordered adult offenders: Test of the level service/risk- need -responsivity instrument*	N	Y	N/A	Y	
5769	Caudy, Durso & Taxman (2013): *How well do dynamic needs predict recidivism? Implications for risk assessment and risk reduction*	N	Y	N/A	Y	
6	Coulson, G., Ilacqua, G., Nutbrown, V., Giulekas, D., & Cudjoe, F. (1996): *Predictive utility of the LSI for incarcerated female offenders.*	N	Y	N/A	Y	Not clear if population is mentally disordered, excluded during data extraction stage.
7	Craig, L.A., Beech, A., & Browne, K.D. (2006): *Cross-validation of the Risk Matrix 2000 sexual and violent scales.*	N	Y	N/A	Y	Incorrect population, deleted during data extraction stage.

5341	de Vogel & de Ruiter (2004): *Differences between clinicians and researchers in assessing risk of violence in forensic psychiatric patients*	Y	Y	N/A	N	
5016	de Vogel, Ruiter, Hildebrand, Bos & van de Ven (2004): *Type of discharge and risk of recidivism measured by the HCR-20: A retrospective study in a Dutch sample of treated forensic psychiatric patients*	Y	N	N/A	Y	
9	de Vogel, V., de Ruiter, C., van Beck, D., & Mead, G. (2004): *Predictive validity of the SVR-20 and Static-99.*	Y	N	N/A	Y	Compares actuarial with SPJ approach, excluded during data extraction stage.
10	Desmerais, S.L., Nicholls, T.L., Don Read, J. & Brink, J. (2010): *Confidence and accuracy in assessments of short-term risks presented by forensic psychiatric population.*	Y	N	N/A	Y	Looks at an SPJ assessment, excluded during data extraction stage.
3965	Dickinson, Polaschek & Casey (2013): *Can the quality of high-risk violent prisoners' release plans predict recidivism following intensive rehabilitation? A comparison with risk assessment instruments*	N	Y	N/A	Y	

12	Douglas, K.S. & Ogloff, J.R.P. (2003): *The impact of confidence on the accuracy of structured professional and actuarial violence risk judgements in a sample of forensic psychiatric patients*	Y	N	N/A	Y	Compares actuarial to SPJ, excluded during data extraction stage.
11	Doyle, M. & Dolan, M. (2006): *Predicting community violence from patients discharged form mental health services*	Y	N	N/A	Y	Looks at SPJ assessment, excluded during data extraction stage
3673	Edwards & Grace (2014): *The development of an actuarial model for arson recidivism*	N	N	N/A	Y	
5035	Eke, Hilton, Meloy, Mohandie & Williams (2011): *Predictors of recidivism by stalkers: a nine-year follow-up of police contacts*	N	N	N/A	Y	
5427	Elbogen, Williams, Kim, Tomkins, & Scalora (2001): *Gender and perceptions of dangerousness in civil psychiatric patients*	Y	Y	N/A	N	
6006	Endrass, Urbaniok, Held, Vetter & Rossegger (2009): *Accuracy of the static-99 in predicting recidivism in Switzerland*	N	Y	N/A	Y	

7361	Gentry, Dulmus & Therio (2005): *Comparing sex offender risk classifications using the static-99 and LSI-R assessment instruments*	N	Y	N/A	Y
4317	Gerace, Curren & Muir-Cocchrane (2013): *Multidisciplinary health professionals' assessments of risk: How are tools used to reach consensus about risk assessment and management?*	N	N	N/A	N
4104	Girard & Wormith (2004): *The predictive validity of the level of service inventory-Ontario revision on general and violent recidivism among various offender groups*	N	Y	N/A	Y
4013	Glover, Nicholson, Hemmati, Bernfeld & Quinsey (2002): *Comparison of predictors of general and violent recidivism among high-risk federal offenders*	N	Y	N/A	Y
4638	Green & Baglioni (1997): *Judging the suitability for release of patients from a maximum security hospital by hospital and community staff*	Y	Y	N/A	N
7220	Grubin (2011): *A large-scale evaluation of risk matrix 2000 in Scotland*	N	Y	N/A	Y
4491	Hanson (2006): *Does static-99 predict recidivism among older sexual offenders?*	N	Y	N/A	N

5372	Harris, Rice, Quinsey, Lalumiere, Boer & Lang (2003): *A multisite comparison of actuarial risk assessments for sex offenders*	N	Y	N/A	Y
5903	Hilton, Harris, Popham & Lang (2010): *Risk assessment among incarcerated male domestic violence offenders*	N	Y	N/A	Y
5972	Hilton, Harris, Rice, Houghton & Eke (2008): *An indepth actuarial assessment for wife assault recidivism: The domestic violence risk appraisal guide*	N	Y	N/A	Y
3837	Hilton & Harris (2009): *How nonrecidivism affects predictive accuracy : Evidence from a cross-validation of the Ontario Domestic Assault Risk Assessment (ODARA)*	N	Y	N/A	Y
3715	Holsinger, Lowenkamp & Latessa (2006): *Exploring the validity of the level of service inventory-revised with native American offenders*	N	Y	N/A	Y
4010	Hood, Shute, Feilzer & Wilcox (2002): *Sex offenders emerging from long-term imprisonment : A study of their long-term reconviction rates and of parole board members judgements of their risk*	N	Y	N/A	Y
4285	Hsu, Caputi & Byrne (2009): *The level of service inventory-revised (LSI-R)*	N	Y	N/A	Y

7176	Jung, Pham & Ennis (2013): *Measuring the disparity of categorical risk among various sex offender risk assessment measures*	Y	Y	N/A	N
3917	Kingston, Yates, Firestone, Babchishin & Bradford (2008): *Long-term predictive validity of the risk-matrix 2000: A comparison with the static-99 and the sex offender risk appraisal guide*	N	Y	N/A	Y
3973	Labrecque, Smith, Lovins & Latessa (2014): *The importance of reassessment : How changes in the LSI-R risk score can improve the prediction of recidivism*	N	Y	N/A	Y
5263	Langevin (2006): *An actuarial study of recidivism risk among sex killers of adults and children: Could we have identified them before it was too late?*	N	Y	N/A	Y
6101	Langton et al (2007): *Actuarial assessment of risk for re-offense among adult sex offenders: Evaluating the predictive accuracy of the static-2002 and five other instruments*	N	Y	N/A	Y
4241	Langton, Barbaree, Hansen, Harkins & Peacock (2007): *Reliability and validity of the static-2002 among adult sexual offenders with reference to treatment status*	N	Y	N/A	Y
4683	Lindsay et al (2008): *Risk assessment in offenders with intellectual disability*	N	Y	N/A	Y

5183	Looman (2006): *Comparison of two risk assessment instruments for sexual offenders*	N	Y	N/A	Y
5074	Looman & Abracen (2010): *Comparison of measures of risk for recidivism in sexual offenders*	N	Y	N/A	Y
4115	Louw, Strydom & Esterhuyse (2005): *Prediction of violent behaviour: Professionals appraisal*	N	Y	N/A	N
5089	Lyall & Bartlett (2010): *Decision making in medium security: Can he have leave?*	Y	Y	N/A	N
4439	Manchak, Skeem & Douglas (2008): *Utility of the revised level of service inventory (LSI-R) in predicting recidivism after long-term incarceration*	N	Y	N/A	Y
4251	Manchak, Skeem, Douglas & Siranosian (2009): *Does gender moderate the predictive utility of the level of service inventory-revised (LSI-R) for serious violent offenders?*	N	Y	N/A	Y
3668	Mills & Gray (2013): *Two-tiered violence risk estimates: A validation study of an integrated-actuarial risk assessment instrument*	N	Y	N/A	Y
4588	Nunes, Firestone, Bradford, Greenberg & Broom (2002): *A comparison of modified versions of the static-99 and the sex offender risk appraisal guide*	N	Y	N/A	Y

5877	Oleson, VanBenschoten, Robinson & Lownekamp (2011): *Training to see risk: Measuring the accuracy of clinical and actuarial assessments among federal probation officers*	N	Y	N/A	Y	
3854	Parent, Guay & Knight (2011): *An assessment of long-term risk of recidivism by adult sexual offenders : one size doesn't fit all*	N	Y	N/A	Y	
3664	Rettenberger & Eher (2013): *Actuarial risk assessment in sexually motivated intimate-partner violence*	N	Y	N/A	Y	
3650	Rettenberger, Matthes, Boer & Eher (2010): *Prospective actuarial risk assessment: A comparison of five risk assessment instruments in different sexual offender subtypes*	N	Y	N/A	Y	
24	Rice, M.E., Harris, G.T., & Lang, C. (2013): *Validation of and revision to the VRAG and SORAG: The violence risk appraisal guide- revised (VRAG-R).*	Y	Y	N/A	Y	Cannot distinguish mentally disordered sample from others, excluded during data extraction stage.

5413	Seifert, Jahn, Bolten & Wirtz (2002): *Prediction of dangerousness in mentally disordered offenders in Germany*	Y	Y	N/A	N
5279	Seifert, Moller-Mussavi & Wirtz (2005): *Risk assessment of sexual offenders in German forensic institutions*	Y	Y	N/A	N
6223	Seto (2005): *Is more better? Combining actuarial risk scales to predict recidivism among adult sexual offenders*	N	Y	N/A	Y
3833	Smallbone & Wortley (2008): *Criterion and predictive validity of the static-99 for adult males convicted of sexual offences against children*	N	Y	N/A	Y
5691	Smid, Kamphuis, Wever & Van Beek (2014): *A comparison of the predictive properties of nine sex offender risk assessment instruments*	N	Y	N/A	Y
7337	Sreenivasan et al (2007): *Predicting the likelihood of future sexual recidivism: Pilot study findings from a California sex offender risk project and cross-validation of the static-99*	N	Y	N/A	Y
3709	Stadtland et al (2005): *Risk assessment and prediction of violent and sexual recidivism in sex offenders: long-term predictive validity of four risk assessment instruments*	N	Y	N/A	Y

3929	Stalans, Hacker & Talbot (2010): *Comparing nonviolent, other-violent and domestic batterer sex offenders. Predictive accuracy of risk assessments on sexual recidivism*	N	Y	N/A	Y
5756	Varela, Boccaccini, Murrie, Caperton & Gonzalez (2013): *Do the static-99and static-99r perform similarly for white, black and latino sexual offenders?*	N	Y	N/A	Y
3852	Wakeling, Howard & Barnett (2011): *Comparing the validity of the rm-2000 scales and OGRS3 for predicting recidivism by internet sexual offenders*	N	Y	N/A	Y
3240	Wilcox, Beech, Markall & Blacker (2009): *Actuarial risk assessment and recidivism in a sample of UK intellectually disabled sexual offenders*	N	Y	N/A	Y
Google 2	Hilton et al (2004): *A brief actuarial assessment for the prediction of wife assault recidivism: The Ontario Domestic Assault risk assessment*	N	Y	N/A	Y
Experts 1	McKee, Harris & Rice (2007): *Improving forensic tribunal decision: The role of the clinician*	Y	Y	N/A	N
Experts 2	McMillan, Hastings & Coldwell (2004): *Clinical and actuarial prediction of physical violence in a forensic intellectual disability hospital: A longitudinal study*	N	Y	N/A	Y

Final papers for review 5	Fuller & Cowan (1999): *Risk assessment in a multi-disciplinary forensic setting: Clinical judgement revisited*	Y	Y	N/A	Y	Excluded due to QAS of 54%, did not meet 60% cut-off for final inclusion and review.
Final papers for review 9	Hilton & Simmons (2001): *The influence of actuarial risk assessment on clinical judgements and tribunal decisions about mentally disordered offenders in maximum security*	Y	Y	N/A	Y	Excluded due to QAS of 54%, did not meet 60% cut-off for final inclusion and review.

APPENDIX C. Quality Assessment Checklist.

Cohort Quality Assessment.

Question	Yes (2)	Unclear (1)	No (0)	Comments
Screening Questions				
Does the study address a clearly focused question?				
Were appropriate methods used to answer the research question?				
Additional Questions				
Was the cohort recruited in an acceptable way? e.g. representative of a mentally disordered offending population				
Was exposure to clinical judgement or actuarial assessment accurately measured?				
Was recidivism defined as an outcome?				
Was the level of recidivism accurately measured?				
Have confounding variables been considered?				
Were confounding variables accounted for in the design and/or analysis?				
Was the follow-up period sufficient?				
Results				

Are the results clearly stated?				
Are the results significant?				
Applicability of Findings				
Are the results from this study transferable? (High internal and external validity?)				
Have limitations been discussed?				
COLUMN TOTAL				
TOTAL SCORE	**/26**			
PERCENTAGE				

Case-control Quality Assessment.

Question	Yes (2)	Unclear (1)	No (0)	Comments
Screening Questions				
Does the study address a clearly focussed research question?				
Were appropriate methods used to answer the research question?				
Selection Bias				
Were the cases recruited in an acceptable way?				
Were the controls accepted in an acceptable way?				
Are sample demographics described clearly and comprehensively?				
Are the case and control groups comparable in respect of confounding variables?				
Have potential confounding variable been adequately accounted for in design and/or analysis?				
Measurement & Outcome Bias				
Were assessment methods used the same across cohorts? e.g. clinical judgement and/or actuarial assessment				
Were the outcomes selected comparable to those seen in other studies?				
Is recidivism clearly defined as an outcome?				
Attrition Bias				
Were the participants blind to the research and outcomes?				
Were the cohorts followed up for an appropriate length of time?				
Statistics				
Was the statistical analysis used appropriate?				
COLUMN TOTAL				

TOTAL SCORE	**/26**
PERCENTAGE	

APPENDIX D. Data Extraction Forms

Author			
Title			
Date of Publication			
Name of Publication Source (Journal)			
Eligibility of Study		**Y**	**N**
P	Males Females Male & Female Mixed Population Age 18 to 64 Offenders Mentally Disordered Offenders Mental Health/ Psychiatric Population Population in secure forensic setting		
I	Risk Assessment/ Actuarial Assessment Clinical Judgement/ Professional Judgement Risk Assessment vs. Clinical Judgement		
C	N/A		
O	Recidivism/ Risk Reduction/ Repeat Offending Validity/ utility of assessment method Recidivism and Validity of Assessment method		
Continue to next stage?		**Y**	**N**
Methodology			
Research Question			
Study Design			
Recruitment Process			

Participant Characteristics	
Sample Size	
Outcomes Measured	
Standardised Measures Used	
Statistical Analysis	
Statistical Test Used	
Were confounding variables assessed/controlled for?	
Results	
What were the results?	
What were the conclusions drawn?	
Limitations of the study	
Strengths of the Study	
Applicability of findings	

Quality Rating Score	

APPENDIX E. Reference List of Included Studies.

Bengtson, S. (2008). Is newer better? A cross-validation of the Static-2002 and the Risk Matrix 2000 in a Danish sample of sexual offenders. *Psychology, Crime & Law, 14(2),* 85-106.

Bengtson, S. & Langstrom, N. (2007). Unguided clinical and actuarial assessment of re-offending risk: A direct comparison with sex offenders in Denmark. *Sexual Abuse: Journal of Research and Treatment,19,* 135-153. doi: 10.1007/s11194-007-9044-5.

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

Ferguson,A.M., Ogloff, J.R.P., & Thomson, L. (2009). Predicting recidivism by mentally disordered offenders using the LSI-R:SV. *Criminal Justice and Behaviour, 36(1),* 5-20. doi: 10.1177/0093854808326525.

Grann, M., Belfrage, H., & Tengstrom, A. (2000). Predictive validity of the VRAG and the historical part of the HCR-20. *Criminal Justice and Behaviour, 27(1),* 97-114.

Hanson, R.K., Helmus, L., & Thornton, D. (2010). Predicting recidivism amongst sexual offenders: A multi-site study of the Static-2002. *Law and Human Behaviour, 34,* 198-211. doi: 10.1007/s10 979-009-918 0-1.

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

Harris, G.T., Rice, M.E., & Cormier, C.A. (2002). Prospective replication of the violence risk appraisal guide in predicting violent recidivism among forensic patients. *Law and Human Behaviour, 26(4),* 377-394.

Huss, M.T., & Zeiss, R.A. (2004). Clinical assessment of violence from inpatient records: A comparison of individual and aggregate decision making across risk strategies. *International Journal of Forensic Mental Health, 3(2),* 139-147.

Kroner, C., Stadtland, C., Eidt, M., & Nedopil, N. (2007). The validity of the violence risk appraisal guide (VRAG) in predicting criminal recidivism. *Criminal Behaviour and Mental Health, 17,* 89-100. doi: 10.1002/cbm.644.

McNiel, D. E., Sandberg, D. A., & Binder, R. L. (1998). The relationship between confidence and accuracy in clinical assessment of psychiatric patients' potential for violence. *Law and Human Behavior, 22,* 1998.

Monahan, J., Steadman, H.J., Applebaum, P.S., Robbins, P.C., Mulvey, E.P., Silver, E., Roth, L.H., & Grisso, T. (2000). Developing a clinically useful actuarial tool for assessing violence risk. *British Journal of Psychiatry, 176,* 312-319.

Odeh, M. S., Zeiss, R. A., & Huss, M. T. (2006). Cues They Use: Clinicians' Endorsement of Risk Cues in Predictions of Dangerousness. *Behavioral Sciences & the Law, 24,* 147-156. doi: 10.1002/bsl.672.

Snowden, R.J., Gray, N.S., Taylor, J., & MacCulloch, M.J. (2007). Actuarial prediction of violent recidivism in mentally disordered offenders. *Psychological Medicine, 37,* 1539-1549.

Tengstrom, A. (2001). Long-term predictive validity of the historical factors in two risk assessment instruments in a group of violent offenders with schizophrenia. *Nordic Journal of Psychiatry, 55(4),* 243-250.

APPENDIX F. Full Explanation of Analytic Approach

There are different types of Thematic Analysis (TA) that can be conducted, as suggested within the research literature, Braun and Clarke (2006) suggest that an inductive or deductive approach can be used. Inductive TA generates the analysis from the bottom (the data) up and is not usually shaped by existing theory. Braun and Clarke (2013) however suggest that to some extent it is shaped by the researcher's standpoint, disciplinary knowledge and epistemology. In the current research study the participants were considered to be 'experts' in the field of the assessment of offenders, due to the fact that they were qualified clinical and forensic psychologists. It is likely that this inevitably shaped the data to some extent. In addition the author was also a trainee forensic psychologist and as such has experience of conducting IOW, this may have also biased the collection of data to an extent. To ensure that this was not the case, the author did not take an active role in the focus groups. For the purpose of this study an inductive approach was used, this means the themes that the analysis generated are strongly linked to the data itself. This was deemed most appropriate by the author as it does not attempt to fit the codes into a pre-existing framework and as there is little research in this area this made most sense. This method was selected over a deductive approach (also known as confirmatory) because a deductive approach is guided by existing theory and theoretical concepts and is not exploratory. As this research study was inherently exploratory, aiming to explore factors relating to IOW rather than confirm or explain pre-existing theories in relation to this topic area, an inductive approach was considered by the author as being most appropriate. Inductive TA is not hypothesis-driven, the researcher is required to read and reread the data looking for themes or ideas within the data which assist in answering the research question, and this is the analytic approach which was utilised.

One of the main strengths of using this approach to qualitative analysis is that it is extremely flexible. As Braun and Clarke (2013) suggest, it can be used to answer almost any type of research question, with the exception of language practices. In addition it can be used with a range of sizes of datasets (although these are usually small) and the sample size used in the current research is considered to be sufficient for the use of this type of qualitative analysis. It was considered the most appropriate method for this research study as the author attempted to explore and resolve issues surrounding difficulties in clinicians' everyday practice. In addition Braun and Clarke (2006) highlight that it can be useful for producing qualitative analyses suited to informing policy development (p.97). This is therefore relevant to the current research which intends to assist the research base in understanding how IOW is conducted and guide future research and everyday practice relating to working towards the development of a policy or protocol for IOW.

There is evidence that some qualitative researchers believe TA be a limited approach to use for qualitative analysis (Braun & Clarke, 2006; Braun & Clarke, 2013). They highlight that it is 'something and nothing' and lacks the substance of other approaches such as Grounded Theory (GT) and Interpretive Phenomenological Analysis (IPA). The literature suggests that one of the benefits of using thematic analysis is that it is a flexible, useful research tool, which can potentially provide rich and detailed, yet complex data (Braun & Clarke, 2006). Its flexibility makes it particularly appealing for use in this context as it is an area that is not widely researched and it is largely unknown what findings will be. TA was chosen over IPA and GT by the current author because unlike these analytic methods TA is not theoretically bounded. IPA is related to phenomenological epistemology (Smith & Osborn, 2003), and is related to understanding peoples everyday experience of reality in great deal. In contrast the goal of GT is to generate a plausible theory of the phenomena that is being researched. The current research

aims are looking at neither of these in relation to IOW and therefore TA was deemed by the author to be the most applicable. Both GT and IPA along with TA share a search for themes and patterns across a data set and therefore to some extent they overlap. However, as highlighted by Braun and Clarke (2006), TA does not require the theoretical and technological knowledge of an approach such as GT and therefore offers a more accessible form of analysis.

For the purpose of this research study it was deemed by the author that TA would provide a rich thematic description of the entire data set, this will provided a clear sense of the predominant and salient themes that arose from the focus groups. Therefore the themes that were identified, coded and analysed are an accurate indication of the content of the entire data set, in the case of this study the understanding and use of IOW. In the context of thematic analysis, a theme is something that captures an important component about the data set in relation to the broader research question and it represents some form of patterned response by participants or meaning. The author is aware that this may have resulted in some depth and complexity in relation to the analysis being lost, however it was decided that it would ensure a rich description of the data is maintained. Braun & Clarke (2006), state that this is a useful method when an under-researched area is being investigated or when participants' views on a topic are not known (p.11), as was the case in the current research study. As identified by Guest et al (2012) TA involves the use of systematic guidelines to conduct the analysis, Braun and Clarke (2013) in their book outline seven individual steps in TA, and this is the process that the current author used in the analysis of data. This process was chosen as it assisted the author in clearly setting out the different stages of analysis and also ensured the method of TA was being followed correctly, reducing researcher bias and adding to the validity of the research. TA met the current studies aims as it allows for an essentialist interpretation of the data (Braun & Clarke, 2013; Strauss & Corbin, 1998), a more descriptive form of analysis which has allowed the

author to tell the story of the data. This was opposed to a more constructionist interpretation of the data which provides a more interpretive view of the meaning of data, such as that found in IPA.

APPENDIX G. Research Summary

Title: An exploration of practitioners' understanding and use of index offence "work".

Chief Investigator: Chloe Whatson – Trainee Forensic Psychologist (University of Birmingham)

Supervisors: Dr Jessica Woodhams (University of Birmingham); Dr Megan Wilkinson Tough (Fromeside Medium Secure Unit)

Rationale: Often in forensic secure settings, as part of the assessment procedure, practitioners are asked to carry out 'index offence work', with offenders or patients. However, there is no standardised procedure or definition for this and therefore what this actually means is not clear. It is thus understandable that practitioners working within these settings may interpret this work in different ways, resulting in different procedures and therefore outcomes. West and Greenhall (2011) suggest that index offence work should be a core task of any forensic clinician engaged in the assessment of offenders/patients. It makes sense that working with an offender in regard to their index offence could provide a better understanding of crime scene actions and offence motivations. Thus, in turn, this could help guide treatment planning and improve risk assessments for individual patients/offenders. Despite this, at present, the research base in this area is very limited and there is no standardised approach to aid practitioners in this work. Therefore the proposed research will aim, with the assistance of qualified clinical and forensic psychologists working within your service, to start to develop the research base and work towards developing some guidelines in regard to the processes used in this work.

Method: The research will require qualified clinical and forensic psychologists working within the service to be approached by the researcher and given the opportunity to take part in this research. Participants would be required to complete a short questionnaire relating to the types of work they carry out with service users. Following this they will take part in a focus group facilitated by the researcher, with several other colleagues where they will be asked to discuss a number of topics relating to index offence work and provide information regarding the work that they complete with service users. The focus groups will be video-recorded and later transcribed by the researcher. All focus groups would be carried out by the researcher at the site, therefore there would be no requirement for staff to travel and would be scheduled at a time suitable to fit in with their busy work schedules. Each focus group lasting up to 90 minutes.

All participants' information will be kept confidential and no one person will be identifiable from their information they provide. The name of the service can also be kept anonymous if this were preferred. Any personal information identifying participants (e.g., consent forms) will be kept securely for the duration of the research and 10 years post-publication, in accordance with the Data Protection Act. This is also the case for the transcriptions of the focus groups. Participants will have the right to withdraw during the focus groups themselves and this will be made clear to them. It will also be made clear that they do not have to answer any question that they do not wish to answer. Following the completion of focus groups, participants will be unable to withdraw as their contributions provide the context for all other participants' contributions and thus their removal would hamper the analysis. Participants will be made aware of this. Participants will be asked not to refer to any service user by name during the focus group. However, if this occurs by accident, this information will be removed during transcription. Following transcription the original video-recordings will be destroyed. Finally, all participants will agree to keep any information discussed within the focus group confidential

to the group. All participants are able to request to see the transcripts from the focus group in which they took part. Upon completion of the research, participants can request to see the results of the study.

Ethical approval has already been awarded by the University of Birmingham on the provision that R&D approval is also obtained. This will be sought once sites confirm that they are happy for this research to take place.

APPENDIX H. Participant Questionnaire

Date:

Location:

Please complete this questionnaire prior to the start of the focus group and return to the facilitator.

Please circle the answer that is applicable to you.

1. **Working as a psychologist, within your service are there policies in place that require you to carry out specific index offence work?**

   YES/ NO/ UNSURE

2. **Do you carry out some form of index offence work/analysis when working with service users either on a one to one basis or as part of a group?**

   YES/ NO/ UNSURE

3. **Within your service is there a pre-prescribed process that everyone adheres to for completing such work?**

   YES/ NO/ UNSURE

4. **Does index offence work/ analysis take up a large proportion of the work that you carry out with service users?**

   YES/ NO/ UNSURE

5. **How many hours per week would you estimate that you spend on work surrounding index offence work?**

   Number of hours:

6. **Do you feel that index offence work/analysis is important as part of service users' treatment pathway?**

   YES/ NO/ UNSURE

7. **Do you feel well equipped to complete index offence work/analysis with service users?**

   YES/ NO/ UNSURE

APPENDIX I. Semi-Structured Interview Schedule

1. Guidelines for the NHS as set out by the Department of Health (2008) suggest that as practitioners working in secure settings, you should carry out some form of 'index offence work/analysis'. As a practitioner what does this entail for you?

2. What do you think are the goals of index offence work? (Further question prompts - e.g risk reduction, processing of trauma, integration of events etc)

3. Do you carry out this work on a one to one basis with the service user or within a group setting with several service users? (Further question prompts – why choose one or other? benefits/difficulties of both?)

4. How would you define or describe the work you carry out with service users in regard to their index offence? (Further question prompts – focus on formulation or treatment or combination?)

5. In your opinion are index offence work and index offence analysis different things? (Further question prompts – how do they differ?)

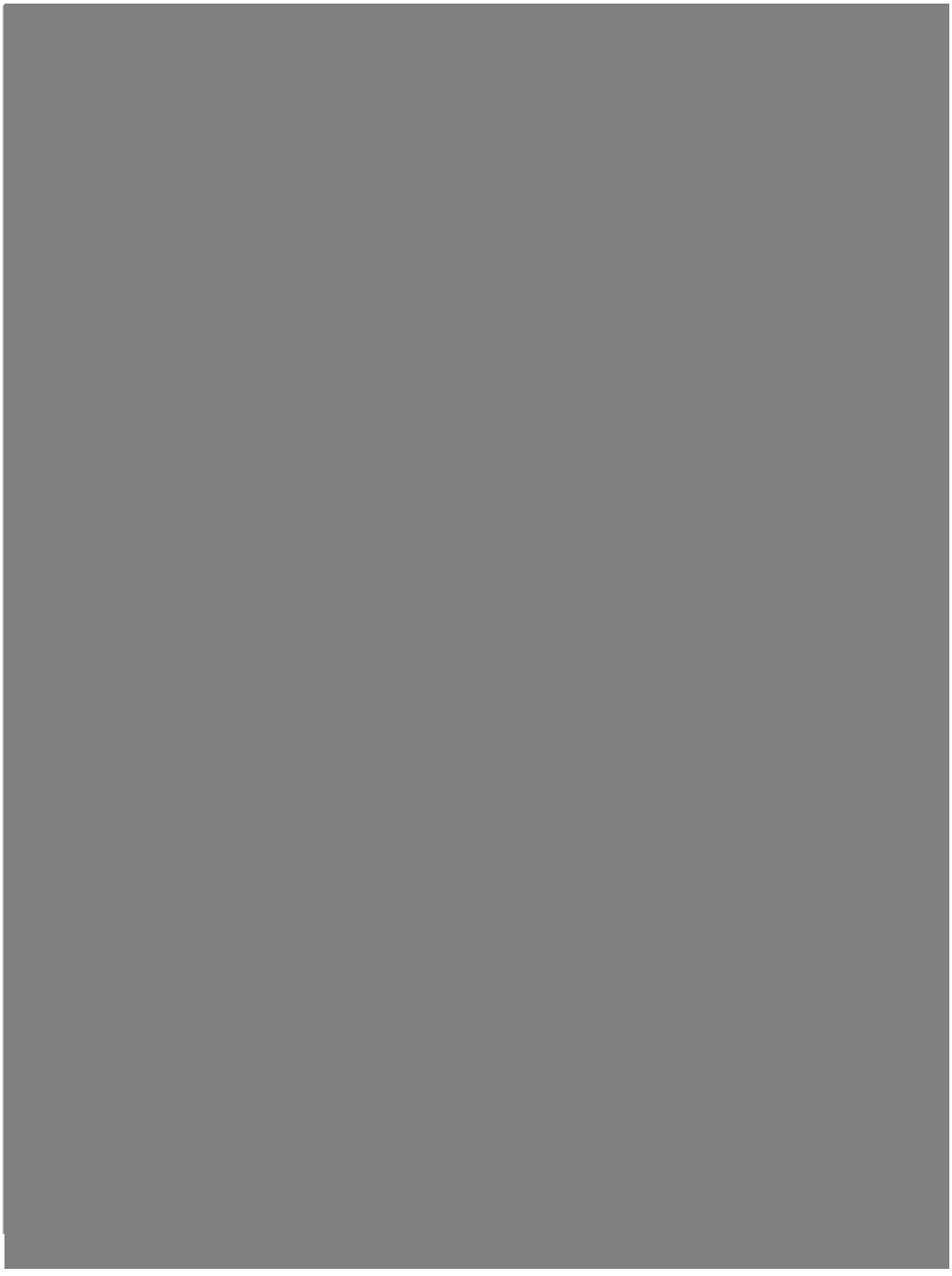
6. Does the work carried out with service users surrounding index offence differ depending upon service user and/or offence type? Or is there a pre-defined process/protocol that you work to with all service users? (Further question prompts – why?)

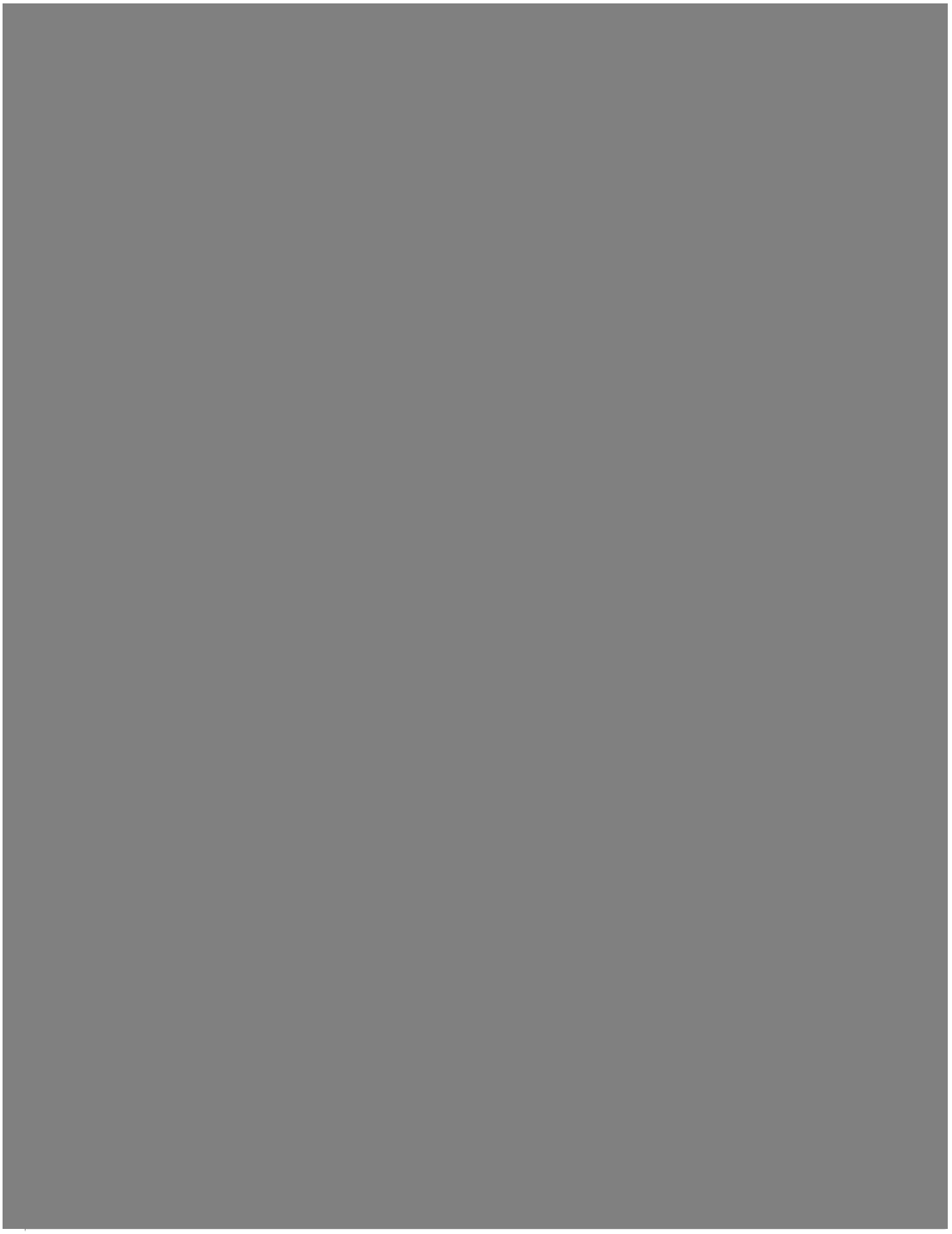
7. What tools do you use to assist you in this work, clinical judgement, actuarial tools or a combination of both?

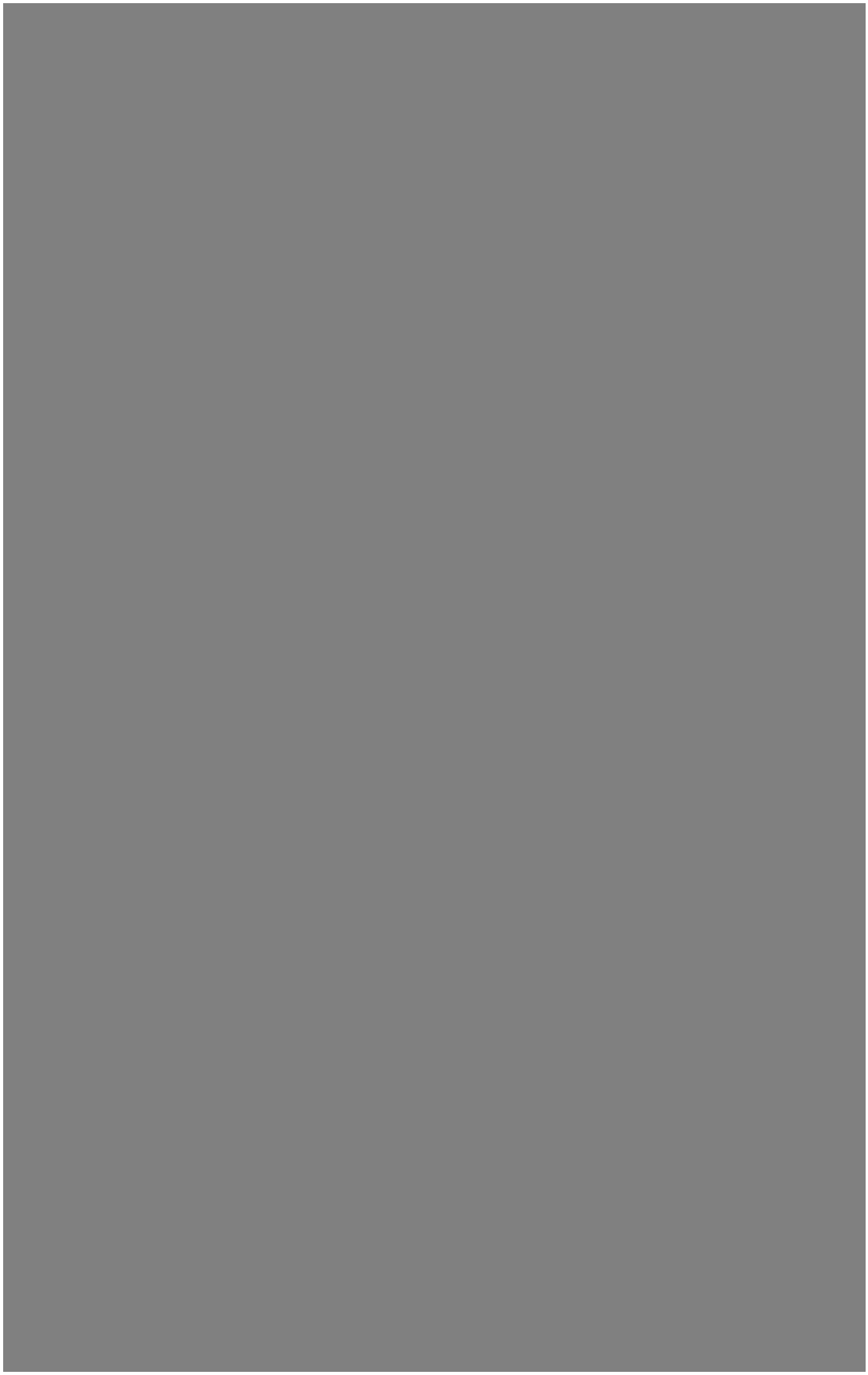
8. How does index offence work fit into treatment planning for the service user and how does it assist their recovery? (Further question prompts – link to goals identified earlier e.g risk reduction etc)

9. What challenges are there personally when engaging in this work? How do you overcome these?

APPENDIX J. Ethical Approval







APPENDIX K. Participant Information Sheet

PROJECT TITLE

An exploration of practitioners' understanding and use of Index Offence 'Work'.

INVITATION

You are being asked to volunteer to take part in a research study on Index Offence 'Work'. Both clinical and forensic psychologists working within secure settings are required by guidelines to carry out Index Offence 'Work' as part of their assessment with patients/offenders. At present there is little information within the literature as to what this work involves for practitioners and also no standardised guidelines outlining procedures to aid practitioners in their work. The aim of this research therefore is to explore this area and develop the research base further as well as working towards developing guidelines on the process which would assist practitioners in this area of work.

My name is Chloe Whatson and I am a Trainee Forensic Psychologist currently studying at the University of Birmingham on the Doctorate in Forensic Psychology Practice. I will be carrying out this research along with supervision from Dr Jessica Woodhams (HCPC registered forensic psychologist and chartered psychologist University of Birmingham), Dr Alison Lauder (chartered clinical psychologist, Fromeside medium secure unit) and Dr Megan Wilkinson-Tough (chartered clinical psychologist, Fromeside medium secure unit). Ethical approval has been gained from the University of Birmingham as well as from your Trust.

WHAT WILL HAPPEN

In this study, if you agree to participate, you will firstly be asked to complete a short questionnaire relating to the nature of Index Offence 'work'. Following this you will be asked to take part in a small focus group (composed of approx. 4 people) with some of your colleagues. The focus group will be videotaped to assist the researcher in creating an accurate transcription. Once the transcript has been made, the video will be destroyed.

During the focus group, you will be asked questions by the group facilitator (Chloe Whatson) relating to the 'index offence work' that you carry out with service users as part of your work with them. This will involve you providing information in regard to the processes and procedures you use as well as information on any psychometric tools and other methods that you may use. It will take the format of a discussion between yourself and your colleagues. All information discussed within the group will be confidential and you will be asked not to discuss other people's participation or information they provide within the group with anyone else. In addition during the group please refrain wherever possible in identifying specific service users. If you refer to a service user by name in error, this information will be redacted as appropriate.

Once the questionnaires and focus groups are completed, the information you and others provide will be analysed using quantitative and qualitative analysis. Researchers will identify common themes in procedure and the work carried out and will then write this up as a Doctorate thesis and potentially as conference presentations/journal articles, in order to inform the

research base in this area and work towards developing some guidelines surrounding the process of this work in order to assist practitioners like you in their future work. You will not be identifiable in any write up of this research.

TIME COMMITMENT

The questionnaire should take you no longer than ten minutes to complete. The focus group which you will take part in will happen on one day and will typically last for 90 minutes. The group facilitator (Chloe Whatson) will travel to you and carry out the focus group at your place of work and to fit in around your work schedule. A date and time will be arranged in advance.

PARTICIPANTS' RIGHTS

You can decide not to answer the questionnaire or decline to answer any question, without any explanation required and without any negative consequences, however post-study we are unable to offer you the option to withdraw your questionnaire responses as this data will have been analysed.

You can decide to stop taking part in the focus group at any time during the focus group itself without having to provide an explanation to the researcher and without penalty. Also, you can decline to answer any question without there being any negative consequences for you. We cannot offer you the option of withdrawing from the study at a date post-focus group because that will involve the removal of all other participants' data also. You also have the right to request to see any of your information at any time, throughout the research study.

You have the right to have your questions about the procedures answered (unless answering these questions would interfere with the study's outcome). If you have any questions as a result of reading this information sheet, you should ask the researcher before the study begins. You can also contact the researcher at anytime during or after the study with further questions that you may have.

BENEFITS AND RISKS

There are no known benefits or risks for you in taking part in this study although we hope that your contribution will enable us to start developing a knowledge base in this area that can be used by other practitioners in the future.

COST, REIMBURSEMENT AND COMPENSATION

Your participation in this study is voluntary and your time will not be reimbursed since this study is not funded but is part of Chloe Whatson's degree course.

CONFIDENTIALITY/ANONYMITY

Any personal information (consent forms) will be kept securely by the researcher for 10 years post the studies publication. The data that is collected will not contain any personal information about you other than the methods and procedures which you use in your 'index offence work'

with service users. The information you provide will not be linked with the identifying information you supplied (e.g., name, address, email). Whilst quotes may be included in analysis, all quotes used in the write-up for illustrative purposes will not identify any individuals. In addition any other data collected such as the transcripts from the focus groups will be kept on a password protected computer in a password protected file.

The data will be used in doctoral research therefore will form part of the thesis that will be publically available. In addition, there is a possibility that it may be published via conference presentations or journal articles. Therefore individuals within the University and the larger psychological community may have access to the research; however no identifying information regarding participants will be detailed in the write-up.

FOR FURTHER INFORMATION

Myself, Dr Jessica Woodhams or Dr Alison Lauder will be glad to answer your questions about this study at any time. You may contact us using the details below. If you want to find out about the final results of this study, you should contact us stating your interest and once the analysis is completed we will be able to send you details of the research findings.

CONTACT DETAILS

Principal Investigator: Chloe Whatson (Trainee Forensic Psychologist) –
[redacted]

Supervisors: Dr Jessica Woodhams (Forensic Psychologist) –
[redacted]

Dr Megan Wilkinson-Tough (Clinical Psychologist) –
[redacted]

APPENDIX L. Participant Consent Form

University of Birmingham

Consent Form for participation in Index Offence Research

Part a – your details

Your full name	
Job Title	
Your employment address and contact telephone number	

Part b – your declaration and signature

I understand that my participation in this research is voluntary and that I will be required to complete a short questionnaire and also to speak about the work that I carry out with service users and discuss this with some of my colleagues. I understand that focus groups will be video-taped and later transcribed in order for data to be analysed.

I understand that I have the right to withdraw from the questionnaire and focus group at any time and that I am unable to withdraw from either post-focus group as this would involve the removal of all other participants' data and questionnaire data will have been analysed. I understand that all of my personal information will remain confidential and that information I provide during this research will be available for me to look at if I so desire. I also understand that any information I do provide whilst participating in this study will be kept securely by the researcher for a period of 10 years.

I understand that filling in and signing this form gives you permission to use the information I provide in this research study. I understand that it will be used to inform and develop research being undertaken by researchers at the University of Birmingham and Fromeside medium secure unit. All information used will have personal identifying information removed to ensure that I cannot be identified.

Your signature:	
Date:	
Witness signature:	
Date:	

Notes for the participant

By signing this form, you are agreeing to participate in research carried out by researchers at the University of Birmingham in conjunction with researchers at Fromeside medium Secure Unit. You agree to information you provide being used to inform the research as well as the knowledge that this research will potentially be published and will be seen by tutors at the University of Birmingham. All access to this information remains confidential.

You will have the right to withdraw from the focus groups without reason at any time without negative consequence to yourself.

Any information used in the research will have personal details and identifiers removed in order that you will not be identifiable. You will have the right to request to view the information that you provide in this study and can also contact researchers if you would like to know the research findings.

APPENDIX M. Step-by-step Summary of the Data Treatment Process

1. Transcription

Bird (2005) argues that transcription is a "key phase of data analysis within interpretive qualitative methodology" (p.227). Focus groups were transcribed using orthographic transcription (verbatim). This was used by the author to ensure a thorough transcription of spoken words was achieved, this took into account all pauses and punctuation in order to retain the information needed from the verbal discussion, so that it remained true to its original nature (Braun & Clarke, 2006). It was important to ensure that transcription was consistent across the four focus groups and therefore the orthographic notation system outlined by Braun & Clarke (2013) allowed for this.

2. Reading and familiarisation

The transcription process itself contributed to the familiarisation of the data. However, following transcription the author then re-read all transcriptions in order to immerse themselves within the data. This involved the watching of the video-recordings and also reading and re-reading the transcripts in order for the author to begin noticing things of interest. In order to maintain a record of such key words, trends within the data and ideas of interest, the author directly identified these on paper versions of the transcripts. As recommended by Braun and Clarke (2013) this process was observational rather than systematic and allowed for the author to identify triggers which would assist in the development of later analysis. This stage was not based on a systematic engagement with the data, it was important for the author to be aware of this and not use such information as the sole basis for developing the analysis further. By doing this the author has ensured that the data is not biased as it is likely that the initial ideas identified and reflected what the author brought to the data and the aspects that were salient to them.

3. **Coding-complete**

Complete coding was utilised by the author in the analysis of the data as opposed to selective coding. Complete coding is not influenced by existing theories or frameworks and instead aims to identify all data relevant to answering the research question (Braun & Clarke, 2013). Due to the limited research evidence in relation to IOW this method was deemed to be most applicable. The coding used in this analysis was data-driven and this therefore means that the themes identified were dependent upon the data that was collected. This process involved the author aiming to identify anything and everything of interest or relevance to the research question within the entire dataset.

Data was initially coded using Nvivo 10 computer software, this was used as it automatically collated data that was coded using the same labels, which was deemed useful for the author for the latter stages of the process. This ensured a log was kept of each coded data item and allowed for the development of a provisional codebook. Initial codes were identified and provided with a label for a feature of the data that was potentially relevant to answering the research question. Data extracts, whether that be a word, phrase or larger section of data were coded in as many ways as possible. A combination of data-derived and researcher-derived codes were utilised by the author, as Braun and Clarke (2013) suggest the separation between semantic and latent codes is not pure and often in practice research can have both elements. Every time the author identified something that was potentially relevant, it was coded and this was done with the entire dataset in a systematic manner. Data that didn't contain anything relevant to the research question was not coded at all.

In line with guidelines for conducting TA (Guest, Namey & MacQueen, 2012; Braun & Clarke, 2013), once the first coding of the dataset had been completed by the author, the entire

dataset was revisited. This ensured that the coding of data was not idiosyncratic and that the majority of codes were evident in one or more data item.

4. Searching for Themes

Braun and Clarke (2006) highlight that this stage involves, collating information into candidate themes. Through reviewing the codes and collated data relating to each code, similarities and overlaps between codes were identified. This involved the author actively making choices about how to 'shape' the data and identifying patterns within it. At this stage some of the codes were promoted to themes, due to the fact they were large, rich and complex enough (Charmaz, 2006). Themes and subthemes were structured in a hierarchical and lateral manner, with some overarching themes, themes and sub-themes. The candidate theme were structured laterally with up to five themes and then hierarchically with two layers (themes and sub-themes). This was deemed the most appropriate way to organise the results during the analysis phase.

Themes and sub-themes identified at this stage were considered to be provisional or candidate themes and the author was willing to revisit and refine them at later stages in the analysis. As the current study adopted an essentialist perspective in regard to data interpretation, the author when searching for candidate themes which identified thoughts, feelings and behaviours of participants in relation to their completion of IOW. A large degree of interpretation in relation to the data items was not necessary, as the author also being a trainee forensic psychologist was able to understand the processes about which participants spoke. As recommended by Frith and Gleeson (2004), due to the analysis being provisional at this stage, codes that were discarded due to not answering the research question were kept in a

miscellaneous category in case they were required as new themes were developed at a later stage in the analysis.

Braun and Clarke (2013) suggest that the number of themes developed is dependent upon the purpose of what the analyst is trying to achieve. In this case the author has attempted to provide a meaningful overview of the data and therefore it was considered important to demonstrate the breadth and diversity of the data, meaning one or two themes were insufficient.

5. Reviewing Themes

Once initial candidate themes had been identified, the next stage of analysis was for the author to ensure that candidate themes fitted well with the coded data and the dataset that was collected. This meant reviewing all the coded and collated data to ensure that each candidate theme was applicable to these, distinct from one another and that the concepts of coded items was consistent. During this process the author was required to tweak themes, moving them and codes around and on occasion discarding or collapsing themes in order for the analysis to better fit the data and best answer the research question. Once a coherent set of candidate themes had been developed, that effectively answered the research question, the author then further reviewed these by revisiting the entire uncoded dataset. This involved re-reading all transcripts to confirm that the themes developed captured the content and meaning of the dataset in order to sufficiently answer the research question. During this phase candidate themes were further reviewed and revised.

As Braun and Clarke (2013, p.234) suggest "…analysis is not a linear process. It is recursive, it goes back and forth on itself, and you have to be prepared to go backwards, and take a different route, to ultimately move forward". This was a continuous process, where the analysis was constantly evolving and only ended once a set of coherent and distinctive themes

had been developed. This was when the author had an understanding of how themes fitted together in order to provide an overall narrative of the data.

6. Defining and Naming Themes

Once the iterative process of reviewing theme content and all codes had been assimilated, it was important for the author to define themes and identify their unique qualities. Therefore the boundaries of themes and sub-themes were identified by the author and definitions for each were refined. These were explicitly delineated within the codebook, as suggested by a number of authors (Braun & Clarke, 2006; Braun & Clarke, 2013; Guest et al, 2012). Themes and sub-theme names were reviewed and the author ensured that the names that were applied, were relevant and appropriate and captured the central organising concept for each theme.

Following the completion of this stage, the codebook was made available to the author's academic peer for use in establishing interrater reliability.

7. Writing and Finalising Analysis

The final stage as identified by Braun and Clarke (2013) is concerned with the write-up and presentation of the data and findings. There were a number of factors that had to be considered by the author at this stage. Firstly, the consideration of whether to treat the data illustratively or analytically. As this research study took an essentialist approach it was decided that the analysis should be presented more illustratively, providing a descriptive account which attempts to recount the narrative of the data. However, the author is aware that at points the semantic and latent approaches may be combined where a more interpretative style of analysis has been conducted. Secondly, a decision had to be made relating to the order in which themes were presented and discussed. The author had attempted to provide a logical presentation of the analysis which fits with the story of the data itself and relates to the overall research questions. Please see the results section of the research report, where results of the analysis are presented.

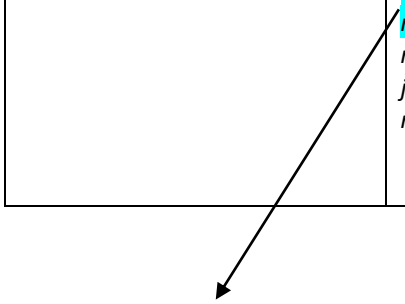
APPENDIX N. Section of Raw Data (Annotated Transcript)

Issues regarding semantics

Participant 9 Increasing insight into risk	*"Ye obviously reducing someone's risk and someone having an understanding of why why it happened and what sort of increases the risk of it happening again and decreases that I spose as well thinking about someone's protective factors erm and as well so things that and trying to bolster those as part of the work".*
Participant 8	*"Which is exactly why it's so difficult to define I think when somebody asks you what a piece of what what does a piece of and I kind of you know I thought about this a lot recently what does it look like and it looks like a combination of all those things, it's about understanding risk factors, er you know understanding their protective factors so the piece of work is not just about kind of reducing risk factors it's also about improving and bolstering protective factors so the piece of work it's just it's kind of all-encompassing isn't it really very wide and you know it's it kind of it's almost like it runs intrinsically through almost every piece of work that you do because it's not fair to assume that somebody's mental illness was necessarily related to their index offence just because they're in hospital we can't that's not always the case is it erm but a lot of the time it is so I think any work that you that directly relates to their mental health is part of that index offence work really (Participant 10 – "Ye definitely, definitely") simple as".*
Participant 7 ***Targeting criminogenic factors***	*"And I always think that this is kind of the or what I view as the idealised outcome of index offence work which I was thinking about offence paralleling behaviour and if someone's index offence is relational in nature if they could get to a point where they can start to notice where they might be enacting something so I was thinking about some of the men* [redacted] *and someone with sexual offences and he engages in lots of offence paralleling behaviour if you can get to a point where he can actually notice that he might be using someone for example in a sadistic way and gain some pleasure from talking to them about their index offence for example that would be like the ideal outcome is someone gains that level of awareness relationally that they're re-enacting something and that they need to do something different, that would be amazing and I don't know if we really get that far, but…".*
Participant 9 ***Need for an MDT approach***	*"And I spose if that is happening on the ward cuz I work with a man where it is I think it's putting the support in place for like systemically for the nursing staff and for them to understand a little bit about that risk formulation and supporting them about managing that on the ward and containing that person really even if you can't work with that person directly???"*
Participant 8	*"So part of the work can be about spending time with staff so that can a)understand and b)be able to kind of you know hold and contain somebody if their distressed and if it is kind of in relation to trauma so again it's just another example of how wide the piece of work can become isn't it?"*

Organisational issues/influences

Participant 10	*"And I've worked with people where actually the mental health awareness has been the key thing in reducing risk cuz that is the key factor in reducing their risk as that is the key factor in increasing their risk and so kind of a big piece of work on mental health awareness has been the thing to help them obviously for reducing risk and public protection etc erm but has made a massive difference and actually technically what you're doing is just educating somebody about how to stay well but that has a massive affect then on their risk".*



"Bringing mental health and risk together can be difficult"