

CURRENT ISSUES IN THE TREATMENT OF SEXUAL OFFENDERS

by

CAROLINE ROBERTSON

A thesis submitted to the Centre for Forensic and Criminological Psychology
of the University of Birmingham
for the degree of ForenPsyD

Centre for Forensic and Criminological Psychology
School of Psychology
College of Life and Environmental Sciences
The University of Birmingham
January 2010

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Abstract

This thesis investigates the debate around sex offender treatment efficacy. Numerous methods are utilised to explore this topic, including a meta-analysis (N =15,931), empirical research (N =322) and a single case study. Chapter 1 reviews the efficacy of sex offender treatment in relation to study design, treatment type, and treatment setting. Results indicate a positive effect of treatment in reducing both sexual and general recidivism for treated versus untreated offenders. However, treatment effects varied greatly according to the study design used, with no significant effect of treatment found for randomised controlled trials. Within Chapter 2, survival analysis and logistic regression are used to examine the impact of treatment dose ('Risk Principle') on reconviction and within-treatment change. Results indicate that whilst controlling for Risk Matrix 2000 (Thornton, Mann, Webster, Blud, Travers, Friendship & Erikson, 2003) classification, treatment dose does not influence treatment outcome. The results are discussed in light of the need to consider the way that sexual offenders interact with the amount of treatment received. Chapter 3 uses a single case design to explore assessment and low-dose intervention with an internet offender. The case study explores practice based issues, including the difficulty in applying pre-existing knowledge of contact sexual offenders to internet offenders. Chapter 4 provides a critique of Risk Matrix 2000 (Thornton et al., 2003). Chapter 5 discusses the practical and theoretical implications of this thesis, explores limitations of the thesis, and provides recommendations for future research.

Acknowledgements

First and foremost, I would like to thank my academic supervisor Anthony Beech, for providing me with invaluable guidance and support over the past five years. I am so grateful for the opportunities and advice you have given me. Also to my supervisor James Bickley, you have been a huge source of support over the past two years and have helped me a great deal in developing my academic and professional skills.

Special thanks go out to the Lucy Faithfull Foundation for their financial support and assistance with my career development. I have been incredibly fortunate to be surrounded by such supportive colleagues and managers, all of whom I would like to say thank you to.

None of this would have been possible without the love and support of the people most important to me. My wonderful family and friends have managed to keep me going with their endless encouragement and faith in me, for which I will always be grateful. Finally, to my wonderful husband-to-be, Pete, what can I say. You have been my rock throughout the last seven years of studying and your patience has never failed to amaze me. Thank you for always believing in me, here's to 12.08.10.

Table of Contents

Thesis Section	Page Number
Introduction	1
Structure of Thesis	6
Statement of Authorship	8
Chapter 1: Literature review	9
A meta-analysis of treatment outcome studies: Comparisons of treatment designs and treatment delivery	10
Chapter 2: Empirical research study	44
Chapter 2 rationale	45
An exploration into the impact of treatment dose for sexual offenders	46
Chapter 3: Single case study design	80
Chapter 3 rationale	82
Case study: An assessment and low-dose intervention with an internet offender	83
Chapter 4: Psychometric test critique	138
Risk Matrix 2000	139
Implications for empirical research study and single case study design	152
Chapter 5: Discussion	154
Conclusions	165
References	167
Appendices	193

List of Tables

Table Number	Title	Page
Chapter 1		
1	Details of all studies used in the meta-analysis	27-29
2	Treatment effectiveness for reducing sexual offence recidivism	31
3	Treatment effectiveness for reducing general offence recidivism	34
4	Treatment effectiveness for reducing sexual offence recidivism by treatment type and setting	36
5	Treatment effectiveness for reducing general offence recidivism by treatment type and setting	38
Chapter 2		
1	Background information on sample	56
2	Means, standard deviations, and cut-offs for each of the measures used in the study	61
3	Percentage of reconvictions over a ten year period by treatment length and risk category	65
4	Cox regression analysis using RM2000 and treatment dose to predict sexual offence recidivism	66
5	Percentage of sexual offenders demonstrating clinically significant change at the end of treatment according to treatment length and Risk Matrix 2000 category	68
6	Logistic regression analysis using static risk classification and treatment dose to predict clinically significant change	69
Chapter 3		
1	Sentencing advisory panel risk levels – Indecent child images (Sentencing Advisory Panel, 2002)	93
2	Summary of pre-intervention psychometric assessment	105

List of Figures

Figure Number	Title	Page
Chapter 3		
1	Figure 1 Stop it now referral progress	95
2	Functional analysis of childhood factors	112
3	Functional analysis of adolescence factors	113
4	Functional analysis of sexual offending behaviour	114
5	Diagrammatic representation of client P's presenting problems	120

List of Appendices

Appendix 1	Details of search terms used in Chapter 1 meta-analysis
Appendix 2	Details of clinically significant change method applied in Chapter 2
Appendix 3	Details of assignment of psychometric measures to Thornton's (2002) dynamic risk domains
Appendix 4	Copy of consent form used in Chapter 3 case study
Appendix 5	Copy of interview guidelines used in Chapter 3 case study
Appendix 6	Summary of clinical history obtained for Chapter 3 case study
Appendix 7	Copy of Risk Matrix 2000 scoring used in Chapter 3 case study
Appendix 8	Description of psychometric tests used in Chapter 3 case study
Appendix 9	Contact log with client P used in Chapter 3 case study
Appendix 10	Supervision log associated with Chapter 3 case study
Appendix 11	Risk Matrix 2000 scoring guidelines (RM2000/S, RM2000/V & RM2000/C)

Introduction

The successful treatment of sexual offenders not only aims to prevent harm being caused to the victims of sexual offences, but also promotes a safer society and assists those who have committed sexual offences in being rehabilitated. Interventions and efforts to provide treatment to sexual offenders have been developed in a number of well documented stages over the last 60 years (Wood, Grossman & Fichtner, 2000). Notably, the last two decades have seen a significant increase in the development and evaluation of programmes offered to sexual offenders.

With an increase in the treatment options available to sexual offenders, a great deal of knowledge has developed regarding what constitutes a successful treatment approach, and how sexual offender treatment should be implemented. The type of treatment that has been revealed as most promising in terms of reducing recidivism is the cognitive-behavioural approach (Brown, 2005). This approach has consistently yielded the most positive effects of treatment in reviews and meta-analyses in this area (e.g., Hanson et al., 2002). Researchers are also relatively confident regarding what constitutes successful treatment methods. Research has provided support for programme characteristics that discriminate more effective programmes from least effective programmes (Brown, 2005). More specifically, research has highlighted that successful treatment approaches are those that adhere to the Risk, Need and Responsivity principles (Andrews & Bonta, 2006). These can be summarised as follows;

“The Risk Principle assists in deciding who might profit most from intensive rehabilitation programming. The Need Principle suggests the appropriate targets of change for effective rehabilitation. Responsivity has to do with the selection of the appropriate modes and style of service” (Andrews, 1989, p. 8).

Evidence for the application of these principles in offender treatment has been supported by meta-analyses based on a large number of various types of treatment

programmes (Andrews, Bonta & Hogue, 1990), and more recently, have been supported in their application within sexual offender treatment programmes (Hanson, Bourgon, Helmus, & Hodgson, 2009). The application of these principles and their theoretical bases has profoundly influenced the policies of key treatment providers, for example HM Prison Service in the UK, and has promoted a shift in emphasis from the containment of prisoners to the rehabilitation of prisoners (Friendship & Thornton, 2001). The fact that these principles seem to be important in the treatment of general and sexual offenders therefore appears to be a reasonably robust finding (Brown, 2005).

However, despite the strong empirical bases of the Risk, Need, Responsivity (R-N-R) model, there are a number of limitations which have been documented over recent years (e.g., Ward, Mesler & Yates, 2007). Ward et al. note that the R-N-R model, with its focus solely on risk reduction, may cause difficulties in motivating offenders to change. Additionally, Ward et al. note that the focus on risk reduction, using a model which is essentially psychometric in nature, may in turn lead to important variables in the change process being ignored. Such variables include the individual's sense of personal identity and agency, the impact of the therapeutic alliance upon rehabilitation, the importance of non-criminogenic needs (e.g., personal distress and/or low self-esteem), and contextual or ecological factors. Ward et al. also highlight that the R-N-R model does not account for that fact that as human beings, sexual offenders naturally seek and require certain goods in order to live fulfilling and personally satisfying lives.

Ward et al. propose that the limitations of the R-N-R model can be addressed by a 'dialogue' with other rehabilitation theories, namely the Good Lives Model (GLM) of offender rehabilitation (e.g., Ward & Stewart, 2003). In sum, the GLM argues that as human beings, those who commit sexual offences are goal orientated beings who are predisposed to seek a number of 'primary goods'. Primary goods can be defined as states of mind, personal

characteristics, activities or experiences that are sought for their own sake and if achieved, are likely to increase psychological well-being. The model also utilises the term ‘secondary goods’, the means by which primary goods are obtained. The GLM essentially argues that it is not the primary goods sought after that are unacceptable (e.g., sexual gratification or feelings of intimacy), but the methods by which they are sought. Ward and Stewart (2003) argue that in terms of risk management, we need to remain aware of internal or external obstacles that may frustrate or block the acquisition of primary human goods and provide the individual with the conditions to obtain and secure these goods.

In their review of the RNR model, Ward et al. (2007) conclude that an integration of the model with the GLM may assist in developing theories of rehabilitation and implementing these into effective practice. Ward et al. also note that through increasing the dialogue between these two models, the potential for the R-N-R model to be used as a “one size fits all” approach (and hence ignoring its own Responsivity principle) is minimised. What both approaches do agree upon is that in order for sexual offender treatment and rehabilitation to be effective, we need to responsive to the individual needs of the offender. Additionally, the principles underlying both the R-N-R and the Good Lives approaches are crucial in terms of considering the issue of treatment efficacy and that factors which impact upon whether or not treatment is effective with sexual offenders.

In addition to the debate surrounding which models of sexual offender rehabilitation are best suited for use in clinical practice, a number of important questions also still remain regarding the efficacy of sexual offender treatment. There is still considerable debate regarding the effectiveness of sexual offender treatment programmes (Hanson et al., 2009) and more importantly, what works and with whom. These questions are vitally important, especially given the potentially devastating consequences for victims if incorrect decisions regarding treatment are made. Additionally, in a climate in which resources are ever-

increasingly scarce, it is vital that researchers make efforts to answer these questions so that resources can be directed in cost-effective ways that also best protect the public.

The debate into the efficacy of sexual offender treatment has many facets, and debates have ranged from questions into whether treatment works *per se* (e.g., Marques, Wiederanders, Day, Nelson & van Ommeren, 2005), to what type of treatment works (e.g., Hanson et al., 2002) and what factors might influence or mediate the effectiveness of treatment (for a review of this research, please see Harkins & Beech, 2007b). The lack of consensus stems from many factors, however, it is often argued to be rooted in the difficulties with the evaluation and measurement of effectiveness of sexual offender treatment programmes (Hanson et al., 2002).

It is commonly argued in the field of sexual offender treatment that the strongest research designs are those in which offenders are randomly assigned to treatment (e.g. Marques et al., 2005). However, the opinions of researchers in the field differ greatly with regard to the appropriateness of this method, with some clearly arguing against this form of research design (e.g., Marshall & Marshall, 2007). Practical and ethical difficulties associated with randomised studies, to name but a few, include the limited scope for undertaking randomised studies in criminal justice settings (e.g., Her Majesty's Prison Service in the United Kingdom) and the 'ethical minefield' associated with deliberately withholding treatment from a group of high risk sexual offenders in aid of research (Marshall & Marshall, 2007). Given these difficulties, it seems unsurprising that randomly designed studies of sex offender treatment effectiveness are rarely implemented in the field (Farrington & Jolliffe, 2002). Using a randomised controlled trial, Marques et al. (2005) did not find an overall treatment effect for a current cognitive-behavioural sexual offender treatment programme. Marques et al. note that "in the context of growing optimism about the benefits of sexual offender treatment, this study's message is, 'Not so fast, we are still far from understanding

how and when treatment works'..." (p. 99). However, Hanson et al. (2002) reported a positive effect of current cognitive-behavioural treatments in a meta-analysis of 43 studies using various study designs.

What all reviewers of sexual offender treatment do however appear to agree on is that more and better studies of sexual offender treatment efficacy are needed. Additionally, reviewers also appear to agree, that in addition to looking at whether treatment for sexual offenders works *per se* (i.e., does it reduce recidivism), we need to continually assess the factors which may mediate the effects of treatment and consider where our resources are best directed.

It is within this context of uncertainty and the need for continuing research into the area of sexual offender treatment that this thesis is based. More specifically, the content of this thesis aims to contribute to the ongoing need for further research into sexual offender treatment and draw together current research undertaken up until this point in time.

Specifically, the thesis aims to deliver the following:

- To provide an updated review of the sexual offender treatment efficacy literature in the form of a meta-analysis.
- To expand upon the outcome of the meta-analysis through a preliminary investigation into the effect of treatment dose on sexual offender treatment efficacy. This is an area which is relatively under-researched but that has important implications for the way in which resources are utilised.
- To use an individual case study design to explore the effects of a low-dosage intervention for a sexual offender and to explore the potential factors that influence treatment efficacy on an individual basis.

- To provide a critique of Risk Matrix 2000 (Thornton et al., 2003), a measure utilised within this thesis and a measure frequently employed within sex offender treatment research.
- To discuss the question of ‘does sex offender treatment work’ in light of the above.

Structure of the thesis

This thesis is comprised of five linked components. In Chapter 1, the literature around sexual offender treatment outcome research is reviewed in the form of a meta-analysis. This meta-analysis explores the effects of treatment according to the design utilised in the research. The meta-analysis also compares treatment effectiveness according to the way in which treatment was delivered. The discussion explores the need to consider the impact of study design when exploring treatment efficacy. The results are also discussed in light of other areas that merit consideration when evaluating the efficacy of sexual offender treatment.

Chapter 2 provides an empirical research study into the concept of treatment dose and the impact of this on treatment outcome. Measures of sexual re-offending and within-treatment change are used to explore this area. The discussion focuses upon the need to further explore the concept of treatment dose for sexual offenders and the impact this may have on treatment outcome.

Chapter 3 utilises a single case study design to explore the case of an internet sexual offender receiving a low-dose, psycho-educational model of intervention. Key issues discussed include treatment provision for sexual offenders at the pre-conviction stage, low-dose treatment provision and factors that may have mediated the effects of treatment for this individual. The case study aims to provide a practice based example of the core issues raised in the treatment of sexual offenders.

Chapter 4 provides an overview and critique of Risk Matrix 2000 (Thornton et al., 2003), a measure utilised in chapters 2 and 3 of this thesis.

Chapter 5, the concluding chapter, explores the results highlighted within this thesis, draws overall conclusions, discusses some of the methodological limitations of the research conducted within this thesis and suggests areas for future research. The question of ‘does sex offender treatment work’ is also explored within Chapter 5.

Statement of Authorship

Chapter 1 contains material that has been submitted for publication to *Sexual Abuse: A Journal of Research and Treatment*. Therefore, the authorship on this article indicates collaborative working. I am the senior author of this paper and my supervisor Anthony Beech is named as an author. Nick Freemantle, The University of Birmingham, is also a named author of this paper. There are no other co-authors in the content of this thesis. I would like to thank Dr. Karl Hanson and two anonymous referees for their helpful comments on an earlier draft of the paper detailed in Chapter 1.

Chapter 1

Literature Review

A Meta-Analysis of Treatment Outcome Studies: Comparisons of Treatment Designs and Treatment Delivery

Abstract

Sixty-one treatment studies ($N = 15,931$), employing a number of designs (all using a control group), were examined using a random effects meta-analysis model. Odds ratio (OR) were employed to assess treatment efficacy, defined as the ratio of the odds of an event in treated/control groups. Results indicated a positive effect of treatment for both sexual ($OR = .49$, 95% $CI = .39$ to $.62$, $p < 0.0001$) and general recidivism ($OR = .54$, 95% $CI = .44$ to $.66$, $p < 0.0001$). Randomised control trials showed no significant effect for sexual ($OR = .49$, 95% $CI = .15$ to 1.60) and general recidivism respectively ($OR = .70$, 95% $CI = .40$ to 1.24). Significant effects were found for incidental designs, and where treatment dropouts were used as a control group. Assignment based on need indicated a negative effect of treatment. Results indicated an advantage of systemic and cognitive-behavioural approaches in reducing both sexual, and general, recidivism. The results also highlight the importance of considering study design when evaluating treatment, and the importance of considering the impact study design may have on reported treatment efficacy.

Introduction

Demonstrating treatment effectiveness is probably the thorniest issue in the field at the present time, and there is no consensus as to whether treatment works or not (see Brooks-Gordon, Bilby, & Wells, 2006; Harkins & Beech, 2007a; Marshall & Marshall, 2007; Seto et al., 2008, for current discussion on this issue). There are a number of factors that have contributed to the lack of accord. The majority of these factors appear to relate to difficulties in the measurement of treatment effectiveness itself, and the methods by which treatment outcome research in this area is conducted. Some of the better-known methodological barriers in sex offender treatment research are as follows: (1) studies vary in their use of what they term as recidivism including reconviction, re-incarceration and/or committing a different

type of sexual offence than that carried out previously; (2) reconviction/recidivism rates are relatively low, therefore, large sample sizes are required for statistically convincing differences to be observed between treated and untreated samples. But outcome studies are typically small; (3) drawing firm conclusions from treatment outcome studies may be problematic given the likelihood that discrepancies exist between conviction rates and actual rates of sexual re-offending (Friendship & Thornton, 2001).

But probably the main issue that causes dissension in the field is which treatment designs have the potential to demonstrate that treatment works, or does not work, and the nature of control groups employed in such studies. Here, comparison groups have included, for example, (i) those who have been unable to complete treatment for geographical reasons (Marshall & Barbaree, 1988); (ii) treatment drop-outs (Hall, 1995a); (iii) treatment unavailability (Proctor, 1996)¹. However, some have argued (see for example, Harris, Rice & Quinsey, 1998; Quinsey, Harris, Rice, & Lalumière, 1993; Rice & Harris, 1997, 2003), that unless a randomised control trial (RCT) approach is employed (where treatment participants are assigned, by chance, to receiving/not receiving treatment and where it is assumed pre-existing differences between the two groups will be controlled for, and randomly distributed), the question can never be properly answered.

The scientific rigour of the RCT approach has led to the assertion that this is the *gold standard* design for evaluating any type of treatment (see for example, Egger, Smith & Altman, 2005), including sex offender treatment (Quinsey et al., 1993). Indeed, as RA Fisher noted, “It may be said that the simple precaution of randomisation will suffice to guarantee the validity of the test of significance, by which the result of the experiment is to be judged” (Fisher RA, 1935, p. 21). However, despite the scientific quality provided by RCTs,

¹ typically researchers will try to address the threat of pre-existing group differences through matching participants on a number of variables (e.g., risk, age, see for example, Hanson et al., 2004).

typically in medical interventions, they are difficult to apply practically within criminal justice settings, and are not without their own set of problems. For example, for a RCT to be effectively implemented the treatment for individuals in the control group will have to be deliberately withheld for a number of years to create a ‘non-intervention’ control group. This raises ethical and legal issues for major institutional systems, given the possible consequences of denying a high risk sexual offender treatment (Harkins & Beech, 2007a; Marshall & Marshall, 2007). In addition, unless the sample is sufficiently large, the random allocation of clients to groups cannot ensure that the groups are equally matched (Marques & Murphy, 2004). Therefore, some have argued that RCT designs are unsuitable in the field due to the difficulties associated with the method, and that less rigorous designs may be the optimal position (Marshall & Marshall, 2007). For example, techniques such as minimisation² may be used to improve the efficiency of randomised designs where there are known differences in subject prognosis on the basis of characteristics which can be described at baseline.

As for an attempt to provide clear guidelines on the quality of research designs, the Collaborative Outcome Data Committee (CODC) has recently published a document on this topic (CODC Guidelines, 2007a; 2007b). Although these guidelines state a clear preference for RCTs, the authors point out that it is highly unlikely a ‘definitive’ study will ever provide a conclusion to the ongoing debate within the field. Therefore, the CODC Guidelines suggest that it is only through the accumulation of results from diverse research methodologies that a more definitive conclusion can be drawn regarding the effectiveness of treatment.

With regard to combining evidence from different types of studies, in order to answer such questions, meta-analysis is becoming increasingly recognised as a useful tool. Meta-

² a largely nonrandom method of treatment allocation for clinical trials, which aims to ensure treatment and control groups are balanced with respect to predefined patient factors as well as for the number of patients in each group (Scott, McPherson, Ramsay & Campbell, 2002). The first subject is allocated truly randomly, whilst the following subjects are allocated to treatment or control groups on the basis of chosen characteristics.

analysis is the process by which various studies' results are combined in order to yield an overall statistic that summarises the effectiveness of a set of studies (Egger et al., 2005). Specifically, meta-analysis involves the process of calculating a summary statistic and then combining these statistics into a weighted average (Egger et al., 2005). Thus, meta-analysis provides a means with which to assess the consistency of results across studies. It also indicates the importance of including unpublished studies, the so called 'file drawer problem' (Rosenthal, 1979) where, according to Rosenthal, 5% of published journal articles show Type 1 errors, while 'file drawers' are filled with 95% of the studies that show non-significant results.

As regarding the effectiveness of sex offender treatment using RCT studies, Kenworthy, Adams, Bilby, Brooks-Gordon, and Fenton (2004), have conducted a meta-analysis of nine identified RCTs, with over 500 offenders. The conclusions from this meta-analysis were limited, and results ranged from one study demonstrating no benefit from group therapy (however this was psychodynamic treatment), to another indicating that a cognitive approach resulted in reduced re-offending. A further meta-analysis of RCTs conducted by Brooks-Gordon, Bilby and Wells (2006), using the same nine studies, concluded that cognitive-behavioural treatment (CBT) reduced re-offence rates at one year but increased re-arrest rates at 10 years. Hence, merely relying on RCTs suggests somewhat inconclusive evidence for treatment.

As for assessing treatment across a wider range of experimental designs, three studies in the 90s (i.e., Hall, 1995a; Alexander, 1999; Gallagher, Wilson, Hirschfield, Coggeshall & MacKenzie, 1999) are early examples of meta-analytic approaches in the area combining all available studies. All have been criticised for methodological reasons. Hall analysed 12 studies (N =1313), in which any comparison groups had been used, finding a small, significant treatment effect, involving re-arrest rates for treated (9%), compared to untreated

offenders (12%), over an average follow-up period of 6.9 years. Hall also noted that medical treatments and CBT were both superior to behavioural treatment. However, the strongest reported treatment effects were derived from comparisons between treatment completers and dropouts, and when the dropout studies were removed from the analysis, significant effects of treatment no longer remained (results reported by Rice & Harris, 1997). Alexander (1999) reported an analysis of 79 studies (N =10,988). When offenders were categorised by offence type, treated offenders had lower recidivism rates than untreated offenders in all categories³. However, samples of treated and untreated sexual offenders were often derived from different studies, follow-up periods were unclear, as were recidivism criteria. These problems introduce the possibility that the observed differences in recidivism rates could have been due to these differences (Hanson et al., 2002).

Gallagher et al. (1999) examined 25 treatment comparison studies, where both published and unpublished research studies were included. A significant effect for cognitive-behavioural, but not for hormonal treatments, was found. However, similarly to previous studies, a number of sources of potential bias were present within this meta-analysis (e.g., the inclusion of treatment drop-outs as comparison groups).

As a response to the lack of clarification provided by these studies, Hanson et al. (2002) conducted a meta-analysis with the aim of addressing the methodological flaws identified. Here, Hanson et al. included all credible studies of psychological treatment for sexual offenders identified by May 2000. Forty-three studies were included in this analysis (N =9,534) of which 23 were published and 20 were unpublished, with an average follow up of 46 months. Studies were required to have a comparison group (incorporating either those who had received no treatment, or alternatively, those who had received treatment that was judged to be inadequate or inappropriate). A significant effect of treatment was found for

³ other than the category where offence type was not specified

sexual recidivism (12.3% for treated and 16.8% for untreated samples). A similar effect of treatment was also demonstrated for general recidivism (27.9% for treated and 39.2% for untreated sexual offender samples). Hanson et al. found that 'older treatment' options (i.e., non-behavioural/non-CBT) appeared to have little effect in reducing recidivism. However, studies of current approaches to sexual offender treatment (i.e., cognitive-behavioural) were reported to have positive treatment effects and show reductions in recidivism. A similar set of results and conclusions has been presented more recently by Lösel and Schmucker (2005). These authors incorporated 69 studies into a meta-analysis (N =22,181). This meta-analysis also identified a positive effect of treatment, with treated sexual offenders (11.1%) recidivating at significantly lower level than comparison groups (17.5%). Again, cognitive-behavioural therapy demonstrated the most robust effects of treatment.

However, while both studies appear to have provided evidence for the positive effects of sexual offender treatment, both Hanson et al. (2002) and Lösel and Schmucker (2005) acknowledge that more evidence is required in order to produce firmer conclusions regarding this debate. Research studies can be interpreted very differently and experts regularly disagree upon the value of studies given the methodological flaws often present within this type of research (Hanson, 1997a). One aspect of this debate that researchers appear to agree continuously upon is that further research is required in order to continue to contribute to this area of research. Therefore, the consensus would seem to be that there is both a continuing need for research in this area and that an increased confidence in outcome studies can be only achieved when similar relationships are replicated in many studies (Hanson & Morton-Bourgon, 2007).

Hence, the aim of the current research is to provide an updated meta-analysis in the light of new findings, with the following predictions:

1. There will be decreases sexual recidivism in those undertaking sex offender treatment compared to those who were in an untreated control group. An associated research question was whether there were differences in effectiveness related to the type of design used in the study (i.e., RCT, incidental assignment, treatment attendance versus treatment refusal, treatment completion versus drop-out, treatment dropout versus treatment refusal, assignment based on need). Additionally, an associated research question was whether there were differences in effectiveness between adult and adolescent samples.
2. There will be decreases in general recidivism in those undertaking sex offender treatment compared to those who were in an untreated control group, as per Hanson et al.'s (2002) previous findings. An associated research question was whether there were differences in effectiveness related to the type of design used in the study. Additionally, an associated research question was whether there were differences in effectiveness between adult and adolescent samples.
3. There will be differences in sexual recidivism rates across types of treatment undertaken. Specifically, CBT interventions will reduce recidivism more than non-CBT interventions. An associated research question was whether there were differences in effectiveness related to the setting in which treatment was delivered.
4. There will be differences in general recidivism rates across types of treatment undertaken. Specifically, CBT interventions will reduce recidivism more than non-CBT interventions. An associated research question was whether there were differences in effectiveness related to the setting in which treatment was delivered.

Method

Study selection

The initial selection of the studies used in the meta-analysis reported was obtained from those included within the Hanson et al. (2002) meta-analysis (K =43) given the comprehensive nature of this previous study. This involved data extraction from the Hanson et al. paper, the reference list, and also contacting authors and researchers in order to obtain original papers detailed within this research. The current research extended this comprehensive meta-analysis by conducting a new literature search to identify suitable studies published after May 2000 (the cut-off point for this was April 2009), or studies published before May 2000, which may have been missed by the Hanson et al. study.

Search terms applied within the Hanson et al. meta-analysis were utilised and additional search terms were generated for the purposes of this study (see Appendix 1 for a list of search terms used within the research). Using the specified list of search terms, internet computer searches of a range of relevant databases were conducted (detailed within Appendix 1). Manual and internet computer searches were conducted of relevant journals, which are also listed in Appendix 1. A search of other additional sources of information was also carried-out, for example, manual searches of relevant reference lists were conducted and relevant articles were followed up. Searches were also conducted of the internet sites of relevant institutions, organisations and departments of corrections (each of which are listed in Appendix 1). Additionally, experts within the field of sexual offender treatment were contacted in order to account for any unpublished data or studies, and to identify any data that may have been missed through other search methods. Once studies were identified, titles were scanned for initial relevance. If a study appeared to be related to the research question, the abstract was read in detail, and if deemed relevant, the full report was accessed and analysed for suitability against the chosen inclusion criteria detailed below.

Inclusion criteria

Inclusion criteria were defined in order for an assessment to be made regarding the suitability of each study identified within the literature search. The inclusion criteria were formulated using examples set through previous meta-analyses undertaken within the field (e.g., Hanson et al., 2002; Lösel & Schmucker, 2005). The following characteristics were decided upon as eligibility criteria for primary studies.

Sexual offender sample

The study had to include the primary sample of sexual offenders (i.e., those convicted of a sexual offence).

Measure of recidivism as outcome

It was necessary for recidivism to be included as an outcome measure in primary studies. The dependent variable of recidivism therefore ranged from lapse behaviour to incarceration. Therefore, as in the Lösel and Schmucker (2005) meta-analysis, the decision was made to exclude those studies focusing exclusively on measures not specifically related to recidivism, for example, changes in measures of personality, or hormone levels.

Comparison procedure

In order to meet the inclusion criteria, primary studies had to incorporate a control group design, that is the study contained a comparison of the recidivism rates (including sexual or general) of a sample of treated sexual offenders with a comparison group of untreated sex offenders. This comparison sample could either be an untreated control group or a group of offenders receiving treatment deemed to be inappropriate, inadequate or that differed from the evaluated program in content, intensity and specificity.

Recidivism criteria

The same recidivism criteria must have been used for both groups in order for equal comparisons to be made. Additionally, recidivism rates must have been reported for approximately the same follow-up period. In terms of follow up periods, most research studies provided an average length of follow up. Where this was not provided (e.g., a minimum and maximum length of follow up was provided), a median length of follow up was calculated. Additionally, if different follow up periods were reported for both treatment, and comparison groups, the average of these two figures was calculated. A decision was also made to include analyses of both sexual and general recidivism within the meta-analysis to account for the evidence suggesting that sexual offenders will often re-offend with a non-sexual offence as opposed to a sexual offence (Hanson & Bussière, 1998). It was hoped that this decision would also help to partially address the issue of plea-bargaining and the masking of sexual reconvictions through official recording of data (Quinsey et al., 1993).

Sample size

The combined sample size used by each primary study had to be at least 10 (5 individuals in each group).

Type of treatment

The programs must have provided predominantly psychological treatment. These studies were also categorised according to the type of treatment administered, as follows: (1) cognitive-behavioural; (2) systemic; (3) psychodynamic; (4) behavioural treatment; (5) unknown; (6) mixed. Studies using medical interventions alone were excluded, unless the medical intervention was administered alongside a form of psychological intervention.

Country of origin

No restrictions were made as to where the studies were conducted. All studies were reported in English with the exception of one French language study (Martin, 1998).

Time of publication

Those studies available before May 2000 were obtained via the Hanson et al. (2002) meta-analysis. The search therefore aimed to obtain studies conducted as of May 2000. However, any studies conducted before this arising within the search were also reviewed for eligibility.

Coding Procedure

In order to analyse the data effectively and account for the influence of each individual study design on the reported effects of treatment, each research study was coded in accordance with the research design used. In order to maintain consistency with the Hanson et al. (2002) meta-analysis, the same coding procedure was applied within this meta-analysis. To this end the original authors of the Hanson et al. (2002) meta-analysis were contacted in order to access detailed study coding.

A number of issues arose through this process. Since the publication of Hanson et al. a number of previously included studies had been updated. It was therefore necessary to update a number of these studies for the purpose of the current research. Some of the original studies from the Hanson et al. meta-analysis were therefore excluded and replaced with updated versions. For example, the Marques, Day, Nelson and West (1994) was replaced with the more recent publication of the same study by Marques, Wiederander, Day, Nelson, and van Ommeren (2005). Personal communication with authors was also utilised for identifying studies with different references/authors, but where these were in fact duplicates of the same sample.

Studies were assigned to one of six research design categories, as follows:

(1) Randomised Controlled Trial (RCT)

This category is where a randomly assigned treated group of offenders was compared to a randomly assigned group receiving no treatment, or alternative treatment deemed to be inappropriate or inadequate.

(2) Incidental Cohort

This category incorporated a group of treated offenders being compared against a group of offenders receiving no treatment or alternate treatment who had been incidentally assigned to these conditions. For example, studies were included within which comparison groups of sexual offenders were drawn from scenarios such as: (i) offenders released before the implementation of treatment programs; (ii) offenders matched using criminal records; (iii) offenders who received an earlier version of the treatment program; (iv) offenders who received no treatment; (v) offenders receiving treatment judged to be lower in quality (e.g., programs are unavailable or for offenders who had insufficient time on their sentence to complete a treatment program).

(3) Volunteers versus refusers

Studies within which any treatment attendance (including drop-outs) was compared to those who refused to partake in treatment.

(4) Completers versus dropouts

Studies comparing treatment completers to those who dropped out of treatment.

(5) Drop-outs versus refusers

Studies comparing treatment drop-outs versus those who refused to participate in treatment.

(6) Assignment based on need

Studies in which those assigned to treatment based on need were compared to those deemed not to need treatment (i.e., higher risk individuals were offered treatment, lower risk individuals were not, and comparisons are made between higher risk ‘treated’ compared to lower risk untreated individuals).

These categories therefore followed a general structure in order to separate those studies from which pre-existing group differences would not be expected (e.g., RCTs), those in which group equivalence was not assured but that reasonable steps had been taken to ensure this (e.g., incidental assignment, matched controls), and studies in which differences may reasonably be expected (e.g., through the inclusion of treatment drop-outs).

Index of treatment effectiveness: Odds ratio

The starting point of any meta-analysis involves the selection of a summary statistic or effect measure⁴ (Egger, Smith, & Altman, 2005). Egger et al. (2005) recommend that all measures of effect should be accompanied by confidence intervals. For the current meta-analysis, the odds ratio was employed. The odds ratio summary statistic is defined as the ratio of the odds of an event occurring in two groups (Egger et al., 2005). The use of odds ratios is recommended with the use of dichotomous data (Fleiss, 1994; Lipsey & Wilson, 2001) since it is a symmetrical measure for which measurement error may be assessed with considerable accuracy, and therefore within this research, odds ratios were used to analyse the

⁴ i.e., which measure will be used to describe the observed treatment effect in each trial

dichotomous data reflecting the outcomes of ‘did recidivate’ or ‘did not recidivate’. Odds ratios as an index of treatment effectiveness are also noted to be relatively unaffected by arbitrary design features such as the proportion of offenders in the treatment and comparison groups, or the overall recidivism rate (Fleiss, 1994). Odds ratios were also chosen following the example of previous meta-analyses conducted within this field of research employing this summary statistic (e.g., Hanson et al., 2002; Lösel & Schmucker, 2005; Reitzel & Carbonell, 2007).

Odds ratios (OR) were derived with the use of 2 x 2 tables detailing the recidivism data outcomes of both the treatment and comparison groups. ORs were then calculated as follows, from Fleiss (1994):

$$OR = \frac{\text{recid}t / \text{nonrecid}t}{\text{recid}c / \text{nonrecid}c}$$

These odds ratio calculations followed those outlined and utilised by Hanson et al. (2002). *Recid_t* refers to the number of recidivists in the treatment group, *nonrecid_t* is the number of non-recidivists in the treatment group, *recid_c* is the number of recidivists in the comparison group and *nonrecid_c* is the number of non-recidivists in the comparison group. ORs cannot be calculated should there be no events in either of the groups, therefore, following the recommendations of Fleiss (1994), in such events a value of .5 was added to each cell of the 2 x 2 contingency table. This method therefore enabled the analysis of empty cells. If a study reported a different set of results for different offender types or risk groups, effect sizes were calculated separately and then averaged to a single effect size. Interpretation of ORs were as follows: (i) a value of 1.0 would indicate no difference between the groups being compared;

(ii) values below 1.0 are indicative of treatment having a positive effect; (iii) values above 1 indicate treatment having a negative effect.

Data Synthesis

In order to perform the data synthesis or meta-analysis, *Stats Direct* was used (www.camcode.com). There is no fixed set of rules to be employed when deciding which model of meta-analysis to apply (e.g., Fleiss & Gross, 1991). But due to the nature of the current research, and the high chance of significant between study variability, the decision was made to employ a random effects model of meta-analysis, as suggested by DerSimonian and Laird (1986). A random effects model makes the assumption that for each study, the true effects are random observations drawn from a common population distribution, that is, that there is no single 'fixed' treatment effect (Egger et al., 2005). Additionally, a random effects model leads to larger confidence intervals and relatively more weight being given to smaller studies versus other models of meta-analysis in the presence of heterogeneity (systematic differences in treatment effects between studies) (Egger et al., 2005). In order to test for the likely potential for a high level of heterogeneity between individual studies, a Q test of homogeneity was applied (Hedges & Olkin, 1985), whereby a statistically significant result is indicative that between trial variability is more than would be expected by chance alone.

Data description

A total of 61 studies were included within the meta-analysis. Of these 61 studies, only one study did not use sexual recidivism as an outcome measure and employed only measures of general recidivism, leaving a total of 60 to incorporate into the sexual recidivism category. Forty-six of the studies included used general recidivism as an outcome measure. From the total of 61 studies, a sample of 15,931 offenders was examined, where 8214 had received

treatment, and 7717 belonged to a comparison group (i.e., did not receive any treatment, or had received treatment deemed to be inadequate). The dates of studies ranged from 1976 through to 2009, with 1998 as the median year of publication. Of those studies collected, 39 were published and 22 unpublished. The majority of studies were based on American (N =25), or Canadian (N =23) samples. Smaller samples were also derived from the United Kingdom (N =9), New Zealand (N =2), The Netherlands (N =1), and Australia (N =1). The total sample of each individual study collected ranged from between 14 and 2557. The mean total sample size was 261, with a median value of 172 participants.

The majority of studies focused on adult male sexual offenders, however, six of the included studies focused on adolescent treatment. Treatment within the included studies was based in the following settings: community (N =20); correctional institutions (i.e., prisons, secure units, residential treatment services) (N =33); both community and correctional settings (N =8).

The type of treatment offered within each study selected was categorised by the following: cognitive-behavioural (N =45); psychotherapy (N =7); systemic (N =4); behavioural (N =2); mixed (N =2)⁵; unknown (N =1). The date range of delivery of all treatments fell between 1965 and 2001. Of the studies included, 5 were RCTs, 31 were incidental assignment studies, 7 compared any treatment attendance versus treatment refusers, 11 compared treatment completers versus drop-outs, 1 study compared treatment dropouts versus treatment refusers, and 6 studies examined assignment based on need.

As well as differences in study design, individual studies also varied in terms of the way in which recidivism was defined. Definitions included those studies that looked

⁵ The first study falling into the 'mixed' category (Washington Institute for Public Policy, 1998) was reported by authors to consist of a combination of treatment techniques including group therapy, psycho-educational classes, behavioural treatment, drama therapy and family involvement. The second study falling into the 'mixed' category (Wilson, Picheca & Prinzo, 2005) was reported to contain a balance of cognitive-behavioural and systemic approaches to treatment.

specifically at reconviction (N =25), those using re-arrest (N =13), those looking at broader definitions of recidivism (N = 22) such as readmission to institutions, and the remaining study (N =1) which examined recidivism using the presence and commission of relapse behaviours. Follow up periods averaged at 54 months (*SD* =38.2) for the studies in which these were reported (N =55). Information on each of the studies is shown in Table 1.

Table 1: Details of All Studies Examined in the Analyses

Study	Year	Treatment Type	Study design	Follow up (months)	SAMPLE SIZE		RECIDIVISM RATES (SEXUAL)		Odds Ratio (sexual)	RECIDIVISM RATES (GENERAL)		Odds Ratio (any)
					Treatment	Comparison	Treatment	Comparison		Treatment	Comparison	
Alberta Hospital	1994, 1998	1	4	60	194	157	8	13	.48	31	53	.37
Allam	1998, 1999	1	2	12	153	74	5	6	.38	20	31	.21
Aytes et al.	2001	1	2	60	170	149	1	7	.12	Not provided	Not provided	n/a
Bakker et al.	1999	1	2	96	238	283	26	54	.52	Not provided	Not provided	n/a
Barnes & Peterson	1999	1	3	36	147	138	4	12	.29	12	36	.25
Berlin et al.	1991	3	4	60	257	206	13	31	.30	Not provided	Not provided	n/a
Bluglass	1980	3	6	60	43	57	Not provided	Not provided	n/a	18	19	1.44
Borduin et al	1990, 2000	2	1	36	8	8	1	6	.05	2	4	.33
Borduin et al	2009	2	1	106	24	24	2	11	.11	7	14	.29
Bremer	1992	3	4	<i>u</i>	66	27	12	3	1.78	Not provided	Not provided	n/a
Clearwater	2000	1	2	60	245	218	32	52	.48	Not provided	Not provided	n/a
Cooper	2000	1	4	84	41	23	1	14	.02	12	4	1.97
Craissati & McClurg	1997	1	2	24	43	14	2	0	1.51	1	2	.14
Dwyer	1995	1	4	96	125	55	8	9	.35	11	11	.39
Florida	1989, 1977	3	6	12	39	199	2	10	1.02	5	36	.67
Friendship et al.	2003	1	2	24	647	1910	17	86	.57	54	315	.46
Guarino-Ghezzi & Kimball [a]	1998	1	2	12	33	25	0	1	.24	10	12	.47
Hall	1995b	1	3	12	24	6	0	1	.07	5	0	3.67

Hanson et al.	2004	1	2		403	321	85	70	.96	228	194	.85
Hanson et al.	1992, 1993	3	2		102	89	38	29	1.23	64	54	1.09
Harkins	2004	1	3	60	53	53	3	6	.47	17	18	.92
Hedderman & Sugg	1996	1	2	24	133	191	6	17	.48	5	38	.16
Hersh et al.	1999	1	4	<i>U</i>	285	100	0	0	.35	108	59	.42
Huot	1999	1	3	84	92	159	15	30	.84	Not provided	Not provided	n/a
Kramer	1985	4	2		37	19	2	6	.12	5	9	.17
Lab et al. [a]	1993	1	6	24	46	109	1	4	.58	12	22	1.40
Lindsay	1998	1	2	<i>U</i>	7	7	0	4	.06			
Martin	1998, 1999	1	2	36	65	56	4	12	.24	9	26	.19
Marques et al.	2005	1	1	78	259	225	57	45	1.13	42	26	1.48
Marshall & Barbaree	1988	1	5	48	68	58	9	29	.15	Not provided	Not provided	n/a
Marshall et al	1991	1	2	84	17	23	4	8	.58	Not provided	Not provided	n/a
Marshall et al	2008	1	2	122	94	86	1	4	.22	4	11	.30
McGrath et al	1998	1	2	60	71	32	1	5	.08	7	11	.21
McGrath et al.	2003	1	2	84	56	90	3	27	.13	17	32	.79
McGuire	2000	1	4	<i>u</i>	54	14	0	1	.08	2	1	.50
Missouri	1988, 1989	1	4	48	105	156	5	20	.34	24	64	.43
Nathan et al	2003	1	2	29	201	283	11	59	.22	Not provided	Not provided	n/a
Nicholaichuk et al.	2000	1	2	79	296	283	43	93	.35	47	48	.92
Nutbrown & Stasiak	1987	3	2	36	59	11	0	0	.19	12	6	.21
Perkins	1987	1	2	<i>u</i>	62	12	20	2	2.38	30	7	.67
Proulx et al	1988, 1996	1	4	72	117	55	29	13	1.06	43	31	.45
Proctor	1996	1	2	48	54	54	2	8	.22	Not provided	Not provided	n/a
Rattenbury	1986	5	6	72	69	69	15	10	1.64	27	25	1.13
Rice et al.	1991	4	6	72	51	85	26	24	2.64	33	46	1.55
Robinson	1995	1	1	72	189	46	0	0	.25	57	21	.51
RHC Pacific	1998, 1989, 1995	1	6	24	38	29	3	0	5.82	13	6	1.99

Romero	1983	3	1		148	83	20	6	2.01	82	50	.82
Looman et al	1998, 1976, 1989	1	2	60	95	95	25	30	.77	Not provided	Not provided	n/a
Ruddijs & Timmerman	2000	1	2	24	56	56	3	1	3.11	17	46	.09
Nickolaichuk	1996	1	2	24	26	35	3	1	4.43	Not provided	Not provided	n/a
Scalora & Garbin	2003	1	2	54	76	118	2	35	.06	Not provided	Not provided	n/a
Schweitzer & Dwyer	2003	1	2	61	196	164	6	8	.62	20	15	1.13
Seager et al	2004	1	4	24	109	37	5	12	.10	Not provided	Not provided	n/a
Taylor	2000	1	2	84	700	142	56	10	1.15	462	104	.71
Ternowski	2004	1	2	66	224	43	16	6	.47	20	8	.43
Twin Rivers	1995, 1998	1	3	36	209	97	4	8	.22	31	24	.53
Walker	2000	2	4	24	44	62	0	0	1.40	0	7	.08
Seto & Barbaree	1999, 1998	1	3	36	312	85	19	5	1.04	41	13	.84
Washington	1998, 1995	6	3	84	321	306	35	43	.75	64	159	.23
Wilson et al	2005	6	2	53	60	60	3	10	.26	17	26	.52
Worling & Curwen [a]	2000	1	2	74	58	46	3	6	.36	11	13	.59

Note; *a* denotes study of adolescent sexual offenders; *u* denotes information unavailable

Key:

Treatment Type: 1 = cognitive-behavioural; 2= systemic; 3 = psychodynamic; 4 = behavioural; 5 = unknown; 6 = mixed

Study Design: 1= RCT; 2 = incidental cohort; 3 = attendance versus treatment refusal; 4= completers versus dropouts; 5= dropouts versus treatment refusal; 6= assignment based on need.

Results

Overall comparisons made across the studies using odds ratio analyses for both sexual and general recidivism are reported below. Comparisons are also reported for the types of design for sexual and general recidivism. Recidivism levels by type of design are also reported.

Test of Hypothesis 1: Sexual recidivism

A significant advantage for treated versus control groups was demonstrated for sexual recidivism ($OR = .49$, $CI = .39$ to $.62$, $p < .0001$), with 9.27% of the treated group recidivating compared to 17.71% of the control group. However, significant between study variability was demonstrated ($Q = 170.21$, $df = 59$, $p < .0001$). A significant effect of treatment was also demonstrated ($OR = .48$, $CI = .36$ to $.63$, $p < .0001$) for the 36 strongest study designs (i.e., incidental/RCTs). However, again significant variability between these studies was demonstrated ($Q = 96.36$, $df = 35$, $p < .0001$). Table 2 shows sexual recidivism rates, as well as odds ratios, confidence intervals, the Q statistic and sample broken down by specific study designs.

Table 2: Treatment Effectiveness for Reducing Sexual Offence Recidivism

Research design	Odds Ratio	95% CI	Q Statistic	n (k)
Randomised Control Trial	.49	.15-1.60	15.02	1014 (5)
Incidental cohort	.46***	.35-.61	72.53***	9666 (31)
Completers vs. dropouts	.35**	.18-.66	27.43**	2289 (11)
Attendance vs. refusal	.60*	.40-.91	7.77	2002 (7)
Dropouts vs. refusal	.15***	.05-.38	n/a	126 (1)
Assignment based on need	1.92**	1.16-3.18	2.97	734 (5)
Sample type				
Adult only	.51***	.40-.65	161.09***	15,420 (54)
Adolescent only	.21**	.08-.49	3.46	411 (6)

Note: OR is the odds ratio (full description provided under Method: Index of Treatment Effectiveness). 95% CI refers to the 95% confidence interval for the odds ratio. Q refers to the measure of between study variability, n refers to the total number of offenders included in each analysis and k refers to the number of studies in each analysis.

* $p < .05$, ** $p < .01$, *** $p < .0001$

It can be seen from Table 2 that, when studies were compared by treatment design, the following pattern emerged. No significant advantage for treated versus controls was identified for the five studies employing a random assignment design ($OR = .49$, $CI = .15$ to 1.60), with the confidence intervals suggesting that this design could potentially either work well or be very poor. Between study variability was no more than would be expected by chance alone, which is unsurprising given that there were only five studies of this type. For the 31 studies employing an incidental design, the results indicate a highly significant

reduction in sexual recidivism ($OR = .46$, $CI = .35$ to $.61$, $p < .0001$) between treated and untreated participants. Not unexpectedly there was significant variability between the studies ($Q = 72.53$, $df = 30$, $p < .0001$), however, the confidence intervals indicate that for most studies employing this design, a significant treatment effect is likely to be found. For the 11 studies examining those who completed treatment in comparison to those who dropped out of treatment, sexual recidivism was shown to be significantly lower in those completing treatment ($OR = .35$, $CI = .18$ to $.66$, $p < .01$). Again, variability between studies was more than would be expected by chance alone ($Q = 27.43$, $df = 10$, $p < .05$). Only one study was identified which examined comparisons of treatment dropouts and treatment refusers in terms of sexual recidivism. This study identified significantly reduced sexual recidivism rates for those dropping out of treatment when compared to those refusing treatment ($OR = .15$, $CI = .05$ to $.38$, $p < .0001$). For studies examining the sexual recidivism rates for those volunteering for treatment versus those who refused treatment ($N = 7$), a significant effect of treatment was identified for treatment volunteers in comparison to treatment refusers ($OR = .60$, $CI = .40$ to 0.91 , $p < .05$). Between study variability did not exceed what would be expected through chance alone. For the five studies comparing assignment based on need a significantly higher rate of sexual recidivism was found in the treated group compared to offenders considered not to need treatment ($OR = 1.92$, $CI = 1.16$ to 3.18 , $p < .01$). Variability between studies was not significant. There was no evidence of a systematic difference between the results of the randomised trials, and the other study designs ($p = .85$) although the small number of randomised trials and their low power mediates against strong conclusions being drawn from this finding.

Table 2 also includes information relating to treatment effects for both adult and adolescent samples. As can be seen in Table 2, when separated from adult samples, treatment studies examining adolescents identified a significant advantage for treated versus control

groups for sexual recidivism ($OR = .21, p < .01, CI = .08 \text{ to } .49$), with 5.29% of the treated group recidivating compared to 26.36% of the control group. Between study variability was not significant, however this sample only incorporated six studies. A significant effect of treatment was also demonstrated ($OR = .51, CI = .40 \text{ to } .65, p < .0001$) for the 54 studies examining adult only samples for sexual recidivism, with 9.71% of the treated group recidivating compared to 16.74% of the control group. However, significant variability between these studies was demonstrated ($Q = 161.09, df = 53, p < .0001$).

Test of Hypothesis 2: General recidivism

A significant advantage for treated versus control groups was demonstrated for general recidivism ($OR = .54, CI = .44 \text{ to } .66, p < .0001$), with 24.71% of the treated group recidivating compared to 35.59% for the control group in the 46 studies ($N = 12,458$) where general recidivism was reported. As for best study designs (incidental/RCTs), from the 26 studies analysed of these types, a significant effect of treatment was demonstrated ($OR = .48, CI = .37 \text{ to } .62, p < .0001$). As might be expected, variability between studies exceeded that to be expected by chance alone ($Q = 88.55, df = 25, p < .0001$). Table 3 shows general recidivism rates, as well as odds ratios, confidence intervals, the Q statistic and sample by each study design.

Table 3: Treatment Effectiveness for Reducing General Offending Recidivism

Research Design:	Odds Ratio	95% CI	<i>Q</i> Statistic	<i>n</i> (<i>k</i>)
Randomised Control Trial	.70	.40-1.24	10.30	1014 (5)
Incidental Cohort	.44***	.33-.59	73.17***	7272 (21)
Completers vs. dropouts	.43***	.33-.55	6.99	1587 (8)
Attendance vs. refusal	.48*	.26-.87	21.62**	1751 (6)
Dropouts vs. refusers	n/a	n/a	n/a	n/a
Assignment based on need	1.29	.92-1.81	2.74	834 (6)
Sample:				
Adult only	.53***	.43-.66	153.27***	12,047 (40)
Adolescent only	.65	.36-1.15	6.76	411 (6)

Note: OR is the odds ratio (full description provided under Method: Index of Treatment Effectiveness). 95% CI refers to the 95% confidence interval for the odds ratio. *Q* refers to the measure of between study variability, *n* refers to the total number of offenders included in each analysis and *k* refers to the number of studies in each analysis.

* $p < .05$, ** $p < .01$, *** $p < .0001$

It can be seen from Table 3 that when general recidivism rates were examined by the research design employed, no advantage for treated versus untreated offenders was identified in RCT designs, while between study variability was more than would be expected by chance alone. For the 21 studies employing an incidental design to examine general recidivism rates, results identified significant reductions in sexual recidivism for treated versus untreated offenders ($OR = .44$, $CI = .33$ to $.59$, $p < .0001$). Incidental studies did however demonstrate more

variability than would be expected by chance alone ($Q = 73.17, df = 20, p < .0001$). For the eight studies examining those who completed treatment versus those who dropped out of treatment, general recidivism was shown to be significantly lower in those completing treatment ($OR = .43, CI = .33 \text{ to } .55, p < .0001$). A significant effect of treatment was identified in the six studies examining the general recidivism rates of those volunteering for treatment versus those who refused treatment ($OR = .44, CI = .26 \text{ to } .87, p < .05$). When assignment was based on need, in the six studies identified, those referred to treatment demonstrated significantly higher rates of general recidivism compared to offenders deemed not to need treatment ($OR = 1.29, CI = .92 \text{ to } 1.81, p < .05$). There was no evidence of a systematic difference between the results of the randomised trials, and the other study designs ($p = .26$) although the small number of randomised trials and their low power mediate against strong conclusions being drawn from this finding.

Table 3 also includes information relating to treatment effects for both adult and adolescent samples for general recidivism. As Table 3 demonstrates, treatment studies examining adolescents failed to identify a significant advantage for treated versus control groups for general recidivism ($OR = .65, p < .14, CI = .36 \text{ to } 1.15$), with 25.05% of the treated group recidivating compared to 35.51% of the control group. Between study variability was not significant, however this sample incorporated only six studies. A significant effect of treatment was however demonstrated for the 40 studies examining adult only samples for general recidivism ($OR = .53, CI = .43 \text{ to } .66, p < .0001$), with 24.65% of the treated group recidivating compared to 35.61% of the control group. However, significant variability between these studies was also demonstrated ($Q = 153.27, df = 39, p < .0001$).

Test of Hypothesis 3: Sexual recidivism levels by treatment delivery

The data were also analysed according to the effect of the type of treatment and treatment setting upon sexual recidivism. Table 4 shows sexual recidivism rates, as well as odds ratios, Confidence Intervals, the *Q* statistic and sample by treatment type and setting

Table 4: Treatment Effectiveness for Reducing Sexual Offence Recidivism by Treatment Type and Setting

Treatment Type:	Odds Ratio	95% CI	<i>Q</i> Statistic	<i>n</i> (<i>k</i>) = 60
CBT	.44***	.34-.56	113.73***	13,298 (46)
Systemic	.11**	.03-.43	1.99	170 (3)
Mixed	.54	.20-1.39	2.06	747 (2)
Behavioural	.63	.03-12.85	10.42*	192 (2)
Psychodynamic	.94	.43-2.03	15.41*	1286 (6)
Unknown	1.63	.67-3.95	n/a	138 (1)
Setting:				
Community	.33***	.19-.58	55.4***	3124 (19)
Corrections	.57***	.42-.76	95.83***	10,856 (33)
Both	.47*	.26-.86	18.17*	1851 (8)

Note: OR is the odds ratio (full description provided under Method: Index of Treatment Effectiveness). 95% CI refers to the 95% confidence interval for the odds ratio. *Q* refers to the measure of between study variability, *n* refers to the total number of offenders included in each analysis and *k* refers to the number of studies in each analysis.

* $p < .05$, ** $p < .01$, *** $p < .0001$

It can be seen from Table 4 that overall, a significant advantage for treated versus control groups was demonstrated for sexual recidivism within those studies using CBT and systemic approaches. Significant variability was found for CBT studies, but not for systemic approaches (although here there were only three studies in this sample). Other treatment

studies using psychodynamic approaches, mixed approaches, or behavioural approaches failed to demonstrate a significant treatment effect.

Table 4 also provides results according to the setting in which treatment was delivered. Significant advantages for treated versus control groups were demonstrated for sexual recidivism in relation to treatment delivered in correctional settings ($OR = .57, p < .01, CI = .42$ to $.76$), community settings ($OR = .33, p < .0001, CI = .19$ to $.58$) and treatment delivered in both settings ($OR = .47, p < .05, CI = .26$ to $.86$). Treatment delivered in community based settings demonstrated most significance in reducing sexual recidivism. However, for all three analyses, significant between study variability was demonstrated.

Test of Hypothesis 4: General recidivism levels by treatment delivery

The data was again analysed according to the effect of the type of treatment and treatment setting upon general recidivism. Table 5 shows general recidivism rates, as well as odds ratios, confidence intervals, the Q statistic and sample by treatment type and treatment setting. It can be seen from Table 5 that overall, a significant advantage for treated versus control groups was demonstrated for general recidivism within those studies using CBT approaches, systemic and mixed designs. Significant study variability was found in the CBT studies, but not for systemic approaches (although there were only three studies of this type in the sample) or mixed approaches (again, there were only two studies of this type in the sample). Behavioural approaches failed to reach significance, again significant study variability was found in these two studies. The results of the psychodynamic and unknown approaches suggest that overall treatment had little effect to that hoped for in treatment.

Table 5: Treatment Effectiveness for Reducing General Offence Recidivism by Treatment Type and Setting

Treatment Type	Odds Ratio	95% CI	<i>Q</i> Statistic	<i>n</i> (<i>k</i>) = 46
CBT	.52***	.42-.65	104.50***	10, 381 (33)
Systemic	.26*	.09-.70	.64	170 (3)
Mixed	.32**	.14-.69	3.58	747 (2)
Behavioural	.55	.06-4.75	8.34*	192 (2)
Psychodynamic	.84	.54-1.32	6.54	830 (5)
Unknown	1.13	.56-2.25	n/a	138 (1)
Setting:				
Corrections	.63***	.50-.79	63.66***	8453 (24)
Community	.53*	.32-.86	50.81***	2154 (14)
Both	.34***	.23-.51	16.53*	1851 (8)

Note: OR is the odds ratio (full description provided under Method: Index of Treatment Effectiveness). 95% CI refers to the 95% confidence interval for the odds ratio. *Q* refers to the measure of between study variability, *n* refers to the total number of offenders included in each analysis and *k* refers to the number of studies in each analysis.

* $p < .05$, ** $p < .01$, *** $p < .0001$

Table 5 also provides results according to the setting in which treatment was delivered. Significant advantages for treated versus control groups were demonstrated for general recidivism in relation to treatment delivered in correctional settings ($OR = .63$, $p < .0001$, $CI = .50$ to $.79$), community settings ($OR = .53$, $p < .05$, $CI = .32$ to $.86$) and treatment delivered in both settings ($OR = .34$, $p < .0001$, $CI = .23$ to $.51$). Treatment delivered in correctional settings demonstrated most significance of the three in reducing general recidivism. However, for all three analyses, significant between study variability was demonstrated.

Discussion

Similarly to previous comprehensive reviews (i.e., Hanson et al., 2002; Lösel & Schmucker, 2005) significant reductions were found for both sexual and general recidivism in treated versus control groups. Therefore, Hypotheses 1 and 2 were somewhat supported, although this evidence was derived largely from studies of lower quality and specifically without the protection of randomisation. However, comment must be made about the associated research question regarding adolescent and adult samples. For both adult and adolescent samples, significant reductions in sexual recidivism were found in treated versus control groups. However, for general recidivism, significant effects of treatment in treated versus control groups were found for adult but not adolescent studies. Thus it is possible that whilst adolescent treatment is potentially effective at significantly reducing sexual recidivism rates, the same cannot be said for general recidivism. It is, however, important to note that very limited conclusions can be drawn from the small samples of six adolescent studies included in the analysis.

The results of the current study indicate that the design of individual studies may impact upon the reported effectiveness of treatment. Specifically, the overall results of the RCTs in the study did not identify any significant effect of treatment⁶. While positive effects of treatment (in both sexual and general recidivism) were identified in studies employing incidental treatment designs, and data using treatment dropouts as a control group, both designs are open to potentially substantial systematic bias. Unsurprisingly, assigning offenders to treatment based on need indicated a negative effect of treatment, in that higher risk individuals received treatment, while lower risk cases did not. The tests for interaction between the randomised trials and the studies of alternative designs for both sexual and

⁶ but this finding must be qualified in that the confidence intervals in the analysis suggests that this design could potentially either show poor, or very good, outcomes

general recidivism were not significant. This latter finding may not be surprising given the small numbers of lower powered randomised trials included in the review, and may simply reflect low statistical power of the test for interaction, rather than the absence of such a systematic difference.

Caution should, however, be exercised regarding the positive results from the study designs that examined completers versus dropouts, attendance versus treatment refusal, and dropouts versus treatment refusal. There may be a number of reasons for these results, for example, evidence suggests that those dropping out of treatment programs may be likely to possess personality characteristics associated with an increased risk of recidivism (e.g., Marques et al., 1994; McGonaghy, 1999), in that treatment refusal may be a marker of a general uncooperative, antisocial lifestyle which are also associated with an increased risk of recidivism (Hanson & Morton-Bourgon, 2005). In short, the use of both treatment refusals and treatment dropouts as a basis for comparison may increase the possibility of between group differences in recidivism occurring for reasons other than the presence or absence of treatment. Therefore, it is not possible to draw reliable conclusions with the use of these data. It may however be useful for future research to examine the relationship between treatment refusal and recidivism, given the current results and the potential implications this may have for applied risk assessment.

Evidence also suggests that similar caution should be applied to designs that compare the recidivism rates of offenders assigned to treatment based on need compared with those deemed not to need treatment. The initial difference in risk level is not accounted for when assessing data from such designs based on treatment need and subsequently, data from these studies may distort results by indicating the ineffectiveness of treatment in such study designs. It is therefore recommended that if we are to use data incorporating comparisons

based on offender need, the initial level of risk an offender poses be accounted for in order for any identified treatment to be examined in relation to this.

The results of this study also indicate a significant effect of CBT in reducing sexual recidivism. Therefore, support was found for Hypothesis 3. It should be noted that systemic therapy produced significant change, although this result is based on just three studies. The results of unknown treatment suggests that treatment increases the likelihood of sexual recidivism in the treated group, while psychodynamic treatment appears to produce little in the way of recidivism reduction after treatment. A similar pattern occurred in the data for reductions in general recidivism after sexual offender treatment, specifically, significant reductions after CBT. Therefore, support was found for Hypothesis 4. Again it should be noted that systemic therapy produced significant change, although this result is based on just three studies.

Finally, the results of this study also indicate a significant effect of treatment across different treatment settings. Results do however suggest that the setting in which treatment is conducted may impact upon reported treatment efficacy. For sexual recidivism, an advantage of community based treatments was demonstrated over treatment delivered in correctional settings or treatment delivered in both community and correctional settings. However, the reverse was demonstrated for general recidivism; treatment delivered in correctional settings demonstrated an advantage over treatment delivered in the community or treatment delivered in both community and correctional settings.

Limitations of the research

As per any meta-analysis, the quality of the results is dependent upon the quality of the studies examined. Additionally, the level of static and dynamic risk is not reported in the studies. The level of static risk is clearly related to level of recidivism, for example,

Friendship, Mann and Beech (2003), in a large sample of sex offenders (N =2551) reported that treatment efficacy was clearly related to the level of pre-treatment static risk. They found there was little impact of treatment in high risk individuals (as measured by Static-99; Hanson & Thornton, 2000), while for medium risk individuals there was a clear effect of treatment. Similarly, level of dynamic risk has been found to be related to the effectiveness of treatment (see for example Beech, Fisher & Beckett, 1999), with those with higher levels of dynamic risk problems being found to be harder to treat than those with less problems. Measures of individual change within treatment are also typically not reported in the studies. Here, it would be very unusual if all individuals taking part in treatment actually changed. Therefore, the lack of control for risk level and individual treatment impact could have had a very major effect upon the results of the individual studies reported in this analysis with the real level of treatment effectiveness being severely underestimated. It is also important to note that the results are likely to be more reflective of treatment effects with those who have offended against children, rather than rapists, given the general consensus that child abusers are usually overrepresented in sexual offender samples and research. For example, within a sample of 704 sexual offenders reported by Marques et al. (2005), 22% had offended against adults in comparison to 78% who had offended against children.

Suggestions for future research

What is required in future research is methodological soundness, and consistency, in the way in which researchers measure the effectiveness of sexual offender treatment. The CODC (2007a; 2007b) have suggested guidelines for those reviewing the literature and also for those designing new studies or evaluating current programs, which will hopefully result in research in which the possible range of plausible interpretations is minimised, and enable researchers to present results with increasing confidence. Future research should also focus more

specifically on the efficacy of treatment in meeting the different needs of different types of sexual offenders (e.g., child molesters, rapists), in that treatment effects for specific groups are possibly being obscured by the assumption that treatment will affect all sex offenders in a uniform manner.

Conclusions

The current research reports studies identified up until early 2009 and also examined both published, and unpublished, research originating from a variety of samples published in English from around the world. It employed a random effects model and reported literature that was not necessarily published. Consequently, it can be argued that the results are reflective not only of identified studies, but are also representative of a random set of observations drawn from the common population distribution (Fleiss, 1993). Therefore, it is argued that these results lend some support for the efficacy of sexual offender treatment, although randomised controlled trials do not provide conclusive evidence of a treatment effect. Importantly, by incorporating an additional eighteen studies into the meta-analysis since Hanson et al. (2002), the current study has both improved upon and strengthened the conclusions drawn by Hanson et al. seven years ago. The results also highlight the importance of considering study design when evaluating sexual offender treatment. It is important to assess why treatment does not work in particular situations, in order to improve treatment efficacy. Systemic treatment approaches, and CBT, may hold promise for effective interventions.

Chapter 2

Empirical Research

An Exploration into the Impact of Treatment Dose for Sexual Offenders

Chapter 2 Rationale

The aim of this chapter is to provide a preliminary exploration into the impact of treatment dose for sexual offenders upon treatment outcome. Chapter 1 has provided evidence in favour of sex offender treatment efficacy, however, there are a number of limitations with the use of meta-analyses. The meta-analysis, which ultimately aimed to explore the dichotomous question of ‘does treatment work’, did not explore questions pertaining to what works, with whom and in what circumstances. Whilst recognising that these are huge questions and beyond the scope of a single research study, a number of specific limitations of the meta-analysis relate to the research undertaken in Chapter 2. Namely, the data amalgamated from individual outcome studies in Chapter 1 did not account for the initial static or dynamic risk level of each offender, the level of individual change within treatment, nor did it account for recidivism in terms of the amount of treatment received.

Therefore, to supplement the findings outlined in Chapter 1, Chapter 2 aims to extend the exploration into treatment efficacy through examining the concepts of treatment length or ‘dose’, static risk level, within-treatment change and reconviction. Whilst recognising that there are many factors which are likely to influence treatment outcome (see Harkins & Beech, 2007b for a review), treatment dose as a variable is something that has not been well researched. This chapter therefore aims to explore the concept of treatment dose and discuss this in relation to resources, current practice, and how we evaluate sex offender treatment.

Abstract

A sample of 322 sexual offenders who received treatment in the UK were examined according to the hours or 'dose' of treatment received (low, moderate, and high). Data were analysed to assess if the amount of treatment received can assist in predicting both sexual reconviction and/or successful change within treatment, as measured by psychometric assessments. Results indicated that whilst controlling for Risk Matrix 2000 classification, treatment dose alone did not significantly contribute to the prediction of reconviction or successful treatment change. The results are discussed in relation to previous research, and also in relation to other factors that may influence how treatment dose is related to sex offender treatment efficacy.

Introduction

Recent years have witnessed a steady flow of outcome studies and meta-analyses examining the effectiveness of treatment for sexual offenders (e.g., Hanson et al., 2002; Hanson et al., 2009; Lösel & Schmucker, 2005; Marques et al., 2005). Whilst some reviews have concluded that psychological treatment reduces sexual offence recidivism (Gallagher, Wilson, Hirschfield, Coggeshall, & MacKenzie, 1999; Hall, 1995; Hanson et al., 2002; Lösel & Schmucker, 2005), other reviews have remained inconclusive (e.g., Furby, Weinrott, & Blackshaw, 1989; Harris, Rice, & Quinsey, 1998; Rice & Harris, 2003). Additionally, studies exist in which positive effects of treatment were not observed, despite the fact that the reported 'gold standard' of research designs, the randomised controlled trial, had been applied (Kenworthy, Adams, Brooks-Gordon, & Fenton, 2004; Marques et al., 2005). Similar results were also demonstrated in Chapter 1.

Reviewers of sexual offender treatment have however all agreed that more and better studies are needed (Hanson et al., 2009). Generally, researchers note that as opposed to solely

focusing on questions such as ‘if’ sexual offender treatment works or ‘what works’ *per se*, research instead should focus on questions such as what works, with whom, in what circumstances, and what factors might mediate and impact upon treatment effectiveness (Harkins & Beech, 2007b; McGuire, 2002). Research has therefore seen an increase in the application of the Risk-Need-Responsivity (R-N-R) principles (Andrews & Bonta, 2003; 2006; 2007; Hanson et al., 2009) to enable the exploration of such questions.

Andrews and Bonta (2007) describe the use of the R-N-R principles for use in offender rehabilitation. First developed in the 1980’s and ever-increasingly referred to in current research, the principles of Risk, Need and Responsivity are considered one of the most influential models for working with offenders, to date (Ward et al., 2007). The three principles can be described as follows. The principle of *Risk* refers to the need to match the level of intervention to the offender’s static risk of re-offending. Therefore, those offenders deemed to be a high risk of re-offending would be allocated the highest ‘dose’ or most intensive level of treatment. Additionally, it is argued that little to no treatment should be offered to low risk offenders because they are less likely to re-offend even without treatment (Friendship, Mann, & Beech, 2003). The principle of *Need* refers to the process of assessing an offenders criminogenic needs and targeting these within treatment. Factors identified as indicators of dynamic risk for sexual offenders and which have been shown to predict recidivism (e.g., socio-affective functioning; self-management; sexual interest; pro-offending thinking; Thornton, 2002) should be targeted. The principle of *Responsivity* refers to the process of ensuring that any intervention delivered is tailored to the offender’s learning style, strengths, abilities, character and motivation to engage, in order to account for those factors that might enhance or reduce the effectiveness of any treatment delivered (Andrews & Bonta, 2007).

Within general offender treatment, if all three of the R-N-R principles are adhered to the effect size⁷ for positive outcome is reported to be .26. However, if only two principles are adhered to, this effect sizes fall to .18, and .02 for one principle only (Andrews & Bonta, 2003). Reviews and meta-analyses with samples from the general offending population have demonstrated those interventions that are most likely to reduce recidivism are those that meaningfully engage higher risk offenders in the process of changing their criminogenic needs (e.g., Andrews & Bonta, 2006; Andrews & Dowden, 2006).

Until recent years, little research has been conducted which examines whether the R-N-R principles identified as relevant in the treatment of general offending behaviour are applicable to the field of sexual offender treatment. Hanson et al. (2009) conducted a meta-analysis aiming to address this gap in the literature. Using the guidelines of the Collaborative Outcome Data Committee (CODC, 2007a, 2007b), sexual offender treatment studies were rated for their research quality, leaving 22 studies for inclusion in the meta-analysis. Studies were also coded according to their adherence to the R-N-R principles. The sexual recidivism rate of the treatment groups ranged from 1.1% to 33.3%, with an unweighted mean⁸ of 10.9%. The sexual recidivism rate for the comparison groups ranged from 1.8% to 75%, with an unweighted mean of 19.2%. Programmes that adhered to the R-N-R principles showed the largest reductions in sexual and general recidivism. For sexual recidivism, Hanson et al. found that the effectiveness of treatment increased according to the total number of principles adhered to (none, only one, any two, all three, corresponding to odds ratios of 1.17, .64, .63

⁷ effect size is a statistical concept that measures the strength of the relationship between two variables. In meta-analysis, effect size is concerned with different studies and then combines all the studies into single analysis. The measure of effect size in the referenced study (Andrews & Bonta, 2003) is a Pearson product-moment correlation coefficient.

⁸ an average in which all the items have equal weights, i.e., items are not weighted or prioritised proportionally to their importance.

and .21, respectively). The analyses based on the Risk Principle were not statistically significant in any of the analyses conducted.

This suggests that Hanson et al.'s findings are also in line with the general offending literature where it has been found that the Risk Principle is the principle with the least influence on treatment effectiveness for general offenders (Andrews & Dowden, 2006). However, Hanson et al. do note that the Risk Principle in this review was coded only upon the risk level of the offenders participating in specific treatment programs. This did not fully extend to the full meaning of the Risk Principle, which states that interventions should be proportional to an offender's risk of recidivism (Andrews & Bonta, 2006). Therefore, static risk itself was analysed, but was not explored in relation to 'dose'. Hanson et al. (2009) argue that the R-N-R principles should be a primary consideration in the design and implementation of intervention programs for sexual offenders. Hanson et al. do however note that further research is needed in terms of Risk Principle, given the inconclusive findings within this area in this study.

In a review of the factors which can influence the efficacy of sexual offender treatment, the Risk Principle is also discussed (Harkins & Beech, 2007b). Harkins and Beech note that whilst many studies have reported the static risk level of the group they are discussing using actuarial measures (e.g., Looman, Abracen, Nicholaichuk et al., 2000), the treatment outcome is not usually reported in terms of the initial risk level of the offender or the amount of treatment received. A review of more than 200 studies of the effects of correctional treatment found less than 10% of the studies reported on the effectiveness of the intervention separately for lower and higher risk cases (Andrews, Dowden, & Gendreau, 1999).

One exception to the above is Friendship et al. (2003), who examined the recidivism rates of treated versus untreated sexual offenders, in accordance with Static 99 (Hanson &

Thornton, 2000) risk classification. The results highlighted significantly lower rates of sexual and/or violent recidivism in treated participants in the medium–low risk (2.7% treated versus 12.7% untreated) and medium–high risk (5.5% treated versus 13.5% untreated) categories compared to the untreated sexual offenders. Low risk (1.9% treated versus 2.6% untreated) and high risk (26% treated versus 28.1% untreated) groups showed trends in the expected direction. However, these were not significant due to the low base rates in the low risk group, and a (hypothesised) lack of treatment intensity in the high risk group.

The argument for this latter observation is that those offered treatment in Friendship et al.'s study received the original version of the sex offender treatment programme set-up in U.K. prisons in the 1990s comprising of around 90 hours of group-based cognitive-behavioural treatment. This program was subsequently revised in the mid 1990s to comprise of roughly double the amount of treatment hours (i.e., around 180 hours of group-based treatment; Beech & Fisher, 2004). Beech, Fisher, and Beckett (1999) found that for high risk/high psychological need individuals, the original program only produced change in less than 30% of offenders, while the revised program produced change in around 60% of offenders. In addition, this study reported that lower risk offenders responded as well to an 80-hour programme as they did to a 160 hour programme. Additionally, the sexual recidivism rates of those deemed to be low risk are very low indeed (e.g., less than 5% over a 5 year follow up in comparison with a 50% rate for the very high risk group; Thornton et al., 2003). These studies raise the question of whether low risk offenders actually require treatment at all (Mann & Marshall, 2009) and also highlights that both over and under resourcing are necessary considerations when we are considering the Risk Principle or treatment ‘dose’.

Mailloux, Abracen, Serin, Cousineau, Malcolm and Looman (2003) aimed to assess whether we are prescribing the correct ‘dosage’ of treatment to sexual offenders. Using a Canadian sample of 337 offenders, Mailloux et al. classified those receiving treatment into

levels of low, moderate or high intensity, according to the number of treatment programmes completed and/or the intensity of the intervention. Mailloux et al. also assessed whether sexual offenders receiving different doses of treatment differed significantly on actuarial risk classification. Using the Psychopathy Checklist Revised (PCL-R; Hare, 1991; a commonly used assessment tool for Psychopathy) and the Level of Service Inventory (LSI; Bonta & Andrews; 1993; an integrated risk and need assessment instrument for use with general populations of offenders), Mailloux et al. reported that the high intensity groups scored significantly higher on both measures in comparison to the moderate and low intensity groups, therefore providing support for notion that higher risk sexual offenders were receiving higher doses of treatment.

Unsurprisingly, the high intensity group was reported to have attended significantly more sexual offender treatment programmes than the moderate and low intensity groups. However, whilst Mailloux et al (2003) noted that the low intensity group significantly differed from the high and moderate intensity groups in terms of actuarial risk assessment scores and static background variables related to risk, the low intensity group did not significantly differ from the moderate intensity groups in terms of the number of programmes completed. Mailloux et al. proposed that a reasonable explanation for this could have been that the low intensity group were receiving too much treatment, and that this over-resourcing would be better redirected towards the highest risk offenders. Mailloux et al. did however note that the data within their research alone could not wholly justify the conclusion that low intensity offenders may be receiving too much treatment, and noted that the study provided only a preliminary investigation.

In response to the findings of Mailloux et al. (2003), Marshall and Yates (2005) acknowledged the importance of research into treatment dosage given that “there are grounds for supposing that too much treatment may have iatrogenic effects, whereas too little

treatment may be ineffective” (p. 221). However, Marshall and Yates also argue that by using the number of programmes completed as a measure of treatment dosage, the findings were an inaccurate representation of treatment dose received by the sample. They further note that whilst Mailloux et al. did not find any significant differences between the number of treatment programmes completed by the low and moderate intensity group, clear differentiations can in fact be made between these groups receiving differing intensities of treatment if we look at the hours of treatment completed. Using the same sample of offenders, Marshall and Yates note that if the dosage had been classified in terms of treatment hours, the high intensity group received approximately 702 to 1053 hours of the treatment, the moderate intensity group between 238 and 298 hours, and the low intensity group 36 and 91 hours. They therefore note that, if anything, the high intensity group were being over treated.

What these conflicting arguments agree on is that it is positive that attention has been drawn to this issue (Abracen, Looman, Mailloux, Serin & Malcolm, 2005). However, in addition to the fact that both Marshall and Yates (2005), and Mailloux et al. (2003), are presenting conflicting conclusions, this debate also provides no information on the impact of treatment dosage on recidivism rates or within-treatment change. Whilst reconviction data is the most commonly used outcome measure in sex offender treatment evaluation research at present (Falshaw, Friendship, & Bates, 2003), information relating to individual psychological change has also been utilised previously in sex offender treatment research (e.g. Beech & Ford, 2006), using what is known as the clinically significant change method. In brief, the clinically significant change method allows an examination of whether the areas expected to change within treatment are observed to have changed following the completion of treatment. A detailed description of this process is contained the Method section.

These outcome measures have not been previously considered in relation to treatment dose. Additionally, there still appears to be a lack of consensus surrounding the conclusions drawn by Mailloux et al. (2003). As a result of this, the current study aimed to provide further research into treatment dosage for sexual offenders and the Risk Principle, whilst considering the outcome variables of reconviction and within-treatment change.

Hypotheses

This study therefore aimed to investigate the Risk Principle in relation to treatment dosage and treatment outcome for sexual offenders. More specifically, the following hypotheses were explored:

1. There will be significant differences in the static risk assessment classifications of sexual offenders attending high, moderate or low intensity sexual offender treatment programmes. More specifically, those attending high intensity treatment will be significantly more likely than those attending low or moderate intensity treatment to be classified as high or very high risk.
2. There will be significant differences in the reconviction rates of those attending low, moderate and high intensity sexual offender treatment programmes. More specifically, those attending high intensity treatment groups will re-offend at a significantly higher rate than those attending low or moderate intensity treatment groups.
3. When controlling for static risk category, treatment dose will significantly contribute to the prediction of sexual reconviction in sexual offenders. Specifically, the presence of treatment dose as a variable will increase the prediction of reconviction in sexual offenders.
4. Significant differences will exist between the proportion of sexual offenders demonstrating significant changes within treatment according to the dose of treatment

received (short, moderate or high intensity). Specifically, those attending high intensity treatment programmes will be significantly more likely than those attending low or moderate intensity treatment programmes to achieve significant levels of within treatment change.

5. When controlling for static risk category, treatment dose will significantly contribute to the prediction of whether a sexual offender will achieve clinically significant change within treatment. Specifically, the presence of treatment dose as a variable will increase the prediction of within-treatment change in sexual offenders.

Method

Sample and treatment settings

The three separate sources from which the data was collected were delivering group-based cognitive-behavioural treatment in England and Wales. These three data sources were as follows:

- Men who had attended the English and Welsh Prison Service's Core Sex Offender Treatment Programme (SOTP, Mann, 1999) in the mid to late 1990s. These were data reported by Beech, Fisher and Beckett (1999).
- Men who had attended community treatment in the early 1990s, drawn from data reported by Beckett, Beech, Fisher and Fordham (1994)
- Men who had attended community treatment in the late 1990s in the West Midlands, U.K., these data were drawn from Allam (2000).

The sample had attended treatment which varied widely in length. For those that had attended the Core SOTP within the English and Welsh Prison Service (N =100), treatment groups ranged from 74-160 hours. For those that attended community based treatment programmes (N =67), treatment length varied from 47.5-60 hours and 60-1000 treatment hours for

community based Probation Service programmes and specialist residential treatment programmes respectively. The remaining participants (N =155) attended community treatment within the West Midlands Sex Offender Unit, receiving 50 hours of assessment and 150 hours of treatment (Allam, 2000). The average length of treatment received by participants was 124 hours, with a mode of 50 hours.

Participants

The original sample was comprised of 322 adult male sexual offenders who had completed a treatment programme in the U.K. Twenty offenders were removed from the sample because they had refused treatment. An additional three offenders were removed because they had not been released prior to the end of the follow-up period. Dropouts and non-completers were included in the treatment sample. The sample can therefore be described as an 'intent to treat' sample (Deeks, Altman & Bradburn, 2001).

From this sample of 322 adult male sexual offenders, two separate groups were used for data analysis. Within the overall sample, Group 1 was used to explore reconviction as an outcome variable. A number of exclusions were made to the original sample based on the absence of reconviction data, data relating to treatment length, demographic information, criminal history and/or other information required to calculate Risk Matrix 2000 (RM2000; Thornton et al., 2003). This left 250 adult male sexual offenders for whom a complete set of data was available. Group 2 was used to investigate the outcome variable of significant within-treatment change. From the original sample of 322 adult male sexual offenders, a number of exclusions were made based on the absence of psychometric data required to calculate clinically significant change, data relating to treatment dosage, demographic information, criminal history and/or other information required to calculate RM2000 scores.

Group 2 was comprised of 98 adult male sexual offenders for whom a complete set of data was available. Background information on Groups 1 and 2 are summarised in Table 1.

Table 1. Background Information on Participants

Sample Characteristics	Group 1	Group 2
N	250	98
Mean age at time of index offence (SD)	35.2 years (11.1)	36.9 years (10.39)
Offence Type		
Children	81%	98%
Adults	11%	0%
Non-contact	3%	0%
Mixed	5%	2%
Victim gender		
Male	42%	18%
Female	49%	68%
Both	9%	14%
Previous sexual convictions	58%	36%
Previous non-sexual violent convictions	19%	10%
Percentage with previous non-sexual non-violent convictions	57%	39%

Measures and concepts

To provide a measure of the Risk Principle, Risk Matrix 2000 (Thornton et al., 2003) was utilised. RM2000 was designed for use in the U.K. where it is used widely within the Prison, Police and National Probation Services. RM2000 was developed to predict the actuarial risk of sexual re-offending in convicted sexual offenders. Specifically, RM2000 is a statistically derived risk classification instrument, designed for use with males who are 18 years and above, who have been convicted of a sexual offence. RM2000 uses a two stage assessment system. Stage 1 contains three risk items (number of previous sexual appearances, number of criminal appearances and age). The sums of these factors are translated into an initial risk category. Stage 2 then considers four aggravating factors (any conviction for a sexual offence

against a male, any conviction for a sexual offence against a stranger, any conviction for a non-contact sexual offence and single / never been married). The initial risk classification can then be amended accordingly, where the presence of two or four aggravating risk factors raises the initial risk category by one or two levels respectively. The scale yields four risk categories of low, medium, high and very high.

The original RM2000 construction dataset consisted of 647 male prisoners at risk for at least two years. A second validation sample comprised of 429 male prisoners discharged from prison in 1979 was followed-up for 16 years. The Area Under the Curve (AUC) statistic for RM2000 was .75 for Sample 1 and .77 for Sample 2 indicating a good level of predictive accuracy (Thornton et al., 2003). The AUC statistic can be interpreted as the probability that a randomly selected recidivist would have a more deviant score than a randomly selected non-recidivist. This varies from .5 (no predictive accuracy) to 1.0 (perfect predictive accuracy). Typically an AUC score of .7 or above indicates significant predictive power (Craig, Beech & Harkins, 2009).

The concept of 'treatment dose' was captured using the number of hours of treatment completed. The recommendation from Marshall and Yates (2005) that treatment dose is most effectively captured using a treatment hours index was followed. This was then broken down into categories of 'low intensity' (0-75 hours), 'moderate intensity' (76-150 hours) and 'high intensity' (150 or more hours). Where a dichotomous variable for treatment dose was required for the statistical testing, treatment dose was broken down into 'low' (less than 100 hours of treatment), or 'high' (more than 100 hours of treatment).

Outcome measures

Two variables were used as outcome measures, the first of these being reconviction. The dataset used included official reconviction data that had been collected from the Home Office

Offenders Index (OI) and Police National Computer (PNC). The average follow up period for the original sample (N =322) sample was 10.9 years ($SD =17.7$ months, range 7.6 to 13.8 years). For the five year fixed follow up period, reconviction rates were 10%, 13% and 22% for sexual, sexual and/or violent and general offences respectively. For the fixed ten year follow up period, reconviction rates were 21%, 30% and 51% for sexual, sexual and/or violent, and general offending respectively. For Group 1, all 250 men were followed for at least five years and 183 were followed for at least ten years. For Group 2, all 98 men were followed for at least 5 years, with 64 being followed for at least 10 years.

The decision to use reconviction data was based on the fact that this is currently the most commonly used outcome measure in sex offender treatment evaluation research (Falshaw, Friendship, & Bates, 2003). However, it is recognised that reconviction data alone cannot provide information on either individual psychological change and/or whether an individual can be said to have successfully completed treatment. Pawson and Tilley (1994) argue that reconviction studies *per se* ignore the underlying mechanisms of change. It has been suggested that an additional way to examine the effectiveness of treatment is by examining, not only ultimate outcomes of interest (i.e., recidivism rates), but also to examine whether or not treatment brings about change (Harkins & Beech, 2007a).

The second outcome variable therefore aimed to assess whether an individual had shown significant improvement within treatment. This was measured by the concept of 'clinically significant change'. The method of clinically significant change aims to examine treatment change through ascertaining whether someone has moved from a score more likely to be found in a dysfunctional distribution of scores (e.g., child abuser attitudes) to a score more likely to be found in a functional distribution of scores (non-child abuser attitudes). This approach differs from the use of reconviction data because it does not assume that all those who participate in treatment will respond in the same way (Friendship et al., 2003).

In sum, the method allows an examination of whether the areas expected to change within treatment are observed to have changed following the completion of treatment. This methodology has been used in a number of other therapeutic schools to assess the impact of treatment (Kazdin, 2003) and has been widely documented in its implementation (e.g., Hansen & Lambert, 1996; Wise, 2004). The most reported method within the sexual offending and general forensic literature is the clinically significant and reliable change index reported by Jacobson and colleagues (Jacobson, Follette, & Revenstorf, 1984; Jacobson & Traux, 1991). Jacobson reports that in order to assess significant change, it is necessary to determine whether an individual's score, on chosen psychometric assessments after treatment, is within the cut-off point between normal and dysfunctional responding on the particular measure of interest and additionally, and whether the amount of change pre- to post-treatment is statistically reliable. When both of these criteria are met, change can be considered as clinically significant.

Calculating clinically significant change

Within this research, clinically significant change was derived from psychometric data. Responses were examined at an individual level to ascertain whether someone had moved from a score more likely to be found in a dysfunctional distribution of scores (e.g., child abuser attitudes) to a score more likely to be found in a functional distribution of responding (non-child abuser attitudes). In order to assess significant change, two things need to be evaluated: (a) a derived cut-off point between normal and dysfunctional responding on a particular measure of interest, and (b) whether that change is statistically reliable. The relevant cut-off between dysfunctional and functional responding is assessed as follows:

$$\text{cut-off} = \frac{(\text{SD}^1)(\text{MEAN}^2) + (\text{SD}^2)(\text{MEAN}^1)}{\text{SD}^1 + \text{SD}^2}$$

Where MEAN¹ and SD¹ are the mean and the standard deviation of the functional group (i.e., non-offenders) and MEAN² and SD² are the mean and the standard deviations of a dysfunctional group (i.e., child abusers).

The Reliability of Change Index (RC) has been described as:

$$\text{RC} = \frac{(\text{post-treatment}) - (\text{pre-treatment})}{S_E}$$

Here any pre - post change is significant at $p < .05$ if RC is greater than 1.96. The method of calculating S_E is as follows:

$$S_E = \text{SD}_x \sqrt{1 - r_{xx}}$$

Here, r_{xx} is the test-retest reliability of the measure and SD_x is the pre-treatment standard deviation for the measure in the offender sample. Christenson and Mendoza (1986) suggest using the equation $\sqrt{2(SE)^2}$ instead of SE because it reflects the expected difference between an individual's two scores on the same test, as a function of measurement error alone. If the reliable change index (RC), which is essentially an effect size, is greater than 1.64 then the change is significant at the .05 level for a one-tailed test.

Table 2 provides details regarding the mean and SD values each of the 12 measures utilised in this research, together with the cut-off values that an offender had to fall within in order to be judged as having scores both pre and post treatment that are indistinguishable

from non-offenders. Appendix 2 provides further information on each of the measures utilised and the variables used in the equations.

Table 2. Means, Standard Deviations and Cut-Offs for the Measures Utilised in the Study

Measures and Relevant Domain	Non-Offender or Untreated Sample Mean and SD	Pre-Mean (SD)	Post-Mean (SD)	Cut-Off Score for Functional Range
Domain 1: Sexual Interests				
MSI: Paraphilias (N =274)	^c 1.80 (1.803)	2.25 (2.80)	2.53 (3.59)	>1.96
MSI: Sexual Obsessions (N =246)	^c 2.70 (1.909)	3.47 (3.53)	3.33 (3.69)	Honest range 3-9
BAC Emotional Congruence (N =241)	^a 19.0 (10.7)	17.68 (12.41)	15.24 (12.51)	<18.39
Domain 2 : Distorted Attitudes				
MSI Justifications (N =276)	^d 5.60 (3.765)	4.59 (4.08)	3.36 (4.09)	<5.13
BAC Cognitive Distortions (N =242)	^a 13.10 (8.8)	16.38 (11.45)	9.76 (9.56)	<14.52
Victim Empathy Scale (N =121)	^a 18.0 (9.6)	34.51 (23.11)	18.87 (18.93)	<16.91
Domain 3: Socioaffective Functioning				
SRI Under Assertiveness (N =123)	^a 8.80 (5.3)	12.65 (7.95)	10.78 (7.66)	<10.34
IRI Personal Distress (N =123)	^a 7.50 (3.8)	12.21 (5.96)	10.63 (5.75)	<9.33
Locus of Control (N =123)	^b 10.96 (5.61)	15.02 (5.53)	13.64 (6.16)	<13.00
UCLA Loneliness (N =123)	^a 33.8 (7.0)	45.48 (11.63)	42.87 (12.06)	<38.19
Self-Esteem Scale (N =123)	^a 7.1 (1.1)	4.13 (2.68)	4.76 (2.70)	>6.24
MSI Cognitive Distortions (N =279)	^d 6.34 (2.416)	6.49 (3.59)	5.14 (3.53)	<6.40

a The mean and SD for were taken from a non offender sample of 81 prison officers (Beckett, Beech, Fisher, & Fordham, 1994).

b The mean and SD were taken from a non offender sample (Nowicki & Duke, 1974).

c The mean and SD were taken from a non offender sample of 40 men (Nichols & Molinder, 1984)

d The mean and SD were taken from a sample of 31 untreated sexual offenders reported in the MSI manual (Nichols & Molinder, 1984)

The questionnaires that were used to assess treatment change are listed in Table 2. For information on how each measure was assigned to each dynamic risk domain (as described by Thornton, 2002), the reader is referred to Appendix 3 where this process is described. Within the data set utilised, data was available to allow clinically significant change calculations to be carried out for the measures relating to the domains of sexual interests, distorted attitudes, and socio-affective functioning. However, it was not possible to measure changes on the measures for the self-management domain as these were represented by scales from SHAPS and this was only administered at pre-treatment in the dataset available.

An overall treatment effect was calculated identically to that reported by the author of the original data set. The scores on the measures for each domain were combined. A score of one was given if there was a clinically significant change from pre to post-treatment on an individual measure or their score was within the cut off after treatment. If no clinically significant change occurred, the participant was given a score of zero. If the participant's score was clinically significantly worse, they were given a minus one for that measure. For each domain, the following guidelines determined if the participant exhibited overall change. The Sexual Interest domain incorporated three measures, thus a participant could receive a maximum score of three if they exhibited change on all three measures. The participant had to receive at least 2 out of 3 on this domain to be classified as having improved. The Distorted Attitudes domain also incorporated three measures, thus receiving a maximum score of three if they exhibited change on all three measures. The participant had to receive at least 2 out of 3 on this domain to be classified as having improved. The Socio-Affective domain incorporated six measures, where it was required that the individual demonstrate change on 4 out of the 6 measures to demonstrate overall change. Overall, each participant was required to demonstrate clinically significant change on 2 of the 3 domains to have demonstrated overall treatment change.

Data analysis

In order to analyse the data within this study, the Statistical Package for Social Sciences (SPSS; Version 17.0) was utilised. Statistical procedures employed are described as follows.

Survival analysis was employed to analyse the data. Cox regression, a form of survival analysis was chosen as this method can produce a survival function which can then predict the probability that an event of interest has occurred (in this case, re-offending) at a given time if we are able to provide information on predictor variables (in this case, risk classification and treatment dose) and the amount of time an individual has been at risk (length of follow up). This method was therefore able to assess the impact of treatment dose on recidivism whilst controlling for other explanatory variables known to also impact the outcome, such as risk level. The shape of the survival function and the regression coefficients for the predictors are then estimated from observed subjects. The model can then be applied to new cases that have measurements for the predictor variables. This method is also able to incorporate censored data (Tabachnick & Fidell, 2001). This means that information from ‘censored’ subjects, that is, those that do not experience the event of interest during the time of observation (in this case, re-offending) can contribute to the estimation of the model. This was considered important given the low base rate for sexual offence recidivism.

A logistic regression analysis was also conducted to enable RM2000 classification to be controlled for when assessing the impact of treatment dose on predicting clinically significant change. Logistic regression can be used to predict an outcome variable that is a categorical dichotomy from one or more categorical or continuous predictor variables. For logistic regression, the dependent variable is dichotomous. Logistic regression coefficients can be used to estimate odds ratios for each of the independent variables in the model. As with cox regression, predictor variables can be entered in blocks, however, it is recommended that the order in which the blocks are added should be based on previous research or theory

(Field, 2009). Logistic regression has no assumptions about the distributions of the predictor variables (Field, 2009).

Results

Test of Hypotheses 1 and 2

Hypothesis 1 aimed to explore whether there would be significant differences in the static risk assessment classifications of sexual offenders attending high, moderate or low intensity sexual offender treatment programmes. Hypotheses 2 aimed to explore whether there would be significant differences in the reconviction rates of those attending high, moderate and low intensity sexual offender treatment programmes. For these analyses the Group 1 sample was utilised. Table 3 presents results associated with the test of Hypotheses 1 and 2. A Pearson's Correlation analysis was conducted to examine the relationship between treatment dosage and static risk classification. The correlation examined the relationship between the total number of RM2000 points and the total hours of treatment received for each participant. A significant relationship was not observed between treatment dose and actuarial risk classification ($r = .028$, $df = 248$). Hypothesis 1 was rejected.

To further explore the area of treatment dose and risk classification, a $3 \times 4 \chi^2$ analysis was conducted to examine if significant differences existed between the RM2000 classifications (low, medium, high or very high risk) of those attending short, moderate or high intensity treatment programmes. The assumptions of χ^2 (i.e., each person falls into only one cell of the contingency table and expected cell frequencies are greater than 5) were not violated (Dancey & Reidy, 2002). No significant differences were found to exist between the risk classifications of those attending different lengths of treatment ($\chi^2 = 10.47$, $df = 5$, $p = .106$). In a $3 \times 2 \chi^2$, the differences between the number of sexual reconvictions (any sexual reconviction over 10 years) across different treatment lengths were not significant ($\chi^2 = .97$,

$df = 2, p = .616$). Hypotheses 2 was rejected. As expected, significant differences were observed between the sexual recidivism rates according to RM2000 classification ($\chi^2 = 32.20, df = 3, p < .05$). The percentages associated with these results can be seen in Table 3.

Table 3. Percentage of Sexual Reconvictions over a Ten-Year Period by Treatment Length and Risk Category

Treatment Dose	Risk Matrix 2000 Category			
	Low	Medium	High	Very high
Low intensity (N =16/104)	8.8% (N =3/34)	6.2% (N =3/48)	36.3% (N =4/11)	54.5% (N =6/11)
Moderate intensity (N =16/86)	12.5% (N =3/24)	16.2% (N =6/37)	23.5% (N =4/17)	37.5% (N =3/8)
High intensity (N =6/60)	7.4% (N =2/27)	5.2% (N =1/19)	9.1% (N =1/11)	66.6% (N =2/3)

Test of Hypothesis 3

Hypothesis 3 aimed to explore whether, whilst controlling for static risk category, treatment dose would significantly contribute to the prediction of sexual reconviction in sexual offenders. To explore this hypothesis, a cox regression was undertaken. There is no reason to suspect that the assumption of proportional hazards was violated in this sample (Norušis, 2005). This assumes that the observations should be independent, and the hazard ratio should be constant across time; that is, the proportionality of hazards from one case to another

should not vary over time (Walters, 2009). Within the cox regression undertaken, the predictor variables of RM2000 classification and treatment dose (total number of treatment hours) were added in a hierarchical manner, where each additional block was tested for significance having considered the earlier block. It is recommended that the order in which blocks are added is based upon theory (Field, 2009). Given the well established link in research between actuarial risk level and sexual reconviction (e.g., Thornton et al., 2003) it was decided that RM2000 risk classification would be added as the first block.

Table 4. Cox Regression Analysis using Risk Matrix 2000 and Treatment Dose to Predict Sexual Recidivism

	B	SE	eB	p-value	95% confidence interval for eB	
					Lower	Upper
Block 1	-2.098*	.458	21.003	.000	.050	.301
RM2000						
Block 2	-2.100*	.458	21.016	.000	.050	.301
RM2000						
Treatment Dose	.000	.001	.017	.896	.997	1.003

Note: $\chi^2(3) = 40.897$ at Block 1, $P < .001$, $\chi^2(4) = 40.912$ at Block 2, $P < .001$

* $p < .0001$

From Table 4 it can be seen that RM2000 score was shown to be a significant predictor of sexual recidivism. When treatment dose was added as a second block, it was not found to make a significant contribution to predicting sexual recidivism after controlling for RM2000 score. Hypothesis 3 was therefore rejected.

Test of Hypothesis 4

Hypothesis 4 aimed to explore if the proportion of sexual offenders demonstrating significant changes within treatment would vary according to the dose of treatment received (short, moderate or high intensity). To test this hypothesis the Group 2 sample was utilised. In order to explore the area of treatment dose and treatment success as measured by clinically significant change, a series of χ^2 contingency tables were developed. The assumptions of χ^2 (i.e., each person falls into only one cell of the contingency table and expected cell frequencies are greater than 5) were not violated (Dancey & Reidy, 2002). No significant differences were found to exist between those attending short or long treatment (less than 100 hours or greater than 100 hours) and those successfully completing treatment according to the clinically significant change method ($\chi^2 = .90, df = 1, p = .764$). The percentages associated with these results are presented in Table 5, which also provides information on the percentages of participants said to be successfully treated according to RM2000 category. Hypothesis 4 was therefore rejected.

Table 5. Percentage of Sexual Offenders Demonstrating Clinically Significant Change at the End of Treatment According to Treatment Length and Risk Matrix Category

Treatment Length	Risk Matrix 2000 Category			
	Low	Medium	High	Very high
Low intensity	14.2%	31.2%	0%	0%
	(N =2/14)	(N =5/16)	(N =0/0)	(N =0/0)
Moderate intensity	46.6%	37.5%	0%	50%
	(N =7/15)	(N =6/16)	(N =0/3)	(N =1/2)
High intensity	58.3%	7.14%	50%	50%
	(N =7/12)	(N =1/14)	(N =2/4)	(N =1/2)

Test of Hypothesis 5

Hypothesis 5 aimed to explore if, whilst controlling for static risk category, treatment dose would significantly contribute to the prediction of whether a sexual offender would achieve clinically significant change. To test this hypothesis the Group 2 sample was utilised. In order to investigate if treatment dosage could predict whether a participant would successfully complete treatment (i.e., through achieving clinically significant change), a logistic regression analysis was conducted. The data was tested for multi-collinearity using the test for tolerance in which a value of less than 0.1 highlights a potential problem (Field, 2005). Table 6 demonstrates the results from this analysis, where neither RM2000 classification nor

treatment dose were found to significantly predict clinically significant change within treatment. Hypothesis 5 was therefore rejected.

Table 6. Logistic Regression Analysis using Static Risk Classification and Treatment Dose to Predict Clinically Significant Treatment Change

	B	SE	eB	p-value	95% confidence interval for eB	
					Lower	Upper
Block 1	-.446	1.050	.640	.671	.082	5.011
Static risk						
Constant	.000	1.000	.000	1.000		
Block 2	-.663	1.092	.276	.544	.061	4.382
Static risk						
Treatment	-.002	.002	.998	.369	.995	1.002
Dose						
Constant	.438	1.123	1.550	.696		

Note: $\chi^2(3) = 2.24$, n.s at Block 1; $\chi^2(1) = .882$, n.s. at Block 2; for the final equation $\chi^2(4) = 3.122$, n.s.

The following provides a summary of the results presented above. Using a Pearson's correlation, no relationship was observed between the amount of treatment received and static risk assessment score. No significant differences were observed between the dosage of treatment allocated to participants according to their RM2000 classification. Using reconviction rates as an outcome measure, no significant differences were observed for participants who sexually re-offended according to the amount of treatment they received.

Additionally, after controlling for static risk classification, results indicated that treatment dose did not contribute significantly to predicting sexual reconviction. Using the variable of clinically significant change as an outcome measure, results did not show any significant differences for participants achieving clinically significant change according to the amount of treatment received. Finally, neither RM2000 classification nor treatment dosage contributed to the prediction of whether a participant would achieve clinically significant change. Hypotheses 1 to 5 were therefore rejected.

Discussion

The purpose of this research was to investigate the principle of Risk, taken from the Risk, Need and Responsivity principles outlined by Andrews and Bonta (2003). The current study aimed to provide a further exploration of the concept of treatment dosage, following on from the explorations into the Risk Principle provided by Mailloux et al. (2003), Marshall and Yates (2005) and Hanson et al. (2009). The study aimed to examine if the amount of treatment received by a sample of sexual offenders influenced the outcome of treatment, as measured by sexual reconviction and 'successful treatment change'. All of the tested hypotheses were rejected, each of which are explored below.

In relation to Hypothesis 1, no significant differences were identified in the RM2000 classifications of those offenders receiving different doses of treatment. What these results suggest is that for the sample utilised, it is possible that the Risk Principle was not being applied. It is therefore important to consider this possibility when discussing the outcome of the subsequent hypotheses tested. Furthermore, in relation to Hypothesis 1, the results were inconsistent with those presented by Mailloux et al. (2003), who found that there were significant differences in the static risk assessment scores of sexual offenders attending different intensities. However, Mailloux et al. utilised different risk assessment measures,

such as the PCL-R, and additionally, conceptualised treatment dose in a different manner. It is therefore possible to conclude that the relationship between treatment dose and risk may be complicated and dependent on how both variables are defined. Using treatment hours as opposed to the number of programmes completed as a measure of treatment dose, as recommended by Marshall and Yates (2005), did not support the findings of Mailloux et al. However, it is also possible that for Mailloux et al., the sample utilised was one in which the Risk Principle was being directly applied and this may therefore help understand the differences in the result presented by Mailloux et al. and the results of the current study.

With regard to Hypothesis 2, no significant differences were identified in the reconviction rates of those attending low, moderate and high intensity sexual offender treatment programmes. These results could be interpreted in a number of ways. Firstly, the results could be regarded as being consistent with those of Hanson et al. (2009), which did not support the application of the Risk Principle in increasing the effectiveness of sexual offender treatment with regard to reductions in reconviction rates. However, Hanson et al. noted that whilst the Risk Principle analyses were not statistically significant, their analysis looked at actuarial risk classification alone, and did not explore this in relation to treatment dose. Secondly, the results of the current study were also consistent with the literature on general offending interventions, where it has been argued that the Risk Principle is the principle with the least influence on treatment effectiveness (Hanson et al., 2009). However, a third possibility also requires consideration in relation to Hypotheses 1 and 2. If, as previously discussed, the results of Hypothesis 1 were reflective of the fact that the Risk Principle was not being applied for the sample utilised, it is possible that Hypothesis 2 is a reflection of the consequences of not applying the Risk Principle. In sum, Hypothesis 1 and 2, in conjunction with each other, suggest that if the Risk Principle is not applied, treatment may not result in decreased reconviction rates for sexual offenders. Thus, it is possible that

both hypotheses offer partial support for the Risk Principle, which is worthy of further exploration.

With regard to Hypotheses 3 and 5, whilst controlling for RM2000 classification, the variable of treatment dose did not significantly add to the prediction of reconviction or successful change within treatment. However, RM2000 classification was still a significant predictor of sexual reconviction in the presence of the treatment dose variable. Regarding Hypothesis 4, there were also no significant differences between the amount of participants achieving clinically significant change who received different doses of treatment. Again, if we consider that Hypothesis 1 may suggest the Risk Principle was not being applied for the sample in question, there are a number of potential implications arising from these results. In relation to Hypothesis 1, the fact that there were no significant differences between the amount of participants achieving clinically significant change across different doses of treatment, could suggest that if the Risk Principle is not being applied (Hypotheses 1), treatment may not result in increased within-treatment change for sexual offenders. Again, this hypothesis requires further exploration but could be interpreted as partial support for the need to apply the Risk Principle.

When interpreting these results, it is important to note that the sample utilised did not demonstrate significant differences in the static risk assessment classifications of sexual offenders attending high, moderate or low intensity sexual offender treatment programmes. Therefore, it is possible that the Risk Principle was not being applied. Additionally, within this sample, the amount of treatment received by a sexual offender did not add to the prediction of sexual reconviction or ‘successful change’ within treatment. Firstly, this could be interpreted as partial support for the Risk Principle, given that the results suggest that within a sample where the Risk Principle was not applied, the amount of treatment received did not impact upon either sexual reconviction or within-treatment change. Secondly, the

results could also be interpreted that the amount of treatment received by a sexual offender has no bearing on the effectiveness of treatment. However, it is important to acknowledge that the concept of 'treatment dose' (or the Risk principle) utilised within the current research did not examine this in relation to the Responsivity principle or factors associated with this. For example, the current research did not examine responsivity factors such as the way in which sexual offenders interact with and utilise the dose of treatment offered to them, or, the initial levels of motivation to engage within the sample. It would be reasonable to assume that responsivity factors are likely to have influenced the process of treatment, for example, motivation to engage in treatment has been explored as a factor which may have an impact upon treatment success (e.g., Beech & Fisher, 2002; Tierney & McCabe, 2002).

Future research would benefit from further exploration into the Risk principle (e.g., treatment dose), the Responsivity principle (e.g., participant motivation), the interaction between these two variables, and the impact they may have on treatment outcome. It would be reasonable to assume that how a sexual offender utilises the dose of treatment offered to them will be influenced by their motivation to engage or their ability to respond to the process. It is also reasonable to assume that simply offering more of something (i.e., more treatment hours) to sexual offenders that may not be effective, without considering the other variables involved, will not make it effective. It is also reasonable to assume that offering more of something that was not required in the first place (e.g. requesting an offender to repeat a module related to self-esteem when this has not been identified as a deficit) will not make treatment more effective.

Harkins and Beech (2007b) also note that treatment programmes designed to meet the same aims are possibly engaging in different processes of unequal effectiveness. Therefore it is also important to consider variables around group interaction, the change process itself and the therapeutic environment of the group. Beech and Hamilton-Giachritsis (2005) found that

significant treatment change on psychometric measures (i.e., victim empathy, emotional identification with children, cognitive distortions) was associated with group cohesiveness and the extent to which group members felt able to express themselves. It is therefore possible that process issues within sexual offender treatment warrant further exploration in relation to treatment dose, given that the concept of treatment dose alone can say very little about the quality of the treatment itself and how sexual offenders engage with the process. Ultimately, we cannot be sure that two separate sexual offenders both receiving the same amount of treatment are receiving the same thing and engaging in the same way.

The results also did not achieve significance when examining the differences in numbers of participants achieving clinically significant change across different lengths of treatment. Again, it is possible that this is related to the possibility drawn from Hypothesis 1 that these results are a consequence of the Risk Principle not being applied within this sample. Additionally, it is possible that treatment dosage is interacting with a number of other variables not measured in this research, which impact on treatment success. Of interest is the finding that in each category of RM2000 classification and treatment dose, all but one (low risk / high intensity treatment) did not see a majority (i.e., above 50%) achieve clinically significant change. This finding can be interpreted in a number of ways. Firstly, it is possible this is evidence that contradicts the findings of Mailloux et al. (2003) that we are over-prescribing to sexual offenders, given that even after a high intensity programme the highest percentage of participants achieving clinically significant change was 58%. Secondly, it is possible that the psychometric tools utilised are not measuring other facets associated with successful treatment change. Thirdly, it is necessary to recognise that the explanation may relate to the method of clinically significant change being problematic in itself, an issue which is explored further in light of the limitations of the current research.

Limitations and directions for future research

This study was limited by the number of participants in the second sample analysed, given the small number of participants for whom pre and post psychometric data was available to calculate clinically significant change. It is possible that with this reduced sample size due to missing data, there was a loss of statistical power. The ideal requirement for this type of research analysing the application of the Risk Principle would be to compare a complete cohort of offenders for whom the Risk Principle was applied, to a separate cohort of offenders for whom treatment intensity was not matched to risk (Hanson et al., 2009). Unfortunately given the scope of the current study this was not possible to undertake, however, future research of this nature may help to clarify the range of interpretations that the results of the current research are open to. It is also important to note that the sample utilised was compiled mainly of child abusers, rather than those who have offended against adults. The results are therefore likely to be more reflective of child abusers than the sexual offender population per se.

Another factor likely to have influenced the results of this study is the methods used to investigate the impact of treatment dose on treatment outcome. The use of reconviction data was supplemented with data aiming to measure treatment change, given that the use of reconviction data alone has been criticised for failing to take into account information relating to the underlying mechanisms of treatment change. However, supplementing reconviction data with information relating to clinically significant change is also not unproblematic. Beech, Fisher and Bishopp (2004) highlight a number of weaknesses associated with the use of the clinically significant change method. The method cannot be used to measure the overall effectiveness of treatment given that there is no untreated comparison group. Secondly, this method employed the use of psychometric data, and therefore consideration must be given to whether the results may have been compromised by the effectiveness of the psychometric tests themselves. The method relies on the honesty of

the subjects responses to the questionnaires and also upon the quality of the psychometric measures. This research utilised a number of measures usually applied as part of the Sex Offender Treatment Evaluation Project (STEP) Battery (Beech, Fisher & Beckett, 1998). Craig and Beech (2009) note that one of the difficulties in the use of psychometric tests such as those incorporated into the STEP Battery is that they frequently lack standardisation with appropriate norms, thus making comparisons with sexual offenders questionable. For example, many tests used to measure constructs in sexual offenders have been standardised on samples of college students (e.g., Nowicki-Strickland Internal-External Locus of Control Scale; Nowicki, 1976). Craig and Beech also note that a sample size of many hundreds is usually considered the norm from which to make accurate comparisons (Kline, 1986) a figure which the many of the STEP measure norms have not reached. Martinovich, Saunders and Howard (1996) note that any underlying psychometric problems with the measures could in fact be exacerbated by the use of the clinically significant change method.

Additionally it is important to acknowledge that the measurement of clinically significant change and the use of the reliable change index may have been affected by participant's deviancy classification prior to receiving treatment (Beech, 1998). Beech's deviancy concept, which essentially looks at how far an individual deviates from non-offender means on the STEP Battery of psychometric tests, is relevant because this may have affected how much room an individual would have had to move to in order to achieve clinically significant change. Speer (1992) notes that the more deviant the initial scores and the less reliable the instrument, the greater the regression based improvement that may occur. There may therefore have been a bias introduced into the results through failure to account for the initial deviancy classification of each participant. Common sense would tell us that the more deviant the initial score, the more room the offender has to 'move', and therefore, these individuals have an increased chance of demonstrating within-treatment change.

However, despite the criticisms of the methods used within the current study, it is also important to acknowledge that the measures employed within this study are widely used throughout the statutory agencies within the UK and when used appropriately, can provide a firm starting point from which to assess change within sexual offenders. It must however be noted that the full range of theoretical facets which discriminate sexual offenders as heterogeneous a group has not yet been necessarily established and therefore, we may be missing other potentially important factors which have yet to be identified and treated (Beech, Fisher & Bishopp, 2004). This may explain the low numbers of participants in this study who achieved clinically significant change.

As a direction for future research, it is also necessary to acknowledge some of the difficulties with the model of rehabilitation upon which this research was based. The R-N-R model (Andrews & Bonta, 2003), whilst considered to be one of the most influential models in offender rehabilitation, has been criticised for its failure to consider process issues or the role of the therapists in group treatment for sex offenders (Ward, Polaschek, & Beech, 2005), despite the research supporting the importance of these (e.g., Beech & Fordham, 1997). Additionally, the model has also been criticised for being overly restrictive in its focus on the reduction of risk factors, and ignoring the relevance of contextual or ecological factors in offender rehabilitation (Ward & Langlands, 2009). Future research into the effects of treatment dose should consider ways of evaluating how dosage and the Risk Principle are related to not only what risk factors are being eliminated (e.g., pro-offending thinking), but also how the implementation of positive goods into offenders lives (e.g., positive social support network) may assist offenders in capitalising on the amount of treatment they are offered.

Summary and conclusions

The current research aimed to assess the impact of treatment dose in relation to sexual reconviction and treatment change. Within this sample, the results may be suggestive of the consequences of not applying the Risk Principle, i.e., that without applying the Risk Principle, the amount of treatment received does not impact upon either sexual reconviction or within-treatment change. Additionally, treatment dose as a stand-alone variable did not have a significant impact on treatment outcome. It is argued that these results are not hugely surprising, as clinical practice would suggest that if we were to solely look at the amount of treatment a sexual offender received, other areas such as the nature and quality of the intervention itself would perhaps be lost. Therefore, if the amount of treatment received itself is not having an effect on treatment outcome, we should then be directed to explore this in relation to the content and quality of services that are actually being delivered. What the current research suggests is that simply increasing the dose of treatment does not necessarily address a greater level of treatment need, and subsequently, treatment dose as a concept may be more useful to consider in relation to individual case formulation.

It is argued that this study has provided a basis from which future research can expand upon, in order to investigate which factors may or may not interact with treatment dose in order to produce a positive effect of treatment. In addition to investigating treatment dose in relation to the criminogenic factors that have been reduced or ‘eliminated’ (e.g., pro-offending thinking), it is recommended that the amount of treatment received by sexual offenders is also explored in relation to the concepts associated with the Good Lives Model of offender rehabilitation (Ward & Gannon, 2006). For example, it would be useful to explore how treatment dose might encourage or interact with positive changes offenders are making within their lives. Further investigation of this question is important in terms of resource allocation but also, in terms of what practitioners can do to maximise scarce

resources and best capitalise on the hours of treatment we are currently able to offer sexual offenders. Future research should also explore treatment dose in relation to process issues and motivation to engage.

Chapter 3

Case Study

An Assessment and Low-Dose Intervention with an Internet Offender

Chapter 4

Critique and Use of a Psychometric Instrument

Risk Matrix 2000 (Thornton et al., 2003)

Introduction

Within the field of forensic psychology, professionals are often called upon to present their understanding of the risk a client may or may not pose of undertaking or repeating offending behaviour (Craig, Browne & Beech, 2008). To do this, professionals will often utilise actuarial risk assessment instruments. The field of forensic psychology has therefore witnessed an increase in the development of actuarial risk assessment instruments for sexual offenders (Beech, Fisher & Thornton, 2003). Actuarial risk assessment instruments are empirically and statistically derived tools which seek to highlight the variables predictive of recidivism and convert these into a score. This score is designed to provide objective, empirically based guidance as to the risk the offender poses of re-offending. Examples of such instruments designed to assess sexual offenders include the Rapid Risk Assessment for Sex Offence Recidivism (RRASOR; Hanson, 1997b), Static-99 (Hanson & Thornton, 2000), and Risk Matrix 2000 (Thornton, Mann, Webster, Blud, Travers, Friendship & Erikson, 2003).

Description of tool

This review will examine an actuarial risk assessment instrument known as Risk Matrix 2000 (RM2000; Thornton et al., 2003). RM2000 was developed with the aim of providing a tool to predict the risk of sexual or violent re-offending in sexual offenders. More specifically, RM2000 is a statistically derived risk classification instrument, designed for use with males who are 18 years and above, who have been convicted of a sexual offence. At least one of these offences should have been committed after the age of 16. RM2000 uses factual information about an offenders past history to develop a score, which in turn can be converted into a category that relates to the rates of reconviction for sexual or violent

offences. RM2000 scores can be used to derive one of four risk categories; low, medium, high and very high.

RM2000 consists of three scales known as RM2000 Sexual (RM2000/S), RM2000 Violent (RM2000/V) and RM2000 Combination (RM2000/C). RM2000/S is a prediction scale for sexual re-offending, and RM2000/V a prediction scale for non-sexual violent re-offending in sexual offenders. The RM2000/C is a combination of the RM2000/S and RM2000/V and predicts the risk of further sexual or non-sexual assaults. RM2000/S uses a two step assessment system. Step 1 contains three risk items (number of previous sexual appearances, number of criminal appearances and age). The sums of these factors are translated into an initial risk category. Step 2 considers four aggravating factors (any conviction for a sexual offence against a male, any conviction for a sexual offence against a stranger, any conviction for a non-contact sexual offence and single / never been married). The initial risk classification can then be amended accordingly, where the presence of two or four aggravating risk factors raises the initial risk category by one or two levels respectively. The RM2000/V scale includes the assessment of three variables (age upon release, amount of prior violence, history of burglary). This scale also yields four risk categories of low, medium, high and very high. A copy of RM2000 and the associated scales is provided in Appendix 11.

Background of RM2000

RM2000 was designed for use in the U.K. where it is used widely as part of general forensic practice. It is routinely employed within the Prison Service, Police and the National Probation Service throughout England and Wales. During the 1990's, the Prison Service, Police and National Probation Service commonly utilised Structured Anchored Clinical Judgement (SACJ; reported in Grubin, 1998). This three stage risk assessment instrument was then

revised into a reduced format (SACJ-Min; reported in Grubin, 1998) and provided the basis for the development of RM2000. The underlying premise of this redevelopment was to provide an actuarial measure that held predictive accuracy but that did not have an overly complicated scoring system. The development of RM2000 was also heavily influenced by key developments in research at the time. Those developing RM2000 referred to the work Hanson and Bussière (1998) who examined 61 follow-up studies to identify the factors most strongly related to recidivism among sexual offenders (N =23,393).

Reviewing RM2000

This critique aims to evaluate whether RM2000 is suitable in terms of meeting the overall aims of the tool, that is, to predict the risk of sexual and violent re-offending in convicted sexual offenders. This review will consider RM2000 in light of empirical research, the tools adherence to the psychometric characteristics that ideally should be held by a robust test, and most importantly within the context of risk assessment, the predictive accuracy of the measure. Essentially, since the aim of the tool lies within prediction, the quality of the research designed to assess the tools predictive accuracy should be of utmost importance (Vincent, Maney & Hart, 2009). In addition to the concept of predictive accuracy, Kline (2000) reports that a robust assessment measure with psychometric properties must firstly be theoretically driven. Kline (1986) also argues that a good psychometric test should include a data level of at least interval or ratio scale, have high levels of reliability, appropriate norms and high levels of validity. This review will address these areas in turn in order to review RM2000 and its clinical utility as a risk assessment instrument.

Theoretical basis

It is argued that a good psychometric test must be theoretically driven (Kline, 2000). Since actuarial risk prediction instruments are essentially atheoretical in nature (Craig, Browne & Beech, 2008), this aspect of the criteria associated with a good psychometric test is difficult to meet. However, this also needs to be considered in light of the fact that RM2000 was developed within a firm base of empirical evidence into static risk factors associated with re-offending in sexual offenders. Thornton et al. (2003) note that the accumulated data of 61 follow up studies (N =23,393; Hanson & Bussière, 1998) significantly contributed to the development of RM2000. Thornton et al. note that this piece of research was felt to be a complete guide to the literature and provides both precise and representative estimates of the predictive accuracy of factors associated with sexual offence recidivism (Craig et al., 2008).

In a review of research looking at static factors associated with sexual recidivism, Craig et al. (2008) note that a consistent pattern of static risk factors associated with sexual re-offending can now be reliably identified. These include static factors such as prior criminal history, prior sexual offending, and non-contact sexual offences, to name but a few. It is this firm body of empirical research (e.g., Hanson & Bussière, 1998; Thornton & Travers, 1991; Hanson & Thornton, 2000) upon which the development of RM2000 items is based. It can therefore be argued that the RM2000 items are based upon empirical research on static factors associated with sexual offender recidivism. Therefore, whilst RM2000 is essentially atheoretical in nature, its empirical bases are well grounded and defensible. However, the atheoretical nature of the RM2000 does limit the use of the measure in specific contexts. For example, RM2000 provides no guidance as to the psychological factors which underlie risk and therefore provides no indication of how risk can be reduced or when a reduction in risk has taken place (Craig et al., 2008).

Appropriate norms and standardisation

A further characteristic of a robust psychological test is that of appropriate norms, which are essential in the practical application of the measure itself (Kline, 1986). Nunnally (1970) has argued that “the collection of people who constitute a normative population is determined by the use to which the scores will be put” (pp. 69). In terms of risk assessment, norms for a scale allow one to judge both what proportion of a population would fall into the different risk categories, and also allow judgement regarding how a particular person scores on the scale relative to others (Thornton et al., 2003). Thornton et al. also note that whilst norms are central to the interpretation of many traditional psychometric tests (for example, the assessment of personality constructs), their importance is less significant for prediction scales, where the recidivism rate associated with a particular risk category is considered a necessary precondition for the development of an actuarial risk assessment measure.

In the development of norms for RM2000, Thornton et al. (2003) used a sample which they argued was representative of all those serving sentences for sexual offences in a national, UK based prison system. This sample consisted of the men released from prison in 1979 (N =429) who had been incarcerated for sexual offences. Thornton et al. note the benefits of this sample being that it is representative of sexual offenders being released from the prison system *per se*, thus avoiding the bias often associated with ‘selected’ samples, such as sexual offenders who have completed a specific treatment programme. Given the norms were developed on an untreated sample, the variable of treatment also will not have impacted upon the reported recidivism rates.

However, a number of precautions are attached to the norms reported by Thornton et al. (2003). The norms provided relate to one particular jurisdiction (England and Wales), one particular era (late 1970’s) and to the follow up period specified (19 years). Thornton et al. note that varying any of these elements is likely to impact upon the recidivism rates, and

therefore the impact upon the accuracy of RM2000 prediction. It is also noted by Thornton et al. that recidivism rates will vary according to whether a sexual offender is sentenced within the community or custody, with the norms for RM2000 reflecting the latter group. Thornton et al. therefore advise those using the RM2000 to establish their own norms according to the context within which they are referring to or working within. The norms provided by Thornton et al. are also likely to underestimate the rates of actual re-offending given the low detection rate of sexual offences. Hanson, Thornton and Price (2003) estimate the detection rate is likely to be around .10, that is, on average, 10 arrests or detections for every 100 victims of sexual offending. In sum, the norms provided by Thornton et al. highlight a general limitation of actuarial measures, in that they require practitioners to compare the profile of an individual to aggregated group data. Given that the individual may share some but not all of the qualities of the original sample (Craig et al., 2009), this can have an impact upon the predictive accuracy of the scale in question and requires consideration by those who are applying the instrument in clinical practice.

Type of data

When describing the characteristics of a robust psychometric assessment, Kline (1986) also argues that the level of data used with the scale should be interval scale at least. The raw scores of RM2000 are interval level data. Therefore, RM2000 raw scores lend themselves to parametric statistical analysis (Kline, 2000).

Validity

The concept of validity relates to whether a test measures what it has been designed to measure (Nunnally, 1970). In this case, risk of sexual or violent re-offending in sexual offenders. The different facets of validity will be referred to and assessed in turn.

Convergent validity is an indication that the tool can yield similar results to other tools which purport to measure the same construct. This can be demonstrated through correlating the measure with another measure designed to measure a similar construct. In this case, the construct in question is that of risk of re-offending in sexual offenders. Kingston, Yates, Firestone, Babchishin and Bradford (2008) assessed the convergent validity of RM2000 by assessing its correlation with Static-99 (Hanson & Thornton, 2000) and the Sex Offence Risk Appraisal Guide (SORAG; Quinsey, Harris, Rice & Cormier, 2006). All RM2000 subscales were significantly correlated with the Static-99, accounting for 25% to 58% of the variance. RM2000 was also significantly correlated with the SORAG, accounting for 37% to 50% of the variance. Large, positive correlations were identified, however, these results should also be considered in light of the fact that the three measures considered also contain similar items (Kingston et al., 2008). Further research into this area is recommended by Kingston et al.

Predictive validity is regarded as strong evidence of the overall validity of a psychometric test (Kline, 1986). It relates to the correlations obtained between the measure administered on one occasion and some later related criterion. In the field of risk assessment, there is a general consensus that the predictive accuracy of actuarial risk assessment instruments is best indexed through the Receiver Operating Characteristics (ROC) analysis' Area Under the Curve (AUC) statistic (Hanley & McNeil, 1982; Mossman, 1994; Rice & Harris, 1997; Quinsey, Harris, Rice & Cormier, 1998). This index has the advantage of not being distorted by variations in the base rate of recidivism (Beech, Fisher & Thornton, 2003). The AUC statistic can be interpreted as the probability that a randomly selected recidivist would have a more deviant score than a randomly selected non-recidivist. This varies from .50 (no predictive accuracy) to 1.0 (perfect predictive accuracy). The AUC itself is a plot of the 'hit rate' (the percentage of re-offenders correctly identified as high risk on assessment)

against the 'false alarm rate' (the percentage of those correctly identified as low risk who did not go on to re-offend) for each score on the scale in question based on a contingency table design (Craig, Beech & Harkins, 2009). Typically an AUC score of .70 or above indicates significant predictive power. It has been noted that the average predictive accuracy for actuarial instruments is an AUC score of .68, averaged over 50 findings (Hanson, Morton & Harris, 2003).

Thornton et al. (2003) validated the RM2000/S on two U.K. samples for a period of 16 years, using treated (N =647) and untreated (N =429) sexual offenders. Thornton et al. obtained AUC scores of .77 and .75, for the treated and untreated groups respectively, in predicting sexual reconviction. For the RM2000/V, two samples were followed-up over 10 years (N =311) and between 16 and 19 years (N =429). RM2000/V obtained AUC scores of .78 and .80 for the two samples respectively, in predicting non-sexual violent reconviction. Thus, for RM2000/S and RM2000/V, moderate predictive utility was demonstrated by the authors of the scales. Further evidence of the predictive validity of RM2000 has been provided within studies which additionally, have also aimed to cross validate the tool. This evidence is described below.

Cross validation involves independently testing the prediction scale on data that was not used in the construction of the scale. The AUC score is also used as a common index when cross validating an actuarial measure. Following on from the development sample, Thornton et al. (2003) analysed data from a separate cross validation sample. Sample one (N =647) consisted of sexual offenders released from custody in England and Wales who had participated in the National Sex Offender Treatment Programme. Sample two (N =429) consisted of all offenders released from prison in 1979 who had been detained for a sexual offence. Sample three (N =311) consisted of adult males released from prison in England and Wales in 1980, following a sentence of at least four years for a sexual offence. The follow up

periods for samples one, two and three were an average of 3.7 years, a maximum of 19 years and a 10 year period respectively. Samples one and two were used to cross validate the RM2000/S (AUC = .77 and .75 respectively), samples one, two and three for the RM2000/V (AUC = .80, .85 and .78 respectively) and sample one and two for the RM2000/C (AUC = .81 and .74 respectively). Thornton et al. therefore note that good predictive accuracy was demonstrated for the three RM2000 scales across samples taken from different era's and from treated and untreated groups of sexual offenders.

However, the cross validation of a measure should also be reported by those independent of the development sample and also across different settings and jurisdictions. Craisatti and Beech (2005) cross validated RM2000/S on a community sample of 80 rapists and 230 child molesters. Here, RM2000/S significantly predicted any type of recidivism (AUC = .70) or pre-offence behaviour such as victim contact (AUC = .65). Craig, Beech and Browne (2006) also cross validated the scale in a number of studies. Craig et al. cross validated RM2000/S and RM2000/V against the Sexual Violence Risk-20 (SVR-20; Boer, Hart, Kropp & Webster, 1997) and Static-99 (Hanson & Thornton, 2000), two alternative widely used actuarial risk assessment measures. Craig et al. analysed data from a U.K. based sample of 85 sexual offenders and 46 non-sexual violent offenders, using an average follow up of eight years and seven months. RM2000/S and RM2000/V consistently obtained the highest AUC indices for predicting sexual reconviction over two, five, and ten year follow up periods compared to the other measures. RM2000/V demonstrated moderate predictive accuracy for predicting sexual reconviction (AUC = .66, .68, and .65 for two, five and ten year follow ups respectively) and large predictive effects for violent recidivism over two, five, and ten year follow up periods with AUC scores of .87, .86 and .86 respectively. For RM2000/S, moderate effects were demonstrated for predicting sexual recidivism over 2, 5, and 10 years (AUC = .60, .68, and .59 respectively) and moderate effects were also reported

for predicting violent recidivism (AUC= .54, .64, .64 for 2, 5 and 10 year follow ups respectively). It was noted from these results that RM2000/V was a better predictor of sexual re-offending than RM2000/S. This research supports the use of RM2000/S and RM2000/V as risk assessment measures in predicting sexual and non-sexual violent reconvictions in sexual offenders, however, it is worth noting that RM2000/C was not reported on within this research.

When assessing cross validation, it is also important to consider whether the measure has been investigated outside of the original country from which the validation sample was taken. Looman and Abracen (2009) cross validated the RM2000 using a North American sample of 419 sexual offenders, followed up over an average period of 7.1 years. The results support the predictive validity of the all three RM2000 scales, using the largest validation sample to date outside of the UK. In predicting sexual recidivism, RM2000/S, RM2000/V and RM2000/C reached AUC scores of .66, .60 and .66 respectively. However, Looman and Abracen note that none of the RM2000 scales reached significance in the prediction of sexual recidivism in child molesters although for rapists, both RM2000/S and RM2000/V scales predicted sexual recidivism better than would be expected by chance. There would therefore appear to be a need for further research to cross validate RM2000 taking into the account the predictive accuracy across different offender types.

In a further independent cross validation of RM2000, Kingston et al. (2008) examined a sample of 351 North American community based sexual offenders for an average of 11.4 years. Preliminary support for RM2000 was again demonstrated. Moderate predictive accuracy was demonstrated for RM2000/V (AUC = .64, .70, .74 for sexual, violent and general recidivism respectively), for RM2000/S (AUC = .65, .65, .69 for sexual, violent and general recidivism respectively) and for RM2000/C (AUC = .66, .69, .73 for sexual, violent and general recidivism respectively). All RM2000 scales significantly predicted all types of

recidivism, however, Kingston et al. note that Static-99 and the SORAG evidenced superior predictive accuracy within this research. This research provides support for the predictive validity of RM2000 with an independent, North American sample of sexual offenders. It is however worth noting that Kingston et al. reported higher AUC scores for RM2000 in predicting general recidivism versus violent or sexual recidivism and therefore question whether RM2000 is a measure more suited to this type of outcome. The authors recommend further research into this area.

Reliability

Reliability is the degree of self-consistency a test holds and the degree to which a test produces repeatable results. Reliability is argued to be a prerequisite of validity (Nunnally, 1970) and when looking at traditional psychometric tests, is reliant upon three sub-constructs: internal consistency, test-retest reliability and inter-rater reliability (Kline, 1986). However, when applying these criteria to static risk assessment measures such as RM2000, the concepts of internal consistency and test-retest reliability are made redundant. For test-retest reliability, given RM2000 only assesses static, unchangeable factors, provided the factual information remains unchanged (e.g., age at commencement of risk), the test will yield the same risk category. For the concept of internal consistency, this concept is more applicable to psychometric tests in which an abstract concept or personality construct is being measured, as opposed to risk prediction instruments, in which the presence of absence of factual events is noted. These two constructs will therefore not be assessed in relation to RM2000.

Inter-rater reliability is the correlation or consistency of a subjects score as rated by two or more independent raters. This is usually measured using Pearson's correlation coefficient, a measure of the proportion of variance of one variable that is shared with the other variable. A Pearson's correlation coefficient of ± 0.30 is considered moderate, whilst

± 0.50 is considered large (Guilford, 1956). Inter-rater reliability has been reported within a number of studies. Looman and Abracen (2009) reported a Pearson's correlation of .81 within their research on a sample of 419 offenders. Knight and Thornton (2007) reported a similar correlation coefficient of .82 using RM2000/S, with at least 281 cases scored by at least two raters. The available evidence therefore suggests that RM2000 demonstrates acceptable levels of inter-rater reliability. However, to date, no information has been made available by the test authors outlining the ways in which reliability may vary according to administrator training and scoring conditions (for example, with the use of large databases). Given that RM2000 is now routinely employed within the U.K., this information is necessary in order to ensure the test is both reliable and valid.

Limitations of actuarial risk assessment measures

It is clear that within the field of sexual offender risk assessment, actuarial methods are considered as important tools in assessing an offender's level of risk (Mandeville-Norden & Beech, 2009). Although actuarial methods are clearly important in assessing an offender's level of risk, they do however pose a number of problems. Actuarial assessments, including RM2000, do not offer any guidance on the particular factors that may contribute to recidivism and therefore may not assist in attempting to reduce risk through treatment (Mandeville-Norden & Beech, 2009). Problems have also been highlighted in applying aggregate group data to an individual whose characteristics may vary from those in the original sample, as highlighted with the use of RM2000 in Chapter 3 of this thesis. Therefore, when using RM2000, norms can only be applied to individuals who share characteristics with the original cohort (Craig et al., 2008). It is also unclear at present as to the extent to which RM2000 can be applied to other offender groups such as those with learning disabilities or those for whom their offence did not exist at the time from which the original cohort was

drawn, namely internet sexual offenders. Beech et al. (2003) also note a number of difficulties with actuarial measures generally which are applicable to RM2000. These limitations include the problem that actuarial measures yield a probability, not a certainty of future recidivism, they are likely to underestimate the true level of re-offending rates, they will ignore any unusual factors relative to that particular case, they fail to account for short term, acute risk, and that they also fail to consider the factors that need to be addressed in treatment in order to reduce risk.

Thornton et al. (2003) also note difficulties with the practical application of RM2000, given that there are theoretical grounds which suggest that other measures will perform better than RM2000 in specific circumstances. RM2000 does not include direct assessment of complex psychological risk factors such as Psychopathy and other mental health diagnoses. Alternative actuarial measures such as the SORAG include the measurement of risk factors such as Psychopathy and therefore may possess advantages for certain populations, for example in the field of Mental Health, that RM2000 is likely to be less suited to.

Conclusions

RM2000 provides a scientific, empirically based tool, upon which to base estimations of the level of risk a sexual offender poses of re-offending, sexually or violently. RM2000 offers a cost effective and resource efficient way of measuring risk. It is easily scored and computerised, and therefore is easily implemented on a day to day basis. The nature of the tool also lends itself easily to further empirical research. Thornton et al. (2003) provide appropriate norms for the tool, however these are limited in breadth across different contexts and require further development. The convergent validity of RM2000 has been partially supported, however, contradictory findings from Craig et al. (2006) and Kingston et al. (2008) require clarification from further research. A number of research studies have

provided support for the predictive validity of the RM2000 scales, which has also been demonstrated within large, independent samples outside of the UK. Craig et al. (2008) do however note that independent research into RM2000 is however still limited and requires development. More specifically, further research should assess the level of validity of RM2000 from setting to setting, across different types of offender and also across the three different RM2000 scales. The reliability of the measure has been reported as part of validation research (e.g., Looman & Abracen, 2009) but has not yet been reported by the test authors in terms of how this may be affected by different variables such as training or scoring context.

Whilst in need of further development, the psychometric properties of RM2000 have been demonstrated as positive and promising for risk prediction in sexual offenders. Actuarial measures are designed to provide absolute predictions of a specified behaviour within a specified time period and are not designed to manage dynamic change based upon motivation, insight or intervention (Craig et al., 2008). It is concluded here that the potential shown by RM2000 is worthy of further research and that as with any other actuarial measure of risk, RM2000 is best utilised in conjunction with the assessment of dynamic risk factors to provide a more accurate assessment of a sexual offender's risk of re-offending.

Implications for Chapters 1, 2 and 3

It is important to acknowledge how this critique of RM2000 impacts upon the information and results presented in previous chapters. The strengths of the tool, such as its strong empirical research base, justify the use of the tool in both Chapters 2 and 3. With reference to Chapter 1, the meta-analysis did not account for the static risk classification of the sample. Whilst RM2000 arguably has robust psychometric properties, the same can be said for many other actuarial measures (e.g., Static-99). Individual research studies therefore tend not to use

the same actuarial measures when reporting initial level of static risk. This gives rise to a lack of consistency across studies and subsequently, combining these studies and including actuarial risk levels in meta-analyses of treatment efficacy may prove difficult.

With reference to Chapter 2, it is important to acknowledge that whilst RM2000 was used as a measure of actuarial risk, there are other actuarial measures which may have demonstrated a different relationship with treatment dose, for example, Static 99. This problem was highlighted within the Mailloux et al. (2003) research, which using different measures of actuarial risk, presented conflicting conclusions to the current research. This may warrant further exploration. Additionally, RM2000 does not provide a measure of dynamic risk or criminogenic need, factors which are likely to impact upon and interact with the amount of treatment an offender receives.

Chapter 3 also utilised RM2000. The single case design provides an example of current limitations of RM2000 and the validation sample. Practitioners are witnessing a rise in requests to assess both those who have sexually offended on the internet, and in requests to work with those at the preconviction stage. At present, RM2000 has not been validated on or designed for use with either sample and practitioners therefore have to interpret actuarial measures with extreme caution, meaning that any conclusions drawn are limited. This may also create difficulties with forensic practice, for example, defensibility of the measure in a court setting. The problems raised with RM2000 in the single case design also highlight the difficulties noted regarding applying group data to individuals.

Chapter 5

Discussion

Thesis aims

This thesis aimed to investigate the topic of treatment efficacy for sexual offenders. There is certainly no debate regarding the importance of this question in terms of resource allocation, and the impact that this area has upon past and future victims of sexual offences. This thesis has also used methods of meta-analysis, single case design, empirical research and a psychometric critique to construct the following discussion around research into treatment efficacy. A summary of each chapter now will precede the discussion around how this thesis and its comprising chapters have contributed to the sex offender treatment efficacy debate.

Summary of findings

Chapter 1 provided a review of the sex offender treatment outcome literature using a meta-analysis design of 61 studies. Treatment outcome was explored in relation to the study design used, the treatment approach taken, and the treatment setting. Chapter 1 provided support for previous comprehensive reviews of sexual offender treatment (i.e., Hanson et al., 2002; Lösel & Schmucker, 2005), in that significant reductions were found for both sexual and general recidivism in treated versus untreated sex offenders. However, the majority of evidence was derived largely from studies of lower quality and without the protection of randomisation. The results also highlighted that the design of individual studies may impact upon the reported effectiveness of treatment. Most importantly, no significant effect of treatment was found for RCT's. In terms of treatment approach, the results also indicated a significant effect of CBT and systemic approaches in reducing sexual recidivism. Chapter 1 discussed the difficulties in sex offender treatment evaluation, the impact of study design on reported outcomes, and the possibility of treatment effects being masked through the exclusion of information relating to static and dynamic risk. Recommendations were made for the use of the CODC guidelines in treatment outcome research in order to encourage consistency.

Chapter 2 expanded upon the issues raised in Chapter 1 through an exploration of the impact of treatment dose on treatment outcome. The results did not demonstrate a relationship between the amount of treatment received and the static risk classification of the sample, which was explored as a possibility that the Risk Principle was not applied within this research sample. The results did not demonstrate a relationship between treatment dose and treatment outcome, as measured by sexual reconviction and within-treatment change. The results were explored in light of the possibility that the results were reflective of the consequences of not applying the Risk Principle, i.e., that if the Risk Principle is not applied, the amount of treatment received will not impact upon sexual reconviction or within-treatment change. In addition to offering partial support for the application of the Risk Principle, Chapter 2 discussed the difficulties in measuring treatment outcome, the limitations of the R-N-R model, and the importance of considering alternative factors that might influence upon how an individual interacts with the amount of treatment provided.

Chapter 3 explored the efficacy of sex offender treatment using a practice-based, single case design. This chapter outlined assessment and intervention with an internet offender who was at the pre-conviction stage. Whilst single case designs cannot be representative of a heterogeneous sex offender population, Chapter 3 highlighted a number of pertinent issues arising in the treatment of sexual offenders. The case study incorporated discussion around the heterogeneity of the sex offender population and its subgroups, the importance of individual case formulation, the impact of motivation to engage on treatment outcome and the relevance of the Good Lives model of offender rehabilitation in treatment.

Chapter 4 provided a description and critique of RM2000, an actuarial risk assessment tool that had been commonly referred to and utilised within preceding chapters. The limitations of the measure were acknowledged, alongside recognition of the developing base of empirical research that RM2000 is supported by.

Theoretical and practical applications

Arguably, the findings within this thesis have a number of implications within the sex offender treatment efficacy debate. These will now be discussed, and followed by an acknowledgement of some of the limitations of this thesis.

Chapter 1 lends support for the ‘cautious optimism’ previously reported regarding sexual offender treatment efficacy. More specifically, the meta-analysis provided support for the developing evidence base that reports treated sexual offenders are significantly less likely to re-offend than untreated sexual offenders (Hanson et al., 2002; Lösel & Schmucker, 2005). However, whilst meta-analyses can demonstrate evidence that is cautiously supportive of sex offender treatment, they provide little information regarding what treatment approach it is we are being optimistic about, and for which individuals this optimism is relevant to. Completion of large scale evaluations usually equates to focussing on a range of very different sexual offenders who are likely to be participating in different treatment programmes across the world. It is therefore argued that these studies, no matter how large the sample, only extend as far to tell us that some treatment is likely to work with some sexual offenders.

This thesis also considered the Risk principle with regard to the amount of treatment received by sexual offenders, and the impact of this on treatment outcome. The fact that no significant differences were found between the RM2000 classifications of the receiving different doses of treatment raises the possibility that the Risk principle was not being applied in the sample utilised. This in turn raises the possibility that if the Risk principle was not applied, the amount of treatment received did not impact upon recidivism or within-treatment change. This lends partial support to the Risk principle and warrants further exploration. Further exploration could also expand upon on the work of Hanson et al. (2009) in terms of continuing to examine if studies are adhering to the RNR principles, and crucially, the impact of adherence or non-adherence upon treatment outcome. Additionally, the finding that that

treatment dose itself was not related to reconviction or within-treatment change lends support to the argument that we need to further explore who is receiving this ‘dose’ of treatment and what this ‘dose’ consists of. As with Chapter 1, perhaps looking at treatment *per se*, and exploring the impact of generic treatment programmes for a vast range of offenders, has masked the impact that the amount of treatment received may have on particular individuals. The research undertaken in Chapters 1 and 2 is not dissimilar to many other studies being conducted in area of sex offender treatment efficacy at present. Namely, that there is a heavy focus on the concept of ‘treatment’ itself and also a focus on dichotomous outcomes (i.e., re-offending), but little attention is paid to the actual content of this treatment and the process of individual change. Brown (2005) notes that the lack of studies reporting on treatment content itself its both “surprising” and “worrying” (pp. 228). Both Chapters 1 and 2 suggest that as researchers, perhaps we have become so focussed on researching the question of ‘does treatment work’, that the concepts of what treatment is comprised of, and who is receiving this, have perhaps been lost.

Chapter 3 aimed to readdress the balance using a single case design to explore some of the above issues. Theory of sexual offending suggests there are multiple pathways to sexual offending (e.g., Ward & Siegert, 2002), and therefore sexual offenders will commit a range of different offences and have a range of different treatment needs. Therefore, within Chapter 3, the heterogeneity of the sexual offender population was exemplified, as was the need to be responsive to the individual needs of those we are treating, through individualised assessments and case formulation. Chapter 3 also highlighted specific examples of those issues raised in Chapters 1 and 2, such as the need to account for motivation to engage and the need to look at what works with whom. It is possible that after sentencing, as an internet offender, P will be requested to complete the i-SOTP. Modules included on the i-SOTP such as those addressing intimacy deficits and relationship skills are would be likely to be highly

beneficial for P and contribute to the reduction of future risk of re-offending. However, other modules, for example those on collecting behaviours and compulsivity, or those addressing globalised cognitive distortions about children are likely to be less relevant for P and will not necessarily contribute to risk reduction.

Thus, when exploring treatment dose and treatment efficacy at an individual level, it is likely that separate components of the treatment programme are likely to contribute to risk reduction and 'treatment success' differently. Similarly, for P, it may be that in terms of dose, if we were to measure the effectiveness of a generic programme as whole, the effects of this may look very different versus the measurement of the effectiveness of different treatment modules, according to what needs were identified for P within the initial formulation. Marshall (2009) also notes that if a sexual offender already has a particular skill, it may in fact be counter-therapeutic for them to go through all of the exercises for these topics in a manualised programme.

Chapter 4 provided a review and critique of RM2000 (Thornton et al., 2003). This Chapter provided an example of an actuarial measure which if used reliably and consistently, can aid sexual offender treatment research. However, issues around the suitability of RM2000 for a heterogeneous sex offender population were raised.

In light of the above discussions, it is argued that the contents of this thesis embody and exemplify a number of difficulties and tensions within the sex offender treatment research field at present. These tensions are argued to present significant blocks to researchers and practitioners trying to answer the question of 'does sex offender treatment work?'. These will now be highlighted and followed with recommendations for further research.

The first tension identified within this thesis relates to the conflicting concepts of sound research versus good clinical practice. The empirical research within Chapters 1 and 2

tells us that consistency in approach (e.g., the use of treatment manuals) would make for methodologically sound research and would allow conclusions to be more definitive. Clearly, if all clinicians were delivering the same thing, researchers could be much clearer about what they are measuring. However, a practice-based example outlined in Chapter 3 highlights the needs to adapt the length and content of our treatment to suit the individual needs and heterogeneity of the sex offender population. Therefore, it is noted that this thesis has highlighted existing tensions between the concepts of methodological soundness (the need for programme integrity as an aid to empirical research) and the responsivity principle (adapting treatment content and style to individual needs).

These identified tensions are in fact mirrored within a recent debate from Marshall (2009), and Mann (2009), around the concept of manualisation. Mann argues that “treatment fidelity...is essential for the evaluation of an intervention” (p.123) and proposes the way to ensure this is through the manualisation of programmes. However, Marshall argues that the restrictions imposed by a manual do not allow for the flexibility required to respond to the individual needs of sexual offenders.

Marshall’s (2009) argument around the need to respond to offenders’ individual treatment requirements also highlights difficulties with the concept of ‘sexual offender’ often used in the question ‘does sex offender treatment work?’. Chapter 3 highlights not only the diversity of the behaviour that the term ‘sexual offender’ encompasses, but also the diverse nature of the pathways into offending behaviour (e.g. Ward & Siegert, 2002) and the impact that offender’s personality characteristics can have upon treatment efficacy. It is therefore recommended that future research focus more explicitly on breaking down this question to identify which programmes are effective with which offenders. It is argued that using the generic term ‘sexual offender’ in evaluation research, by its very nature, will only enable

researchers to draw limited conclusions regarding this population unless the term is broken down to account for the heterogeneous nature of this population.

The second key tension identified within this thesis focuses around the concept of 'treatment' itself. It is argued that similarly to the term 'sexual offender', attempts to measure the term 'treatment' are stunted by the fact that this term is far too large in size to conceptualise and measure. It is recommended that the term 'treatment' itself, in a research sense, be reconsidered. Of utmost importance is the recognition within future research that 'treatment' itself is actually made up of smaller components (Shaughnessy & Zechmeister, 1997). At present, these smaller components are rarely evaluated and are often masked within larger treatment outcome studies such as those in Chapter 1. It is possible that the lack of attention to individual treatment components has stemmed from the difficulties in measuring the different goals of these modules. Both Chapters 2 and 3 highlighted the difficulties with the current battery of psychometric measures used to measure treatment change in sexual offenders. The measures are often transparent (Brown, 2005) and are also underdeveloped for their use with different sub-populations of sexual offenders, such as internet offenders. Therefore, in addition to the recognition of treatment as a package of individual components, it is also recommended that future research focus on the development of reliable and valid ways in which to measure key treatment goals. As Chapter 4 highlighted, a great deal of time and effort has been devoted to the development of sound actuarial measures for sexual offenders to measure static risk factors, however, the measurement of dynamic risk factors (e.g., through psychometric assessment) still requires a great deal of research to enable us to rely on such measures.

It is argued that by using this approach, the consistency required for evaluation could still be applied through the use of a manual or 'treatment guide', that is, a manual from which treatment providers can choose relevant modules according to the formulation of their clients

treatment needs. Thus, the focus on individual treatment components would also allow treatment programmes to be more flexible and responsive to the individual needs identified from the case formulation and assessment. Such an approach would enable a broader range of questions to be answered, such as what components of treatment best impact upon treatment change, how different treatment components impact upon the reduction of dynamic risk, and how each treatment component is related to re-offending. It is also argued that through reducing the size of evaluations to individual treatment modules, researchers would have more room to evaluate the impact of other factors believed to impact upon sex offender treatment efficacy, such as group cohesion, therapist characteristics, style of treatment delivery and motivation to engage.

The third and final tension identified within this thesis focuses around the concept of ‘treatment efficacy’, that is, ‘does sex offender treatment *work*’. The concept of treatment efficacy is often defined through the use of reconviction data, which by its very nature (i.e., dichotomous data), can tell us nothing about why an intervention has or has not worked, or indeed, how well it has worked. It is argued that expanding the concept of treatment ‘working’ to look further at how treatment enables sexual offenders to make positive changes in their lives (that ultimately also reduce the likelihood of re-offending) would be of benefit to the field. In line with this, it is recommended that sexual offender treatment evaluation utilise key principles of the Good Lives model of offender rehabilitation within its evaluation systems (e.g., Ward, Mann & Gannon, 2007).

For example, Chapter 2 highlighted the fact that the empirical research conducted may have been more meaningful if it was examined in relation to how the dose of treatment received is utilised outside of the treatment room, that is, how offenders are using the treatment offered to implement positive changes in their lives. Additionally, client P in Chapter 3 may have responded very differently to the intervention offered if he had access to

a positive social support network or meaningful activities outside of treatment. Measuring reconviction data alone, which is common practice within sexual offender treatment research, ignores important factors associated with how a sexual offender might respond to the treatment offered, such as their strengths, resiliencies and access to human goods⁹ outside of treatment sessions. For client P, he had received information around important treatment goals for him such as intimacy, social skills and relationships, but within his life outside of the treatment room, had limited opportunity to practice and implement these skills. Common sense would tell us that this will have an impact on the likelihood of P repeating his offending behaviour in the future.

It is possible that innovative projects such as Circles of Support and Accountability can offer ways of collecting such data in order to supplement reconviction data or psychometric data. In turn, it is argued that a focus on other facets of ‘treatment *efficacy*’ (i.e., looking at additional ways of researching ‘does treatment *work*’, rather than reconviction data alone) will assist in informing the sex offender treatment efficacy debate. Circles of Support and Accountability offer opportunities for sexual offenders to be linked up with a group of trained volunteers who, upon the offender being released from prison, provide them with opportunities to rehabilitate, to receive support, and to be held accountable for their actions. Additionally, volunteers are trained to help the offender make positive changes in their lives and practically apply what they have learned within treatment into real life settings. Such projects would enable subtle data to be collected which may inform researchers about how life *after* treatment might impact upon treatment efficacy itself. However, with this recommendation also comes an acknowledgement that for the general public, the idea of reducing sexual harm by assisting sexual offenders to make positive life

⁹ Ward and Stewart (2003) propose that humans naturally seek and require certain goods in order to live fulfilling and personally satisfying lives.

changes is wholly unpalatable, and therefore it is envisaged that attempts at such research may be steered by a level of political sensitivity.

Thesis limitations

It is argued that this thesis has highlighted a number of practical recommendations for further sexual offender treatment research, however, it is also important to acknowledge the limitations of the research conducted. Within Chapter 1, the limited time and resources available for data-analysis did not allow for study designs to be coded by a second reviewer. This may have introduced bias into the classification system. Additionally, it is also important to acknowledge that within Chapter 1, bias was also introduced into the research through the inclusion of studies such as those with unequal follow-up periods and those including treatment drop-outs. Whilst recognising the impact that these studies will have on the validity of the study, the inclusion of these studies did however allow an exploration of how such studies influence the reported effects of treatment.

Within Chapter 2, it is important to recognise the limiting effect of the relatively small sample size. It is also important to acknowledge the difficulties with the use of the clinically significant change method and the psychometric measures used to apply this method. However, it is also important to note that these were the most appropriate measures available at the time of this research. Chapter 3 was also limited by the use of psychometric tests, however, in this chapter, the difficulties were in applying measures that have not been designed for use with internet offenders. These difficulties were overcome by supplementing the information gathered with clinical information from client P, however, it is clear that further research is necessary to overcome this problem in clinical practice. Chapter 4, in discussing the use of the RM2000, also highlighted the limitations of the use of this measure which has been employed throughout this thesis, for example, the difficulties in applying the

measure to populations which may differ from the original development sample. Comment must also be made that the findings and discussions presented relate predominantly to those who have offended against children, as opposed to those who have offended against adults.

It is also important to acknowledge an area which this thesis does not directly explore, but which is considered to be ever-increasingly relevant to the treatment and rehabilitation of sexual offenders. In line with the Good Lives model of offender rehabilitation (e.g., Ward, Mann & Gannon, 2007), this thesis has not explored sex offender treatment in relation to how treatment efficacy may be influenced by strengths, resiliencies, and what sex offenders are putting into their lives as well as what they are reducing. Within Chapter 3, a more in-depth assessment focussed on the Good Lives Model may have helped to inform or even predict how client P would respond to and implement the intervention offered at the time.

Conclusions

We can be cautiously optimistic about the efficacy of sexual offender treatment, however, most meta-analyses at present can only tell us that some treatment works with some sexual offenders. Additionally, treatment outcome is not solely the product of how much treatment a sex offender receives. Approaches to working with different types of sexual offenders (e.g., internet offenders), require further development, especially in the area of actuarial and psychometric assessment measures. Therefore, treatment research should pay less attention to the 'does it work' question, and instead, focus upon what components of treatment work, and with whom. More attention should be paid to the content of treatment programmes and how sexual offenders engage with treatment modules according to individual need. To reduce the identified tensions between the responsivity principle and the need for methodological rigour in treatment research, it is recommended that future studies focus upon individual treatment components and the ways in which the efficacy of these are measured. In conclusion, it is

recommended that the question of ‘does sex offender treatment work’ be broken down to refocus on who we are evaluating (i.e., what type of offender), what we are evaluating (i.e., what treatment component) and to expand the use of reconviction data as our primary outcome of interest to include a focus on what changes sexual offenders are making in their lives during and after treatment.

References

Studies included in the Chapter 1 meta-analysis are marked with an asterisk. Additionally for a number of studies included within the meta-analysis original papers could not be obtained (e.g., unpublished data). In these cases, mark with two asterisks, the data provided by Hanson et al., (2002) was utilised within the meta-analysis.

Abracen, J., & Looman, J. (2004). Issues in the treatment of sexual offenders: Recent developments and directions for future research. *Aggression and Violent Behavior, 9*, 229-246.

Abracen, J., Looman, J., Mailloux, D., Serin, R., & Malcolm, B. (2005). Clarification regarding Marshall and Yates's critique of "Dosage of treatment to sexual offenders: Are we over prescribing?". *International Journal of Offender Therapy and Comparative Criminology, 49*, 225-230.

Alexander, M. (1999). Sexual offender treatment efficacy revisited. *Sexual Abuse: A Journal of Research and Treatment, 11*, 101-116.

*Allam, J. (1998). *Community-based treatment for sex offenders: An evaluation*. Birmingham, UK: University of Birmingham.

*Allam, J. (1999). *Effective practice in work with sex offenders: A re-conviction study comparing treated and untreated sex offenders*. West Midlands Probation Service Sex Offender Unit. Available from Jayne Allam WMPS SOU, 826 Bristol Road, Selly Oak, Birmingham, West Midlands, B29 6NA.

Allam, J. (2000). *Community-based treatment for child sex offenders: An evaluation*. Unpublished doctoral dissertation, University of Birmingham, U.K.

Allam, J.M. & Browne, K.D. (1998). Evaluating community based treatment programmes for men who sexually abuse children. *Child Abuse Review, 7*, 13-29.

- Andrews, D.A. (1989). Recidivism is predictable and can be influenced: Using risk assessments to reduce recidivism. *Forum of Correction Research (online)*, 1(2), 11-18. Available from: http://www.csc-scc.gc.ca/text/pblct/forum/e012/12j_e.pdf.
- Andrews, D.A. & Bonta, J. (2003). *The psychology of criminal conduct* (3rd Ed.). Cincinnati, OH: Anderson
- Andrews, D.A. & Bonta, J. (2006). *The psychology of criminal conduct* (4th Ed.). Cincinnati, OH: Anderson.
- Andrews, D.A., Bonta, J., & Hogue, R.D. (1990). Classification for effective rehabilitation. *Criminal Justice and Behavior*, 17, 19-52.
- Andrews, D. A., & Dowden, C. (2006). Risk principle of case classification in correctional treatment. *International Journal of Offender Therapy and Comparative Criminology*, 50, 88–100.
- Andrews, D. A., Dowden, C., & Gendreau, P. (1999). *Clinically relevant and psychologically informed approaches to reducing criminal recidivism: A meta-analytic study of human service, risk, need, responsivity and other concerns in justice contexts*. Unpublished manuscript.
- *Aytes, K.E, Olsen, S.S., Zakrajsek, T., Murray, P., & Ireson, R. (2001). Cognitive behavioral treatment for sexual offenders: An examination of recidivism. *Sexual Abuse: A Journal of Research and Treatment*, 13, 223-231.
- **Bakker, L., Hudson, S., Wales, D., & Riley, D. (1999). *“And there was light”*: An evaluation of the Kia Marama Treatment Programme for New Zealand sex offenders against children. Christchurch, NZ, Psychological Service, Department of Corrections New Zealand.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191-215.

- *Barnes, J.M., & Peterson, K.D. (1997). *The Kentucky Sex Offender Treatment Program*. Commonwealth of Kentucky, Justice Cabinet, Department of Corrections, Division of Mental Health. Kentucky State Reformatory, Kentucky.
- Bartholomew, K., & Horowitz, L. M. (1991). Attachment styles among young adults: A test of a four-category model. *Journal of Personality and Social Psychology*, 61, 226-244.
- Beckett, R.C. (1987). *The children and sex questionnaire*. Available from Forensic Psychology Services, 17 Piplely Furlong, Littlemore, Oxford, UK, OX4 4JW.
- Beckett, R.C., Beech, A., Fisher, D. & Fordham, A.S. (1994). *Community-based treatment for sex offenders: An evaluation of seven treatment programmes*. London: Home Office Publications Unit.
- Beckett, R.C., & Fisher, D. (1994). *Assessing victim empathy: A new measure*. Paper presented at the 13th Conference of the Association for the Treatment of Sexual Abusers (ATSA), San Francisco, CA., USA.
- Beckett, R.C., Fisher, D., & Gerhold, C.K. (2000). *Internet pornography victim empathy questionnaire*. Unpublished measure. Available from Forensic Psychology Services, 17 Piplely Furlong, Littlemore, Oxford, UK, OX4 4JW.
- Beech, A.R. (1998). A psychometric typology of child abusers. *International Journal of Offender Therapy and Comparative Criminology*, 42, 319-339.
- Beech, A. R., Elliott, I. A., Birgden, A., & Findlater, D. (2008). The internet and child sexual offending: A criminological review. *Aggression and Violent Behavior*, 13, 216–228.
- Beech, A.R. & Fisher, D. (2002). The rehabilitation of child sex offenders. *Australian Psychologist*, 37, 206-214.
- Beech, A., & Fisher, D. (2004). Treatment of sex offenders in the U.K. in prison and probation settings. In H. Kemshall & G. McIvor (Eds.), *Managing sex offender risk: Research highlights*, 46 (pp. 137–164). London: Jessica Kingsley Publications.

- Beech, A.R., Fisher, D., & Beckett, R.C. (1998). *An evaluation of the Prison Sex Offender Treatment Programme*. U.K. Home Office Occasional Report. Home Office Publications Unit, 50, Queen Anne's Gate, London, SW1 9AT, England. Available electronically from www.homeoffice.gov.uk/rds/pdfs/occ-step3.pdf
- Beech, A., Fisher, D., & Bishopp, D. (2004). *Pre-post change designs*. Paper presented at the 23rd Annual Conference of the Association for the Treatment of Sexual Abusers, Albuquerque, NM, USA, October 2004.
- Beech, A.R., Fisher, D., & Thornton, D. (2003). Risk assessment of sex offenders. *Professional Psychology: Research and Practice*, 34, 339-352.
- Beech, A., & Fordham, A. S. (1997). Therapeutic climate of sexual offender treatment programs. *Sexual Abuse: A Journal of Research and Treatment*, 9, 219-237.
- Beech, A. R., & Hamilton-Giachritsis, C. E. (2005). Relationship between therapeutic climate and treatment outcome in a group-based sexual offender program. *Sexual Abuse: A Journal of Research and Treatment*, 17, 127.
- Beech, A.R., & Ward, T. (2004). The integration of etiology and risk in sex offenders: A theoretical model. *Aggression and Violent Behavior*, 10, 31-63.
- *Berlin, F.S., Hunt, W.P., Malin, M.H., Dyer, A., Lehne, G.K., & Dean, S. (1991). A five-year plus follow-up survey of criminal recidivism within a treated cohort of 406 pedophiles, 111 exhibitionists and 109 sexual aggressives: Issues and outcomes. *American Journal of Forensic Psychiatry*, 12, 5-28.
- Berliner, L. (2002). Commentary on Hanson et al. (2002). *Sexual Abuse: A Journal of Research and Treatment*, 14, 195-197.
- Bickley, J. (2004). *Comprehensive assessment framework for young people who sexually abuse*. Unpublished manuscript. Available from The Lucy Faithfull Foundation, 46-48 East Street, Epsom, Surrey, UK., KT17 1HB.

- *Bluglass, R. (1980). Indecent exposure in the West Midlands. In D. West (Ed.), *Sex offenders in the criminal justice system* (pp. 171-180). Cambridge, UK: Cambridge Institute of Criminology.
- 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*. Canada: Mental Health, Law, and Policy Institute, Simon Fraser University, Vancouver, BC.
- Bonta, J., & Andrews, D.A. (2007). *Risk-Need-Responsivity Model for offender assessment and rehabilitation*. Corrections Research, Public Safety and Emergency Preparedness Canada, Ottawa, Canada. Available from http://www.publicsafety.gc.ca/res/cor/rep/_fl/crp2007-05-en.pdf
- *Borduin, C.M., Henggeler, S.W., Blaske, D.M., & Stein, R.J. (1990). Multisystemic treatment of adolescent sexual offenders. *International Journal of Offender Therapy and Comparative Criminology*, 34, 105-113.
- *Borduin, C.M., Schaeffer, C.M., & Heilblum, N. (2000). *Multi-systemic treatment of juvenile sexual offenders: A progress report*. Paper presented at the 6th International Conference on the Treatment of Sexual Offenders, Toronto.
- *Borduin, C.M., Schaeffer, C.M., & Heilblum, N. (2009). A randomized clinical trial of multisystemic therapy with juvenile offenders: Effects on social ecology and criminal activity. *Journal of Consulting and Clinical Psychology*, 77, 26-37.
- *Bremer, J.F., (1992). Serious juvenile sex offenders: Treatment and long-term follow-up. *Psychiatric Annals*, 22, 326-332.
- Brooks-Gordon, B., Bilby, C., & Wells, H. (2006). A systematic review of psychological interventions for sexual offenders: I. Randomised controlled trials. *Journal of Forensic Psychiatry and Psychology*, 17, 442-466.
- Brown, S. (2005). *Treating sex offenders. An introduction to sex offender treatment programmes*. Devon, UK: Willan.

- Burke, A., Sowerbutts, S., Blundell, S., & Sherry, M. (2001). Child pornography and the internet: Policing and treatment issues. *Psychiatry, Psychology and Law*, 9, 79-84.
- Carich, M.S. & Calder, M.C. (2002). *Contemporary treatment of adult male sex offenders*. Lyme Regis, Dorset: Russell House Publishing.
- *Clelland, S.R., Studer, L.H., & Reddon, J.R. (1998). Follow-up report of rapists treated in a forensic psychiatric hospital. *Violence and Victims*, 13, 79-86. [Alberta Hospital].
- Collaborative Outcome Data Committee. (2007a). *Sexual offender treatment outcome research: CODC guidelines for evaluation Part 1: Introduction and overview* (Corrections Research User Report No. 2007-02). Ottawa, Ontario: Public Safety Canada.
- Collaborative Outcome Data Committee. (2007b). *The Collaborative Outcome Data Committee's guidelines for the evaluation of sexual offender treatment outcome research Part 2: CODC guidelines* (Corrections Research User Report No. 2007-03). Ottawa, Ontario: Public Safety Canada.
- *Cooper, H.M. (2000). Long term follow up of a community-based treatment program for adolescent sex offenders. *Masters Abstracts International*, 45(03). (UMI No. MR21542).
- Craig, L., & Beech, A.R. (2009). Psychometric assessment of sexual deviance. In A.R. Beech, L.A. Craig & K.D. Browne (Eds.). *Assessment and treatment of sexual offenders: A handbook* (pp. 89-108). Wiley-Blackwell: UK.
- Craig, L.A, Beech, A.R, & Harkins, L. (2009). The predictive accuracy of risk factors and frameworks. In A.R. Beech, L.A. Craig, & K.D. Browne (Eds.), *Assessment and treatment of sex offenders: A handbook* (pp. 53-76). Chichester: Wiley-Blackwell.
- Craig, L.A, Browne, K.D., & Beech, A.R. (2006). Cross validation of the Risk Matrix 2000 Sexual and Violent Scales. *Journal of Interpersonal Violence*, 21, 612-633.
- Craig, L.A, Browne, K.D., & Beech, A.R. (2008). *Assessing risk in sex offenders: A practitioner's guide*. Chichester: Wiley.

- Craig, L.A., Browne, K.D., Stringer, I., & Hogue, T.E. (2008). Sexual reconviction rates in the United Kingdom and actuarial risk estimates. *Child Abuse and Neglect*, 32, 121-138.
- Craisatti, J., & Beech, A.R. (2005). Risk prediction and failure in a complete urban sample of sex offenders. *Journal of Forensic Psychiatry and Psychology*, 16, 24-40.
- *Craissati, J., & McClurg, G. (1997). The challenge project: A treatment program evaluation for perpetrators of child sexual abuse. *Child Abuse and Neglect*, 21, 637-648.
- *CS/RESORS. (1991). *An evaluation of community sex offender programs in the Pacific Region* (No. 21803-0-A602/01-XSB). Vancouver, BC: Author.
- Dancey, C.P., & Reidy, J. (2002). *Statistics without maths for psychology using SPSS for Windows* (2nd Ed). London: Pearson Education.
- Davidson, J. (2007). *Current practice and research into internet sex offending*. Report prepared on behalf of the Risk Management Authority (Scotland). Available from <http://www.rmascotland.gov.uk/currentprojects.aspx>.
- Davis, M.H. (1980). A multi-dimensional approach to individual differences in empathy. *JSAS Catalogue of Selected Documents in Psychology*, 10, 85.
- Deeks, J.J., Altman, D.G., & Bradburn, M.J. (2001). Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In M. Egger, G.D. Smith & D.G. Altman (Eds.), *Systematic reviews in health care: Meta-analysis in context* (pp. 285-312). London: BMJ Books.
- DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7, 177-186.
- Division of Clinical Psychology (2001). *The core purpose and philosophy of the profession*. Leicester: British Psychological Society.
- *Dwyer, S.M., (1997). Treatment outcome study: Seventeen years after sexual offender treatment. *Sexual Abuse: A Journal of Research and Treatment*, 9, 149-160.

- Egger, M., Smith, G.D., & Altman, D.G. (2005). *Systematic reviews in health care: Meta-analysis in context*. (5th Ed). London: BMJ Publishing Group.
- Elliott, I. A., & Beech, A. R. (2009). Understanding online child pornography use: Applying sexual offense theory to internet offenders. *Aggression and Violent Behavior*, 14, 180-193.
- Elliott, I.A., Beech, A.R., Mandeville-Norden, & Hayes, E. (2009). Psychological profiles of internet sexual offenders: Comparisons with contact sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 21, 76–92.
- Falshaw, L., Friendship, C., & Bates, A. (2003). Sexual offenders – measuring reconviction, reoffending and recidivism. *Home Office Research Findings*, 183. Home Office Research, Development and Statistics Directorate.
- Field, A. (2005). *Discovering statistics using SPSS* (2nd Ed). London: Sage.
- Field, A. (2009). *Discovering statistics using SPSS* (3rd Ed). London: Sage.
- Finkelhor, D. (1984). *Child sexual abuse. New theory and research*. New York: Free Press.
- Fleiss, J.L. (1993). The statistical basis of meta-analysis. *Statistical Methods in Medical Research*, 2, 121-145.
- Fleiss, J.L. (1994). Measures of effect size for categorical data. In H. Cooper & L.V. Hedges (Eds.), *The handbook of research synthesis* (pp. 245-260). New York: Sage.
- Fleiss, J.L., & Gross, A.J. (1991). Meta-analysis in epidemiology, with special reference to studies of the association to environmental tobacco smoke and lung cancer: A critique. *Journal of Clinical Epidemiology*, 44, 127-139.
- *Friendship, C., Mann, R.E., & Beech, A.R. (2003). Evaluation of a national prison-based treatment program for sexual offenders in England and Wales. *Journal of Interpersonal Violence*, 18, 744-759.

- Friendship, C., & Thornton, D. (2001). Sexual reconviction rates for sexual offenders discharged from prison in England and Wales: Implications for evaluating treatment. *British Journal of Criminology*, 41, 285-292.
- Furby, L., Weinrott, M. R., & Blackshaw, L. (1989). Sex offender recidivism: A review. *Psychological Bulletin*, 105, 3-30.
- Gallagher, C.A., Wilson, D.B., Hirschfield, P., Coggeshall, M.B., & MacKenzie, D.L. (1999). A quantitative review of the effects of sex offender treatment of sexual reoffending. *Corrections Management Quarterly*, 3, 19-29.
- Gannon, T.A., Ward, T., & Collie, R. (2007). Cognitive distortions of child molesters: Theoretical and research developments over the past two decades. *Aggression and Violent Behavior*, 12, 402-416.
- Glasgow, D. (2001). *Affinity*. Available from Pacific Psychological Assessment Corporation, <http://www.pacific-psych.com>.
- **Gordon, A., & Packard, R. (1998). The impact of community maintenance treatment on sex offender recidivism. Paper presented at the 17th annual meeting of the Association for the Treatment of Sexual Abusers, Vancouver, BC. [Twin Rivers].
- Grubin, D. (1998). Sex offending against children: Understanding the risk. *Police Research Series, Paper 99*. Available from the Research and Statistics Directorate, Home Office, 50 Queen Anne's Gate, London, SW1H 9AT, England. Available electronically from: www.homeoffice.gov.uk/rds/prgpdfs/fprs99.pdf.
- Grubin, D. (2004). The risk assessment of sex offenders. In H. Kemshall & G. McIvor (Eds.), *Research highlights in social work 46: Managing sex offender risk*. London: Jessica Kingsley.
- *Guarino-Ghezzi, S., & Kimball, L.M. (1998). Juvenile sex offenders in treatment. *Corrections Management Quarterly*, 2, 45-54.

- Guilford, J.P. (1956). *Psychometric methods* (2nd Ed.). New York: McGraw-Hill.
- Hall, G.C.N. (1995a). Sexual offender recidivism revisited: A meta-analysis of recent treatment studies. *Journal of Consulting and Clinical Psychology*, 63, 802-809.
- *Hall, G.C.N. (1995b). The preliminary development of theory-based community treatment for sexual offenders. *Professional Psychology: Research and Practice*, 26, 478-483.
- Hanley, J.A., & McNeil, B.J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143, 29-36.
- Hansen, N., & Lambert, M. (1996). Clinical significance: An overview of methods. *Journal of Mental Health*, 5, 17-24.
- Hanson, R.K. (1997a). How to know what works with sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 9, 129-145.
- Hanson, R.K. (1997b). *The development of a brief actuarial scale for sex offence recidivism*. (User report No. 1997-04). Ottawa: Department of the Solicitor General of Canada. 1
- Hanson, K., Bourgon, G., Helmus, L., & Hodgson, S. (2009). *A meta-analysis of the effectiveness of treatment for sexual offenders: Risk, need, and responsivity*. Corrections Research, Public Safety Canada, Ottawa, Canada. Available from: <http://www.publicsafety.gc.ca/res/cor/rep/2009-01-trt-so-eng.aspx>
- *Hanson, R.K., Broom, I., & Stephenson, M. (2004). Evaluating community sex offender treatment programs: A 12-year follow-up of 724 offenders. *Canadian Journal of Behavioural Science*, 36, 85-94.
- Hanson, R.K. & Bussière, M.T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism Studies. *Journal of Consulting and Clinical Psychology*, 66, 348-362.

- Hanson, R.K., Gordon, A., Harris, A.J.R., Marques, J.K., Murphy, W., Quinsey, V.L., & Seto, M.C. (2002). First report of the Collaborative Outcome Data Project on the effectiveness of psychological treatment for sex offenders. *Sexual Abuse: A Journal of Research and Treatment*, 14, 169-197.
- Hanson, R. K., & Harris, A. J. R. (2000). Where should we intervene? Dynamic predictors of sex offense recidivism. *Criminal Justice and Behavior*, 27, 6–35.
- Hanson, R.K., Morton, K.E., & Harris, A.J.R. (2003). Sexual offender recidivism risk: What we know and what we need to know. *Annals of the New York Academy of Sciences*, 989, 154–166.
- Hanson, R. K., & Morton-Bourgon, K. (2004). *Predictors of sexual recidivism: An updated meta-analysis*. Corrections Research. Public Safety and Emergency Preparedness Canada, Ottawa, Canada. Available from:
http://www.publicsafety.gc.ca/res/cor/rep/_fl/2004-02-pred-se-eng.pdf
- Hanson, R.K., & Morton-Bourgon, K.E. (2007). *The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis*. Corrections Research. Public Safety and Emergency Preparedness Canada, Ottawa, Canada.. Available from:
www.publicsafety.gc.ca/res/cor/rep/_fl/crp2007-01-en.pdf
- *Hanson, R.K., & Nicholaichuk, T. (2000). A cautionary note regarding Nicholaichuk et al. (2000). *Sexual Abuse: A Journal of Research and Treatment*, 12, 289-293. [Clearwater].
- *Hanson, R.K., Steffy, R.A., & Gauthier, R. (1992). *Long-term follow-up of child molesters: Risk prediction and treatment and treatment outcome* (User Report No. 1992-02). Ottawa: Corrections Branch, Ministry of the Solicitor General of Canada.
- *Hanson, R.K., Steffy, R.A., & Gauthier, R. (1993). Long-term recidivism of child molesters. *Journal of Consulting and Clinical Psychology*, 61, 646-652.

- Hanson, R.K. & Thornton, D. (2000). Improving risk assessments for sex offenders: A comparison of three actuarial scales. *Law and Human Behavior*, 24, 119-136.
- Hanson, R.K., Thornton, D., & Price, S. (2003). *Estimating sexual recidivism rates: Observed and undetected*. Paper Presented at the ATSA Conference, October 9, 2003, St. Louis, Missouri.
- Hare, R. D. (1991). *The Hare Psychopathy Checklist-Revised*. Toronto, Canada: Multi-Health Systems.
- Harkins, L. (2004). Recidivism and within-treatment change among treated sex offenders and matched comparison subjects. *Masters Abstracts International*, 43(03), pp.993. (UMI No. MQ95168).
- Harkins, L. (2008). *Sex offender treatment: Risk, need, responsivity, and process issues*. Unpublished doctoral thesis. University of Birmingham, UK.
- Harkins, L., & Beech, A. (2007a). Measurement of the effectiveness of sexual offender treatment. *Aggression and Violent Behavior*, 12, 36-44.
- Harkins, L., & Beech, A. (2007b). A review of the factors that can influence the effectiveness of sexual offender treatment: Risk, need, responsivity and process issues. *Aggression and Violent Behavior*, 12, 615-627.
- Harris, G.T., Rice, M.E., & Quinsey, V.L. (1998). Appraisal and management of risk in sexual aggression: Implications for criminal justice policy. *Psychology, Public Policy and Law*, 4, 73-115.
- Hayes, E. & Middleton, D. (2005). *Internet Sexual Offending Treatment Programme: Theory Manual*. Crown Copyright. National Offender Management Service.
- *Hedderman, C., & Sugg, D. (1996). Does treating sex offenders reduce re-offending? Home Office *Research and Statistics Directorate: Research Findings*, 45. London: Home Office. Available from: www.homeoffice.gov.uk/rds/pdfs/r45.pdf.
- Hedges, L.V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York: Academic Press.

- *Hersh, K., Baldwin, K., & Gray-Little, B. (1999). Treatment completion and recidivism among incarcerated sex offenders. Paper presented at the annual conference of the Association for the Treatment of Sexual Abusers, Orlando, FL, September 1999.
- Howitt, D. & Sheldon, K. (2007). The role of cognitive distortions in paedophilic offending: Internet and contact offenders compared. *Psychology, Crime and Law*, 13, 469-486.
- *Huot, S. (1999). Minnesota study of sex offender recidivism. Unpublished data.
- Jacobson, N.S., Follette, W.C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy*, 15, 336-352.
- Jacobson, N.S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12-19.
- Johnstone, L. & Dallos, R. (2006). *Formulation in psychology and psychotherapy: Making sense of people's problems*. London: Routledge.
- Kazdin, A. E. (2003). Clinical significance: Measuring whether interventions make a difference. In A. E. Kazdin (Ed.). *Methodological issues and strategies in clinical research*, (3rd Ed) (pp. 691-710). Washington, DC: American Psychological Association.
- Keltner, A.A., Marshall, P.G., & Marshall, W.L. (1981). Measurement and correlation of assertiveness and social fear in a prison population. *Corrective and Social Psychiatry*, 27, 41-47.
- Kenworthy, T., Adams, C. E., Brooks-Gordon, B., & Fenton, M. (2004). *Psychological interventions for those who have sexually offended or are at risk of offending* (CD004858). Cochrane Database of Systematic Reviews, Issue 3. Chichester, UK: Wiley. Available from: <http://mrw.interscience.wiley.com/cochrane/clsysrev/articles/CD004858/frame.html>.

- Kingston, D.A., Yates, P.M., Firestone, P., Babchishin, K., & Bradford, J.M. (2008). Long-term predictive validity of the Risk Matrix 2000: A comparison with the Static-99 and the Sex Offender Risk Appraisal Guide. *Sexual Abuse: A Journal of Research and Treatment*, 20, 466-484
- Kline, P. (1986). *A handbook of psychological testing: Introduction to psychometric design*. Cambridge University Press.
- Kline, P. (2000). *The handbook of psychological testing* (2nd Ed). New York: Routledge.
- Knight, R.A. & Thornton, D. (2007). Evaluating and improving risk assessment schemes for sexual recidivism. A long term follow up of convicted sexual offenders. Available electronically from <http://www.ncjrs.gov/pdffiles1/nij/grants/217618.pdf>.
- **Kramer, S.P. (1985). *Sex offender treatment and tracking: Problems, perspectives and outcomes in the Utah criminal justice system*. Unpublished manuscript.
- *Lab, S.P, Shields, G., & Schondel, C. (1993). Research note: An evaluation of juvenile sex offender treatment. *Crime and Delinquency*, 39, 543-553.
- *Lindsay, W.R., Marshall, I., & Neilson, C. (1998). The treatment of men with a learning disability convicted of exhibitionism. *Research in Developmental Disabilities*, 19, 295-316.
- *Lindsay, W.R., Neilson, C.Q., & Morrison, F. (1998). The treatment of six men with a learning disability convicted of sex offenses with children. *British Journal of Clinical Psychology*, 37, 83-89.
- *Lindsay, W.R., & Smith, A.H.W. (1998). Response to treatment for sex offenders with intellectual disability: A comparison of men with one and two-year probation sentences. *Journal of Intellectual Disability Research*, 42, 346-353.
- Lipsey, M.W., & Wilson, D. (2001). *Practical meta-analysis*. Newbury Park, CA: Sage Publications.

- Looman, J. & Abracen, J. (2009). Comparisons of measures of risk for recidivism in sexual offenders. *Journal of Interpersonal Violence* (in press). Retrieved on 01/08/09 from <http://jiv.sagepub.com.ezproxid.bham.ac.uk/cgi/rapidpdf/0886260509336961v1?ck=nck>
- *Looman, J., Abracen, J., & Nicholaichuk, T.P. (2000). Recidivism among treated sexual offenders and matched controls: data from the Regional Treatment Centre (Ontario). *Journal of Interpersonal Violence*, 15, 279-290.
- Lösel, F., & Schmucker, M. (2005). The effectiveness of treatment for sexual offenders: A comprehensive meta-analysis. *Journal of Experimental Criminology*, 1, 117-146.
- Lucy Faithfull Foundation (2009). *'Inform Plus' Programme Manual*. The Lucy Faithfull Foundation, Nightingale house, 46-48 East Street, Epsom, Surrey, U.K., KT17 1HB.
- Mailloux, D.L., Abracen, J., Serin, R., Cousineau, C., Malcolm, B., & Looman, J. (2003). Dosage of treatment to sexual offenders: Are we over-prescribing? *International Journal of Offender Therapy and Comparative Criminology*, 47, 171-184.
- Mandeville-Norden, R. & Beech, A.R. (2009). Development of a psychometric typology of child molesters: Implications for treatment. *Journal of Interpersonal Violence*, 24, 307-325.
- Mann, R.E. (2009). Sexual offender treatment: The case for manualisation. *Journal of Sexual Aggression*, 15, 121-131.
- Mann, R.E. & Marshall, W. L. (2009). Advances in the treatment of adult incarcerated sex offenders. In A.R. Beech, L.A. Craig & K.D. Browne (Eds.). *Assessment and treatment of sexual offenders: A handbook* (pp. 329-348). Wiley-Blackwell: UK.
- Marlatt, G.A. (1989). How to handle the PIG. In: D.R. Laws. (Ed), *Relapse prevention with sex offenders* (pp. 227-235). New York: Guilford Press.
- Marques, J.K., Day, D.M., Nelson, C., & West, M.A. (1994). Effects of cognitive- behavioral treatment on sex offenders recidivism: Preliminary results of a longitudinal study. *Criminal Justice and Behavior*, 21, 28-54.

- Marques, J.K., & Murphy, W. (2004). Random assignment designs. Paper presented at the 23rd Annual Conference of the Association for the Treatment of Sexual Abusers, Albuquerque, NM, October 2004.
- *Marques, J. K., Weideranders, M., Day, D. M., Nelson, C., & van Ommeren, A. (2005). Effects of a relapse prevention program on sexual recidivism. Final results from California's Sex Offender Treatment and Evaluation Project (SOTEP). *Sexual Abuse: A Journal of Research and Treatment*, 17, 79-107.
- Marshall, W.L. (2009). Manualization: A blessing or a curse? *Journal of Sexual Aggression*, 15, 109-120.
- *Marshall, W.L., & Barbaree, H.E. (1988). The long-term evaluation of a behavioral program for child molesters. *Behavior Research and Therapy*, 26, 499-511.
- *Marshall, W.L., Eccles, A., & Barbaree, H.E. (1991). The treatment of exhibitionists: A focus on sexual deviance versus cognitive and relationship features. *Behaviour Research and Therapy*, 29, 129-135.
- Marshall, W. L., Eccles, A., & Barbaree, H. E. (1993). A three-tiered approach to the rehabilitation of incarcerated sex offenders. *Behavioral Sciences and the Law*, 11, 441-455.
- Marshall, W.L., & Marshall, L.E. (2007). The utility of the random controlled trial for evaluating sexual offender treatment: The gold standard or an inappropriate strategy? *Sexual Abuse: A Journal of Research and Treatment*, 19, 175-191.
- *Marshall, L.E., Marshall, W.L., Fernandez, Y.M., Malcolm, P.B., & Moulden, H.M. (2008). The Rockwood Preparatory Program for sexual offenders: Description and preliminary appraisal. *Sexual Abuse: A Journal of Research and Treatment*, 20, 25-42.
- Marshall, L.M & Yates, P.M. (2005). Comment on Mailloux et al.'s (2003) study "Dosage of treatment to sexual offenders: Are we over-prescribing?". *International Journal of Offender Therapy and Comparative Criminology*, 49, 221-224.

- **Martin, I. (1998). *Efficacite d'un programme cognitive-behavioral institutionnel pour delinquents sexuels*. Unpublished doctoral dissertation. Montreal: Department of Psychology, University of Montreal.
- Martinovich, A., Saunders, S., & Howard, K.I. (1996). Some comments on "Assessing Clinical Significance". *Psychotherapy Research*, 6, 124-132.
- McGonaghy, N. (1999). Methodological issues concerning evaluation of treatment for sexual offenders: randomization, treatment dropouts, untreated controls, and within-treatment studies. *Sexual Abuse: A Journal of Research and Treatment*, 11, 183-193.
- McConaughy, E. A., Prochaska, J. O., & Velcier, W. F. (1983). Stages of change in psychotherapy: Measurement and sample profiles. *Psychotherapy: Theory, Research and Practice*, 20, 368-375.
- * McGrath, R.J., Cumming, G., Livingston, J.A., & Hoke, S.E. (2003). Outcome of a treatment program for adult sex offenders: From prison to community. *Journal of Interpersonal Violence*, 18, 3-17.
- *McGrath, R.J., Hoke, S.E., & Vojtisek, J.E. (1998). Cognitive-behavioral treatment of sex offenders. *Criminal Justice and Behavior*, 25, 203-225.
- *McGuire, T.J. (2000). Correctional institution based sex offender treatment: A lapse behavior study. *Behavioral Sciences and the Law*, 18, 57-71.
- McGuire, J. (2002). Integrating findings from research reviews. In J. McGuire (Ed.), *Offender rehabilitation and treatment: Effective programs and policies to reduce reoffending* (pp. 3-38). Chichester: Wiley.
- Middleton, D., Elliott, I.A., Mandeville-Norden, R., & Beech, A.R. (2006). An investigation into the application of the Ward and Siegert Pathways Model of child sexual abuse with internet offenders. *Psychology, Crime and Law*, 12, 589-603.

- Middleton, D. & Mandeville-Norden, R. & Hayes, E. (2009). Does treatment work with internet sex offenders? Emerging findings from the Internet Sex Offender Treatment Programme (i-SOTP). *Journal of Sexual Aggression*, 15, 5-19.
- Middleton, D. & Shirley, M., (2008). *Implications for research for the assessment and treatment of internet sexual offenders*. Paper presented at the 18th International conference for the National Association for the Treatment of Abusers, Cardiff, Wales.
- Mossman, D. (1994). Assessing predictions of violence: Being accurate about accuracy. *Journal of Consulting and Clinical Psychology*, 62, 783-792.
- *Nathan, L., Wilson, N.J., & Hillman, D. (2003). *Te Whakakotahitanga: An evaluation of the Te Piriti Special Treatment Program for child sex offenders in New Zealand* (ISBN 047825201). New Zealand: Department of Corrections.
- *Nicholaichuk, T.P. (1996). Sex offender treatment priority: An illustration of the risk/need principle. *Forum on Corrections Research*, 8, 30-32.
- *Nicholaichuk, T., Gordon, A., Gu, D., & Wong, S. (2000). Outcome of an institutional sexual offender treatment program: A comparison between treated and matched untreated offenders. *Sexual Abuse: A Journal of Research and Treatment*, 12, 139-153. [Clearwater].
- Nichols, H.R., & Molinder, I. (1984). *The Multiphasic Sex Inventory Manual*. Available from Nichols and Molinder, 437 Bowes Drive, Tacoma, WA 98466, USA.
- Norušis, M.J. (2005). *SPSS 13 advanced statistical procedures companion*. Harlow, UK: Prentice Hall Publishers.
- Nowicki, S. (1976). *Adult Nowicki-Strickland Internal-External Locus of Control Scale*. Test Manual available from S.Nowicki, Jr., Department of Psychology, Emory University, Atlanta, GA 30322, USA..
- Nowicki, S., & Duke, M.P. (1974). A locus of control scale for non-college as well as college adults. *Journal of Personality Assessment*, 38, 136–137.

- Nunnally, J.C. (1970). *Introduction to psychological measurement*. New York: McGraw-Hill Publishing.
- *Nutbrown, V., & Stasiak, E. (1987). Research monograph: A retrospective analysis of O.C.I cost effectiveness 1977-1981. *Ontario Correctional Institute*, 2, 1-16.
- O'Brien, M. D., & Webster, S.D. (2007). The construction and preliminary validation of the Internet Behaviours & Attitudes Questionnaire. *Sexual Abuse: A Journal of Research and Treatment*, 19, 237-256.
- Osborn, J. & Beech, A.R. (2006). *The Suitability of Risk Matrix 2000 for Use with Internet Sex Offenders*. Unpublished M.Res Thesis, University of Birmingham, UK.
- Pawson, R. & Tilley, N. (1994). What works in evaluation research? *British Journal of Criminology*, 34, 291-306.
- Paulhus, D.L. (1999). *Paulhus Deception Scales user manual*. New York: Multi-Health Systems.
- *Perkins, D. (1987). A psychological treatment programme for sexual offenders. In B.J. McGurk, D.M. Thornton, & M. Williams (Eds.), *Applying psychology to treatment: Theory and practice* (pp. 192-217). London: Her Majesty's Stationary Office.
- Prochaska, J.O. & DiClemente, C.C. (1984). *The trans-theoretical approach: Crossing traditional boundaries of change*. Homewood, IL: Dorsey Press.
- *Proctor, E. (1996). A five-year outcome evaluation of a community-based treatment program for convicted sexual offenders run by the probation service. *Journal of Sexual Aggression*, 2, 3-16.
- **Proulx, J., Ouimet, M., Pellerin, B., Paradis, Y., Mckibben, A., & Aubut, J. (1998). *Post-treatment recidivism rates in sexual aggressors: A comparison between dropout and non-dropout subjects* (No. 98-22). Montreal, Quebec : Les Cahiers de L'ecole de Criminology, University of Montreal,.

- Quayle, E. (2009). Assessment of internet sexual abuse. In M. Calder (Ed.). *Sexual abuse assessments: Using and developing frameworks for practice*. Lyme Regis, Dorset: Russell House Publishing.
- Quayle, E. & Taylor, M. (2002). Paedophiles, pornography and the internet: Assessment issues. *British Journal of Social Work*, 32, 863-875.
- Quayle, E. & Taylor, M. (2003). Model of problematic internet use on people with a sexual interest in children. *CyberPsychology and Behavior*, 6, 93-106
- Quinsey, V.L., Harris, G.T., Rice, M.E., & Cormier, C.A. (1998). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.
- Quinsey, V.L., Harris, G.T., Rice, M.E., & Cormier, C.A. (2006). *Violent offenders: Appraising and managing risk* (2nd Ed). Washington, DC: American Psychological Association.
- Quinsey, V.L., Harris, G.T., Rice, M.E., & Lalumière, M.L. (1993). Assessing the treatment efficacy in outcome studies of sex offenders. *Journal of Interpersonal Violence*, 13, 621-644.
- **Rattenbury, F.R. (1986). The outcomes of hospitalized and incarcerated sex offenders: A study of offender types, recidivism rates, and identifying characteristics of the repeat offender. Unpublished doctoral dissertation, Chicago, IL: School of Loyola University of Chicago.
- Reitzel, L.R., & Carbonell, J.L (2006). The effectiveness of sexual offender treatment for juveniles as measured by recidivism: A meta-analysis. *Sexual Abuse: A Journal of Research and Treatment*, 18, 401-421.
- ***Report on treatment programs for sex offenders: Summary and recommendations*. (1984). In conjunction with the Governors Task Force on Sex Offenders and Their Victims, Florida Mental Health Institute, Tampa, FL. [Florida].
- Rice, M.E. & Harris, G.T. (1997). Cross-validation and extension of the Violence Risk Appraisal Guide for child molesters and rapists. *Law and Human Behavior*, 21, 231-241.

- Rice, M.E., & Harris, G.T. (1997). The treatment of adult offenders. In D.M. Stoff, J. Breiling and J.D. Maser (Eds.). *Handbook of antisocial behavior* (pp. 425-435). Toronto: Wiley.
- Rice, M.E., & Harris, G.T. (2003). The size and signs of treatment effects in sex offender therapy. In R.A. Prentky, E.S. Janus, & M.C. Seto (Eds.), *Sexually coercive behavior: Understanding and management. Annals of the New York Academy of Sciences*, 989, 428-440.
- *Rice, M.E., Quinsey, V.L., & Harris, G.T. (1991). Sexual recidivism among child molesters released from a maximum security psychiatric institution. *Journal of Consulting and Clinical Psychology*, 59, 381-386.
- *Robinson, D. (1995). *The impact of cognitive skills training on post-release recidivism among Canadian federal offenders* (No. R-41). Ottawa, OM, Correctional Service Canada, Correctional research and Development.
- *Romero, J.J., & Williams, L.M. (1983). Group psychotherapy and intensive probation supervision with sex offenders. *Federal Probation*, 47, 36-42.
- Rosenthal, R. (1997). The "file drawer" problem and tolerance for null results. *Psychological Bulletin*, 86, 638-641.
- *Ruddijs, F., & Timmerman, H. (2000). The stichting ambulante preventie projecten method: A comparative study of recidivism in first offenders in a Dutch outpatient setting. *International Journal of Offender Therapy and Comparative Criminology*, 44, 725- 739.
- Russell, D., Peplau, L.A., Cutrona, C.A. (1980). The revised UCLA Loneliness Scale: Concurrent and discriminant validity evidence. *Journal of Personality and Social Psychology*, 39, 472-480.
- *Scalora, M.J. & Garbin, C. (2003). A multivariate analysis of sex offender recidivism. *International Journal of Offender Therapy and Comparative Criminology*, 47, 309-323.
- *Schweitzer, R., & Dwyer, J. (2003). Sex crime recidivism: Evaluation of a sexual offender treatment program. *Journal of Interpersonal Violence*, 18, 1292-1310.

- Scott, N.W., McPherson, G.C., Ramsay, C.R., & Campbell, M.K., (2002). The method of minimization for allocation to clinical trials: a review. *Controlled Clinical Trials*, 23, 662-674.
- *Seager, J.A., Jellicoe, D., & Dhaliwal, G.K. (2004). Refusers, dropouts and completers: Measuring sex offender treatment efficacy. *International Journal of Offender Therapy and Comparative Criminology*, 48, 600-612.
- Segal, S.V. & Marshall, W.L. (1985). Discrepancies between self efficacy predictions and actual performance in a sample of rapists and child molesters. *Cognitive Therapy and Research*, 10, 363-376.
- Sentencing Advisory Panel (2002). *The Panel's Advice to the Court of Appeal von Offences involving Child Pornography*. Crown Copyright. Retrieved, March 21, 2009 from http://www.sentencing-guidelines.gov.uk/docs/advice_child_porn.pdf.
- *Seto, M.C., & Barbaree, H.E. (1999). Psychopathy, treatment behavior, and sex offender recidivism. *Journal of Interpersonal Violence*, 14, 1235-1248.
- Seto, M.C., Marques, J., Harris, G.T., Chaffin, M., Lalumière, M.L., Miner, M.H., Berliner, L., Rice, M.E., Leib, R., & Quinsey, V. (2008). Good science and progress in sex offender treatment ate intertwined: A response to Marshall and Marshall (2007). *Sexual Abuse: A Journal of Research and Treatment*, 20, 247-255.
- ***Sex offender programs: Program manuals, descriptions, evaluations, and research*. (1998). Regional Health Centre (Pacific). [RHC Pacific].
- ***Sex offenses in Washington State; 1998 update*. (1998). Washington State Institute for Public Policy, Olympia, WA. [Washington].
- Shaughnessy, J.T. & Zechmeister, E.B. (1997). *Research methods in psychology* (4th Ed.). New York: McGraw-Hill International Editions.
- **Smiley, W.C., & Mulloy, R. (1995). *Treatment of sex offenders: Does intervention work?* Unpublished Manuscript. [RHC Pacific].

- Smith, M.L., & Glass, G.V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 32, 752-760.
- *Song, L., & Lieb, R. (1995). Washington state sex offenders: Overview of recidivism studies. Olympia, WA: Washington State Institute for Public Policy. [Twin Rivers; Washington].
- Speer, D.C. (1992). Clinically significant change: Jacobson and Truax (1991) revisited. *Journal of consulting and Clinical Psychology*, 60, 402-408.
- *Studer, L.H., & Reddon, J.R. (1998). Treatment may change risk prediction for sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 10, 175-181. [Alberta Hospital].
- **Studer, L.H., Reddon, J.R., Roper, V., & Estrada, L. (1994). *Sexual offender recidivism: An evaluation of the Phoenix program*. Alberta Hospital, Edmonton, AB. [Alberta Hospital].
- Sturmey, P. (1996). *Functional analysis in clinical psychology*. Chichester: Wiley.
- Sullivan, J. (2001). The spiral of sexual abuse: A conceptual framework for understanding and illustrating the evolution of sexually abusive behaviour. *NOTA News*, 41, 17-21
- Sullivan, J., & Beech, A. (2004). Assessing internet offenders. In M. Calder (Ed.), *Child sexual abuse and the internet: Tackling the new frontier* (pp. 69-83). Lyme Regis, Dorset: Russell House Publishing Ltd.
- Tabachnick, B., & Fidell, L.S. (2001). *Using multivariate statistics*. (4th Ed.). London: Allyn and Bacon.
- *Taylor, R. (2000). A seven-year reconviction study of HMP Grendon Therapeutic Community. *Home Office Research, Development and Statistics Directorate. Research Findings*, 115. London: Home Office. Available to download from: www.homeoffice.gov.uk/rds/pdfs/r115.pdf.
- Terry, K.J., & Mitchell, E.W. (2001). Motivation and sex offender treatment efficacy: Leading a horse to water and making it drink? *International Journal of Offender Therapy and Comparative Criminology*, 45, 663-672.

- *Ternowski, D.R. (2004). Sex offender treatment: An evaluation of the Stave Lake Correction Centre Program. *Dissertation Abstracts International*, 66(06), 3428B. (UMI No. NR03201).
- *The Missouri Sexual Offender Program. (1989). *Updated figures for Missouri sex offender study*. Jefferson City, MO: Department of Corrections, Division of Classification and Treatment. [Missouri].
- Thornton, D. (2000). *Self-esteem Scale*. Unpublished manuscript. Available from Dr. David Thornton, Sand Ridge Secure Treatment Centre, P.O Box 700, 11 North Road, Mauston, WI 53948.
- Thornton, D. (2002). Constructing and testing a framework for dynamic risk assessment. *Sexual Abuse: A Journal of Research and Treatment*, 14(1), 139–153.
- Thornton, D. (2007). Scoring guide for Risk Matrix 2000.9/SVC. Retrieved 10th January 2009 from [http://www.cfcg.bham.ac.uk/Extras/SCORING%20GUIDE%20FOR%20RISK%20MATRIX%202000.9-%20SVC%20-%20\(ver.%20Feb%202007\).pdf](http://www.cfcg.bham.ac.uk/Extras/SCORING%20GUIDE%20FOR%20RISK%20MATRIX%202000.9-%20SVC%20-%20(ver.%20Feb%202007).pdf)
- 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 (Eds.), *Understanding and managing sexually coercive behavior, Annals of the New York Academy of Sciences*, 989, 225-235.
- Thornton, D. & Travers, R. (1991). *A longitudinal study of the criminal behaviour of convicted sexual offenders*. Proceedings of the Prison Psychologist's Conference: HM Prison Service.
- Tierney, D. W., & McCabe, M. P. (2002). Motivation for behavior change among sex offenders: A review of the literature. *Clinical Psychology Review*, 22, 113-129.
- Tierney, D.W & McCabe, M.P. (2005). The utility of the trans-theoretical model of behavior change in the treatment of sex offenders. *Sexual Abuse: A Journal of Research and Treatment*, 17, 153-170.

- Vincent, G.M., Maney, S.M., & Hart, S.D. (2009). The use of actuarial instruments in sex offenders. In F.M. Saleh, A.J. Grudzinskas, J.M. Bradford, & D.J. Brodsky (Eds.). *Sex Offenders: Identification, Risk Assessment, Treatment and Legal Issues*. New York: Oxford University Press.
- **Walker, D.W. (2000). The treatment of adult male child molesters through group family intervention. In A.P. Colmenarez, & R.H. Serreno (Eds.), *Violencia* (pp.191-197). Caracas, Venezuela: Universidad Central de Venezuela.
- Walters, F.J. (2009). *What is a Cox Model?* (2nd Ed). Hayward Medical Communications, Hayward Group Limited. Available from:
http://www.medicine.ox.ac.uk/bandolier/painres/download/whatis/COX_MODEL.pdf.
- Ward, T., & Gannon, T. (2006). Rehabilitation, etiology, and self-regulation: The good lives model of sexual offender treatment. *Aggression and Violent Behavior*, 11, 77–94.
- Ward, T., Hudson, S.M. & Keenan, T. (1998). A self-regulating model of the sexual offense process. *Sexual Abuse: A Journal of Research and Treatment*, 10, 141-157.
- Ward, T. & Keenan, T.S. (1999). Child molester's implicit theories. *Journal of Interpersonal Violence*, 14, 821-838.
- Ward, T. & Langlands, R. (2009). Repairing the rupture: Restorative justice and the rehabilitation of offenders. *Aggression and Violent Behavior*, 14, 205-214.
- Ward, T., Mann, R., & Gannon, T. (2007). The good lives model of offender rehabilitation: Clinical implications. *Aggression and Violent Behavior*, 12, 87-107.
- Ward, T., Mesler, J., & Yates, P.M. (2007). Reconstructing the Risk-Need-Responsivity model: A theoretical elaboration and evaluation. *Aggression and Violent Behavior*, 12, 208-228.
- Ward, T., Polaschek, D., & Beech, A.R (2006). *Theories of sexual offending*. Chichester: Wiley.
- Ward, T. & Siegert, R.J. (2002). Toward a comprehensive theory of child sexual abuse: A theory knitting perspective. *Psychology, Crime and Law*, 8, 319-351.

- Ward, T., & Stewart, C. (2003). Criminogenic needs and human needs: A theoretical model. *Psychology, Crime & Law*, 9, 125–143.
- Weirzbicki, M., & Pekaririk, G. (1993). A meta-analysis of psychotherapy drop-out. *Professional Psychology: Research and Practice*, 24, 190-195.
- Wessler, R.A., & Wessler, R.L. (1980). *The principles and practice of rational emotive therapy*. San-Francisco, CA: Josey Bass.
- *Wilson, R.J., Pichea, J.E., & Prinzo, M. (2005). Circles of Support and Accountability: An evaluation of the pilot project in south-central Ontario (User Report R-168). Ottawa, Ontario: Correction Service Canada.
- Wincze, J. P. (2000). Assessment and treatment of atypical sexual behavior. In S. R. Leiblum & R. C. Rosen (Eds.), *Principles and Practice of Sex Therapy* (3rd Ed.). New York: Guilford.
- Wise, E.A. (2004). Methods for analyzing psychotherapy outcomes: A review of clinical significance reliable change and recommendations for future directions. *Journal of Personality Assessment*, 82, 50-59.
- Wolf, S.C. (1985). A multi-factor model of deviant sexuality. *Victimology*, 10, 359–374
- Wood, R.M., Grossman, L.S., & Fichtner, C.G. (2000). Psychological assessment, treatment, and outcome with sex offenders. *Behavioral Sciences and the Law*, 18, 23-41.
- *Worling, J.R., & Curwen, T. (2000). Adolescent sexual offender recidivism: Success of specialized treatment and implications for risk prediction. *Child Abuse and Neglect*, 4, 965-982.

Appendices

Appendix 1: Search terms, Databases, Journals, Websites Searched

Search Terms:

The following lists contain those used within the literature search:

sex(ual) offender, rapist, child molester, pedophile, pedophilia, exhibitionist, exhibitionism, sexual assault, incest, voyeur, frotteur, indecent exposure, sexual deviant, paraphilia(c), and, treatment, outcome, recidivism, recidivist, outcome, recidivate, reoffend, reoffense, relapse, failure.

Additional search terms used for current research:

Randomised control(led) trial, Randomised clinical trial, SOTP, Intervention, STEP, SOTEP.

Databases:

PsychLit; PsychInfo; PsychNet UK; Bath Information Data Service; Biological Abstracts; The Cochrane Library; The Cochrane Register of Trials; Medline; Sociofile; National Criminal Justice Reference System; Science Direct; Ingenta; Web of Science

Journals:

Aggression and Violent Behavior; American Journal of Forensic Psychiatry; American Psychologist; Behavior Research and Therapy; Behavioral Sciences and the Law; British Journal of Clinical Psychology; British Journal of Criminology; Canadian Journal of Behavioural Science; Child Abuse and Neglect; Corrections Management Quarterly; Crime and Delinquency; Criminal Justice and Behavior; Forensic Update; Forum on Corrections Research; International Journal of Forensic Mental Health; International Journal of Offender Therapy and Comparative Criminology; Journal of Consulting and Clinical Psychology; Journal of Experimental Criminology; Journal of Forensic Psychiatry and Psychology; Journal of Intellectual Disability Research; Journal of Interpersonal Violence; Journal of Sexual Aggression; Legal and Criminological Psychology; Professional Psychology: Research and Practice; Psychology, Crime and Law; Research in Developmental Disabilities; Sexual Abuse: A Journal of Research and Treatment

Websites:

www.sexual-offender-treatment.org (Journal of the International Association for the Treatment of Sexual Offenders; IATSO). Searched 16/02/2008

www.csom.org. (Centre for Sexual offender Management – U.S Department of Justice). Searched 16/02/2008

www.csc-scc.gc.ca (Correctional Service Canada). Searched 16/02/2008

www.publicsafety.gc.ca (Public Safety Canada). Searched 16/02/2008

www.homeoffice.gov.uk (Home Office). Searched 16/02/2008

www.hmprisonservice.gov.uk (H.M.Prison Service). Searched 16/02/2008

www.cochrane.org (The Cochrane Collaboration). Search 16/02/2008

Appendix 2

The r_{xx} , SE, and $\sqrt{2(S_E)^2}$ for each measure

Where r_{xx} = the test-retest reliability of the measure and SD_x is the pre-treatment standard deviation for the measure in the offender sample.

MEASURE	SD _x	r _{xx}	SE	$\sqrt{2(S_E)^2}$
SRI Under assertiveness N =123	7.95	a.80	3.56	5.03
IRI Personal Distress N =123	5.96	b.68	3.37	4.76
Locus of Control N = 123	5.53	c.83	2.28	3.22
UCLA Emotional Loneliness N =123	11.63	a.70	6.37	9.01
Self-Esteem Scale N =123	2.68	a.75	1.34	1.89
MSI Cognitive Distortions N =279	3.59	d.84	1.44	2.04
MSI Paraphilias N =274	2.80	d.71	1.51	2.13
MSI: Sexual Obsessions N =246	3.53	d.80	1.58	2.23
BAC Emotional Congruence N =241	12.41	a.63	7.55	10.68
MSI Justifications N =276	4.08	d.78	1.91	2.70
BAC Cognitive Distortions N =242	11.45	a.77	5.49	7.76
Victim Empathy Scale N =121	23.11	a.95	5.16	7.30

aThe r_{xx} (=test-retest reliability) was taken from Beech (1998)

bThe r_{xx} (=test-retest reliability) was taken from Davis (1980)

cThe r_{xx} (=test-retest reliability) was taken from Nowicki & Duke (1974)

dThe r_{xx} (=test-retest reliability) was taken from Simkins et al. (1989)

Appendix 3: Information on how each measure was assigned to Thornton's (2002) dynamic risk domains

The following is information extracted from Harkins (2008). Chapter 2 utilised the dataset developed by Harkins (2008), hence why this dataset was used as a basis for justifying how each psychometric measure was allocated to each dynamic risk domain. The following provides a description of the process followed by Harkins (2008). Harkins (2008) utilised a Principal components analysis with Varimax rotation to examine the factor structure of the variables. These analyses were based on a subset of the total cases which had completed all of the psychometric tests (N =132). Inspection of the scree plot and eigenvalues produced by a principal components analysis suggested that a four-factor solution was optimal. Principal Axis Factoring was then used to extract four factors, which together accounted for 65.3% of the variance in the data. Varimax rotation was applied to increase the interpretability of the factor loadings. Items which have loadings greater than .4 were selected. The expected dynamic risk domains were accounted for reasonably well by the four factors. Factor 1 (accounting for 21.44% of the variance) contained measures related to the Socio-affective Domain. These measures were SHAPS Anxiety scale, SHAPS Depression scale, SRI Underassertiveness, IRI Personal Distress, Locus of Control, UCLA Emotional Loneliness, Self-esteem scale, and MSI Cognitive Distortions. Factor 2 (accounting for 17.46% of the variance) contained items related to the Self-Management Domain: SHAPS Aggression, SHAPS Impulsivity, SHAPS Psychopathic Deviate, and SHAPS Hostility. Factor 3 (accounting for 14.25% of the variance) contained measures related to the Sexual Interests Domain, particularly MSI Paraphilias, MSI Sexual Obsessions, and BACS Emotional Congruence. Items related to the Distorted Attitudes Domain loaded onto Factor 4 (accounting for 12.08% of the variance). These measures were MSI Justifications, BACS Cognitive Distortion, and Victim Empathy Scale. The list of measures used and factor loadings are included in Table 1 below.

Table 1: Factor Loadings from the Principal Components Analysis of the Psychometric Measure

Measure	Factors &	Dynamic	Risk	Domains
	Factor 1: Socioaffective	Factor 2: Self- management	Factor 3: Sexual interests	Factor 4: Distorted attitudes
SHAPS Anxiety	.79	.45		
SHAPS Depression	.72	.47		
SRI underassertiveness	.69			
IRI Personal Distress	.63			
Locus of Control	.60			
UCLA Emotional Loneliness	.57			
Self-Esteem	-.54			
MSI Cognitive Distortions	.53		.43	.42
SHAPS Aggression		.89		
SHAPS Impulsivity		.84		
SHAPS Psychopathic Deviate	.45	.67		
SHAPS Hostility	.49	.63		
MSI Paraphilias			.84	
MSI Sexual Obsessions			.80	
BACS Emotional Congruence			.71	
MSI Justifications				.78
BACS Cognitive Distortion				.77
Victim Empathy Scale				.63

Note. Factor eigenvalues were: Factor 1, 3.87; Factor 2, 3.14; Factor 3, 2.56; Factor 4, 2.17; accounting for 21.49 %, 17.46%, 14.25%, and 12.08% of the variance respectively and accounting for a total 65.29% of the variance.

Appendix 4: Consent form and information sheet

Information sheet

My name is Caroline Robertson and I am a Trainee Forensic Psychologist employed by the Lucy Faithfull Foundation. I am currently completing a Doctorate in Forensic Psychology Practice at the University of Birmingham in order to complete my training.

As part of my training, I am required to complete a 'case study' of a client I am working with at the Lucy Faithfull Foundation. This will involve completing a piece of academic work in which I write up and discuss involvement with an individual client. This will involve undertaking an initial assessment (such as questionnaires and interviews), following the clients progress through the Inform Plus programme and evaluating the clients progress upon completion of the programme. This will then be written up as an academic piece of work. Upon completion of the case study, it will be sent to the University of Birmingham for marking.

The case study will be completely anonymous and will not include any information from which the client could be identified. The case study is also entirely voluntary and is not a requirement of attendance on the Inform Plus programme. Additionally, should the client at a later point decide they do wish to participate in the case study, they are free to withdraw consent at any point. This can be done by contacting Caroline on [REDACTED]. The content of the case study will not be discussed or referred to within the Inform Plus programme unless the client should choose to do so. The client will have the option to receive feedback during and upon completion of the case study.

Confidentiality and child protection: Should any involvement with the client as part of the case study and/or the Inform Plus programme indicate that the client is a risk to themselves or to others, or that other offences have been committed that the police are not aware of, there will be a duty to inform the relevant authorities.

Consent form: Case Study

Please read this form carefully and sign at the bottom to indicate that you agree to the following:

- ✓ I have read and understood the case study information sheet
- ✓ I understand that my permission to participate in the case study is entirely voluntary and I can withdraw this consent at any point by contacting Caroline Robertson on [REDACTED].
- ✓ I understand that the case study is not a requirement of the Inform Plus programme

Thank you for your participation

Signed.....

Date.....

Appendix 5: Offence-based interview guidelines (Quayle & Taylor, 2002)

1. Action

- What is the number of total hours that the individual spends on-line in any one week and the proportion of this time that was spent in contact with others sexually interested in children or in downloading images?
- What has been the level of general disruption in their lives that being online has played – particularly in relation to work or real life social relationships?
- Has there been a reduction where appropriate with sexual interest in their partner?
- Has there been emotional withdrawal from family members or friends?
- In there a preoccupation with accessing the internet such that there are ongoing difficulties in concentrating?
- How many internet media are being accessed – chat rooms, web sites, email, newsgroups?
- What did they do with each and what level of pleasure is associated with these activities?
- What nicknames are used and what do they mean to the person?
- How is material retrieved from the internet saved and organised (in particular, how is it stored, how are fields labelled, what changes are made to existing file names?)
- How much time is spent off line with collected material, either editing or sorting or for use as an aid to masturbation?
- Have images been exchanged with others – how has this been done, what volume and what purpose did this serve?
- Have images been created through scanning from existing pictures or by digital camera?
- Have fantasies been acted out with real children (which may or may not have been of an explicitly sexual nature)?
- What are the person's social networks and levels of emotional support?
- What level of social isolation is present?
- Has there been any contact in real life with people (adults or children) met online?

2. Reflection

- What level of preoccupation is there with regard to re-living past experiences?

- How much time is spent thinking about their latest internet experience? (chat or image) or planning the next?
- Are details of other online people kept and reflected on?
- Does the person keep making promises to stop going online and then breaking them?
- Are there difficulties in concentrating on or keeping offline commitments?

3. Excitement

- Does the individual take risks in terms of accessing the material (either because of others in the house or the same room storing it)
- Have images been downloaded while children were in the room or close proximity?
- Have images been shared with others offline?
- Is there a sense of excitement in anticipation of going online or a sense of frustration or irritation when blocked from doing so?
- Does the person chat to others about real or imagined sexual encounters with children?
- Is there self representation as other individuals? (either same or other sex or age?)
- What attempts have been made to contact children through the internet?

4. Arousal

- What level of masturbation is associated with online activities?
- Does masturbation take place online or offline?
- What has been the increase or change in sexual activities since accessing the internet?
- Does the individual engage in virtual sexual relationships with others (adults or children) for example through IRC?
- Has there been a change in the kinds of texts or images accessed (age or other characteristics of the child, types of images and level of victimisation)?
- Does arousal happen to other non-child images?

Appendix 6: Summary of developmental and personal history

Appendix 7: RM2000 scoring for client P

Appendix 8: Details of psychometric assessments used in case study

1. Paulhus Deception Scale (Paulhus, 1999).

This is a psychometric assessment designed to determine the validity of responses on self-report instruments. The measure is split into two scales. The impression management subscale aims to assess for faking or lying. The self-deceptive enhancement scale aims to assess the extent to which the individual has insight. The questionnaire consists of 40 items in which respondents are asked to rate using a five-point scale. Within the field of sexual offender treatment, those who score highly on the PDS would be prone towards denial and possibly a denial of underlying motives.

2. Self-Esteem Scale (Thornton, 2000).

This eight-item questionnaire devised by Thornton (2000) was used to measure self-esteem. Offenders answer true or false to questions regarding how they feel about themselves. Thornton reports the scale has high internal reliability (Cronbach alpha 0.8). The higher the score, the higher the individual's self-esteem.

3. Social Response Inventory (Keltner, Marshall & Marshall, 1981).

The Social Response Inventory is a self-report questionnaire designed to measure assertive behaviour in a variety of social situations. The questionnaire provides a number of scenarios and then asks the respondent to pick which response they would be most likely to make. The measure consists of 22 situations with a choice of five responses covering both under and over-assertive response types. The higher the over-assertiveness score, the higher the tendency for the individual to react over-assertively in a number of social situations. The higher the under-assertiveness score, the higher the tendency for the individual to react under-assertively in a number of social situations.

4. Emotional Loneliness Questionnaire (Russell et al., 1980).

The Emotional Loneliness Questionnaire consists of 20 items where the respondent is required to indicate how often they feel in a variety of everyday situations. The authors report high internal consistency ($\alpha = .94$). The higher the score, the more emotionally lonely the individual is reporting themselves to be.

5. Internal Personal Reactivity Index (Davis, 1980).

The Interpersonal Reactivity Index is a measure of general empathy and can be used to determine if an offender lacks empathy generally or if empathy deficits are more likely to be offence/victim specific. The 28 item measure is split in four subscales: Perspective taking (the ability to assume cognitively the role of another); Empathic concern (feelings of warmth and compassion for another); Fantasy (the ability to relate to fictional characters); Personal distress (anxiety and negativity as a result of feelings of the distress of another).

6. Locus of Control Questionnaire (Nowicki, 1976).

This is a 40 item questionnaire and aims to assess the extent to which respondents believe that events are based on the way in which they behave and the extent to which they believe events are controlled externally. The higher the score, the more externally controlled the respondent believes their lives to be.

7. Victim Empathy Questionnaire (Internet Pornography; Beckett, Fisher & Gerhold, 2000).

The original Victim Empathy Questionnaire (Becket & Fisher, 1994) was designed to measure sex offender's view of the impact of offending on their victim. The scale is reported to have high internal reliability (alpha coefficient =0.9). The higher the score on this measure, the less able the offender is to demonstrate appropriate victim empathy skills. This scale was then adapted for use within internet child pornography offenders, although to date this scale has not yet been adequately normed. Caution must therefore be applied to the interpretation of this measure.

8. Children and Sex Questionnaire (Beckett, 1987).

This measure is made up of two scales, a cognitive distortions scale and an emotional congruence scale. The cognitive distortion scale is a fifteen item scale designed to assess an individual's beliefs about children and their sexuality. Questions are based on a four point scale and include items such as 'children can lead adults on'. The scales reports high internal reliability (alpha =.09). The higher the score, the greater number of cognitive distortions regarding children and sex. The emotional congruence scale contains fifteen items and assesses the extent to which an individual can understand, relate to and identify with what they believe to be the thoughts, feelings and concerns of children. The higher the score, the higher the individual's self reported emotional congruence with children.

9. Internet Behaviours and Attitudes Questionnaire (O'Brien & Webster, 2007).

The Internet Behaviours and Attitudes Questionnaire is still in development and therefore still has not been adequately normed. Following an initial pilot with 123 community-based internet offenders, the measure reported good internal consistency and factor analysis identified a two factor solution; distorted thinking and self-management. The measure is split into items which measure internet behaviours (yes/no responses) and attitudinal items measures on a 5 point scale.

10. Relationship Questionnaire (Bartholomew & Horowitz, 1991).

The Relationship Questionnaire (RQ) is a single item measure made up of four short paragraphs, each describing a prototypical attachment pattern as it applies in close adult peer relationships. Participants are asked to rate their degree of correspondence to each prototype on a 7-point scale relating to attachment styles of Secure, Fearful, Preoccupied and Dismissing. These ratings (or "scores") provide a *profile* of an individual's attachment feelings and behaviour. The RQ was designed to obtain *continuous* ratings of each of the four attachment patterns, and this is the ideal use of the measure. The RQ can also be used to categorise participants into their best fitting attachment pattern. The highest of the four attachment prototype ratings can be used to classify participants into an attachment category.

Appendix 9: Contact log with client P

Date	Contact Type	Session Outline
18/04/08	Call via Stop It now helpline (1)	Initial self referral from P after arrest. Information gathering.
22/04/08	Call via Stop It Now helpline (2)	Information gathering. Initial discussions around offending behaviour
28/04/08	Call via Stop It now helpline (3)	Information gathering. Initial discussions around offending behaviour
14/05/08	Individual session (2 hours)	Initial assessment
11/07/08	Individual session (2 hours)	Initial assessment
23/07/08	Group Session 1 (2.5 hours)	Introductions / offence analysis
30/07/08	Group Session 2 (2.5 hours)	Offence analysis
06/08/08	Individual session (2 hours)	Individual offence analysis
06/08/08	Group Session 3 (2.5 hours)	Fantasy and arousal
13/08/08	Group Session 4 (2.5 hours)	Addictions, compulsions and collecting behaviour
20/08/08	Group Session 5 (2.5 hours)	Disclosure, relationships and social skills
27/08/08	Individual session (2 hours)	Relationships and social skills
27/08/08	Group Session 6 (2.5 hours)	Relationships and social skills
03/09/08	Group Session 7 (2.5 hours)	Victim empathy
17/09/08	Group Session 8 (2.5 hours)	Relapse prevention and the legal system
17/09/08	Individual session (2 hours)	Relapse prevention and good lives
24/09/08	Group Session 9 (2.5 hours)	Relapse prevention
01/10/08	Group Session 10 (2.5 hours)	Relapse prevention and endings
02/10/08	Individual session (2 hours) – failed to attend	Post assessment and relapse prevention/good lives review

Appendix 10: Supervision log

Date (Duration)	Duration	Activity / Outcome / Future Actions / Lessons Learnt / Reflection on Development and Professional Practice
09/01/08	1 hour	Supervision: <i>Prof. Tony Beech.</i> Explored opportunities within LFF to access case study client and how this may be done.
25/01/08	1 hour	Supervision: <i>Simon Sauze (placement manager).</i> As above – discussion around opportunities within daily practice to access case study client.
30/01/08	1 hour	Supervision: <i>Dr. James Bickley:</i> Discussed option for case study and approaches to asking pre-conviction clients to take part in case study. Discussed how to best implement single case design at LFF. Discussions around differences in use of measures to inform intervention or use of assessment measures with a view to also using these as outcome measures. Discussion around how it is not suitable to use an assessment measure to measure an outcome if you are never expecting that thing to change in the first place. Discussed possible case study clients (e.g., inform plus group) and issues around gaining consent.
29/04/08	1 hour	Supervision: <i>Prof. Tony Beech.</i> Discussed identified case study client (internet offender) and use of IBAQ as pre and post measure for group. Explore identified difficulties with this measure and how best to overcome these.
01/05/08	2 hours	Supervision: <i>Dr. James Bickley.</i> Exploration of case study client and contact so far. Discussed my initial observations of P and how I will use these initial observations to shape my assessment. Explore assessment measures and importance of being able to justify use of particular psychometric measures (e.g., not applying measures unnecessarily).
13/05/08	1.5 hours	Supervision: <i>Prof. Tony Beech.</i> Discussed case study client and the positive and negatives of working with a Pre-conviction. Considered the opportunity to develop skills in encouraging offenders to begin thinking and reflecting upon behaviour at an early stage, often for the first time.
27/05/08	1.5 hours	Supervision: <i>Prof. Tony Beech.</i> Discussed assessment processes for case study client and offence focused interview guidelines.
11/06/08	1 hour	Supervision: <i>Dr. James Bickley.</i> Began to develop working hypothesis of P in discussions with supervisor and explored how to put this into formal structure.

Placement supervisor signature.....

Date (Duration)	Duration	Activity / Outcome / Future Actions / Lessons Learnt / Reflection on Development and Professional Practice
27/06/08	1 hour	<p>Supervision: Prof. Tony Beech. Discussed psychometric assessment of internet offenders generally and possible options available. Discussed use of STEP measures but lack of evidence at present to enable reliable interpretations to be drawn from use with internet offenders. Discussed use of IBAQ, specifically designed for use with internet offenders but poorly developed at present. Also discussed more clinical forms of assessment including an assessment of an individuals offence pathway (e.g., through the use of the self regulation model) and the use of a clinical interview. Further action 1) contact current researchers in the field of internet offending re current assessment processes 2) further research 3) obtain informed consent from client in order to determine assessment needs more specifically.</p>
27/06/08	1.5 hours	<p>Supervision: Dr. James Bickley. Discussed possible methods of assessment for internet offenders. Discussed psychometric assessment and the importance (and ethical issue) surrounding making sure than any measures I administer I am able to justify why I have chosen to administer them. What am I expecting to change? Importance of being able to justify why a particular measure has been applied to a particular client (does previous research justify this? Clients presenting difficulties). Importance of not administering psychometric assessments e.g., STEP battery because they are ‘usually administered’. Highlights to me the ‘less is more approach’ and more importantly - ensuring there is a clear distinction between academic needs for a case study and client needs for ethical assessment and treatment. Further action: continue research into internet offending assessment methods and obtain informed consent, develop assessment plan upon identification of client.</p>
10/07/08	1hour	<p>Supervision: Prof Tony Beech. Supervision focused around case study client and consideration of appropriate assessment measures. Discussed importance of remaining aware not to assess for only what interventions are available but also for interventions which are necessary but not necessarily offered by LFF. Discussed IBAQ – although not yet standardised, considered use of this measure in providing qualitative information to inform assessment and contribute towards formulation – i.e., using psychometric information qualitatively and not over-focussing on scores. Also discussed the importance of not using psychometric measures unnecessarily and carefully considering why I would chose to administer a given measure. Encouraged me to think critically about what measures I consider administering and why.</p>
31/07/08	2 hours	<p>Supervision: 1 hour, Dr. James Bickley (telephone supervision). Supervision: 1 hour, Prof. Tony Beech (telephone supervision). Issues discussed in both sessions included the following:</p> <p>Discussed models of formulation and which to currently apply to case study-discussed benefits of CBT model being as this is most common placed and the basis upon which I am currently working with this client. Although it is important for me remain aware that this will not always be the case and to familiarise myself with other psychological approaches to formulation and not limit myself to one approach.</p>

Placement supervisor signature.....

Date (Duration)	Duration	Activity / Outcome / Future Actions / Lessons Learnt / Reflection on Development and Professional Practice
28/08/08	1 hour	Supervision: <i>Dr. James Bickley.</i> Discussion around case study client and my observations of P in group and any impact this may have on my formulation.
28/08/08	1 hour	Supervision: <i>Prof Tony Beech.</i> Telephone supervision. Explored current case study client with whom I am currently delivering Inform Plus programme to – discussed what I feel his needs are and what I feel they are likely to be upon completion of the programme (i.e., what areas do I feel he will still need to address). Considered ways in which to communicate this to client. Discussed limitations of group work given that intervention cannot be individually tailored by for P, considered benefit of group work on developing social skills for example.
02/10/08	1 hour	Supervision: <i>Prof. Tony Beech.</i> Discussion around client’s failure to maintain contact and possible reasons for this. Reflections on process and possible factors which may have contributed to client’s failure to attend.
10/10/08	1.5hours	Supervision: <i>Dr. James Bickley.</i> Discussion around case study and write up of this. Reflections on process and how this has contributed to developing role as a forensic psychologist.

Placement supervisor signature

Appendix 11: Copy of RM2000 scales and scoring template (Thornton et al., 2003).

RM2000/S

Step One: Scoring Risk Factors

Circle the number of points that apply for each risk factor

- Age 18-24 = 2 points; 25-34 = 1 point; Older = 0 points
- Sexual Appearances 1 = 0 points; 2 = 1 point; 3,4 = 2 points; 5+ = 3 points
- Criminal Appearances 4 or less = 0 points; 5 or more = 1 point

Step One: Categorization

Circle the total number of points from the previous table, and the corresponding Category and Label.

Points	Category	Label
0	I	Low
1-2	II	Medium
3-4	III	High
5-6	IV	Very high

Step Two: Aggravating Factors

Circle the number of points that apply for each aggravating factor.

- Male Victim of Sex Offence No = 0 points; Yes = 1 point
- Stranger Victim of Sex Offence No = 0 points; Yes = 1 point
- Single (Never in Marital Type Relationship) No = 0 points; Yes = 1 point
- Non- Contact Sex Offence No = 0 points; Yes = 1 point

Step Two: Revised Risk Category

Put the risk category up one (e.g., from I to II or from II to III, or from III to IV) if two or three aggravating factors apply, and up two categories (e.g., from I to III, or from II to IV) if four aggravating factors apply.

Circle the Revised Risk Category and Label.

Revised risk Category	I	II	III	IV
Label	Low	Medium	High	Very high

RM2000/V

Risk Factor	Points Assigned
Age	18 to 24 = 3 points; 25 to 34 = 2 points; 35 to 44 = 1 point; Older = 0 points
Violent Appearances	0 = 0 points; 1 = 1 point; 2-3 = 2 points; 4+ = 3 points
Burglary	None = 0 points; Any = 2 points

Enter the number of points accrued above in the table below and circle the corresponding Risk Category and Label.

Points	Risk category	Label
0-1	I	Low
2-3	II	Medium
4-5	III	High
6 or more	IV	Very High

RM2000/C

Assign C-scale points from each of the V and S scale Categories

S or V Categories	I	II	III	IV
C Points Assigned for S scale	0	1	2	3
C Points Assigned for V scale	0	1	2	3

C-scale Labels derived from C-Scores: Circle the C-scale score and Label that applies.

Score on C-Scale	Label
1	Medium
2	Medium
3	High
4	High
5	Very High
6	Very High