

**OFFENCE SUPPORTIVE COGNITION IN YOUNG
SEXUAL OFFENDERS: ROLE, RELEVANCE &
PSYCHOMETRIC MEASUREMENT**

by

LOUISE ELIZABETH BOWERS

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

Centre for Forensic and Criminological Psychology
School of Psychology
College of Life and Environmental Science
University of Birmingham
October 2015

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This unpublished thesis/dissertation is copyright of the author and/or third parties.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

Offence supportive cognition (OSC) is an important theoretical and clinical concept in researching and treating adult sexual offenders. Much less is known about the role, relevance and measurement of OSC with younger sexual offenders, and this thesis aims to address that gap. Chapter 1 presents an introduction to OSC research, highlighting issues with measurement. Chapter 2 presents a systematic review of literature that has examined OSC in younger sexual offenders, with the aim of establishing whether OSC is a treatment need in this group. Chapter 3 examines the psychometric properties of the Children and Sex Questionnaire-Adolescent Version (CASQ-AV; Beckett 1995), a measure of child abuse supportive beliefs that is in widespread use with younger sexual offenders. Chapter 4 aims to establish the reliability and validity of the CASQ-AV using data from a large sample of young adult sexual offenders (aged 18 to 21 years) serving prison sentences. Chapter 5 draws together the findings from previous chapters, highlighting that overall, the role of OSC in younger sexual offender populations is poorly understood, the relevance of this concept as a treatment need may have been overestimated, but psychometric measurement is possible. The implications of these findings for research and practice are discussed.

Dedication

I dedicate this thesis to all the young men who have committed sexual offences and to all of the therapists who work tirelessly, and often without thanks, to assist them to make positive changes to their young lives. I hope that this work will inform future practice and help these young men on their journey towards a positive and offence-free future.

Acknowledgements

My first thanks must go to my supervisors, Dr Jessica Woodhams and Dr Daz Bishopp. Jess inspired me with her brilliance; she was always there for me and made me believe I could do this. To Daz, who has a brain the size of a planet, a dark sense of humour, and made psychometric theory and statistics both understandable and fun. I also owe gratitude to Vicky Wallace, librarian extraordinaire, who was on speed dial during my systematic review, and Sue Hanson, whose knowledge of this qualification and ability to give support is endless.

I would like to thank Dr Helen Wakeling, Dr Adam Carter and Georgia Barnett at NOMS who supported me in negotiating NOMS NRC and gave me access to the data that forms a significant part of this thesis. I must also thank HM Prison Service staff;

_____ ,
for helping me to access a test-retest sample. And I must not forget the lads at the two YOI's who gave up their free time to complete questionnaires not once but twice, simply to help make things better for others. Huge thanks also to Emily Daley, the research assistant you want in your life when things go wrong.

I could not have completed this thesis without the support of all my friends and family, but special thanks must go to Dr Caroline Friendship and Dr Emma Donnelly, who gave up their time to engage in peer review of this work. Special thanks must also go to the Wimbledon ladies and the UHS gang, who had to put up with high frequency moaning, tears and no-shows but never gave up on me.

And finally to my family, who have endured so much over the last two years. There have been some low points but you have always been there for me. To Mum,

who never doubted for one minute that one day I would become a Doctor – just wishing Daddy could be here to see it. To my darling son Ben, who is always so helpful but went the extra mile in teaching himself APA referencing and checking my syntax. To my darling daughter Nella, who always puts a smile on my face and knew exactly when to make me leave the office and take her shopping. Finally, to my darling husband Stuart, who brought Frank into our lives at what I thought was the worst possible moment of this doctoral journey – but clearly he knows me better than I know myself, and the timing was perfect – just like Frank.

The thesis was copy edited for conventions of language, spelling and grammar by Oxbridge Proofreading.

Contents

Abstract.....	ii
Dedication.....	iii
Acknowledgements.....	iv
List of Tables	ix
List of Figures.....	x
List of Appendices	xi
CHAPTER 1: INTRODUCTION TO THE THESIS	1
Introduction.....	1
The role of OSC in sexual offending	3
The relevance of OSC	5
Measurement of OSC.....	8
The study population.....	9
Aim of the thesis	10
Summary of chapters	11
CHAPTER 2: OFFENCE SUPPORTIVE COGNITION AS A TREATMENT NEED FOR YOUNG MALES WHO HAVE COMMITTED SEXUAL OFFENCES: A SYSTEMATIC REVIEW.....	13
Abstract.....	13
Introduction.....	14
OSC as a treatment need for adult males who sexually offend.....	15
OSC as a treatment need for young people who sexually offend	18
The psychometric measurement of offence supportive cognition	20
The current review	21
Aims and objectives	22
Method	23
Overview of search strategy.....	23
Search terms	24
Screening and selection of studies (applying the inclusion/exclusion criteria)	26
The screening and selection tool (SST)	27
Quality assessment	31
Data extraction	33

Results.....	34
Can young sexual offenders be distinguished from non-offenders or other groups of non-sexual offenders on measures of OSC?	48
Are there differences in OSC between different sub-groups of young sexual offenders?.....	51
Do scores on measures of OSC change following sexual offender treatment?	52
Is there an association between scores on measures of OSC and recidivism for young sexual offenders?.....	54
Do valid and reliable measures of OSA&Bs exist for young sexual offenders? ..	55
Discussion.....	57
Main findings of the review	57
Strengths and limitations of the review.....	64
Implications for practice and future direction.....	64
Conclusions.....	65
CHAPTER 3: THE CHILDREN & SEX QUESTIONNAIRE – ADOLESCENT	
VERSION (CASQ-AV): A PSYCHOMETRIC CRITIQUE.....	67
Introduction.....	67
Overview of the tool	70
Background to the measure.....	70
Development of the measure.....	71
Critical features and scoring of the measure	72
Psychometric characteristics of the measure.....	73
Conclusions	83
CHAPTER 4: A PSYCHOMETRIC VALIDATION OF THE CASQ-AV WITH	
YOUNG ADULT SEXUAL OFFENDERS AGED 18 TO 21 YEARS	86
Abstract.....	86
Introduction.....	87
The clinical importance of OSC: Two dominant paradigms	87
Psychometric measurement of the beliefs that support child sexual abuse with adult sexual offenders	92
Measuring child abuse supportive beliefs in younger sexual offenders	98
The current study.....	100

Specific aims and hypotheses	101
Method	102
Participants	102
Measures	105
Procedure.....	107
Treatment of the data	108
Ethical considerations	109
Results.....	110
Factor structure.....	110
Reliability.....	118
Validity.....	120
Discussion.....	123
CHAPTER 5: GENERAL DISCUSSION.....	138
Key findings from Chapter 2: A Systematic Review	138
Key findings from Chapter 3: A Psychometric Critique	143
Key Findings from Chapter 4: Research.....	144
Implications for Practice and Research	145
Strengths and Limitations of the Thesis	147
Summary and Conclusions	148
REFERENCES	149
APPENDICES	173

List of Tables

Table 1. Characteristics and quality assessment scores for the 13 included studies	35
Table 2: Comparison groups in each study	46
Table 3: Sample size, sample composition and measure used for each analysis	109
Table 4: Factor loadings from the CASQ-AV rotated four-factor solution	114
Table 5: Mean scores, alpha and test-retest coefficients for the CASQ-AV measure and components	119
Table 6: Correlations of CASQ-AV with PRI (socially desirable responding) and CDI sub-scale (concurrent validity)	120
Table 7: CASQ-AV component and total scores for each risk category	121
Table 8: Mean pre- and post-treatment CASQ-AV component and total scores and effect sizes	123

List of Figures

Figure 1: Flowchart of the review selection process.....30

List of Appendices

Appendix 1: Examples of search terms and syntax developed to access data on the OvidSP and Proquest platforms.....	173
Appendix 2: Screening and selection tool (SST).....	177
Appendix 3: List of excluded studies (n=24) and reasons for exclusion	178
Appendix 4: Quality assessment form for case control studies.....	182
Appendix 5: Quality assessment form for case series studies	185
Appendix 6: Data extraction form	187
Appendix 7: The CASQ-AV	191
Appendix 8: Items contained in the two sub-scales of the original CASQ-AV	192
Appendix 9: HM Prison Service Consent Form.....	193
Appendix 11: CASQ-AV Three Factor Solution	197
Appendix 12: Items Contained in the CASQ-AV Four Component Solution	202

CHAPTER 1: INTRODUCTION TO THE THESIS

Introduction

It is an unpalatable truth that sexual offending is not the preserve of adults and that both children and adolescents (aged 10 to 17 years) and young adults (aged 18 to 21 years) commit sexual crimes. However, it is difficult to estimate the incidence of sexual offending by young people. Estimates vary depending on the definition of sexual offence and the source of data (for example, official or crime survey data). Recent figures suggest that in England and Wales, in the 12 months ending March 2014, approximately 400¹ adolescents aged between ten and 17 years and approximately 400 young adults aged between 18 and 20 years were convicted of a sexual offence, compared with 4,900 adults aged 21 or older (Ministry of Justice, Office for National Statistics, 2014). This means that young people committed 16.3% of all sexual offences resulting in a conviction during that time period. The Crime Survey for England and Wales (CSEW) does not routinely ask victims about the age of the person who committed the sexual offence against them. Where this information was requested (between 2007 and 2012), 30% of perpetrators were reported to be aged under 20 years (Ministry of Justice, Home Office, Office for National Statistics, 2013).

Estimates of recidivism in young people who have sexually offended vary across studies owing to different measures of recidivism and length of follow-up. Evidence suggests that few young people who commit their first sexual offence as a juvenile will re-offend sexually and recidivism rates are lower for young people than for adults (National Criminal Justice Association (NCJA), 2014). Two recent studies, a meta-

¹ Data reported to the nearest 0.1 thousand

analysis and a cohort study, reported an official recidivism rate of approximately 7% over five years (Caldwell, 2010; Hargreaves & Francis, 2014). This suggests that the vast majority of young men who commit a sexual offence do not go on to be reconvicted for a sexual crime; the use of official data to measure recidivism (for example, arrest, charge or conviction), however, is known to underestimate the number of crimes sexual offenders commit (Falshaw, Bates, Patel, Corbett, & Friendship, 2003; Langevin et al., 2004). The impact of sexual crimes on victims is undeniable (Maniglio, 2009; Paras et al., 2009) and there is no reason to suggest that the youth of perpetrators lessens the impact of these types of offences. In most countries, considerable time and resources are thus expended on assessing and treating young sexual offenders in order to encourage them to desist and to prevent them from becoming the adult sexual offenders of the future.

Over the past 20 years, Donald Andrews and James Bonta have made empirically-based recommendations about how to organise offending behaviour interventions in ways that reduce recidivism (Andrews & Bonta, 2010). Commonly referred to as the Risk-Need-Responsivity (RNR) model, this body of work proposes that, for interventions effectively to reduce re-offence, they must be organised around three core principles: they must target higher *risk* individuals; they must target criminogenic *needs*; and they must be delivered in a way that is *responsive* to the general and specific needs of the individual. Criminogenic needs are the dynamic features of an individual or situation that require intervention if reduced re-offence risk is to be achieved. They are therefore often referred to interchangeably as *treatment needs* or *risk factors* (Hoge, 2015). One particular criminogenic treatment need, offence supportive cognition (OSC), is the focus of this thesis.

The role of OSC in sexual offending

OSC in one form or another is implicated in the aetiology and/or maintenance of sexual offending in most, if not all, contemporary multifactorial theories of sexual offending (Ward, Polaschek, & Beech, 2006). In addition, there are several single factor theories that focus specifically on the role of different types of cognitive phenomena in the offence pathway. In multifactorial integrated theories of sexual offending presented by Marshall and Barbaree (1990), Marshall and Marshall (2000) and Ward and Beech (2006), distorted beliefs and attitudes about women, sex and relationships formed during early childhood and consolidated by negative adolescent experiences are emphasised to play a crucial role in later sexual offending. Finkelhor (1984) presents a (four) pre-condition model of sexual offending. He argues that child sexual offenders have self-serving thoughts when contemplating offending, which help them to overcome internal inhibitions and give themselves permission to offend.

Abel, Becker and Cunningham-Rathner (1984) and Abel et al. (1989) conducted pioneering work on a single-factor theory of OSC in the mid-to-late 1980s, analysing the statements adult sexual abusers² made to professionals about their offending. They noticed that their explanations often included statements that rationalised, excused and justified their abusive behaviour; for example, ‘I didn’t hurt him’ and ‘I was teaching her about sex’. They proposed that these statements served to reduce feelings of guilt and anxiety, thereby allowing the abusive behaviour to continue. According to this theory, self-statements play a maintenance role in offending behaviour (Abel et al., 1984; Abel et al., 1989). Abel and his colleagues referred to these statements as ‘cognitive distortions’ and this body of work became known as the ‘cognitive distortion’

²Although Abel and colleagues developed their theory of cognitive distortion on men who abused children, they later applied their theory to men who committed rape and found similar cognitive distortions (Abel, Becker, & Skinner, 1987).

hypothesis. This theory dominated research and treatment efforts for many years, and directed treatment efforts towards ameliorating the justifications, minimisations and excuses sexual offenders made for their crimes (e.g., Murphy 1990). Clinical experience suggests that, when asked to explain their offending, young sexual offenders also use cognitive distortions. The extent to which this has been empirically demonstrated is reviewed in Chapter 2.

Only in 1999 was an alternative conceptualisation of OSC proposed, one that emphasised the aetiological role of cognitive phenomena in sexual offending. Ward and Keenan proposed that sexual offenders hold offence-supportive ‘implicit theories’, defined as coherent structures containing a set of assumptions and maladaptive beliefs about themselves, others and the world around them, and that they use to explain human behaviour in different circumstances (for example, ‘the world is dangerous’, ‘children are sexual beings’, ‘women are sex objects’) (Ward, 2000; Ward & Keenan, 1999). Five implicit theories have been identified for men who offend against children (Ward & Keenan, 1999) and adults (Polaschek & Ward 2002). These overlap somewhat. For example, the ‘dangerous world’ implicit theory – that the world is inherently dangerous and made up of hostile and threatening people – has been identified for both child abusers and adult rapists (Ó Ciardha & Ward, 2013). Because sexual offenders are thought to use these implicit theories to interpret their social world, they view their surroundings in an offence-congruent way, making offending more likely. Implicit theories are thought to originate in childhood and play an aetiological role in sexual offending. This hypothesis should therefore apply equally well to adults and younger people who sexually offend, although this has not been empirically tested. A fuller

description of this theory follows in Chapter 4, together with an examination of the extent to which younger sexual offenders demonstrate evidence of implicit theories.

The single-factor schema-based model of sexual assault (Mann & Beech, 2003) is similar to the implicit theory hypothesis. In this theory, a range of offence-supportive schemas (defined as stable structures that contain beliefs, attitudes and assumptions) are thought to develop in response to negative early life events and, when activated, interact with other risk factors (for example, offence-related sexual interests) to make sexual offending seem appealing. However, unlike in the implicit theory hypothesis, the specific nature of the maladaptive schemas is unidentified. This might be why this theory has received less research attention than cognitive distortion theory and is less well validated than the implicit theory hypothesis (Ó Ciardha & Ward, 2013).

The relevance of OSC

Which particular aspects of OSC represent relevant criminogenic needs (and should therefore be targeted in treatment) is a cause for much debate (Gannon & Ward, 2009; Marshall, Marshall, & Kingston, 2011; Maruna & Mann, 2006; Ó Ciardha & Gannon, 2011). This debate, has in part, been fuelled by semantic and definitional confusion over the term ‘cognitive distortion’, which has become inextricable from research on OSC with sexual offenders. The term is used to describe such disparate concepts as belief systems, justifications, perceptions, judgments, excuses, defensiveness, rationalisations, minimisations and denial of offending (Ó Ciardha & Gannon, 2011). It has been argued that the concept’s broadness renders it theoretically and clinically meaningless (Drake, Ward, Nathan, & Lee, 2001; Maruna & Mann, 2006; Ó Ciardha & Gannon 2011; Ó Ciardha & Ward 2013). Ó Ciardha and Gannon (2011) provide a useful

conceptualisation to discriminate between the various forms of OSC and alleviate confusion about how ‘cognitive distortion’ should be defined. According to this conceptualisation, a distinction is made between *cognitive structures* (beliefs, attitudes, schemas, and implicit theories) sexual offenders might have; *cognitive processes* (how information is processed and the influence of cognitive structures on this); and *cognitive products* (the thoughts/statements that are made after information is processed via cognitive structures). This conceptualisation of OSC has been adopted in this thesis, as it can accommodate the range of cognitive phenomena referred to in the literature. For example, in this model, the distorted statements that offenders make about their offending are viewed as products of underlying beliefs, attitudes, schemas, or implicit theories.

In this study, OSC is defined broadly, to include any cognitive structures (attitudes, beliefs, implicit schemas) or cognitive products (thoughts, statements) that are theoretically or empirically associated with sexual offending³. In defining OSC in this broad way it is accepted that there is little evidence that the post-hoc excuses, justifications, rationalisations and minimisations sexual offenders employ when asked to account for their offending are linked to recidivism (Maruna & Mann, 2006). It would, however, be unhelpful to exclude cognitive content from this thesis, for two reasons. First, it is not possible to measure cognitive structures directly. Rather, attitudes, beliefs, schemas or implicit theory must be inferred from statements made by the individual and much empirical research on OSC relies on cognitive content (statements) gathered via interview or self-report questionnaires (Ó Ciardha & Gannon, 2011). Second, it is often difficult to establish whether statements made by sexual

³ Although cognitive processing can support offending, this thesis focuses on the *direct* measurement of cognitive phenomena and therefore processing will not be explored.

offenders are indicative of underlying beliefs or schemas, or represent post-hoc justifications for behaviour about which they feel guilty and shameful (Maruna & Mann, 2006; Gannon, & Ward, 2009). To illustrate, an individual who states ‘But girls do come on to men in that way’ may be attempting to justify his offending and make it seem more socially acceptable. Equally, this statement could indicate that the individual has interpreted the child’s behaviour in that way because he has an implicit theory about children as sexual beings and a belief that children want to have sex with adults.

One aspect of OSC that has been extensively studied is Offence Supportive Attitudes and Beliefs (OSA&Bs) related both to the abuse of children (for example, children are sexually provocative, they enjoy sex with adults and are not harmed by this) and adults (for example, some women deserve to be raped, some women enjoy being raped). There is good evidence, for adults at least, of an association between engaging in sexual offender interventions and modifications to OSA&Bs as measured by self-report questionnaires (Beggs, 2010; Nunes, Pettersen, Hermann, Looman, & Spape, 2014). However, if OSA&Bs are a criminogenic need (Andrews & Bonta, 2010), this positive treatment change would be associated with reduced re-offending: empirical findings to support this hypothesis are scarce and conflicting (Nunes et al., 2014). The extent to which scores on OSA&B measures change following intervention with younger populations and the degree to which any of these changes translate into reductions in recidivism is systematically reviewed in Chapter 2. A recent meta-analysis (46 studies with a sample size of 13,728) demonstrated that offence supportive attitudes do predict recidivism in a small but meaningful way (Helmus, Hanson, Babchishin, & Mann, 2013). This study included data from younger populations; such data was not, however, separately analysed and age was treated only as a moderating variable. Thus,

this study did not contribute considerably to knowledge about the relevance of OSA&Bs for younger sexual offenders, a theme revisited in Chapter 2.

Measurement of OSC

OSC is measured in both treatment and research settings and is usually assessed in one of three ways. First, self-report questionnaires are frequently used to assess OSA&Bs (for example, The RAPE and MOLEST scales; Bumby, 1996); less commonly, schemas (for example, My Life Questionnaire; Mann & Hollin, 2010) and implicit theories (for example, The Implicit Theories Scale; Goddard, 2006). Self-report questionnaires are quick, inexpensive and simple to administer and score, but have been criticised for being subject to socially desirable responding and a limited ability to predict reconviction outcome (Wakeling & Barnett, 2014). Second, clinical rating scales are used to assess the presence or absence of OSA&Bs. This is usually part of a wider risk assessment protocol, for example the ‘attitudes domain’ on the Structured Assessment of Risk and Need (SARN: Ministry of Justice 2009) or Item 3 on the Estimate of Adolescent Sexual Offence Recidivism (ERASOR; Worling, 2001). When making a clinical rating, assessor use their judgment about the degree to which an OSA&B is present or absent, using information from interviews with the individual and others, a collateral review and, occasionally, behavioural observation. The range of data sources is a strength of this method, but scoring criteria are often limited and judgment may be subject to external bias; for example, how likeable the individual is. Third, indirect measures of OSC have been developed to assess a range of cognitive phenomena (schemas, implicit theories, attitudes and beliefs). In this type of assessment, the participant does not provide a view on the presence or absence of the attribute (in this

case, OSC). Rather, it is inferred by the examiner based upon time of participant response to certain stimuli (Snowden, Craig, & Gray, 2011). The strength of this assessment method is the assessment of cognitive phenomena of which the participant is not consciously aware, reducing the potential for socially desirable responses. However, such studies also report mixed findings and statistical differences between response times of sexual offenders and non-sexual offenders/non-offenders are inconsistent (Beech, Bartels & Dixon, 2013; Keown, Gannon & Ward, 2010). This suggests that measuring OSC indirectly also has validity issues. The current thesis focuses on the psychometric measurement of OSC; thus indirect measures are not discussed further.

A range of questionnaires have been developed to measure the attitudes and beliefs that support both the sexual abuse of children and adults (Nunes et al., 2014). Fanniff and Becker (2006) suggest that specialist instruments are required to assess all aspects of functioning with younger sexual offender populations. However, after conducting a review of this area they conclude that there was ‘limited evidence for use of specialized assessment instruments during routine evaluations of adolescent sex offenders’ (Fanniff & Becker, 2006, p. 272). A further review of developmentally appropriate assessments, including self-report questionnaires that might now be available for use with younger sexual offenders, has not yet been undertaken.

The study population

Adolescence is a time of major biological, emotional, social and cognitive change. Attempting to assess young people during this stage of development has been described as ‘akin to hitting a moving target’, unlike assessing adults, who have ‘decades of stable behaviour patterns’ (Prentky, Righthand, & Lamade, 2015, p. 8). A growing body of

research suggests that the factors motivating and influencing adolescent (and young adult) criminal behaviour are different to those underpinning adult offending (Tolan, Walker, & Reppucci, 2012). For example, compared to adults, young people are less able effectively to regulate their emotions and control their impulses; they are more likely to take risks and to be influenced by their peers (Gardner & Steinberg, 2005; Steinberg, 2008). In the criminal justice system in England and Wales, young people are considered adults at the age of 18; they are sentenced according to the law as it applies to adults, and the courts are under no obligation to take youth into account as a mitigating factor (Lösel, Bottoms, & Farrington, 2012). Recent research indicates that brain development – particularly higher order cognitive processes and executive functioning (including planning, impulse control and interpretation of emotions), as well as social and emotional development – continues well into the mid-twenties (Prior et al., 2011). Notwithstanding how the law views young adults, it is clear that they do not become mature adults on their 18th birthday. Rather, it appears that they are in a stage of ‘emerging adulthood’ (Arnett, 2000, p.469). For these reasons, young adults (aged 18 to 21 years) were included in this study.

Aim of the thesis

Over the past decade, research on risk factors for sexual offending in younger populations has increased (Prentky et al., 2015). This knowledge base is still relatively small compared to the literature on adults. Despite OSC being a well-researched and reasonably well-established criminogenic phenomenon for adult sexual offenders, its relevance for younger offenders remains relatively uninvestigated. The aim of this study is to identify, assimilate and analyse disparate data on OSC with younger sexual

offenders, exploring its role in offending, the relevance it might have as a criminogenic need for this group, and whether one particular form of OSC (the attitudes and beliefs that support the sexual abuse of children) may be measured in a valid and reliable manner.

Summary of chapters

Chapter 2 presents a systematic review of the literature on OSC in young sexual offenders (both young people who have offended against children and those that have offended against their peers or adults), in order to establish whether OSC represents a treatment need for this group. To answer this question, a narrative synthesis of data from a range of different study types was undertaken, including research that has attempted to discriminate young sexual offenders from non-offenders or other groups of non-sexual offenders, and studies that have attempted to discriminate between sub-groups of sexual offenders. Studies on the effect of treatment on measures of OSC are also included, as are those on the association between OSC and recidivism. The reliability and validity of OSC measures of OSC in the studies is examined. The findings of this review are discussed in relation to the role that OSC plays in the aetiology and maintenance of sexual offending generally, and for sub-groups of young sexual offenders specifically. The implications of the findings for sexual offender interventions and the testing and construction of measures of OSC is discussed.

Chapter 3 examines the psychometric properties of a measure that has a ‘cognitive distortions’ subscale and has been adapted for use with young sexual offenders: the Children and Sex Questionnaire-Adolescent Version (CASQ-AV; Beckett 1995). The CASQ-AV was selected because, according to the literature, it is one of the few OSC

measures adapted for younger populations, and is widely used, both in the UK and in Europe. The findings of this critique are discussed in terms of the use of the CASQ-AV in clinical and research settings. The results of the critique reveal that the psychometric properties of the CASQ-AV have not been adequately established. Given its widespread use, this is concerning.

Chapter 4 presents an empirical paper that aims to establish both the reliability and validity of the CASQ-AV using data from a large sample of young adult sexual offenders (aged 18 to 21 years) serving prison sentences in two young offender institutions in England. The underlying factor structure of the CASQ-AV was tested and interpreted in light of current conceptualisations of OSC and as evidence of the construct validity of the measure. The internal consistency and temporal stability of the measure were investigated to establish whether the CASQ-AV is reliable. The concurrent and discriminant validity of the measure were also tested, and, finally, the sensitivity of the CASQ-AV to treatment change was investigated. The findings of the research are discussed in relation to the future use of the CASQ-AV in clinical and research settings.

Chapter 5 draws together findings from the previous chapters to provide an overall view of the role, relevance and psychometric measurement of OSC in younger sexual offender populations. This chapter also discusses the implications of this study for practice and future research.

CHAPTER 2: OFFENCE SUPPORTIVE COGNITION AS A TREATMENT NEED FOR YOUNG MALES WHO HAVE COMMITTED SEXUAL OFFENCES: A SYSTEMATIC REVIEW

Abstract

Empirical support exists for OSC as a treatment need for adult male sexual offenders. In this systematic review, the role and relevance of OSC as a treatment need for young males (aged ten to 21 years) who have committed sexual offences is explored. General and specific scoping searches were undertaken to assess the need for the current review. The search strategy included searching five major electronic databases and setting up alerts. Manual searches of the reference lists of included studies and two meta-analyses were conducted. 19 experts in the field were contacted for participation. Specific inclusion, exclusion and quality appraisal criteria were applied to each study. 13 studies met the inclusion criteria. OSA&Bs were the form of OSC investigated in all studies. Narrative data synthesis emphasises that young people who commit sexual offences cannot reliably and consistently be discriminated from non-offenders or from other types of offenders on measures of OSA&Bs. Similarly, child abusers and peer-adult abusers cannot be discriminated. However, data comparing OSA&Bs in extra-familial/inter-familial child abusers had greater discriminatory power. Limited evidence suggests that scores on measures of OSA&Bs change following sexual offender treatment, but evidence was insufficient to establish whether recidivist/non-recidivist sexual offenders may be discriminated on the basis of problems with OSA&Bs.

Literature on young sexual offenders relies heavily on questionnaire measures of OSA&Bs, but the measures in current use have been designed for adults and none of them have demonstrated adequate psychometric properties for use with young people.

These results question the role of OSA&Bs in the aetiology of sexual offending for young people and the relevance of the construct as a treatment need for them as a group (although extra-familial child abusers might represent a special case). Existing findings are limited by a weak evidence-base and a lack of psychometrically-sound measures of OSA&Bs for young people. It is critical that the psychometric properties of existing measures of OSA&Bs are established and developmentally-sensitive measures of OSA&Bs constructed. This will allow further investigation of the role of OSA&Bs as a treatment need for young sexual offenders (and specific groups of offenders) and for empirically-informed decisions to be made about whether OSA&Bs should be a target for intervention.

Introduction

A key principle of effective correctional treatments is targeting changeable, dynamic, criminogenic needs (Andrews & Bonta, 2010). Sexual offender treatment programmes that adhere to these principles demonstrate the greatest reductions in recidivism; this is true both for adult and juvenile programmes (Hanson, Bourgon, Helmus, & Hodgson, 2009). In the adult male sexual offender literature, OSC has been extensively examined as a risk factor for sexual offending, both theoretically and empirically (Ward et al., 2006; Ó Ciardha & Ward, 2013). Referring to cognitive distortions in general, ‘many practitioners feel considerable confidence about their importance as a treatment target’ (Ward et al., 2006, p. 116). A broad definition of OSC is employed in this thesis; it includes both cognitive structures (stable OSA&Bs, implicit theories and schemas) and cognitive products (what sexual offenders think and say about their offending). This approach was adopted because cognitive structures cannot be directly accessed; both research on and treatment of OSC rely heavily upon inferences from what offenders say

(Beech et al., 2013). Until recently, sexual offender research was dominated by studies that examined adult male offenders and the risk and treatment needs of younger sexual offenders were often ignored and poorly understood (Viljoen, Mordell, & Beneteau, 2012).

OSC as a treatment need for adult males who sexually offend

If OSC plays an aetiological role in adult sexual offending, it should be possible to reliably distinguish sexual offenders from non-offender samples on measures of OSC. This type of research usually consists of administering self-report questionnaires that propose to measure OSA&Bs to groups of sexual offenders and men from the community who are assumed not to have sexually offended (Gannon & Polaschek, 2006). This type of research has yielded inconsistent results for child abusers (Gannon, Keown, & Rose, 2009) and poor results for rapists (Gannon, Collie, Ward, & Thakker, 2008). Various reasons have been suggested for this lack of discrimination, including offenders engaging in socially desirable responding or ‘faking good’ (Langevin 1991); insensitive measures or lack of construct validity (Gannon & Polaschek, 2006); the prevalence of OSA&Bs in the community as well as in sexual offenders (Abbey, Wegner, Pierce, & Jacques-Tiura, 2012; Gannon & O’Connor, 2011); or, more controversially, OSA&Bs not being a causal risk factor for sexual offenders at all (Benbouriche, Longpré, Guay, & Proulx, 2015; Nunes et al., 2014). OSA&Bs may also only be problematic in subgroups of offenders. In support of this latter proposal, Fisher, Beech and Brown (1999) found that a mixed group of child molesters could not be discriminated from community comparisons (trainee prison officers), but that when child molesters were split into extra-familial and inter-familial sub-groups, the former

had significantly higher levels of cognitive distortions than community comparisons, whereas inter-familial offenders did not.

There is some evidence for offence supportive attitudes being linked to recidivism. Offence supportive attitudes reportedly meet the criteria for an ‘empirically supported’ risk factor with a ‘small but significant relationship with sexual recidivism’ (Mann, Hanson, & Thornton, 2010, p. 11). However, this conclusion is based on one meta-analysis (Hanson & Morton-Bourgon, 2004), in which the relationship between offence supportive attitudes and increased recidivism was context specific. When offenders were assessed for treatment, this relationship remained, but there was no relationship between offence supportive attitudes and recidivism in other adversarial situations, such as when offenders were assessed for court proceedings or community supervision. More recently, a ‘small but consistent relationship’ between attitudes supportive of sexual offending and recidivism was found in a large-scale meta-analysis involving 46 samples and 13,782 participants (Helmus et al., 2013). This review was limited in that it only included seven studies, which had a ‘published’ effect size, only three studies explicitly examined OSA&Bs, and the psychometric properties of measures were not considered.

At the individual level, there is good evidence that adult male sexual offenders show ‘clinically significant and reliable change’ on measures of OSA&Bs following treatment (Nunes et al., 2014). This means moving from a dysfunctional range on a measure to a pre-defined functional level not due to chance (Nunes, Babchishin, & Cortoni, 2011). Such changes have been identified on specific measures of offence supportive attitudes for sexual offenders with a learning disability (Keeling, Rose, & Beech, 2006), a mixed group of offenders receiving treatment in prison in the UK

(Barnett, Wakeling, Mandeville-Norden, & Rakestrow, 2013; Beech & Hamilton-Giachritsis, 2005) and a mixed custodial/community group in Canada (Nunes et al., 2011).

More recently, attempts have been made to link clinically significant changes on criminogenic variables to risk of re-offending for adult male sexual offenders; however, results in relation to OSC have been mixed. For example, one small-scale study of child molesters receiving treatment in the community showed that none of the men who demonstrated clinically significant change on measures of pro-offending attitudes re-offended (Beech & Ford, 2006). In a sample of 3,773 prisoners who had undergone treatment in the UK prison service, however, there was no association between clinically significant change on individual measures of offence supportive attitudes and recidivism; furthermore, changes on the pro-offending attitudes domain as a whole did not predict recidivism (Wakeling, Beech, & Freemantle, 2013). This was also true in a large sample that had undergone treatment in the community in the UK (Barnett, Wakeling, Mandeville-Norden, & Rakestrow, 2012).

Schemas and implicit theories are not usually under conscious control (Beech et al., 2013) and are thus more difficult to measure with self-report questionnaires. Nonetheless, this approach has provided some interesting results. For example, Marziano, Ward, Beech, & Pattison (2006), analysed interview data from child offenders and found evidence for all five of Ward's (1999) implicit theories (children as sexual beings, nature of harm, entitlement, dangerous world and uncontrollability), with those who abused boys being much more likely to see the world as a dangerous place and children as sexual beings than those who abused girls. Interview and questionnaires have also been used to provide evidence of implicit theories in rapists (Polaschek &

Gannon, 2004) and offence-related schemas (dominance and disadvantage) in both child offenders and men who rape adults (Mann & Hollin, 2010). No claims have been made about the links between implicit theories, schemas and reductions in recidivism.

Despite mixed results, it is reasonable to conclude that OSC might play a role in the aetiology, onset and/or maintenance of sexual offending for adult males. In addition, reductions in endorsement of OSA&Bs are possible when offenders engage in offence-focused interventions, and there is evidence that lower endorsement of OSA&Bs is linked to reduced recidivism. Whilst not overwhelming, the evidence supports the continued targeting of OSC in sexual offender treatment.

OSC as a treatment need for young people who sexually offend

The sexual offences young people commit might be similar to those committed by adults, but young people are fundamentally different owing to their cognitive, psychological and neurological immaturity (Tolan et al., 2012). Researchers and practitioners stress that the development of assessment methods and treatment approaches for young sexual offenders must be based on evidence gathered from young people and take account of their developmental immaturity and differences in motivation, behaviour and pathways to offending (Rich, 2009). Despite this, concerns remain that some interventions continue to be based on adult research and adult sex offender treatment programmes (Calleja, 2013; Letourneau & Miner, 2005).

Since 2000, there has been a small but growing number of studies examining potential risk factors and treatment needs for young people who sexually offend (National Criminal Justice Association (NCJA), 2014). However, these studies have produced inconsistent and often contradictory results and they have been criticised for

short follow-up periods and small sample sizes. In addition, the risk factors examined are based on clinical experience and adult research rather than being theoretically driven by an understanding of adolescent sexual offending. Consequently, there is no agreed-upon set of risk factors for young people who sexually offend that can be used to define and set targets for treatment (NCJA, 2014). In terms of OSC, 'attitudes supportive of sexual offending' have been proposed as a 'promising treatment need' based on a narrative review of potential risk factors for sexual offending in young people (Worling & Långström, 2006). A more recent review, taking account of current evidence, has not been undertaken.

The recent meta-analysis of offence supportive attitudes and recidivism (Helmus et al., 2013) contained juvenile samples (15 of the 45 studies) but this data was not analysed independently of the adult data, although age was not found to be a moderating variable. Five of the 15 studies were unpublished dissertations; the other studies used single clinical rating items from the ERASOR (Worling & Curwen, 2001) or JSOAP-II (Prentky & Righthand, 2003) risk assessment tools as the measure of OSC (item scored as present, possibly/partially present or not present). It should be noted that the definition of item 22, 'cognitive distortions' from JSOAP-II includes both attitudes supportive of sexual offending and general pro-criminal attitudes. In addition, the purpose of all the studies that used ERASOR or JSOAP-II was to assesses sexual and non-sexual recidivism for young sexual offenders, therefore almost all of the studies reported data at the subscale level (that is 'deviant sexual interests, attitudes and behaviour' (ERASOR) and 'interventions' (JSOAP-II)) or overall risk level (low, medium, high) rather than at the item level (3 or 22). None of the studies examined OSC in young participants more generally. Thus, it is not possible to conclude from this

study whether offence supportive attitudes are a relevant or important treatment need for young people who sexually offend.

The psychometric measurement of offence supportive cognition

When assessing the needs of sexual offenders and evaluating treatment change using a self-report approach, valid, reliable and appropriately standardised measures should be used (Grady, Brodersen, & Abramson, 2011). A number of valid and reliable psychometric measures for assessing OSA&Bs for both child abusers and rapists have been identified (Grady et al., 2011). Clinical rating scales for the assessment of offence supportive attitudes also exist for adult males; for example, the pro-offending attitudes domain of the Structured Assessment of Risk and Need –Treatment Needs Analysis (SARN-TNA; Thornton, 2002⁴). In this scale interview data, behavioural evidence and scores from questionnaires are combined for use in an explicit scoring guide to provide a perspective on whether a range of specific OSA&Bs are present (believing, women can't be trusted, men should dominate women, men have a right to sex, and child abuse supportive beliefs). This scale has good inter-rater reliability (Webster et al., 2006) but it has been demonstrated that the offence supportive attitude domain of the SARN-TNA does not predict recidivism (Tully, Browne, & Craig, 2014).

By contrast, researchers and practitioners have commented on the limited availability of psychometrically-sound questionnaire measures for young people who sexually offend (Gerhold, Browne, & Beckett, 2007; Oneal, Burns, Kahn, Rich, & Worling, 2008). As already identified, the ERASOR (Worling & Curwen, 2001) and JSOAP-II (Prentky & Righthand, 2003) are structured risk assessments for use with

⁴ The SARN-TNA was previously known as the SRA.

young sexual offenders; both contain one global clinical rating of OSA&Bs. This contrasts with the SARN, in which several specific OSA&Bs are identified.

The current review

There is some evidence to suggest that OSC is a relevant treatment need for adult male sexual offenders, but the extent to which young people who sexually offend are motivated or influenced by OSC has not been systematically evaluated. The domination of the field by research on adult male sexual offenders is problematic for practitioners working with minority groups (Gannon & Alleyne, 2013). In response to these shortcomings Gannon & Alleyne undertook a systematic review of the cognition of female sexual abusers, demonstrating that this was a relevant treatment need for this group. However, they argued that there were differences between males and females, which needed to be taken into account in treatment, and that there was a need for gender sensitive assessments.

An equivalent review for young sexual offenders has not been completed and this review is designed to fill that gap. The review identifies existing research examining cognition that may support sexual offending in young people. The data is systematically evaluated to draw conclusions about whether this group in fact does possess forms of OSC that support their offending, and whether this is a relevant or important treatment need for this group as a whole, or for sub-groups of sexual offenders. This study also reviews research that has examined treatment change for OSC to establish whether this variable might be modified by sexual offender interventions, as well as identifying any evidence that examines the relationship between such changes and recidivism. Evaluating how OSC is measured in this group is

also a critical part of this review. It is proposed that in exploring the role of OSC in the aetiology and maintenance of sexual offending for young people, and establishing how amenable OSC is to treatment-related change, this systematic review will provide information that may be used by researchers and practitioners to develop assessment measures and interventions that are developmentally sensitive and tailored to the specific needs of young people.

Aims and objectives

The aim of the systematic review is to explore the relevance of OSC as a treatment need for young males aged between ten and 21 years who have committed sexual offences.

Specific objectives were:

- To determine whether young sexual offenders can be distinguished from non-offenders or other groups of non-sexual offenders on measures of OSC
- To determine whether there are differences in OSC between different sub-groups of sexual offenders; for example, those who have offended against children as opposed to peers or adults
- To establish whether scores on measures of OSC change following sexual offender treatment
- To establish whether there is an association between scores on measures of OSC and recidivism for young sexual offenders
- To determine whether there are valid and reliable measures of OSC for young sexual offenders

Method

Scoping exercise. Several databases were searched before this review commenced, to establish whether previous reviews of a similar or identical nature existed or were planned. Searches were conducted using the Cochrane Database of Systematic Reviews (CDSR), the Centre for Reviews and Dissemination (DARE), the Campbell Collaboration, and the International Prospective Register of Systematic Reviews (PROSPERO). No existing or planned reviews were identified confirming the need for this review. In addition, a brief scoping search was undertaken on PsycINFO to identify current/key issues in the area of OSC, to establish the quantity and range of data that might potentially be retrieved and to check the feasibility of the review question. The search was conducted using basic free text terms that captured some of the most common ways OSC is referred to in the literature (cognitive distortions and offence supportive attitudes and beliefs):

‘sex* offen*’ AND ‘cognit* distort*’ OR (‘offen* support*’ adj3 (attitude* or belief*))

This basic search retrieved articles relevant to the review question and suggested that sufficient data was likely to be retrieved through a more extensive, comprehensive search.

Overview of search strategy

The search for relevant research was conducted in three stages. First, major electronic databases were searched, including: PsycINFO (1967 - present); EMBASE (1974 - present); MEDLINE (1946 - present); ASSIA (1987 - present) and the National Criminal Justice Reference Service (NCJRS, 1975 - present). All searches were

conducted on 18th and 19th September 2014. After completing these searches, alerts were set up on each database to allow each set of search syntax (see below) to be run on a monthly basis. These alerts provided the title and abstract of any studies that met the search criteria and were published after the initial searches were complete but before the review was submitted. No relevant studies were published after 19th September 2014 and finalisation of this thesis on 25th September 2015. Second, the reference lists of all the full text articles that met the inclusion/exclusion criteria and two retrieved meta-analyses were searched manually for potentially relevant articles. Third, 19 recognised experts in the field of sexual offender cognition and young sexual offenders were contacted to request any pertinent studies that may have been missed (a list of included studies was provided) and to request any studies that were due to be published within the following three months. 15 experts responded.

Search terms

Three of the databases searched (PsychINFO, EMBASE and MEDLINE) were accessed via the OvidSP platform, which uses subject headings to index the contents of the databases. A traditional approach to a systematic search was therefore initially implemented, including mapping search words to subject or MeSH⁵ headings and deciding whether to expand or restrict terms. A relevant subject heading was identified for ‘young sexual offender’ but not for OSC. As an alternative, free text words may be used to search on the OvidSP platform. These words may be located in the title, abstract or main body of the articles stored on the database. It is considered best practice to use a mixture of subject headings and free text searching to maximise the accuracy of search

⁵ Medical Subject Headings; a list of subject headings used for indexing articles in medical literature for example within MEDLINE

results (Dundar & Fleeman, 2014). This strategy was utilised where possible. Two of the databases, ASSIA and NCJRS, were accessed via the Proquest platform, which does not support the use of subject headings. Thus search terms devised for the OvidSP platform had to be modified to take this into account.

As already identified, OSC is a broad concept that may be defined in various ways. In addition, the way in which young sexual offenders are referred to in the literature has evolved over time. To ensure all different ways of defining the key terms were captured in the search, a review of several seminal papers relating to the assessment and treatment of young sexual offenders and OSC (in either adult or young sexual offenders) were examined and a list of key terms created. It became apparent that whilst ‘young’ and ‘sexual offender’ were fairly easy to define using key search terms, it was necessary to make extensive use of ‘adjacency searching’ to search for phrases that related to the concept of OSC. Various versions of the search terms were trialled until the optimal search syntax was developed that appeared to balance the need for specificity (identifying relevant papers) and sensitivity (not having too many irrelevant papers).

The following search terms were used, making use of the Boolean operators ‘OR’ (for synonyms) and ‘AND’ (to combine the three separate search concepts). The search terms were modified to take account of different search conventions used in different databases:

juvenile* OR young* OR adolescen* OR young adult* OR child OR children OR youth* OR teen*

AND

‘sex* offen*’ OR ‘sex* abus*’ OR ‘child molest*’ OR ‘child abus*’ OR ‘child offen*’
OR ‘sex* harm*’ OR ‘rapist*’

AND

(‘offen* support*’ adj3 (thought* or attitude* or belief* or cognit* or schema*)) OR
 (‘offen* relate*’ adj3 (thought* or attitude* or belief* or cognit* or schema*)) OR
 (‘abuse* support*’ adj3 (thought* or attitude* or belief* or cognit* or schema*)) OR
 (‘abuse* relate*’ adj3 (thought* or attitude* or belief* or cognit* or schema*)) OR
 (‘sex* offen*’ adj3 (thought* or attitude* or belief* or cognit* or schema*)) OR
 (‘child molest*’ adj3 (thought* or attitude* or belief* or cognit* or schema*)) OR
 (rape adj3 (thought* or attitude* or belief* or cognit* or myth*)) OR
 (‘pro-offen*’ adj2 (thought* or attitude* or belief* or cognit* or schema*)) OR
 (distort* adj2 (thought* or attitude* or belief* or cognit* or schema*)) OR
 (implicit adj2 (theor* or schema*))

A sample set of search syntax used on the OvidSP platform (PsycINFO) and the Proquest platform (NCJRS) is presented in Appendix 1. Search results were exported into EndNote reference manager.

Screening and selection of studies (applying the inclusion/exclusion criteria)

1,374 search hits were returned. First, all duplicate references were identified and removed ($n = 457$). Second, all titles, abstracts and sources of the remaining studies ($n = 917$) were screened using the screening and selection tool (SST) to remove those irrelevant to the review (see Appendix 2). Third, full text copies were obtained for all remaining references ($n = 33$); the inclusion/exclusion criteria contained in the SST were applied to each paper and reasons for the exclusion or inclusion of each paper were noted. A list of excluded studies and reasons for exclusion ($n = 24$) can be found in Appendix 3. The fourth stage involved applying the SST to articles retrieved from manual searches, which returned one extra paper. The final stage consisted of applying

the SST to articles suggested as relevant by experts, which returned three articles. A diagram of this process is presented in Figure 1.

The screening and selection tool (SST)

The PICOS (Population, Intervention, Comparator, Outcome, Study (design)) framework is a widely used tool adopted by the Cochrane group to define systematic review questions, develop search terms and establish inclusion/exclusion criteria in quantitative studies (O'Connor, Green, & Higgins, 2008). Because this review does not specifically attempt to evaluate intervention efficacy, not all elements of this framework (for example, intervention) were relevant to this review. The SPIDER (Sample, Phenomenon of Interest, Design, Evaluation, Research type) is an alternative framework used in qualitative and mixed methods reviews (Cooke, Smith, & Booth, 2012). Elements of both the PICOS and SPIDER frameworks that best met the question posed in this review were incorporated into the screening and selection tool (see Appendix 2). The main SST inclusion criteria included:

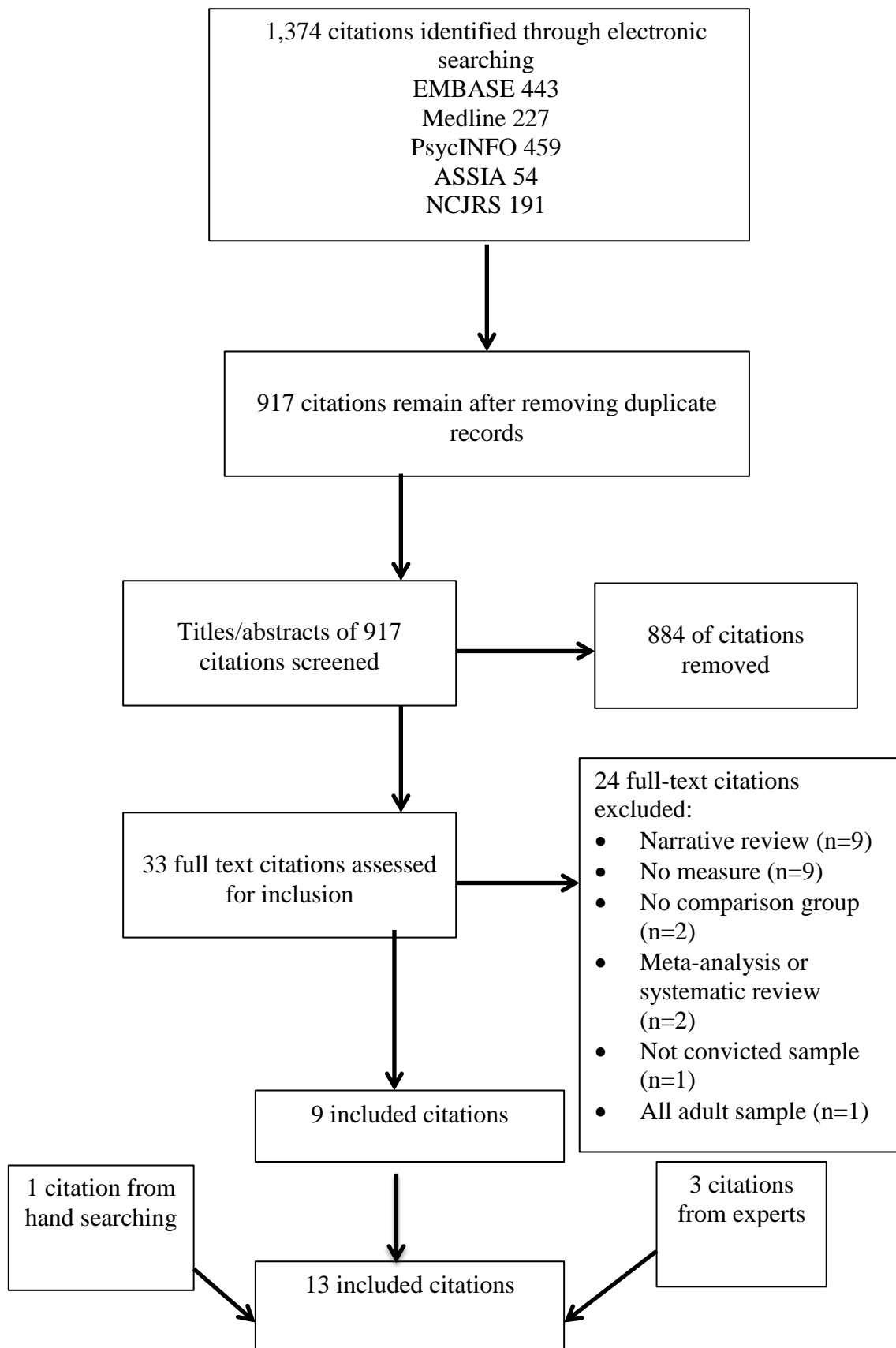
<i>Population:</i>	Male, adolescent and young adult (aged ten to 21 years) sexual offenders
<i>Phenomenon of Interest:</i>	OSC: cognitive structures and content, including thoughts, beliefs, attitudes, schemas, implicit theories, (or distortions thereof) that support sexual offending
<i>Comparison Group:</i>	<ul style="list-style-type: none"> • Non-offenders (community comparison) • Non-sexual offenders (e.g., violent offenders) • Sub-groups of sexual offender (e.g., child molester, peer aggressor)

	<ul style="list-style-type: none"> • Pre- post-intervention • Recidivists, non-recidivists
<i>Outcome:</i>	A comparison between the population of interest and <i>at least</i> one other comparison group on a <i>specific numerical</i> measure of OSC
<i>Research type/design:</i>	Any quantitative: must be published and include a comparison group
<i>Language:</i>	English language only

A decision was made to exclude studies that:

1. did not have a comparison group, as it would have been difficult to drawn conclusions regarding the role of OSC in the sexual offending without such comparisons, and lack of controls is considered a limitation (Gannon & Alleyne, 2013);
2. were unpublished (theses, dissertations, grey literature) due to the lack of robust peer review and difficulty locating these sorts of studies;
3. only used data from ERASOR or JSOAP-II. These are risk assessment tools that contain therapist checklists, with only one item on each tool relating to OSC. They were not designed to measure or examine specific, risk-related variables in depth or to measure treatment change (Oneal et al., 2008), and could not be described as ‘a specific measure of OSC’. Furthermore, as highlighted in the introduction, data from these tools is usually presented at the subscale or overall risk score level rather than at the item level.

When applying inclusion/exclusion criteria, three studies were identified that included young people up to the age of 22 years. These studies were included, as the mean age of participants was less than 18 years and it is likely they would have been younger than 22-years old when they committed their offences.

Figure 1: Flowchart of the review selection process

Quality assessment

Having applied the SST, 13 studies remained. There was limited variability in the research design of the included studies: ten were case control studies, one was a case series design, one used what was referred to as a 'cross-over longitudinal' design, and another study by the same author used the same design but did not use this label. The two cross-over longitudinal studies were treatment evaluations comparing the level of cognitive distortions of several *different* groups of young sexual offenders who were at various stages of treatment. Owing to the emphasis on comparing psychometric test scores between separate groups of young sexual offenders, these studies were treated as a case control design for the purposes of this review.

Some quality assessment tools propose to cater for a range of study designs within one pro-forma; for example, The Social Care Institute for Excellence (SCIE, 2010). Case control and case series studies, however, differ extensively methodologically, and the relative importance of the inherent sources of bias in them is differs accordingly. It was decided that a single, off-the-shelf quality appraisal tool would not provide the flexibility required to test the methodological rigour of both study designs. Thus two separate quality appraisal tools were selected.

The Critical Appraisal Skills Programme (CASP, 2013) offers eight critical appraisal tools, including a systematic review checklist and a cohort study checklist, which may be applied to different study designs. The CASP case control study tool was selected because of the developer's expansive expertise in appraising the applicability, reliability and validity of published research, and the wide use of this particular tool. The criteria contained in this tool were adapted slightly to meet the aims of this particular review (see Appendix 4).

Case series studies have no comparison group. They are thus often viewed as having a weak design, and providing weak evidence for the efficacy of health interventions (Dalziel et al., 2005). There are few quality appraisal tools available to assess this research design but a group of researchers have recently developed such a tool (Moga, Guo, Schopflocher, & Hartstall, 2012). The tool was developed using a modified Delphi technique that included selecting a panel of experts, critically appraising other tools, several rounds of piloting the new tool, and developing a coding dictionary. Although a finalised validated tool has not yet been published, Moga et al. (2012) concluded that their checklist could be used and/or modified to assess the quality of case series studies. This tool was selected to assess the quality of the case series studies in this review, owing to the methodological rigour with which it was developed. The criteria were modified slightly better to meet the aims of this particular review (see Appendix 5). A quality appraisal scoring system was developed as follows:

Yes = 2 (criteria fully met)

Partially = 1 (criteria partly met)

No = 0 (criteria not met)

Can't tell = 0 (unclear/insufficient information)

Scores were summed and a percentage quality score awarded for each article. Where information was missing or could not be deduced from the study, a score of zero was given as it was considered that missing or unclear methodological data reflected poorly on the quality of the study. For example, in one study, it was unclear how or from where the control group had been recruited. In another, it was unclear whether the

control group had been assessed in the same way as the experimental (case study) group, and under the same conditions. To allocate a score of 'one' in these situations could have artificially raised the overall quality score. Scores of zero in the 'can't tell' category were retained as part of the overall quality score rather than omitted, for the same reasons. It is accepted that, where data is missing or unclear, contacting the author(s) of the paper to gather further information or seek clarification is considered best practice (Fleeman & Dundar, 2014). This was not feasible in the time allowed. The number of included studies subject to quality appraisal was relatively low ($n = 13$), so no defined quality score cut-off point for exclusion was applied.

Data extraction

A data extraction form was specially developed to consistently extract relevant information from each study (see Appendix 6). General and specific information was extracted that included study characteristics and results/outcomes that were relevant to the aims of this systematic review. This form focused on the following key variables:

- General information (title, author, year, source, country)
- Study characteristics (study design, aims/objectives, measures used and associated standardisation, reliability and validity)
- Participant characteristics (definition of participants/types of offence(s), age, sample size, location and recruitment)
- Comparison group characteristics (for case control studies only, as above)
- Intervention information (for case series studies only, description of intervention and attrition rate)
- Study results (analysis, results, conclusions)

Results

Overview of Studies. Table 1 presents a summary of the synthesised data for the 13 included studies, allowing evaluation of how young sexual offenders and sub-groups of offenders compare to a range of comparison groups, on a variety of different measures of OSC with varying levels of validity, reliability and standardisation.

Methodological and study characteristics. This was an international set of results, including studies from a wide range of countries: most from the United States (five studies), followed by the UK and Republic of Ireland (three studies). Holland contributed two studies and there was one study each from Australia, Sweden and Canada. There was variability in the date of the studies: the oldest was being published 22 years ago (Epps, Haworth, & Swaffer, 1993) and the most recent in 2012 (Edwards, Whittaker, Beckett, Bishopp, & Bates, 2012). Study design was relatively uniform: ten studies employed a case control design, one was a treatment evaluation employing a case series design (Edwards et al., 2012), and two used a cross-over longitudinal design (Eastman, 2004, 2005), in which three different groups of young people at various stages of treatment were examined.

Table 1. Characteristics and quality assessment scores for the 13 included studies

Author, year & country	Aims of the study and design	Sample / case characteristics	Comparison group(s) - case control studies	Intervention	Measure of OSC, standardisation validity & reliability	Findings (in relation to OSC)	Quality assessment score
1. Beckett (2006) UK and Republic of Ireland	Consideration of sexual and general recidivism in adolescents. Examination of psychological test design with adolescent abusers. Introduction to the Adolescent Sexual Abuser Project (ASAP). Presentation of findings from the ASAP in relation to both general and victim empathy, and cognitive distortions. Case control	235 adolescent child abusers who had committed 'hands on offences' (121 interfamilial and 114 extra familial). Referred from 37 community, residential and juvenile prison treatment programmes in the UK and Republic of Ireland. (Age not specified – but referred to as 'adolescents' therefore must meet 'age' review criteria).	57 'normal adolescents' described as 'an unselected sample of secondary school students' (Age not specified – but referred to as 'adolescents' therefore must meet 'age' review criteria).	N/A	The Children and Sex Questionnaire - Adolescent Version (CASQ-AV; Beckett, 1995) cognitive distortions scale only. Standardised on a group of 97 'normal adolescents' Validity not reported. Reliability - 'good' internal consistency reported, but for the whole questionnaire, which includes the cognitive distortions and emotional congruence with children scales ($\alpha = 0.92$ on the study sample). 'Adequate' test re-test data reported, but for the whole questionnaire (0.76 on the study sample over 14 days)	No significant differences found between adolescent child abusers and community comparisons on the CASQ-AV cognitive distortions scale, although the community comparisons scored higher ($M = 17.19$) than the young sexual offenders ($M = 16.34$) indicating they endorsed more cognitive distortions. 30 % of extra familial adolescent child abusers had a 'high CD score' ⁶ compared to 19 % of intra familial offenders ($p > 0.05$).	47 %
2. Butz & Spaccarelli (1999) USA	To establish if the use of force can be reliably assessed in juvenile sexual offenders, and if it is linked to social competence, delinquent attitudes, sexual deviancy and previous sexual victimisation.	44 'rapists' recruited while undergoing clinical assessment in a residential sexual offender treatment facility between 1993-1996. 'Rapist' - an	57 other sexual offenders drawn from the same residential treatment facility, comprising: 30 'Non-rapists' – did not self-report use of force on the MSI Rape scale and	N/A	Multiphasic Sex Inventory (MSI) adolescent version, Cognitive Distortions/Immaturity (CDI) and Justifications (J) scales (Nichols & Molinder, 1984). Standardisation and validity not reported.	No significant differences between any of the groups on MSI cognitive distortion / immaturity or justifications scale.	56 %

⁶ Defined as more than one standard deviation above 'normal' - normal not defined

Author, year & country	Aims of the study and design	Sample / case characteristics	Comparison group(s) - case control studies	Intervention	Measure of OSC, standardisation validity & reliability	Findings (in relation to OSC)	Quality assessment score
	Case control	individual who self reported the use of force on the MSI 'Rape Scale' scale or whose case history indicated they had used force in their offences. Age range 12-19 years, mean age = 15.02 (<i>sd</i> = 1.6) for whole sample (<i>N</i> = 101).	case file confirmed no use of force. 27 'Deniers' – had information in file that suggested they had used force, but did not report this on the self-report MSI Rape scale.		Reliability - internal consistency ($\alpha = 0.61$ and $\alpha = 0.75$ on this sample)		
3. Eastman (2004) USA	To explore the attainment of specific treatment goals in an adolescent sex offender treatment program. Factors investigated: level of cognitive distortions, level of sexual knowledge, attitude about sexual behaviour, ability to understand the concept of empathy and perception of self-worth. Cross over longitudinal.	40 convicted sexual offenders who were waiting to start a residential treatment programme Age range of whole sample (<i>N</i> = 100) 13-22 years, mean age = 17	40 convicted sexual offenders who had completed treatment and were waiting to be released 20 convicted sexual offenders who had completed treatment and had been in the community for 6-18 months.	No details of the treatment programme reported	MOLEST and RAPE scales (Bumby, 1996) No standardisation or validation data reported. Reliability - internal consistency (Molest $\alpha = 0.97$, Rape $\alpha = 0.96$). Test re-test over two weeks (Molest 0.84 Rape 0.86) ⁷ .	Pre treatment participants differed significantly from post treatment and post release participants on the Rape and Molest scales, with the pre-treatment subjects having the highest scores (indicating a higher levels of endorsement of offence supportive beliefs) followed by the post treatment group and the post release subjects having the lowest scores. Rape scale (<i>M</i> = 76.08 vs. 52.75, vs 49.00 respectively) Molest scale (<i>M</i> = 76.02 vs. 56.38 vs 51.60)	65%

⁷ Data is from Bumby (1996) adult sample

Author, year & country	Aims of the study and design	Sample / case characteristics	Comparison group(s) - case control studies	Intervention	Measure of OSC, standardisation validity & reliability	Findings (in relation to OSC)	Quality assessment score
4. Eastman ⁸ (2005) USA	To investigate whether characteristics of juvenile sexual offenders or factors associated with sex offender treatment are predictive of treatment response (drop out). Factors investigated: demographic factors; background characteristics; cognitive distortions; sexual knowledge; attitude about sexual behaviour; empathy; and self esteem. Cross over longitudinal	19 convicted sexual offenders who did not complete treatment 'drop outs'. Mean age = 14.9. All participants were court ordered to attend treatment either in a juvenile justice or mental health facility	56 convicted sexual offenders waiting for treatment. Mean age = 14.2 years 63 convicted sexual offenders who had successfully completed treatment. Mean age = 15.8	No details of the treatment programme reported	MOLEST and RAPE scales (Bumby, 1996) No standardisation or validation data reported. Reliability - internal consistency (Molest $\alpha = 0.97$, Rape $\alpha = 0.96$). Test re-test over two weeks (Molest 0.84 Rape 0.86) ⁹ .	Cognitive distortions had the strongest potential to distinguish treated from untreated juvenile offenders. The most powerful discriminant function emerging from the analysis, and reflecting 79% of the variance, contained the instruments assessing offender cognitive distortions.	67%
5. Edwards, Whittaker, Beckett, Bishopp, & Bates. (2012) UK	To evaluate the effectiveness of a residential sex offender group work programme for adolescent males by investigating the extent to which a battery of pretreatment psychometric scores and the dynamic domains of the ERASOR risk assessment differ pre- and post-treatment.	34 adolescents who had engaged in sexually harmful behaviour and who had completed residential treatment. Age range pre-treatment 11 years 6 months - 16 years 3 months, mean age = 14 years 3 months ($SD = 1.2$)	N/A	The gateway offence specific programme consists of cognitive behavioural group work delivered in a rolling format. It aims to address the offence, sex and relationships, decision-making, rights	The cognitive distortions scale of the Children and Sex Questionnaire - Adolescent Version (CASQ-AV; Beckett, 1995). Multiphasic Sex Inventory (MSI) adolescent version, Justifications (J) sub-scale (Nichols & Molinder, 1984). CASQ-AV - Standardised on 56 post-treatment	The group as a whole showed a highly significant positive reduction in distorted thinking regarding children and sex following treatment ($p < .0001$). 21% of participants had a score outside the normal range pre-treatment and 100% fell into the 'treated range' post treatment. The group as whole showed positive and significant reduction on the MSI (J)	78%

⁸ It is not possible to tell if there was overlap in the two Eastman samples

⁹ Data is from Bumby (1996) adult sample

Author, year & country	Aims of the study and design	Sample / case characteristics	Comparison group(s) - case control studies	Intervention	Measure of OSC, standardisation validity & reliability	Findings (in relation to OSC)	Quality assessment score
		<p>Age range post treatment: 14 years 8 months - 18 years 8 months, mean age = 7 years ($SD = 1.0$) post treatment.</p> <p>25 young people who did not complete treatment were indistinguishable from the completers on all of the measures prior to the commencement of treatment</p>		<p>and responsibilities, victim issues and relapse prevention. Individual sessions are also provided. The ASAP assessment battery (Beckett, Gerhold, & Brown, 2002) is administered pre- and post-treatment.</p>	<p>adolescent Validity not reported. Reliability - internal consistency ($\alpha = 0.78$) on the adolescent version</p> <p>MSI-J. Standardisation and validity reported as per Nichols and Molinder (1984).</p> <p>Reliability - internal consistency of <i>all</i> MSI subscales are reported as 'adequate' ($\alpha = 0.58 - 0.92$) and ($\alpha = 0.89$ on this sample). Test-re-test of <i>all</i> MSI subscales is reported as 'adequate' (between 0.8 and 0.9) over three months (all figures from Milner et al., 1998).</p>	<p>scale post treatment ($p < 0.002$). 87.5% of the participants had a score outside the normal range pre-treatment and 56% of participants fell into the 'treated range' post treatment</p> <p>Those participants who dropped out of treatment ($n = 25$) were indistinguishable on all of the measures prior to the commencement of treatment from those who managed to complete it successfully.</p>	
6. Epps, Haworth, & Swaffer (1993) UK	<p>To measure and compare attitudes towards women and rape, for male adolescents convicted of violent sexual offences against women, with male adolescents convicted of non-sexual violent offences.</p> <p>Case control</p>	31 (stated in the abstract) 27 (stated in the method) violent sexual offenders - defined as having committed an offence of rape, attempted rape or indecent assault against a women. Cut off age for women not defined.	27 (stated in the abstract) 29 (stated in the method) violent non-sexual offenders - defined as convicted of violence but not having any recorded history of sexual offences. Recruited from Glenthorne and St Charles secure treatment	N/A	Rape Myth Acceptance Scale (RMAS; Burt, 1980). Standardisation, validity, and reliability not reported	No significant differences between the sexual and violent offenders' scores on the BRMAS.	38%

Author, year & country	Aims of the study and design	Sample / case characteristics	Comparison group(s) - case control studies	Intervention	Measure of OSC, standardisation validity & reliability	Findings (in relation to OSC)	Quality assessment score
		Recruited from Glenthorne and St Charles secure treatment centre in Birmingham Age range 14-20 years, mean age =17.3.	centre in Birmingham. No information provided about comparison group.				
7. Kenny, Keogh, & Seidler (2001) Australia	To test a theoretical model of how and why juvenile sexual offenders recidivate, by examining the relationship between sexually deviant fantasies, deviant sexual experiences cognitive distortions, poor social skills and learning problems. It was hypothesised that 'exogenous' variables would have an impact on recidivism via mediating variables – cognitive distortions and sexual fantasies. Case Control	30 'recidivists' - adolescent sexual offenders who had been charged with a sexual offence previously and were awaiting sentence in a two year period between Jan 96 and Jan 98. Mean age = 15.40 (<i>SD</i> = 1.4)	40 'first time offenders' – adolescent sexual offenders who had not been charged with a sexual offence previously and were awaiting sentence in a two year period between Jan 96 and Jan 98 were approached. Mean age = 15.86 (<i>SD</i> = 0.64)	N/A	A clinical rating scale was developed to measure cognitive distortions. Specific questions at interview were combined with information from informants and case file history. Clinical ratings were graded 1-5 (5 = CD's were a major part of the young person's account of his offence(s) and 1 = CD's were a minor part) Standardisation and validation not reported. Reliability – inter-rater reliability for the CD scale was reported as excellent (ICC 0.91) on this sample.	All factors examined had significant associations with the offence category recidivist. However, the inter-correlation matrix was not presented. 'Pertinent results' were described but this did not include any results about the relationship between offence category and the CD scale. Results of structured equation modelling suggested learning problems and deviant sexual experiences have an indirect path to recidivism that are mediated through cognitive distortions and deviant sexual fantasies. Also, cognitive distortions have an indirect path to recidivism through deviant sexual fantasies	85%
8. Racey, Lopez, & Schneider (2000)	To explore the differences between adolescent sex offenders and non-sex offenders on measures	36 convicted sexual offenders. 28 were incarcerated and 8 were accessing treatment in the	38 convicted 'delinquent' non-sexual offenders. 30 were incarcerated and 8 were	N/A	Abel and Becker Cognitions scale (Abel et al.,1984). The questionnaire was amended by replacing 'adult' with 'adolescent' and reducing	There was a significant difference ($p<0.001$) between the scores of the sexual offenders and the control group, with the non-sexual	26%

Author, year & country	Aims of the study and design	Sample / case characteristics	Comparison group(s) - case control studies	Intervention	Measure of OSC, standardisation validity & reliability	Findings (in relation to OSC)	Quality assessment score
USA	of social functioning, non-verbal communication skills, attitudes towards sexual contact with children, and sexual knowledge. Case Control	community Age range 13-18 years, mean age = 15.7, ($SD = 1.10$). For the whole sample (cases and comparisons)	accessing treatment in the community		the number of items from 29 to 10. Standardisation, validity, and reliability not reported	offenders having a <i>more</i> permissive attitude towards sex with children ($M = 19.87$, $SD = 6.72$) than the sexual offenders ($M = 13.11$, $SD = 5.88$).	
9. Tidefors, Goulding, & Arvidsson (2011) Sweden	To describe a group of Swedish adolescent males who have sexually offended with regard to background variables, individual characteristics and offending behaviour. To investigate whether the psychometric measures contained in the ASAP (Beckett, Gerhold, & Brown, 2002) were able to distinguish adolescent males who have sexually offended from adolescent males from the general population. Case Control	45 boys who had committed sexual abuse. Recruited from a range of secure and community settings Age range 13-22 years, mean age = 16.2 ($SD = 1.9$). All participants had committed their offences when aged under 18.	42 boys from a local high school. No other information provided.	N/A	The Children and Sex Questionnaire - Adolescent Version (CASQ-AV; Beckett, 1995). Cognitive Distortions and Emotional Congruence Scale Not standardised on a Swedish sample. Validity not reported. Reliability figures reported from van Outsem et al (2006) for the whole ASAP assessment battery. Reliability - Internal consistency ($\alpha = 0.70 - 0.93$). Test-re-test reliability (0.70 - 0.98)	No significant differences in scores between the sexual offenders and the community comparison group on the CD scale ($p = 0.39$) or EC scale ($p = 0.84$). The community comparison group endorsed <i>more</i> distorted beliefs about sex with children ($M = 18.9$) and reported more emotional congruence with children ($M = 17.3$) than the sexual offenders ($M = 16.8$, $M = 17.8$).	71%
10. Van Outsem et al. (2006) Holland	To present the psychometric properties of the ASAP - D assessment pack (Dutch translated version). To compare the personality	180 'hands on' sexual offenders recruited from a range of inpatient ($n = 30$) and outpatient ($n = 150$) facilities.	153 non-sexual violent offenders. Age 12-20 years, mean = 16.7 ($SD = 1.8$)	N/A	The Children and Sex Questionnaire - Adolescent Version (CASQ-AV; Beckett, 1995). The researchers split the measure into 5 subscales, justifications; child sex	Sexual offenders scored significantly <i>lower</i> on 4/5 of the CASQ-AV scales than the community comparison group: justifications ($p = 0.03$), self-identification as a child ($p = 0.001$), mutual/	79%

Author, year & country	Aims of the study and design	Sample / case characteristics	Comparison group(s) - case control studies	Intervention	Measure of OSC, standardisation validity & reliability	Findings (in relation to OSC)	Quality assessment score
	characteristics as measured by the ASAP-D between a group of juvenile sex offenders, a group of juvenile non-sexual violent offenders and a representative sample of non-delinquent youth. Case Control.	Age range 12-20 years, mean age = 16.4, ($SD = 2.3$).	500 non-delinquent juveniles from local schools. Age range 14 - 18 years, mean age = 15.8 ($SD = 0.8$).		maturity, self-identification as a child; mutual/extraordinary relationship with children; ideation of attractiveness for children. Standardised on 500 community comparisons. Validity - face, construct, convergent, and discriminate reported as 'good'. Predictive validity was poor - correlations were low between self-reported cognitive distortions and therapist ratings of the same construct. Reliability for the whole pack - Internal consistency (α range 0.7 -0.93). Test re-test correlations ranged from 0.70 – 0.98.	special relationship with children ($p = 0.007$), and ideation of attractiveness for children ($p = 0.022$). Sexual offenders scored <i>lower</i> than violent offenders on two scales: justifications ($p = 0.012$), and mutual/special relationship with children ($p = 0.012$).	
11. Van Vugt et al. (2011) Holland	To examine differences between juvenile child abusers and juvenile peer abusers in their level of moral judgment, the existence of distorted beliefs supporting child sexual abuse, and the degree of (general) cognitive	56 child abusers - the victims were more than 5 years younger than the perpetrator. Recruited from a range of inpatient and out-patient facilities.	21 peer/adult abusers - age difference between perpetrator and victim is <i>less</i> than 5 years or victim is older than perpetrator. Recruited from a range of inpatient	N/A	Sex With Children questionnaire (Mann, Webster, Wakeling, & Marshall, 2007) adapted by translation into Dutch and removal of the word 'adults'. Standardisation and validity not reported. Internal consistency was reported as	No significant differences between the child abusers and peer abusers on their scores on the two sub-scales contained within the SWCH. 'Sex with children is harmless'. Child abusers ($M = 1.38$, $SD = 0.56$), peer/adult abusers ($M = 1.65$, $SD = 0.88$). 'Children are sexually	68%

Author, year & country	Aims of the study and design	Sample / case characteristics	Comparison group(s) - case control studies	Intervention	Measure of OSC, standardisation validity & reliability	Findings (in relation to OSC)	Quality assessment score
	distortions. To examine the relation between general cognitive distortions, implicit theories, and the level of moral judgment of the two groups in sexual and own abuse situations. Case Control.	Age range 13 - 22 years, mean age = 17.23 (<i>SD</i> = 2.2).	and out-patient facilities. Age range 15 - 22 years, mean age = 18.29 (<i>SD</i> = 2.24).		'good' ($\alpha = 0.86$ and $\alpha = 0.87$) ¹⁰ .	provocative'. Child abusers ($M = 1.60$, $SD = 0.73$), peer/adult abusers ($M = 1.77$, $SD = 0.96$)	
12. Worling (1995) Canada	To determine whether adolescent sex offenders who assault children differ from those who assault peers/adults in terms of: quality of peer relationships; history of sexual abuse; physical discipline; rape supportive attitudes and beliefs. Case Control.	29 child sexual offenders - offender was at least 4 years older than the victim and the victim was under 12 years. Participants recruited from the SAFE-T (community out-patient) programme. Mean age = 15.76 ($SD = 1.86$).	27 peer/adult sexual offenders – these were the young people who did not meet the definition of a child offender. Recruited from the SAFE-T treatment programme. Mean age = 15.33 ($SD = 1.47$).	N/A	Rape Myth Acceptance Scale (RMAS; Burt, 1980). Standardisation and validity not reported. Reliability - internal consistency ($\alpha = 0.905$) on a large sample ($N = 209$) attending the SAFE-T programme. The sample included adult males and females.	No significant difference between the young people with child victims or adult/peer victims on endorsement of rape myths on the RMAS. Child abusers ($M = 55.32$, $SD = 21.62$), peer/adult abusers ($M = 61.93$, $SD = 16.85$).	76%

¹⁰ It is not clear if internal consistency was measured on this sample or the original development sample.

Author, year & country	Aims of the study and design	Sample / case characteristics	Comparison group(s) - case control studies	Intervention	Measure of OSC, standardisation validity & reliability	Findings (in relation to OSC)	Quality assessment score
13. Zgourides, Monto, & Harris (1997) USA	To explore the relationship between prevalence of prior adult sexual contact, sexual attitudes (rape myths), viewing sexually explicit material, and offender status (sexual offender vs non offender). Case Control.	80 convicted sexual offenders, recruited from a state juvenile facility (46) or under probation supervision (34). Age range 13-19 years, mean age = 16.13.	96 non-convicted males attending an urban high school, selected to 'mirror' the sexual offender sample. Age range 14 -19 years, mean age = 16.33.	N/A	56-item questionnaire (constructed by the authors) including 3 items from the Burt Rape Myth Acceptance Scale (RMAS; Burt, 1980). Standardisation, validity, and reliability not reported.	Offenders were significantly <i>less</i> likely than the community comparison group to endorse rape myths.	65%

Participants and recruitment. Age of participants was specified as in the range of ten to 21 years or a mean age of under 18 years. Only one study included participants under the age of 12 years (Edwards et al., 2012). Where the age range was reported (ten studies), all but one study included young people aged over 18 years. Reasons for the inclusion of three studies with participants aged 22 years are provided under methodology. Samples sizes were generally small. The smallest was a study in which 19 sexual offenders who dropped out of treatment were compared with 56 sexual offenders who were waiting for treatment and 63 who had successfully completed treatment (Eastman, 2005). There were two notable studies in terms of larger sample size. One compared 235 adolescent child abusers with 57 normal adolescents from a local secondary school (Beckett, 2006); the other compared 180 young sexual offenders with 153 violent offenders and 500 non-delinquent youths from local schools (van Outsem et al., 2006). The locations from which participants were drawn varied. Four studies recruited participants exclusively from residential settings (correctional settings, in-patient facilities, community homes); one exclusively from a community out-patient treatment facility; six from both the community and secure settings; and one in which the population from which participants were drawn was unclear (Kenny, Keogh, & Seidler, 2001).

Study focus / aims and comparison groups. Only one study included in this review assessed only OSA&Bs; that is, rape myth acceptance (Epps et al., 1993). In five studies, OSC was analysed as one of the main variables (Beckett, 2006; Eastman, 2004; van Vught et al., 2011; Worling, 1995; Zgourides, Monto, & Harris, 1997). In the rest of the studies, OSC was one factor amongst many examined. One of the most important

inclusion criteria for this review was that studies must contain a comparison group of some description. Table 2 summarises the different comparison groups for each of these studies.

The main aim in approximately half of the studies ($n = 7$) was to compare young male sexual offenders with either non-sexual offenders or non-offending community comparison groups, or to compare sub-groups of sexual offenders on at least one psychosocial variable. The main aim of three studies (Beckett, 2006; Tidefors, Goulding, & Arvidsson, 2011; van Outsem et al., 2006) was to present psychometric data from the Adolescent Sexual Abuser Project assessment battery (ASAP; Beckett, Gerhold, & Brown, 2002), which included OSC measures. All three studies compared the psychometric test scores of sexual offenders to non-offending community comparisons. Additionally, van Outsem et al. (2006) compared young sexual offenders' scores to a group of violent non-sexual offenders, and Beckett compared the scores of subgroups of child molesters (intra- and extra-familial). Edwards et al. (2012) compared the pre- and post-treatment psychometric test scores of 34 adolescent males who had engaged in sexually harmful behaviour. The pre-treatment scores of these 34 adolescents were also compared to 25 young sexual offenders who had not completed treatment. Eastman (2004) compared scores on a range of psychosocial variables for three different groups of adolescent sexual offenders (pre-treatment, post-treatment, post-release) and Eastman (2005) was similar but compared a group of treatment dropouts to a group who were pre-treatment and a different group who were post-treatment. It is not possible to deduce from the study reports if there was any cross over in samples between these studies.

Table 2: Comparison groups in each study

Study	Comparison group												
	Pre-treatment	Post-treatment	Post treatment & release	Treatment drop out	Recidivist	Non recidivist	Generic sex offenders	Child abusers	Inter-familial child abusers	Extra familial child abusers	Peer/adult abuser (rapist)	Non-sexual offenders	Community comparison
Beckett (2006)								✓	✓	✓			✓
Butz & Spaccarelli (1999)							✓				✓		
Eastman (2004)	✓	✓	✓										
Eastman (2005)	✓	✓		✓									
Edwards et al. (2012)	✓	✓		✓									
Epps et al. (1993)											✓	✓	
Kenny et al. (2001)					✓	✓							
Racey et al. (2000)							✓					✓	
Tidefors et al. (2011)							✓						✓
Van Outsem et al. (2006)							✓					✓	✓
Van Vugt et al. (2011)								✓			✓		
Worling (1995)								✓			✓		
Zgourides et al. (1997)							✓						✓

OSC Measures. There was very little variability in terms of the *type* of measures used to evaluate OSC. Kenny et al. (2001) developed a clinical rating scale that made use of semi-structured interview questions, case history and informant data to assess the degree to which cognitive distortions were part of the young person's offence account. All other studies administered self-report questionnaires with a range of different measures utilised. The most frequently used measures included the Beckett (1995) Children and Sex Questionnaire – adolescent version (CASQ-AV; four studies) followed by The Burt (1980) Rape Myth Acceptance Scale (RMAS; three studies). The Bumby (1996) RAPE and MOLEST scales and the Cognitive Distortions/Immaturity and Justifications subscales of the Multiphasic Sex Inventory (MSI; Nicholas & Molinder, 1984) were used in two studies. Other measures included the Sex with Children Questionnaire (SWCH; Mann, Webster, Wakeling & Marshall, 2007) and the Abel and Becker Cognitions Scale (ABCS; Abel et al., 1989). All measures were designed to assess OSA&Bs, except the Justifications subscale of the MSI, which assesses causal explanations and degree of accountability for offending (for example, 'my sex offence happened because I had not had sex before'). In addition, the SWCH measures two specific offence supportive beliefs (sex with children is harmless; children are sexually provocative) that have been mapped onto, and provide evidence for, two of Ward's implicit theories: 'nature of harm' and 'children as sexual beings' (Mann et al., 2007). All questionnaire measures were either adapted from adult versions of the measure, or the adult version was used. There was no evidence that any of the questionnaire measures had been specifically designed or constructed for a population of young sexual offenders. The psychometric properties of the measures are discussed in the key findings section.

Quality of included studies. There was wide variability in the quality ratings for studies included in the review (range 26% to 85%, mean score of 63%). Three studies were of notably poor quality and were awarded scores of less than 50%. In each of these studies five of the 17 quality appraisal criterion could not be scored owing to missing data and poor methodological reporting. The main issues affecting the quality of many studies was the use of small convenience samples, not controlling for confounding variables, and the unavailability of data regarding the psychometric properties of measures.

Narrative data synthesis and key findings. A range of aims, methodologies and participant groups were evident in the studies included and the resulting data was heterogeneous. It was therefore inappropriate to combine and synthesise data in a meta-analysis. Rather, a narrative data synthesis was conducted, emphasising key findings in relation to each of specific aims.

Can young sexual offenders be distinguished from non-offenders or other groups of non-sexual offenders on measures of OSC?

Four studies compared young sexual offenders to community controls. Three used the same measure (CASQ-AV), but results were analysed slightly differently in each study. Beckett (2006) only reported on the ‘cognitive distortions’ subscale of this measure, Tidefors et al. (2011) reported on the ‘cognitive distortions’ and ‘emotional congruence with children’ sub-scales, and van Outsem et al. (2006) reported on five subscales in this measure. All studies reported that the community groups scored *higher* than the young sexual offenders on the CASQ-AV regardless of how it was analysed, indicating that they had a higher level of unhelpful attitudes/beliefs about sex with children than

sexual offenders. In addition, van Outsem et al. (2006) reported that identified differences were significant on four of the five subscales analysed. Zgourides et al. (1997) administered items from the Burt RMAS, to which participants had to respond 'yes' or 'no'. On each item, a statistically significant higher percentage of high school students than sexual offenders answered 'yes', which indicates endorsement of the rape myth. All studies recruited the young sexual offender participants from a mixture of residential and community settings and recruited high school students of a similar age to the sexual offenders to act as community controls. However, only two studies attempted to match high school pupils with young sexual offenders in a meaningful way. In addition, only Beckett, specifically analysed child abuser's responses to the CASQ-AV with the other three studies using a generic group of sexual offenders, even though the CASQ-AV measures child abuse supportive beliefs and the RMAS measures rape myth acceptance. None of these studies reported score ranges and means in a way that allowed deductions about the degree to which groups were endorsing OSA&Bs.

These studies suggest that young sexual offenders cannot reliably and consistently be distinguished from non-offenders on psychometric measures of OSA&Bs. Furthermore, where differences do exist, young people who have come to the attention of the authorities because they have engaged in sexually harmful behaviour, consistently self-reported a *lower* level of OSA&Bs than community comparisons.

Three studies compared young sexual offenders to other types of non-sexual offenders. Each study used a different psychometric measure. Epps et al. (1993) and van Outsem et al. (2006) recruited convicted violent offenders as the comparison group. Racey, Lopez, and Schneider (2000) used convicted 'delinquents', without specifying conviction type. Epps et al. (1993) reported no significant differences between the

endorsement of rape myths by the group of young men who had committed sexual offences against women and violent offenders on the RMAS. Mean scores were not reported; the direction of non-significant differences in scores thus remains unknown. Racey et al. (2000) reported significant differences between a generic group of sexual offenders and the 'delinquent' comparison group on the ABCS, with non-sexual offenders endorsing *more* permissive attitudes/beliefs about having sex with children. Van Outsem et al. (2006) reported that a generic group of young sexual offenders scored significantly *lower* than the violent offenders on two of the five scales of the CASQ-AV (justifications for sexual offending and believing you have a special relationship with children). Data on the other three sub-scales was not reported. Epps et al. (1993) recruited participants from a secure setting; Racey et al. (2000) and Van Outsem et al. (2006) recruited participants from both residential and community treatment settings. None of the studies reported taking potential confounding variables into account. Again, how the groups responded to questionnaires in terms of endorsement of OSA&Bs could not be established from the data presented.

Overall, these studies suggest that young sexual offenders cannot reliably and consistently be distinguished from other types of non-sexual offenders on measures of OSC. An unexpected finding was that young people who had come to the attention of the authorities because they had engaged in sexually harmful behaviour, reported lower levels of OSA&Bs than the non-sexual offenders.

Are there differences in OSC between different sub-groups of young sexual offenders?

Four studies examined sub-groups of sexual offenders. Beckett (2006) compared the scores of intra-familial and extra-familial child abusers on the cognitive distortions scale of the CASQ-AV. 'High scores' were defined as more than one standard deviation above 'normal' however the normal range was not specified. Beckett reported significant differences between the percentage of extra-familial offenders that had a high cognitive distortions score (30%) as compared to intra-familial offenders (19%). Butz and Spaccarelli (1999) categorised a sub-group of young sexual offenders undergoing assessment in a residential treatment centre in a novel way: by the use of force in their sexual offences. They compared young sexual offenders' self-reported use of force on the rape scale of the MSI with their case history and devised a three-way classification system: 'rapist' (self-reported use of force and case history reports use of force); 'non-rapist' (does not self-report use of force and case history confirms this); 'denier' (does not self-report use of force but case history suggests they have used force). No significant differences were reported amongst any groups for the endorsement of offence-related justifications on the Cognitive Distortions/Immaturity or Justification scale of the MSI. Van Vught et al. (2011) and Worling (1995) both compared a sub-group of child abusers to a sub-group of peer/adult abusers and both went to considerable lengths to take account of potential confounding variables in their analysis. Van Vught et al. (2011) administered the SWCH to a mixed inpatient/outpatient sample and Worling (1995) administered the RMAS to a community treatment sample. The results were similar. Van Vught et al. (2011) reported no differences between groups on endorsement of implicit beliefs justifying sex with

children (although, interestingly, the peer/adult abusers mean score was slightly *higher* than that of child abusers). Worling (1995) reported no differences between groups on rape myth acceptance (the peer/adult abusers score was slightly higher than that of child abusers). It is useful to look at the responses of sub-groups of young sexual offenders to offence-specific measures based on these results, but the only sub-groups that could be distinguished were the extra/intra-familial child abusers. The degree of endorsement of OSA&Bs was unestablished in three studies (Beckett 2006; Butz & Spaccarelli 1999; Worling 1995). Van Vught et al. (2011) reported that the SWCH consisted in 18 items scored on a 5-point Likert-type scale; the range of scores was thus 18 to 90. However, mean scores for child abusers and peer/adult abusers were reported to range from 1.38 to 1.7. This either indicates that item scores were averaged to create a scaled score, or there is a reporting error.

Do scores on measures of OSC change following sexual offender treatment?

Three studies examined the effects of treatment on OSA&Bs. Edwards et al. (2012) employed a case series design and overcame some of the problems of lack of control group by calculating clinically significant change: that is, the number of individuals who exhibited problems on each psychosocial variable before treatment commenced and the proportion who had made positive progress and moved into the 'treated range' post-treatment. Significant reductions in distorted thinking were demonstrated post-treatment on the cognitive distortions scale of the CASQ-AV (100% of participants fell into the treated range post-treatment) and the MSI justifications scale (56% of participants fell into the treated range post-treatment). It should be noted that few young people demonstrated problems on the cognitive distortions scale of the CASQ-AV pre-

treatment, and clearly a significant proportion of young people were still presenting with problems as measured by the MSI-J scale post-treatment. It should also be noted that the mean score on the cognitive distortions scale of the CASQ-AV was in the treated range pre-treatment and was classed as 'better than normal' post-treatment. Attrition was dealt with appropriately and data from young sexual offenders who failed to complete treatment were reported to be indistinguishable from the young people who completed treatment on all psychosocial variables including the measures of OSA&Bs. Eastman (2004) chose an unusual design (cross-over longitudinal) that involved comparing scores on Bumby's (1996) Molest and Rape scales for three different groups (pre-treatment, post-treatment and post-release). Confounding variables were controlled for and significant differences were reported between the pre-treatment group and both the post-treatment and post-release group ($M_s = 76.08$ vs 52.75 , vs 49.00 for the Rape scale; $M_s = 76.02$ vs. 56.38 vs 51.60 for the Molest scale). Eastman (2005) used the same measures to assess three slightly different groups (failed to complete treatment, waiting for treatment, and successfully completed treatment), and used discriminant analysis to predict response to treatment. Cognitive distortions distinguished best between treated and untreated juvenile offenders. The most powerful discriminant function emerging from the analysis, and reflecting 79% of the variance, contained the instruments assessing offender cognitive distortions.

There is some evidence that OSA&Bs are amenable to specific sexual offender treatment (Edwards et al., 2012). The quality score of this study is reasonably high (78%), but the sample size was small. Results reported by Eastman (2004, 2005) also provide some support for the positive effect of treatment on OSA&Bs and the retention of these gains once young people are released from a secure setting. But again, sample

sizes were reasonably small and the weak design of this study made disentangling the effects of treatment from extraneous factors problematic. Furthermore it was not possible to establish the proportion of young people who had treatment needs pre-treatment and how many made positive gains post-treatment.

Is there an association between scores on measures of OSC and recidivism for young sexual offenders?

Only one study examined this question (Kenny et al., 2001). The definition of ‘recidivist’ in this study was a young person who was being assessed for court, having being charged with a second sexual offence and ‘non-recidivists’ were the young people being charged for the first time. Cognitive distortions were measured using a clinical rating scale with good inter-rater reliability. ‘Positive associations’ were reported for the variable ‘cognitive distortions’ and the ‘recidivist’ category of offenders, but no specific data was provided. The results of the path analysis provided interesting findings, which suggested cognitive distortions are indirectly linked to recidivism through deviant sexual fantasies, and in their mediating effects on other variables in the pathway to recidivism. However, the baseline data regarding differences between the recidivists and non-recidivists on levels of cognitive distortions was not presented. The authors took account of potential confounding variables between the groups and the study obtained a highest quality assessment score in the review (85%). However, rates of sexual offence detection do not accurately represent re-offending (for example, Falshaw et al., 2003; Langevin et al., 2004) and wrongful charges may occur. Therefore, the potential for ‘recidivists’ to be in the ‘non-recidivist group’ and vice versa is a significant limitation of this study. No conclusions may be drawn regarding whether

young people who recidivate can be discriminated from those who do not on measures of OSC based on only one study, particularly a study that adopted a pathways analysis rather than a comparative approach.

Do valid and reliable measures of OSC exist for young sexual offenders?

Only the CASQ-AV has undergone any form of standardisation on relevant groups of young people. Beckett (2006) reported a standardisation group of 97 'normal adolescents', Edwards et al. (2012) reported on 56 'post-treatment adolescents' and van Outsem et al. (2006) had a large standardisation sample ($n = 500$), but this sample consisted of Dutch youth and does not generalise to young men from other countries. Only one study reported any validation data, again for the Dutch version of the CASQ-AV. Validation data was described as 'good' for the whole ASAP pack, but individual measure and sub-scale data was not provided, and the predictive validity of the CASQ-AV, was singled out as 'poor', with a low correlation between participant self-report and therapist ratings of cognitive distortions.

Internal consistency and temporal stability of the various measures was more widely reported. Kline (2000) suggests that a correlation co-efficient of 0.8 and a sample of at least 100 respondents is the minimum acceptable standard for a test-retest analysis and a co-efficient of 0.7 is adequate and 0.8 is good when measuring internal consistency.

According to Beckett (2006), the CASQ-AV meets this standard in terms of test-retest sample size, and the internal consistency co-efficient is very high ($\alpha = 0.92$) but Beckett reported a test-retest co-efficient slightly below the acceptable figure (0.76). Edwards et al. (2012) reported an adequate internal consistency alpha ($\alpha = 0.78$) for the

cognitive distortions scale of the CASQ-AV but did not report test-retest data, possibly as the sample size was small for use in determining reliability. Van Outsem et al. (2006) report a range of correlation figures for the whole ASAP assessment battery and so it is not possible to tell how the CASQ-AV performs individually on either internal consistency or temporal stability.

The internal consistency of both the cognitive distortions/immaturity and justifications scale of the MSI does not reach the appropriate threshold when measured in young sexual offenders in the study by Butz and Spaccarelli (1999). Edwards et al. (2012) quote a range of coefficients from another study for both internal consistency and test-retest reliability for all MSI scales. Some of these figures are in the acceptable range and some are not, but it is not possible to identify a score that relates to the justification sub-scale. They report a high internal consistency alpha when tested on their own sample. Eastman (2004, 2005) reports the alpha co-efficient and test-retest scores for Bumby's (1996) Rape and Molest scales from the original adult sample; these figures are all in the acceptable /good range. The clinical rating scale devised by Kenny et al. (2001) is reported to have excellent inter-rater reliability. Van Vught et al. (2011) report that the 'nature of harm' and the 'children as sexual objects' SWCH scales both have good internal consistency but it is unclear whether the reported figures are from the original adult validation sample, or from the study sample.

In summary, evidence of standardisation and validation of measures of OSC for young sexual offenders is limited. For reliability, internal consistency figures were reported most frequently, but test-retest data was reported only occasionally. The results are clear: there is currently no reliable, valid measure of OSC for young sexual offenders.

Discussion

Main findings of the review

This study explored the role and relevance of OSC as a treatment need for young males who have committed sexual offences. Given the prominence of this treatment need in the adult male sexual offender literature, it was surprising how few studies were found that addressed this area with young people ($n = 13$). Grey literature was not sought as part of this review, unpublished dissertations were excluded and only studies that used a specific measure of OSC were included (single clinical ratings from ERASOR and JSOAP-II were excluded). It is accepted that this has limited the number of works identified and introduced some publication bias into this review, although, given the wide reporting of unexpected results publication/reporting bias is unlikely to be a significant issue in this review. To counter these limitations, five bibliographic databases were searched; manual searching of reference lists of included studies and two meta-analyses were undertaken; and a large number of experts in the field were contacted. Response from experts was good and they provided three additional studies. This provides confidence that most, if not all, relevant research is included in this review and that conclusions are based upon the synthesis of a comprehensive evidence base.

Results indicate that young sexual offenders cannot reliably and consistently be discriminated from non-sexual offenders on measures of OSA&Bs. Paradoxically, where significant differences were found, these were in the opposite direction as anticipated, with community comparisons self-reporting *higher* levels of distorted beliefs about sex with children and *greater* endorsement of rape myths than sexual offenders. Young sexual offenders cannot reliably or consistently be discriminated from

violent/delinquent youth on measures of OSA&Bs. Significant differences between groups were reported, but these were in the opposite direction as would be anticipated, with non-sexual offenders endorsing more pro-offending attitudes/beliefs than sexual offenders. Young people who abuse children could not be discriminated from those that abuse peers/adults on either measures of child abuse supportive beliefs or pro-rape attitudes but attitudes that support child abuse were found to be significantly more prevalent in a sub-group of young people offended outside of the family (extra-familial). There is some limited evidence that OSA&Bs can be modified via offence specific interventions. One study addressed the issue of recidivism, but provide insufficient information from which to draw conclusions about whether young recidivist sexual offenders and non-recidivists can be discriminated on the basis of problems with OSC. OSA&Bs in young sexual offenders were almost exclusively measured with self-report questionnaires. None of these measures was constructed using relevant samples of young people. This has resulted in both researchers and practitioners using adult measures or adult adapted measures in their work with young people. Where data on validity, reliability or standardisation of measures was provided, none of the measures had a full range of adequate psychometric properties.

In terms of the group and sub-group comparison research, methodological weakness might explain some the unexpected results. First, some studies asked generic groups of young sexual offenders to respond to questionnaires that were offence specific. As we would not necessarily expect child offenders to endorse rape myths in a significant way or peer/adult abusers to hold high levels of child abuse supportive beliefs, this might have diluted the strength of the results. Second, some studies used mixed groups of young sexual offenders from secure facilities and the community.

Arguably, young people who have engaged in sexually harmful behaviour but who remain in the community would be much more likely to be first time offenders and have committed more minor offences than young people who were detained in a secure facility. Therefore, psychologically, the former group may be much more like community non-offenders than the latter. This, too, may have affected the results. Third, few studies controlled for confounding variables, such as cognitive ability, levels of risk, gender of victim, treatment status. Finally, the lack of psychometrically sound measures is cause for concern, as it is not possible to tell whether the adult (adapted) measures have the sensitivity to detect OSA&Bs as exhibited by young people, although it is unlikely that OSC is a completely different concept in this group.

These weaknesses are unlikely fully to account for the lack of discrimination between young sexual offenders and community control groups, or the fact that the community groups consistently scored *higher* than sexual offenders on measures of OSA&Bs. One proposed explanation is that it relates to the differing contexts in which the assessments took place, with Beckett (2006), Tidefors et al. (2011), Edwards et al. (2012) and van Outsem et al. (2006) all emphasising that the young sexual offenders were known to the authorities as having engaged in sexually harmful behaviour and therefore might have felt under pressure to provide socially desirable responses, whereas community controls were given total anonymity. A tendency to respond in a socially desirable way ('fake good') when there is a reason to do so (for example, a parole assessment) has also been noted for adult sexual offenders (Arkowitz & Vess 2003; Gannon, Keown, & Polaschek, 2007). However, an alternative interpretation is that pro-offending attitudes/beliefs are genuinely prevalent in young non-offenders in the community. No matter how objectionable most people say they find the notion of

sexually harming a child or adult, there is evidence to suggest that adult males in the community hold pro-rape attitudes and an interest in sexual aggression (Calhoun, Bernat, Clum, & Frame, 1997) and some an interest in child molestation (Gannon & O'Connor, 2011). There is no reason to suggest that the same could not be true for young males in the community. A further alternative explanation is that neither sexual offenders nor community controls hold offence supportive beliefs. Indeed, it has been noted in the adult research that even when significant differences do exist between questionnaire scores, endorsement of OSA&Bs is often very low, even by sexual offenders (Langevin, 1991; Tierney & McCabe, 2001). Insufficient details were provided in the results sections of the included studies to allow any analysis of the extent to which the groups endorsed offence-supportive items in the questionnaires. Regardless, if young sexual offenders and non-offenders do in fact both hold OSA&Bs or indeed if neither group do, there would be reasonable doubt for the role of this construct in the aetiology and onset of sexual offending for young people. It would also call into question whether this variable is a relevant treatment need for this group.

Turning to the lack of discrimination between young sexual offenders/other offenders and child abusers/peer-adult abusers, the young people in each subgroup came from the same location and were tested in the same conditions, therefore demand characteristics thus cannot account for this. Although mixed secure/community groups were used in some studies, arguably with differing levels of risk, this was also true of comparison groups and is unlikely to have affected the results. Taking all this into account, one explanation for the non-discrimination between young sexual offenders/other offenders is that both groups do in fact hold distorted attitudes/beliefs about sex. If this is the case, this suggests that having pro-offending attitudes does not

make a unique contribution to the pathway to sexual offending, as opposed to other kinds of offending in young people. In two of the three studies examining differences between child abusers and peer/adult abusers on offence-specific measures of child abuse and rape supportive beliefs, confounding variables were controlled for. This suggests that the lack of discrimination between groups could be because both sub-groups of young sexual offenders hold these kinds of beliefs, or neither group does. The same pattern is observed in relation to pro-rape beliefs in adult sexual offenders (Bumby, 1996) but not with child abuse supportive beliefs (Mann et al., 2007). If young people who abuse children hold rape supportive beliefs and young people who abuse peers/adults hold child abuse supportive beliefs, or if neither do, this would challenge theories that propose that child abusers and rapists develop a specific set of implicit theories, schemas, attitudes or beliefs that are linked to the age of their victim and play a role in their specific offence pathway. However, again, results in the studies were reported in such a way that it was not possible to analyse this issue.

Although based on only one study, the finding that extra-familial offenders have higher levels of cognitive distortions about sex with children than intra-familial offenders is consistent with previous results in the adult literature (Fisher et al., 1999; Mann et al. 2007; Seto, Babchishin, Pullman, & McPhail, 2015). This suggests that, for certain sub-types of young people who abuse children, that is, those that go outside of the family to find child victims, OSA&Bs could represent an important treatment need.

As regards treatment change on measures of OSC for young sexual offenders, the study by Edwards et al. (2012) highlights the utility of the clinically significant change methodology. This contrasts with the methodology employed by Eastman (2004, 2005), which on the face of it appears to reveal significant differences on measures of

OSA&Bs for those who have completed treatment, but the nature of the experimental design makes it impossible to establish the percentage of young people with pre-treatment OSC needs and how this changed post-treatment. Additionally, the fact that Eastman compared scores from different groups of young people rather than following the same young people through treatment and into the community disallows support or refutation of the proposal that offence-specific treatment has an impact on pro-offending attitudes/beliefs. Therefore, Edwards et al. (2012), is the only study in the review that can be used to examine the question about whether scores on measures of OSA&Bs change following treatment for young people who sexually offend. This study was assessed as having good methodological rigour, and therefore offers tentative support for the proposal that offence-specific treatment impacts on OSA&Bs in young people. Interestingly, however, and in keeping with the earlier proposal that young people who sexually offend may not have treatment needs in the OSC domain, results from the cognitive distortions scale of the CASQ-AV suggest that few young people had treatment needs in relation to distorted attitudes/beliefs about children and sex pre-treatment.

In terms of measures used, the lack of psychometrically-sound assessment tools has been noted. Whilst internal consistency figures were reported fairly frequently, and test-retest figures were reported occasionally, it is generally accepted that ‘reliability is necessary but not sufficient for validity’ (Kline, 2000, p. 29). A test might demonstrate internal and temporal stability, yet still not measure what it is purports to, which has implications for practitioners and researchers using these measures, and for the weight that can be attached to the findings of this review. The types of measures used were also extremely limited compared with those available for adults; all but one study relied

upon self-report questionnaires. Given the high number of studies that have examined implicit theories in the adult male sexual offender literature, it was surprising that only one study touched on this concept in young people, with most studies taking a more traditional approach in examining ‘cognitive distortions’, attitudes and beliefs.

Despite the methodological limitations of some of the studies reviewed here, and the absence of psychometrically-sound measures of OSC, there is some consistency in the results. Overall, regardless of the method of measurement, there is little evidence that OSA&Bs are a treatment need for young sexual offenders or specific sub-groups of young offenders. The exception is for a distinct subgroup of young people who offend against child victims outside the family. Taken together, these results suggest that, contrary to the proposals in many current prominent theories of sexual offending (Abel et al., 1984, 1989; Marshall & Barbaree, 1990; Marshall & Marshall, 2000; Ward & Beech, 2005; Ward 2000; Ward & Keenan 1999; Mann & Beech, 2003), OSC might not play a significant causal role in sexual offending behaviour for most young people. As noted in the introduction, compared with adults, adolescents and young adults are impulsive risk-takers who struggle to manage their emotions/behaviour and are heavily influenced by their peers (Calleja, 2013). This stage of life is characterised by cognitive, emotional and physical change, and Rich (2009) highlights how, compared with adults, young people at this time of life do not have ‘fixed ideas, interests and motivations’, but are ‘more fluid in every aspect of their lives’ (p. 432). When viewed through a developmental lens, it is easy to see why the established and relatively inflexible implicit theories, schemas, attitudes and beliefs that have been empirically demonstrated to play a role in adult sexual offending, might be less significant in young people. It is possible that young people who sexually offend are much more influenced by the

context in which they find themselves, including their home, family, peers, sexual drive and opportunities to offend, rather than deeper level cognitive structures, which may still be developing. This is not to suggest that OSC has no role to play in the sexual offending of young people and that it should not be addressed in treatment. It may simply be that the relationship between the two is more complicated. For example, OSA&Bs may interact with more established risk factors (for example, offence-related sexual interests) to make sexual offending more likely (Mann & Beech, 2003). This is an area worthy of future research.

Strengths and limitations of the review

The main weaknesses in this review were identified in the introductory paragraph, namely, excluding grey literature and data from risk assessment measures. Additionally, it is accepted that some of the studies reviewed were of poor quality and excluding these may have raised the overall quality and increased the weight that could be attached to the findings. The strengths of this review are that a comprehensive search strategy was used and that only studies with control groups were included. Furthermore, long held assumptions about the role of OSC in young sexual offenders have been challenged, and the potential impact of these findings for practice in the field is large.

Implications for practice and future direction

There is an urgent need to test robustly the psychometric properties of existing measures of OSC for use with young sexual offenders, and/or construct and test measures that are designed specifically for this group. This will enable research to be undertaken that may more effectively examine whether OSC is a treatment need for young people and allow

for young people's individual needs in this area to be more appropriately assessed. Given the prominence of implicit theories as a concept in the adult sexual offender literature, coupled with the fact that the associated theory has a degree of empirical support (Ó Ciardha & Ward, 2013) it would be wise to construct measures that could access these deeper level cognitive structures as they develop. It might also be useful to conduct a review of the use of indirect measures of cognition with young sexual offenders in order to overcome some of the difficulties inherent in self-report questionnaires that are transparent and open to socially desirability. To the best of the author's knowledge, no such studies have yet been conducted with young male sexual offenders. This could be fertile ground for future research. The results of this review have led to the proposal that OSC might not be a significant treatment need for all young sexual offenders, but relevant for some. It will be important that future research examines different sub-groups of young sexual offenders and that the provision of sexual offender interventions for young people take an individualised approach to setting treatment targets for their clients.

Conclusions

OSA&Bs have some empirical support as a treatment need for adults, and it appears that it may have been *assumed* that the same is true for young people who sexually offend. Despite some weakness in the studies included in this review and the review itself, current findings call into question the role and relevance of OSA&Bs as a treatment need for younger sexual offenders as a group (although extra-familial child abusers might be a special case). There is some evidence that OSA&Bs are amenable to intervention, but if this factor was not a significant treatment need in the first place,

addressing this area in treatment may be misguided. There has been too little examination of the links between OSA&Bs using specific assessment measures and recidivism to be able to draw any firm conclusions about this relationship. There are no fully validated, reliable or standardised measures of OSA&Bs for young sexual offenders and it is accepted that this situation may in part explain some of the unusual results identified in this review. It is critical that the psychometric properties of current measures are established and/or developmentally sensitive measures of OSA&Bs are devised and that these measures are constructed and tested using data from relevant samples. Only then will it be possible to finally establish whether OSA&Bs are a treatment need for young sexual offenders (or specific groups of offenders) and whether it should be a core target for intervention.

CHAPTER 3: THE CHILDREN & SEX QUESTIONNAIRE – ADOLESCENT VERSION (CASQ-AV): A PSYCHOMETRIC CRITIQUE

Introduction

In Chapters 1 and 2, it was established that OSA&Bs are a form of OSC that feature prominently in sexual offender theory (Abel et al., 1984; Abel, Becker, & Skinner, 1987; Marshall & Barbaree, 1990; Mann & Beech, 2003; Marshall & Marshall, 2000; Ward 2000; Ward & Beech, 2006; Ward & Keenan 1999). There is some empirical evidence for OSA&Bs as a criminogenic risk factor in adult sexual offending but the systematic review contained in Chapter 2 identified that evidence for the role and relevance of this construct as a treatment need for sexual offending in younger populations is poor. It has been suggested that one of the problems with research into OSC is the reliance on transparent self-report questionnaires as a measure of OSA&Bs (Gannon & Polaschek, 2006; Keown et al., 2010) which are often assumed to be subject to social desirable responding by offenders who are either embarrassed to disclose aspects of their personal functioning that might portray them as ‘deviant’ or ‘risky’, or are worried about the consequences for sentencing or parole if they do (Arkowitz & Vess, 2003; Langevin, 1991; Mills & Kroner, 2006; Kroner & Weeks, 1996). However, evidence also exists that suggests psychometric measures can be used reliably in correctional settings to assess criminogenic treatment needs and predict recidivism (Mathie & Wakeling, 2011; Wakeling & Barnett, 2014). This debate is not yet resolved, but OSA&Bs are a key treatment target for many sexual offender treatment programmes (McGrath, Cumming, Burchard, Zeoli, & Ellerby, 2010) and self-report questionnaires remain the most popular way to measure them (Beech et al., 2013).

Young people are physically, emotionally and cognitively different to adults. Seto

& Lalumière (2010) have commented on how the lack of valid, reliable and standardised assessment measures that take account of these differences has hampered the advancement of knowledge about the risk and treatment needs of young sexual offender compared to adults. An absence of psychometrically robust measures of OSA&Bs that have been developed with or tested on younger sexual offenders was identified as a key area of concern in the systematic review contained in Chapter 2, and it was demonstrated that there is a reliance on questionnaire measures of OSA&Bs that have been constructed and tested on adults. Although some measures have been ‘adapted’ for younger populations, no current measures are developmentally sensitive; that is, none is developed by analysing the content of what young sexual offenders say about their offending.

The Children & Sex Questionnaire – Adolescent Version (CASQ-AV; Beckett, 1995) is a measure of ‘distorted beliefs about children’s sexuality’ (cognitive distortions CD scale) and ‘emotional congruence with children’ (EC scale). It was the measure of OSA&Bs most commonly used in the studies in the systematic review, but its psychometric properties were not well reported. In addition, HM Prison Service has used the CASQ-AV since 2002 as a pre- and post-treatment measure for young adult offenders aged 18 to 21 years who undertake sexual offender treatment in two young offender establishments in England (H. Wakeling, personal communication, January 2014). This questionnaire was adapted from the adult version of the Children and Sex Questionnaire (CASQ; Beckett, 1987), which is used by HM Prison Service and The Probation Service as part of the assessment and treatment process for adult males (aged 21 years and over) undertaking the sexual offenders treatment programme in the UK.

The CASQ is not a formally published test but some of the psychometric

properties of this measure are provided in a number of different publications. In term of reliability, Beech, Fisher, & Beckett (1999) reported that the CD scale of the CASQ had a high level of internal reliability (alpha coefficient of 0.9) and the test-retest reliability was 0.77 for the CD sub-scale and 0.63 for the EC subscale when measured using 45 untreated child sexual offenders. In terms of validity, the CD scale of CASQ demonstrated a significant correlation with the two sub-scales of the Sex with Children Scale (SWCH; Mann et al., 2007), which is a valid and reliable measure of child abuse supportive beliefs, with correlations of between 0.56 and 0.63 reported. This indicates the subscale has a degree of concurrent validity, although the two tests may be measuring slightly different constructs. The discriminant properties of the CASQ are mixed; only extra-familial sexual offenders obtained significantly higher scores on the CD scale than a prison officer, community comparison group (Fisher et al., 1999), but the EC subscale did discriminate between the child offender and community group, with the prison officers reporting significantly higher emotional affiliation with children than the offender group. The measure has poor predictive validity and, in a sample of 3,402 sexual offenders treated in the community, CD and EC scale scores did not predict reconviction for either violent or sexual offences, and change in scores following treatment was also not associated with decreased recidivism (Barnett et al, 2012, 2013). So, in summary the reliability of the CASQ appears adequate and there is some, but not conclusive, evidence of the validity of the test when used with adults.

The CASQ-AV has been chosen for this critique for a number of reasons. First, there is doubt regarding some of the psychometric properties of the adult version of the measure. However, even if these were confirmed, this could not be relied upon as demonstrating that the measure is valid and reliable for use with younger sexual

offender populations. Second, the CASQ-AV is in current use in a wide range of settings in UK and Europe. For example, it is used to assess individual need and progress in treatment (Edwards et al., 2012; Tidefors et al., 2011; Van Outsem et al., 2006) to assess the efficacy of sexual offender interventions overall (Edwards et al., 2012), and in the HM Prison Service, change on the measure following treatment is used to contribute to wider risk assessment, which can influence parole and release decisions (H. Wakeling, personal communication, January 2014). Third, the systematic review suggests that the psychometric properties of the measure may be lacking, which raises issues for defensible decision-making. Taking all of this into account, there is a need and significant potential for providing an evidence-based assessment of the reliability and validity of this measure. This review will therefore examine the psychometric properties of the CASQ-AV, but as offence-supportive cognition is the subject of this thesis, this critique will focus on the cognitive distortions (CD) scale of the measure and the measure overall, but will comment on the emotional congruence (EC) scale, where necessary. This critique will examine the applicability of the CASQ-AV as a pre- and post-intervention measure for practitioners completing treatment needs analyses and risk assessments and also researchers evaluating treatment efficacy overall for adolescent (aged under 18 years) and young adult (aged between 18 and 21 years) sexual offenders.

Overview of the tool

Background to the measure

The CASQ-AV forms part of a larger pack of self-report psychometric measures developed by the Adolescent Sexual Abuser Project (ASAP; Beckett, Gerhold, & Brown, 2002). The ASAP was created to develop a standardised psychometric test

battery that could be reliably used in a range of treatment settings to assess the socio-affective functioning, distorted attitudes and self-management of adolescents (aged 12 to 18 years) who had engaged in sexually harmful behaviour (Beckett, 2006). It was hoped that it could also be used to evaluate treatment efficacy. In addition to the wide use of the ASAP pack in the UK, revised and translated versions of the battery are used in the Netherlands (van Outsem et al., 2006), and in Sweden (Tidefors et al., 2011).

Development of the measure

The original adult version of the CASQ was developed by Beckett (2002) and the items were based on his clinical experience of the distorted comments adult child molesters commonly expressed about children and sex when assessed (R.C. Beckett, personal communication, January 2014). There is no formal published manual for the ASAP battery of tests or the individual measures in the pack (R. C. Beckett, personal communication, January 2014), only an unpaginated 'Background and Description of Measures' document with a scoring guide (Beckett, et al., 2002). This document states that the CASQ-AV was designed to measure sexual offenders' 'distorted beliefs about children's sexuality' and 'emotional congruence with children'. Closer examination of these subscales identified that the CD subscale contains items that relate to a mixture of beliefs that might support the sexual abuse of children. For example, children are sexually mature and motivated to have sex with older people; children are sexually provocative; and children are not harmed by having sex with older people. EC subscale items emphasise personal emotional connection and affiliation with children, for example, 'I prefer to spend time with children'. The CASQ and the CASQ-AV are similar measures, except that three 'lie' items have been eliminated from the CASQ-AV

and, in some questions, the wording has been changed to reflect the age of the respondents. So, for example, 'Children want sexual contact *with adults*, is changed to 'Children want sexual contact with *people of my age*'.

Critical features and scoring of the measure

The CASQ-AV is a self-report questionnaire with 84 items. Respondents are asked to rate the items using a 5-point Likert-type scale depending on how much they agree with the statements presented. Responses are coded as 0 = very true, 1 = somewhat true, 2 = somewhat untrue, 3 = very untrue, and 4 = don't know. Items are recoded at the scoring stage as follows, 4 = very true, 3 = somewhat true, 2 = don't know, 1 = somewhat untrue, and 0 = very untrue. Missing items are given a score of 2. Higher scores are considered to represent a greater degree of endorsement of attitudes supportive of the sexual abuse of children or emotional identification with children.

Although there are 87 items on the CASQ and 84 items on the CASQ-AV, the scoring guide indicates that only 30 items are scored. These 30 items yield the two subscales, 'cognitive distortions' (CD scale) and 'emotional congruence with children' (EC scale), which can be summed to produce an overall score, although given that the two subscales appear to measure different constructs, the usefulness of this is questionable. Example items in the cognitive distortions scale from the CASQ-AV include 'Children know a lot about sex', 'Not all sexual contact between people of my age and children causes harm', and from the emotional congruence scale 'Thinking about children makes me feel good' and 'Some children prefer to be with me than their parents'. There are 44 items that act as 'filler items' so that the critical items are embedded in items that are

not scored and there are also ten ‘lie items’ that can be used to assess the degree of socially desirable responding (R.C. Beckett, personal communication, January 2014).

Psychometric characteristics of the measure

Level of measurement. Kline (2000) highlights that ideally, for true scientific measurement of a construct, a ratio scale should be used where the distances between scores on the variable is known and the scale has a true zero point. However, Kline (2000) accepts that most psychological tests are based on interval level data and although the distances between the intervals on these scales can never be equal, this can be assumed if the test has been constructed in an appropriate way with sufficient breadth of item content. Kline (2000) further argues that making this assumption is both acceptable and sensible as it allows sophisticated analysis to be conducted on measures that produce meaningful results and that have, in turn, impacted on both psychological theory and practice.

The CASQ-AV uses a 5-point Likert-type scale. This is not a true interval scale, but for analysis purposes it is treated as one, for the reasons cited above. Kline (2000) highlights that selecting a middle category of response that indicates uncertainty is a ‘fatally attractive’ answer for some respondents and this might be why Beckett chose to allocate the item response ‘*don’t know*’ a code of 4 (recoded to 2 at the analysis stage). The tendency to use Likert items with an odd number of response anchors has been criticised in research that assesses OSA&Bs, as it allows respondents to take a neutral stance and avoid either agreeing or disagreeing with the offence supportive statements, thus potentially restricting the usefulness of the scale (Bumby, 1996). Regardless, most

measures of OAS&Bs continue to use allow for a 'neutral' response (Gannon et al., 2009).

Reliability. Reliability in classical test theory refers to the consistency of the test, both internally and in terms of its temporal stability over time, which is known as test-retest reliability (Kline, 2000).

Reliability: Internal consistency. Internal reliability is a measure of how related scale items are. The most common way to measure internal consistency is with Cronbach's Alpha (Kline, 1986). It is generally accepted that when testing the reliability of a measure, a co-efficient of 0.7 is adequate and 0.8 is good (Kline, 2000). Whilst some psychological measures report excellent reliability coefficients of over 0.9, Kline cautions that this can indicate a problem with test construction. For example, reliability coefficients may be boosted by having many items that repeatedly ask the same thing, something that Cattell (1973) referred to as 'bloated specifics' – such measures might be too specific and might fail to measure the breadth of content within a variable.

Several internal consistency figures have been reported for the CASQ-AV. The findings for the whole of the ASAP pack report Cronbach's Alpha for the CASQ-AV as 0.92 (Beckett, 2006). This suggests that, despite the CASQ-AV measuring two separate constructs, it has extremely high internal consistency. Internal consistency figures for the CD and EC scale of the CASQ-AV were not reported separately. A recent treatment evaluation study that used the CD scale of the CASQ-AV as a measure of treatment change, yielded an alpha co-efficient of 0.78 for the scale, which is acceptable (Edwards et al., 2012), however, no details of the characteristics of the sample used to secure this

figure were given. The Dutch translated version of the ASAP battery has been subject to reliability testing and Cronbach alphas were reported to range from 0.70 to 0.93 (van Outsem et al., 2006). However, internal consistency measurements for individual tests and scales are not reported. It is therefore not possible to establish whether internal consistency of the scales on the translated versions of the CASQ-AV are in the adequate or excellent range.

In summary, all figures reported for the internal consistency of the different versions of the CASQ-AV are at the minimum in the acceptable range, with some of the estimates in the excellent range. A review of the items on the measure suggests that internal consistency has not been artificially inflated by including a high number of similar items.

Reliability: Test-retest. Test-retest reliability is measured by correlating the test scores from respondents who take the same test on two separate occasions. According to Kline (2000) a correlation co-efficient of 0.8 is the minimum acceptable figure. Kline (2000) suggests that a test-retest sample should consist of at least 100 respondents to minimise statistical error and that testing should be at least three months apart to reduce learning effects. The test-retest reliability coefficient of the CASQ-AV is reported as 0.76 over 14 days (Beckett, 2006), which is slightly below the acceptable figure quoted by Kline (2000). Furthermore, the number of young people who completed the CASQ-AV twice is not reported and the test-retest period of 14 days is significantly below the desired period of at least three months; this might have boosted the correlation if respondents were able to recall some of their answers. Van Outsem et al. (2006) reported that test-

retest correlations on all scales of the ASAP pack ranged from 0.78 to 0.98 when measured over 14 days with 28 secondary school students aged between 14 to 17 years.

Taken together, the test-retest reliability figures reported for the CASQ-AV are almost at the acceptable level, but the lack of information about sample sizes and testing conditions weakens the strength of conclusions that may be drawn about the temporal stability of the measure.

Validity. In classical test theory, a measure is valid 'if it measures what it claims to measure' (Kline, 2000, p. 18), and there are a range of ways the validity of a test may be assessed.

Face validity. Face validity refers to whether a measure *appears* to be measuring what it is designed to measure. However, having a measure with high face validity can be disadvantageous, as respondents can detect what the test is measuring and distort their responses accordingly (Kline, 2000). This is a particular problem when attitudes or aspects of personality that may be negative or undesirable are being measured, as is the case with the CASQ-AV. Examination of the items on the CASQ-AV suggests that it is a transparent measure and it would be obvious to respondents that they are being asked to give their views on attitudes and beliefs about children and sex that are not condoned in wider society. So, it appears that the measure has good face validity but this leaves it vulnerable to social desirability.

Kline (2000) suggests that one way to reduce the tendency of respondents to produce socially desirable responses in measures with high face validity is to insert 'lie items' and provide a cut off point for discounting scores. The CASQ-AV has ten

‘impression management items’ embedded in the measure; however, these items are not scored and so they are not currently being used to identify respondents who might be providing socially desirable responses. In addition, the lie items appear conspicuous (for example, ‘my table manners aren’t as good at home as they are when I am eating out’) and appear to do little to disguise the sexual content of the other items on the questionnaire. The impact of impression management on test validity will be discussed further in the construct validity section.

Concurrent validity (criterion-related). A measure has concurrent validity if it can be demonstrated that it correlates highly with another test that measures the same criterion (Kline, 2000). However, Kline cautions that the choice of the second test may cause difficulties as it should itself be valid and reliable (‘a benchmark test’) before any meaningful conclusions may be drawn about the concurrent validity of the first test. Because of the difficulties inherent in finding other valid and reliable tests for correlational purposes, Kline (2000) suggests that moderate correlations of 0.4 or 0.5 are acceptable. Concurrent validity testing of the CASQ-AV has not been conducted in the UK, Netherlands or Sweden.

Predictive validity (criterion-related). Predictive validity refers to a measure’s ability to predict another criterion (Kline, 2000). When using psychometric test scores in a forensic assessment and or treatment, most researchers aspire to predict recidivism, although the methodological problems inherent in this type of research are immense (Beggs & Grace, 2011; Wakeling & Barnett, 2014). For example, because test scores are usually obtained in a treatment rather than a research setting, this provides a

potential motivation for respondents to present in a favourable light, in addition, follow-up times are often short and base rates for sexual re-offending are typically low, and so demonstrating an association between a particular test and reconviction can be difficult. Evidence of the predictive validity of the CASQ-AV has not been reported in the UK.

The predictive validity of the Dutch version of the CD scale (but not the EC scale) of the CASQ-AV has been investigated, but this study made use of clinician ratings as the predicted criterion (van Outsem et al., 2006). Participant scores on the CD scale of the CASQ-AV at the beginning of treatment were correlated with clinician ratings of participant cognitive distortions (on a scale of 0, 1 or 2), three months into treatment. A correlation of 0.3 was reported, which is substantially lower than some other ASAP measures. This suggests a fairly low association between participants' self-reported attitudes and beliefs on the CD scale of the CASQ-AV and clinician ratings of the same construct. Van Outsem et al. (2006) suggested the reason for this low correlation is because it is more difficult to rate the presence or absence of an abstract concept like cognition compared to an observable behaviour such as reactive aggression. However, it is also noted that the clinical rating was given three months into treatment, when it would be expected that there would have been some change in scores on the measure.

Content validity. Content validity refers to the extent to which items on a measure are representative of the whole pool of items relevant to the construct. Kline (2000) states that this form of validity testing is most suitable for tests of attainment or ability, in which there is good agreement about what the domain of items should consist of and 'there is some clear content to specify' (p. 64). As described above, the item content of

the whole measure was derived from the author's clinical experience with adult males who have abused children. The sample was additionally described as 'paedophilic', having a 'strong underlying sexual interest in children' and 'typically recidivist child sex abusers' (Becket, 2006, p. 230). This sample is likely to represent a very specific type of child sexual abuser and the content of what they told the test author is likely to be influenced by this. Furthermore, sample size and details of how the content of the items was elicited, recorded and analysed was unreported. It is thus unlikely that the content of items on the CD scale reflects the full range of 'distorted beliefs about children's sexuality' that could exist. Furthermore, as the item content is derived from statements made by adult sexual offenders, it is not possible to establish if the items represent the full range of 'distorted' beliefs younger sexual offenders might verbalise when describing their offending.

Construct validity. Construct validity refers to the extent to which the scores on a test measuring a concept fit with the theory about, or psychological nature of, that concept (Kline, 2000). Kline also highlights that the construct validity of a measure cannot be established using only one method and that a range of tests of construct validity should be performed.

How the author's clinical experience with adult child sexual offenders was transformed into questionnaire items has not been reported, but in developing the CASQ, the items were subject to factor analysis (R.C. Beckett, personal communication, January 2014). This procedure resulted in two factors with 15 items each: the CD and EC subscales. Examination of the items that form part of the test but are not scored suggests that they are highly similar in nature to the items that are scored,

possibly reflecting the same constructs. It is not clear why they did not load on, or were not selected, for either of the two derived scales. The semantic and definitional confusion surrounding OSC and in particular the term 'cognitive distortion' has been mentioned. The term 'aetiological cognition' might be a better term for this construct, as it highlights the cognition contribution to sexual offending, but this phrase not been adopted (Ó Ciardha & Gannon, 2011). There is conceptual confusion in the way the CASQ-AV is presented. One subscale is called the 'cognitive distortion' subscale, which at the time of scale development usually referred to justifications, rationalisations and minimisations for offending. When subscale items are examined, however, they appear to represent generalised attitudes and beliefs that could support the abuse of children; for example, 'children want sexual contact with older people and are not harmed by this'.

As already identified, the implicit theories hypothesis provides a theoretically and clinically prominent conceptualisation of how OSC influences sexual offending (Ward & Keenan, 1999; Drake et al., 2001; Maruna & Mann 2006; Ward 2000). Given the importance of this theory, Gannon et al. (2009) examined the extent to which six measures of OAS&Bs, including the CASQ, contained items that were relevant to the five child molester implicit theories identified by Ward and Keenan (1999). They only examined the CD scale of the CASQ but demonstrated that items from this sub-scale map onto only two of the five implicit theories and these were implicit sexual theories (children as sexual beings and nature of harm). The CASQ did not have any items that reflected the non-sexual implicit theories and so it could be argued that the CASQ, and by default the CASQ-AV, does not adequately represent the construct of OSC as defined by Ward and Keenan (1999) in its entirety. Non-sexual implicit theories could

potentially play an important role in sexual offending for young people, and it could be argued that having items that, for example, reflect the world as a dangerous place, could improve the discriminatory power of the test. However, the CASQ was developed prior to publication of the implicit theory hypothesis. Furthermore, given that the CD scale is supposed to assess 'distorted beliefs about children's sexuality' it is unsurprising that there are no items measuring the non-sexual implicit theories 'dangerous world', 'entitlement' and 'uncontrollability'.

In terms of the construct validity of the measure as it applies to younger populations, although some of the words used in the CASQ-AV have been modified to reflect the age of the respondents, the content of the questions that make up both scales remain the same. There does not appear to have been any research undertaken to establish what form of OSC might underpin the sexual abuse of children for adolescent and young adult sexual offenders and how this construct might differ for them. Van Outsem et al. (2006) assessed the construct validity of the Dutch version of the ASAP by using factor analysis to establish whether the scales measured on the ASAP-D fitted with the scale definitions. However, they do not provide specific details of how the CASQ-AV performed.

As described in the introduction, many theories of sexual offender cognition assume that OSA&Bs play a role in sexually abusive behaviour towards children, and if this is true, an appropriate test of the construct validity of the CD scale of the CASQ-AV (and also the discriminate validity) would be to assess if it discriminated between young people who sexually abuse children and those who do not. This hypothesis has been tested in the UK (Beckett, 2006), the Netherlands (van Outsem et al., 2006) and Sweden (Tidefors et al., 2011) and as reported in the systematic review, it has

consistently been found that the scale does not discriminate in this way¹¹. In fact, in the Dutch sample, young sexual offenders scored *lower* on the CD scale than did their non-delinquent counterparts. Differences in the assessment context for the sexual and non-sexual offenders and the strong motivation for the young sexual offenders to present themselves in a favourable light has been proposed as a potential explanation for the lack of discriminant validity, but this has not been tested further. The only support for the discriminate validity of the CASQ-AV comes from the comparison of extra and intra-familial offenders: the former have high levels of cognitive distortions (Beckett, 2006).

Standardisation and norms. Establishing appropriate norms describes the process of standardising a test, and it is essential to have such norms in order for scores on tests to have meaning and be of value to both the respondent and examiner (Kline, 2000). Kline also highlights that the size and the representativeness of the norm group are critical features of standardisation. Beckett (2006) describes a normative comparison group for the CASQ-AV as ‘97 normal adolescents’ (p.231). Given this norm group was compared to 235 child sex offenders, this is an adequate sample size. However, Beckett does not give any further details about this norm group and so it is not known whether they were representative of the sexual offender group in the study; for example, in terms of age, ethnicity, social class and education. The CASQ-AV scoring guide that is used with both adolescents and young adults states that individual scores should be compared to a standardisation sample based on the scores from ‘56 post treatment adolescent child abusers’. This sample appears relatively small and, as no further details are given, again,

¹¹ Tidefors et al. (2011) also reported no significant differences between known sexual offenders and community control group on the EC scale.

it is not possible to establish if the sample was drawn from a representative group. However, given this measure was designed for 12 to 18 year olds, the norm group are likely to be younger than the 18 to 21 year olds who would be undergoing treatment in custody in the UK. In their treatment evaluation study, Edwards et al. (2012) use the same standardisation sample. Tidefors et al. (2011) and van Outsem et al. (2006) standardised the translated version of the CASQ-AV on 500 Dutch 'non-delinquent youth' and 42 Swedish junior high school pupils respectively, however, these groups can not be considered as a normative group for young people in the UK due to the potential impact of cultural differences (Fisher et al., 1999).

Conclusions

The reliability of the adult version of the CASQ is reasonably well established. The CASQ-AV also appears to demonstrate a reasonable degree of reliability for use with participants under 18 years of age, although better reporting of sample characteristics and testing conditions would increase confidence in the results that have been reported. It is generally accepted that 'reliability is necessary but not sufficient for validity' (Kline, 2000, p. 29). In other words, a test may have internal consistency and temporal stability, yet still not be a reliable measure what it is supposed to measure. Despite this assertion, Kline (2000) identifies that it is not uncommon for test constructors to put a great deal of effort into establishing the reliability of a measure at the expense of establishing validity. This is what appears to have occurred with the CASQ-AV, in that the reliability of the measure has been established to a reasonable degree, but an evaluation of the validity of the measure is lacking.

One of the biggest threats to establishing the validity of CASQ-AV appears to be confusion about the concept of OSC and, until this issue is resolved, attempts to establish both the content and construct validity of both forms of the test will be difficult. The fact that neither the adult or adolescent version of the CASQ discriminates between males who offend sexually against children and those who do not represents another threat to establishing the construct validity of the test. It appears that designing and developing a questionnaire for younger people that can, first, define and access child sexual offenders' OSA&Bs; second, is not so transparent that it is subject to the social desirability bias; and, third, can discriminate between offenders and non-offenders, presents a considerable challenge. Given these difficulties with self-report questionnaires, researchers have been seeking other ways to access and assess offence-supportive cognition. Indirect measures of assessing offence-related cognition that are not under conscious control, such as the Implicit Association Task (IAT; Greenwald, McGhee, & Schwartz, 1998) have been developed, but have also encountered problems, including how to establish their construct validity (Beech et al., 2013) and discriminant properties (Keown et al., 2010).

In summary, the CASQ-AV is a measure commonly used to assess OSC in adolescent sexual offenders (aged ten to 18 years). Whilst its reliability has been established for use with this group, its validity has not. In addition, more thought needs to be given to what constitutes a suitable norm group for adolescent sexual offenders. More explicit reporting of the characteristics of this group would also be helpful. It is therefore suggested that the CASQ-AV should be used extremely cautiously in its current form in intervention settings when assessing adolescent sexual offender treatment needs, their progress in treatment and their risk of re-offending. It should also

be used cautiously when evaluating the impact of treatment programmes for this group for the same reasons. The CASQ-AV is also currently used to aid treatment planning, assess progress in treatment and make judgments about risk for young adults aged between 18 and 21 years, who are undertaking sexual offender treatment with HM Prison Service. However, no reliability or validity testing of this measure for use with this group in this setting has been undertaken, and appropriate norms are not available. This is concerning, and something that requires attention in order to provide practitioners and researchers with a more defensible position in using scores from the measure to examine treatment change and reductions in risk for individuals and groups as a whole.

CHAPTER 4: A PSYCHOMETRIC VALIDATION OF THE CASQ-AV WITH YOUNG ADULT SEXUAL OFFENDERS AGED 18 TO 21 YEARS

Abstract

Addressing OSC is a target of cognitive-behavioural treatment programmes for both adults and younger sexual offenders. The administration of self-report questionnaires is the most popular way to assess OSC. Whilst valid and reliable measures exist for adults, the same is not true for younger sexual offenders. This study assessed the psychometric properties of a measure of child abuse supportive beliefs, the Children and Sex Questionnaire – adolescent version (CASQ-AV; Beckett 1995) with 356 sexual offenders aged between 18 and 21 years who were undertaking treatment in a custodial setting. Principal Components Analysis revealed four components: each reflected a child abuse supportive belief and fitted with current conceptualisations of OSC, providing evidence of the construct validity of the measure. Overall, the measure had excellent internal consistency, and test-retest reliability, and was not susceptible to socially-desirable responding. The CASQ-AV correlated significantly with another established measure of OSC demonstrating concurrent validity. Higher-risk offenders showed higher levels of child abuse supportive beliefs than lower-risk offenders, demonstrating that the measure had a degree of discriminant validity, although these differences were only significant for one component. The CASQ-AV was sensitive to treatment change and significant reductions in child abuse supportive beliefs were observed following treatment. The CASQ-AV components matched beliefs underpinning two of Ward and Keenan's (1999) implicit theories, offering support for the proposal that maladaptive implicit theories develop in early life. The four-

component CASQ-AV can be used for research, but the results require replication before being usable for clinical purposes.

Introduction

Sexual Offender Treatment Programmes aim to reduce recidivism by targeting the changeable (dynamic) psychosocial factors associated with the risk of sexual re-offending (Mann et al., 2010). The different types of OSC that are *theoretically* believed to play a role in the origins and maintenance of sexual offending have already been discussed in this study, and modifying OSC in its many forms has been an important target for sexual offender treatment intervention since the early 1990s (Gannon & Polaschek 2006; Maruna & Mann, 2006). However, evidence presented so far suggests that whilst there is some empirical support for OSC as a criminogenic risk factor with adults, with younger sexual offenders the evidence base is weak. Despite this, modifying OSC and particularly OSA&Bs remains a clinically important treatment target in UK sexual offender treatment programmes for both adults (Barnett et al., 2012, 2013) and younger sexual offenders (Edwards et al., 2012). Thus, practitioners and researchers need ways to measure it.

The clinical importance of OSC: Two dominant paradigms

Abel et al. (1984, 1989) are often credited with being the first researchers seriously to consider the role of cognition in sexual offending, introducing the term ‘cognitive distortion’ to the sexual offender literature. Clinicians are used to hearing both adult and younger sexual offenders minimise, excuse, justify and rationalise their offending by making statements such as ‘I was drunk’, ‘she was flirting with me’, ‘I didn’t harm

him'. The cognitive distortions theory therefore had intuitive appeal. The term cognitive distortion was quickly adopted in the treatment room and attempts made to modify these distortions (Marshall, 1994; Murphy, 1990). Abel et al. (1984, 1989) clearly considered these 'distorted statements' as evidence of underlying OSA&Bs. Indeed, they developed a questionnaire (The Abel-Becker Cognition Scale - ABCS; Abel et al., 1989) that contains items reflecting a number of general beliefs that could support the abuse of children (such as sex between a 13-year old child or younger and an adult causes the child no emotional problems). Interestingly, however, this questionnaire also contains offence-specific distorted statements (for example, 'If I tell my young child what to do sexually and they do it, that means they will always do it because they really want to'). This mixing of general OSA&Bs with offence/victim specific statements may have contributed to conflation of the terms 'belief' and 'cognitive distortion' and the more general semantic and definitional confusion of the latter terms described in the introduction. The 'distorted' statements themselves were the focus of treatment in the 1980s and 90s. During this period, therapists focused almost exclusively on the *content* of what sexual offenders said; the goal of treatment was to obtain congruence between victim and offender accounts in order to make the offender take 'full responsibility' for his offence (Salter, 1988). The mechanisms and structures underpinning what was said, however, were neglected (Gannon et al., 2009). Detailed analysis of how treatment was delivered to younger sexual offender populations during this period is not available; however, given that interventions for this population were historically based on models of treatment for adults (Calleja, 2013; Letourneau & Borduin, 2008) it is likely that the content of what young people said about their offending was the target of change for this population, too.

More recently, Maruna and Mann (2006) have highlighted that humans have a propensity for engaging in excuse-making, justifying, minimising and denying after participating in many forms of socially unacceptable behavior. They furthermore propose that this tendency may actually be protective for sexual offenders, by reducing feelings of guilt and shame that can become barriers to successful engagement in treatment. Furthermore, they propose that simply addressing cognitive content might lead to the appearance of treatment change for individuals, but this change may be superficial and simply lead the individual to modify how he describes his offence to others. They argued persuasively that in sexual offender treatment there was a need for 'shifting the focus of cognitive interventions away from individual excuses and toward other aspects of self identity (beliefs, schemas, implicit theories)' (Maruna & Mann, 2006, p. 13).

This move towards identifying, understanding and modifying cognitive structures in treatment, rather than cognitive content, was greatly assisted by the emergence of the implicit theories hypothesis, a new conceptualisation of the role and relevance of OSC in sexual offending (Ward 2000; Ward & Keenan 1999). Ward and Keenan (1999) proposed that sexual offenders have unhelpful or maladaptive (implicit) theories about themselves, others and the world and that these theories both generate and are supported by networks of inter-related beliefs and assumptions. Implicit theories for child sexual offenders are thought to exist in five domains. The first two are sexual and the other three more general: *children as sexual objects* (associated beliefs: children are inherently sexual and enjoy and seek sexual contact with adults); *nature of harm* (beliefs that children are not harmed by having sex with adults or that only extremely violent sex causes them harm); *entitlement* (associated beliefs - some people are

superior to others and these people have a right to have their sexual needs met, even by children); *dangerous world* (associated beliefs - people are inherently hostile and rejecting and either this includes children, therefore they should be controlled by abuse, or it excludes children, rendering them safe sexual partners); and *uncontrollability* (associated beliefs: sexual urges cannot be controlled and external events are to blame for sexual abuse; for example, children's sexual advances, drugs and alcohol).

Child abusers are thought to use these implicit theories to interpret children's behaviour and to understand their desires, motivations and intentions. In other words, they interpret children's behaviour in a way that is consistent with their implicit theories and associated beliefs (Ward, Hudson, Johnston, & Marshall, 1997). For example, in this conceptualisation, a man who has an implicit theory about children as sexual objects may interpret a child coming to sit on his lap as a sexual invitation, rather than as the child being friendly (Ward, 1999). This illustrates how implicit theories (children are inherently sexual) generate beliefs about the motivations and desires of children (children want to have sex with adults) that may underpin the offence-specific statements frequently heard in treatment ('she came onto me', 'she initiated it', 'she wanted to have sex with me'). The implicit theory hypothesis takes account of developmental psychology, and just like other, more adaptive implicit theories, the offence supportive type are thought to originate in childhood. However, it is proposed that that sexual offenders develop maladaptive implicit theories in response to negative life events, such as sexual abuse. Therefore, there is no reason to suggest that the implicit theory hypothesis could not be applied to young adults and adolescents who offend against children. This has not been empirically tested, however.

There is some empirical support for the presence of implicit theories in adult males who abuse children. For example, Marziano et al. (2006) examined the interviews of 22 adult child molesters and found that the five implicit theories identified above accounted exclusively for the majority of ‘cognitive distortions’ in these interviews. However, other research, particularly that using indirect methods of assessment does not support the presence of implicit theories in men who abuse children (Gannon, Wright, Beech, & Williams, 2006; Keown, Gannon, & Ward, 2008). Despite inconclusive evidence, the implicit theory hypothesis has been evaluated as the strongest major theory of cognitive distortion (Ó Ciardha & Ward, 2013).

The influence of the implicit theory hypothesis of sexual offender cognition and the almost identical schema theory (Mann & Beech, 2003) has resulted in implicit schema-based treatment approaches being implemented in sexual offender interventions (Drake et al., 2001; Mann & Shingler, 2006). Most researchers and practitioners now agree that, in order to bring about meaningful change in OSC, underlying cognitive structures, such as unhelpful attitudes, beliefs and schemas should be targeted for change in treatment, rather than surface-level rationalisations and excuses for behaviour (Beech et al., 2013; Dean, Mann, Milner & Maruna, 2009; Mann & Beech, 2003; Maruna & Mann 2006; Ó Ciardha & Gannon, 2011). There is some evidence that these approaches are also being adopted with younger sexual offenders (Calleja, 2013; Edwards et al., 2012; Richardson, Bhate & Graham, 1997). Therefore, practitioners need valid, reliable measures to assess these cognitive structures. Evidence presented in the systematic review suggests that these measures are unavailable for younger sexual offender populations.

Psychometric measurement of the beliefs that support child sexual abuse with adult sexual offenders

In order to accurately assess treatment change in core treatment targets, valid and reliable measures of the constructs underpinning these targets must be used (Beckett, 1994; Beggs 2010; Grady et al., 2011). How to develop valid, reliable measures with good psychometric properties has been the subject of many books (Guildford, 1956; Kline, 1986; Kline, 2000; Nunally, 1978). In brief, psychometrically-sound measures should demonstrate reliability over time (temporal stability) and be internally consistent. They must also demonstrate validity (measure what they claim to measure) in a range of ways, including: appearing to measure what they are designed to measure (face validity); containing a representative pool of items relevant to the construct (content validity); correlating highly with other tests that measures the same or a similar criterion (concurrent validity); predicting other criteria (predictive validity); and the scores on the measure should fit with the psychological theory of that concept (construct validity). Psychometrically-sound tests should also be standardised with appropriate normative groups in order for scores on tests to have meaning and be of value to both the respondent and examiner, although this might be difficult with forensic populations in which the construct of interest often has little relevance to any control group.

Measuring OSA&Bs with sexual offenders typically involves administering self-report questionnaires containing a number of items that reflect relevant offence-related beliefs. The items contained in most measures of child abuse supportive beliefs were derived from the authors clinical and research experience of working with adult men who sexually abuse children (Gannon & Polaschek, 2006). Grady et al. (2011) conducted a review of the measures used with adults to assess the common treatment

targets for sexual offender interventions, including child abuse supportive beliefs. This identified three published scales that could be used to measure the attitudes and beliefs that support child sexual offending: the Abel-Becker Cognition Scale (ABCS; Abel et al., 1989), the MOLEST Scale (Bumby, 1996) and the Cognitive Distortions and Immaturity (CDI) and Justification (JU) subscales of the Multiphasic Sex Inventory (MSI; Nichols & Molinder, 1984). Grady et al. (2011) emphasise that all three measures demonstrated reliability and validity. However, although the ABCS is described as measuring ‘29 common cognitive distortions of child molesters’ (Abel et al., 1989, p. 139), closer examination highlights that it contains a mix of offence-specific justifications and child abuse supportive beliefs. It is unclear which offence supportive constructs it measures. For example, ‘If I tell my young child (step-child or close relative) what to do sexually and they do it, that means they will always do it because they want to’ is an offence-specific statement, whereas ‘Sometime in the future, our society will realise that sex between a child and an adult is all right’ is an example of a belief that having sex with children is acceptable and not harmful. The discriminant validity of the ABCS has also recently been questioned (Benbouriche et al., 2015; Gannon et al., 2009). One of the main issues with the MSI is that, to undertake the assessment, it is assumed that the respondent has committed a sexual offence. It is therefore not possible to examine whether the questionnaire demonstrates discriminant validity; that is, whether it can discriminate between child sexual offenders and other types of sexual offender or non-offenders. In addition, it has been noted that some items appear to assess beliefs about treatment rather than child sexual offending (Bumby, 1996).

Of the three measures identified by Grady et al. (2011), the MOLEST scale

appears most thoroughly tested, demonstrating sound internal consistency and test-retest reliability, convergent validity (when correlated with the ABCS and MSI sub-scales), discriminant validity (with differences on scores between child molesters and both rapists and non-sex offenders), and the measure was unaffected by social desirability (Bumby, 1996). Furthermore, significant reductions in child abuse supportive beliefs were reported in the initial stages of a programme of cognitive restructuring when measured using MOLEST (1996).

Two other commonly used measures of child abuse supportive beliefs not identified by Grady et al. (2011) are the Sex with Children Questionnaire (SWCH; Mann et al., 2007) and the Hanson Sex Attitude Questionnaire (SAQ; Hanson, Gizzarelli, & Scott, 1994). The SWCH (Mann et al., 2007) has been rigorously tested with a large sample and the results replicated in a further sample. Mann et al. (2007) report that it demonstrates excellent internal consistency and test-retest reliability and that the two sub-scales ('harmless sex with children' and 'provocative sexual children') were not subject to social desirability bias. The SWCH demonstrated convergent validity in that the subscale scores correlated with the 'cognitive distortions' scale of the Children and Sex Questionnaire (CASQ; Beckett, 1987) and the CDI scale of the MSI. It demonstrated discriminant validity (significant differences were found between the SWCH scores and different types of sexual offender and non-offenders) and predictive validity (SWCH scores and risk levels as measured by Risk Matrix 2000 differed significantly; Thornton, et al., 2003). However, Walton, Duff, and Chou, (2014) recently re-examined the SWCH, highlighting that the ability of the test to predict sexual and violent reconviction was no better than chance and suggesting that the discriminatory powers of the measure lie in identifying those who disagree more or less

with child abuse supportive statements, rather than those who agree with them.

The SAQ (Hanson et al., 1994) has been subject to less rigorous testing but is reported to have acceptable internal consistency for all subscales when tested on a mixed group of incest offenders, violent offenders and community controls. However, the test-retest reliability and validity of the measure have not been reported. The SAQ (Hanson et al., 1994) was developed for, and tested on, incest offenders, which limits its utility. However, one advantage of this questionnaire over similar measures is that it is the only one with a subscale that measures ‘sexual entitlement’ beliefs, which are thought to underpin the sexual entitlement implicit theory (Ward, 2000; Ward & Keenan, 1999).

There is a degree of consistency in the child abuse supportive beliefs assessed across measures. All measures assess beliefs that children are or can be sexually enticing/provocative, that they want to have sex with adults, that they are mature enough to have sex and that it does not harm them. In addition, the SAQ measures beliefs about sexual entitlement, the acceptability of extra-marital affairs, sexual frustration and positive relationships being sexual.

Given the influence of Ward’s (2000) implicit theory hypothesis on treatment, Gannon et al. (2009) examined a range of current measures of OSC (ABCS; MOLEST; SAQ; CASQ; the CDI scale from the MSI and QACSO; Lindsay, Whitefield, Carson, Broxholme, & Steptoe, 2004) to establish to what extent these assessments measured Ward’s five child abuser implicit theories. Results suggested that the sexual offence specific implicit theories (that is, children as sexual beings and the nature of harm) are well covered in current measures but the non-sexual offence specific theories (uncontrollability, entitlement and dangerous world) were underrepresented. There are

obvious benefits to practitioners of having measures that assess all five of Ward's implicit theories for both treatment planning and post-treatment assessment purposes. Such measures have been developed (Goddard, 2006) but none yet published.

If child abuse supportive beliefs play a role in the development or maintenance of child sexual abuse, it would be expected that individuals who abuse children would endorse items on psychometric measure that relate to these attitudes. However, results regarding the degree to which the measures described above might distinguish (adult) child sexual offenders from other types of offenders and community controls is mixed (Gannon et al., 2009). Even when child abusers can be discriminated from other types of offenders or community controls, Likert-scale scores suggest that child offenders frequently disagree with the child abuse supportive statements: they do not endorse *strongly agree* or *agree*, but simply disagree less than the non-abusers, endorsing *disagree* or *neutral* and their scores are often low and skewed towards disagreement (Arkowitz & Vess, 2003; Langevin 1991; Walton et al., 2014). These results could be interpreted as suggesting that many child sexual offenders do not hold generalised beliefs that support their abuse, or alternatively, measures in current use are not sensitive enough to capture these beliefs.

As already identified, a significant problem confronted when measuring child abuse supportive beliefs is that such views are considered repugnant by society and acknowledging these beliefs may have significant consequences in a forensic setting. A measure needs to be transparent to have good face validity (an important psychometric property) but by being so, it inevitably influences the respondent's answers. There is often an assumption that offenders generally engage in impression management (Mathie & Wakeling, 2011) and a frequently-adopted explanation for why

psychometric measures do not always distinguish between child sexual offenders and other groups, and why their scores are often low, is because child sexual offenders engage in socially desirable responding. To assess the impact of socially desirable responding on measures of child abuse supportive beliefs in both research and treatment settings, measures of this cognitive structure are thus often correlated with a known measure of socially desirable responding (Mann et al., 2007; Mathie & Wakeling, 2011; Hanson et al., 1994). However, this approach has been criticised, and it has been argued that it might be possible for child sexual offenders to produce acceptable responses on impression management items whilst being unwilling to endorse beliefs that support the sexual abuse of children (Gannon et al., 2007; Keown et al., 2010).

Evidence that child sexual offenders might engage in socially desirable responding comes from research in which the context of testing was changed, which led to corresponding changes on measures of child abuse supportive beliefs. For example, a group of child sexual offenders offered anonymity when completing the Child Molester Scale (CMS; Cann, Konopasky, & McGrath, 1995) endorsed significantly more cognitive distortions than a group being assessed for parole (McGrath, Cann, Konopasky, 1998). More recently, Gannon et al. (2007) assessed child abuse supportive beliefs in the same group of extra-familial child sexual offenders in two different settings. In Setting One, participants were free to respond as they wished; in Setting Two, the group were split into two, with one half attached to what they thought was a 'lie detector' and the other free to respond as they wished. In the lie detector setting, the participants endorsed significantly more child abuse supportive beliefs than they had previously and significantly more than the group who were not attached to the fake lie detector.

In summary, a range of measures may be used to assess child abuse supportive beliefs with adult sexual offenders. All have been evaluated with at least some psychometric testing in a relevant population. However, none of the current measures demonstrated a full range of psychometric properties, and the problems with the face validity of the measures, and why these measure do not consistently discriminate between child abusers and non-offending comparison groups, is an ongoing issue that has not yet been adequately resolved.

Measuring child abuse supportive beliefs in younger sexual offenders

One aim of the systematic review was to establish whether there were valid, reliable measures of OSC that could be used with younger sexual offenders. Findings indicated that no such measure exists. All 12 studies that used self-report measures to assess OSC with young sexual offenders used one of the adult measures described above. They were adapted to various degrees; for example, the word ‘adult’ was replaced with ‘child’ in the adolescent version of the CASQ (CASQ-AV; Beckett, 1995) and the number of items in the ABCS questionnaire was reduced from 29 to 10 (Racey et al., 2000). Nonetheless, current measures were developed via clinical experience with adults and have mainly also been tested with adults. The systematic review identified that psychometric data has been published for some measures; however, this was usually evidence for internal and temporal consistency gathered incidentally from the sample rather than as part of a planned examination of the psychometric properties of the test (e.g., Worling, 1995). In addition, psychometric data for the original adult sample was sometimes reported (e.g., Eastman, 2004). The systematic review identified the CASQ-AV (Beckett, 1995) as a widely-used measure of OSC in young sexual offenders both in

the UK and Europe. The critique of the CASQ-AV in Chapter 3 highlighted that there have been attempts to establish the psychometric properties of this measure (Beckett, 2006; van Outsem et al., 2006) but the approach appears to have been piecemeal and the results not always reported in a helpful way. Overall, it was clear that no measures of OSC currently in use with young sexual offenders had undergone the rigorous testing applied, for example, to MOLEST (Bumby, 1996) or SWCH (Mann et al., 2007).

Difficulties discriminating between sexual offenders who offend against children and other types of offenders and/or community controls on measures of child abuse supportive beliefs identified for adult males have also been reported for young sexual offenders (Beckett, 2006; van Outsem et al., 2006). A systematic review of whether young sexual offenders could reliably and consistently be distinguished from non-sexual offenders or non-offenders on measures of OSC was a key aim of Chapter 2. The results were unequivocal: they could not be. Indeed, the systematic review identified that where differences did exist between young sexual offenders and other types of offenders or non-offenders, the scores were in the opposite direction to that anticipated. In the studies reported, most authors suggested this was due to the testing conditions (that is, the young sexual offenders had all been reprimanded for a sexual offence against a child) and therefore that they must have been engaging in socially desirable responding. However, this proposal has not been empirically evaluated. Furthermore, the response patterns, and mean and range of scores on measures of OSA&Bs have not been well reported. It is therefore not possible to say whether, like adult offenders, younger populations predominantly disagreed with the child abuse supportive statements and that this is why it was not possible to discriminate them from others.

The current study

Given the dearth of psychometric measures suitable for use with young sexual offenders, clinicians and researchers are forced to rely on measures of OSC, including measures of child abuse supportive beliefs that are developmentally insensitive and psychometrically unvalidated. This could mean that practitioner assessments of individual treatment change and researcher evaluations of the impact of interventions for young sexual offenders are indefensible. There is a clear need for new measures of OSC based on the content of what young people say about their offending, but, arguably, there is a more urgent need to test the psychometric properties of the measures already in use.

As described above, the CASQ-AV is used widely in the UK and Europe, though its psychometric properties have not been formally established. In particular, as described in Chapter 3, this measure has been used to assess treatment needs and changes in child abuse supportive beliefs since 2002 for young adult sexual offenders (aged 18 to 21 years) who undergo the HM Prison Service Sex Offender Treatment Programme (Mann & Thornton, 1998) in England and Wales. It is important to establish whether the CASQ-AV is psychometrically sound so that practitioners in the Prison Service may apply it with confidence, knowing that the opinion they provide about changes to young peoples' child abuse supportive beliefs are defensible and that researchers can effectively evaluate the impact of this treatment programme on reconviction. This study therefore intends to test and report the psychometric properties of CASQ-AV.

Specific aims and hypotheses

The factor structure of the CASQ-AV will be explored using Principal Component Analysis (PCA) to test its construct validity. Given the theoretical and clinical importance of the implicit theories hypothesis, the claim that implicit theories develop in childhood and therefore should be present for young adult sexual offenders, and the lack of empirical testing of this claim, any sound factor structure identified will be compared to the five implicit theories identified for child abusers by Ward and Keenan (1999). The following specific hypotheses will then be tested:

1. The CASQ-AV is a reliable measure:
 - 1.1 It will have good internal consistency;
 - 1.2 It will have good test-retest reliability;
 - 1.3 It will not be significantly affected by socially desirable responding. In particular, it will not demonstrate a correlation with an adaptation of the Personal Reaction Inventory (PRI; Greenwald & Satow, 1970), which is a measure of self-presentation bias.
2. The CASQ-AV is a valid measure:
 - 2.1 It will show concurrent validity by correlating significantly with another measure of child abuse supportive attitudes and beliefs; namely, the CDI scale of the Multiphasic Sex Inventory - adolescent male form (MSI –J; Nichols & Molinder 1984);
 - 2.2 The CASQ-AV will show discriminant validity, in that higher risk individuals who have committed offences against children will have higher scores on the CASQ-AV than lower risk individuals, as measured by Risk Matrix 2000 (RM2000; Thornton et al., 2003).

3. If it is demonstrated that the CASQ-AV is a reliable and valid test, the child abuse supportive cognitions measured therein will demonstrate sensitivity to treatment change: offenders' scores on this measure will decrease following a specific intervention designed to address this treatment need.

Method

Participants

Primary sample. The primary sample was selected from a pre-existing dataset held by Intervention Services Group (ISG) at the National Offender Management Service (NOMS). This dataset contained CASQ-AV scores on young adult sexual offenders aged 18 to 21 years, who had undertaken the national sexual offender treatment programme (SOTP) for England and Wales at one of two Young Offender Institutions (YOIs). This study aimed to assess the reliability and validity of the CASQ-AV, not to evaluate the efficacy of the SOTP; the programme is therefore not described. The critical feature of the SOTP for the purposes of this study is that it is a cognitive behavioural intervention, and a fundamental component of this intervention is identifying and challenging cognition that supports offending (Mann & Thornton, 1998). Only young adult sexual offenders who have committed a sexual offence against at least one child complete the CASQ-AV. For the purposes of the CASQ-AV, a child sexual offender is defined as a young person whose victim was four or more years younger than themselves at the time they committed the offence. CASQ-AV data was available from young adult sexual offenders who had undertaken the SOTP between 2002 and 2013, resulting in an opportunistic sample of 182 pre-treatment participants. Post-treatment data was available for 174 participants and pre- and post-treatment matched data was available for 160 participants.

The mean age of the primary sample, pre-intervention, was 19 years and 8 months (range 17 years – 21 years 8 months; $SD = 1.01$). Prior to 2003, YOIs were permitted to detain some prisoners who were aged under 18 years. However, only three individuals in the sample were younger than 18-years old. The primary sample was predominantly White British/Irish/other white background (92.8%); the remainder described themselves as black, of mixed heritage, or of other ethnic minority group.

Some demographic data was available for participants, and indicated the presence of social disadvantage prior to being taken into custody: 31 % had divorced parents, 26% were not living in their family home when arrested, 38% had run away from home at least once and 32% reported being sexually abused before the age of 16. Physical punishment was used in the home: 78% had been hit at least once with a hand or slipper, 35% had been punched and 34% had been hit with an object. Psychosocial adjustment problems were also prevalent; 38% of participants had been referred to a psychologist or psychiatrist before the age of 16, 89% regularly drank alcohol and 54% regularly took drugs. In terms of education, training, and employment, 31% of the participants had attended a non-mainstream school (for either learning or behavioural problems) and 65% had been excluded from school at least once. At the time of offending, 23% of participants self-identified as students, 32% were employed (part time or full time), and the majority (55%) was not engaged in purposeful activity. In terms of sex and relationships, 79% of participants had experienced some form of consenting sexual contact with either a female or male, but 42 % were not in any form of relationship (intimate, permanent or casual) at the time of the offending. Having multiple sexual partners was reasonably common in this group (51%) and 69% of participants reported occasional or regular use of pornography.

Previous research has highlighted the importance of comparing child abuse supportive beliefs for sub-groups of child sexual offenders; for example, those who offend within or outside of the family and those who offend against girls, boys or both (Beckett, 2006; Mann et al., 2007; Seto et al., 2015). Owing to changes in the way offence-related data has been recorded over time, this information was only available for 30 participants (16% of the sample). Sub-group analysis was thus not conducted. Resultant sub-group sizes would be too small for meaningful analysis; it would be difficult to generalise conclusions drawn from such a small sub-group to the group as a whole.

Secondary (test-retest) sample. The secondary sample was a new sample, recruited specifically for the purposes of this study. These participants were recruited from the same two YOIs as the primary sample. They had committed at least one sexual offence against a child and were assessed as suitable for the SOTP but had not yet commenced treatment. These participants formed the test-retest sample and completed the CASQ-AV at the beginning of May 2015 and again one month later. The testing conditions for the test-retest sample were the same as for the primary sample. It was a condition of testing that participants did not undergo any form of intervention between the first and second phase of testing. This sample comprised 15 participants, whose average age at the first test phase was 19 years 5 months (range 18 to 21 years; $SD = 0.74$). The small number of participants recruited was due to a number of factors, including falling rates of young adult sexual offenders given custodial sentences, the ready availability of treatment, participants being transferred to other prisons between test and retest, and participants declining to complete the CASQ-AV again at Time Two. Owing to the

difficulties encountered securing the test-retest sample, additional demographic data was not requested for these participants.

Measures

The measures used in this study are those used in a standard psychometric battery completed by young adult sexual offenders pre- and post-treatment in the two YOIs from which the data is drawn. The specific measures are as follows:

The Beckett Children and Sex Questionnaire-Adolescent Version (CASQ-AV)

The Beckett Children and Sex Questionnaire-Adolescent Version (CASQ-AV; Beckett 1995; see Appendix 7) is a self-report measure that forms part of the battery of measures developed by the Adolescent Sexual Abuser Project (ASAP; Beckett, et al., 2002). It is based on the adult version of the Children and Sex Questionnaire (CASQ; Beckett 1987). Both versions are unpublished. The CASQ-AV and CASQ are similar. In the former, the wording has been altered to take account of the age of respondents. The questionnaire is described as a measure of ‘distorted beliefs about children’s sexuality’ and ‘emotional congruence with children’ (Beckett et al., 2002). It has 84 items, but only 30 are scored, forming two 15-item scales ‘cognitive distortions’ (CD scale) and ‘emotional congruence with children’ (EC scale). It is not known how this factor structure was derived. The other 44 questions are 34 ‘filler items’ and ten ‘lie items’ (R.C. Beckett, personal communication, January 2014). Response anchors are 0 = *very true*, 1 = *somewhat true*, 2 = *somewhat untrue*, 3 = *very untrue*, and 4 = *don’t know*. Items were recoded at the scoring stage, as follows: 4 = *very true*, 3 = *somewhat true*, 2 = *don’t know*, 1 = *somewhat untrue*, and 0 = *very untrue*. Missing items were scored as 2. Higher scores represent a greater degree of endorsement of attitudes

supportive of the sexual abuse of children. The psychometric properties of this measure have not been fully established.

The (adapted) *Personal Reaction Inventory* (Greenwald & Satow, 1970) also forms part of the battery of measures developed by the ASAP. It is a self-report measure consisting of 17 items adapted from an original scale by Greenwald and Satow (1970), designed to measure the degree to which individuals provide socially desirable responses. Edwards et al. (2012) report that the internal consistency of the adapted measure when tested on 128 non-offending British adolescent males was 0.82. This measure was used in the current study to assess the relationship between CASQ-AV scores and socially desirable responding.

The Multiphasic Sex Inventory–adolescent male form (MSI –J)

The Multiphasic Sex Inventory–adolescent male form (MSI –J; Nichols & Molinder, 1984) is a true/false self-report measure consisting of 300 items and containing 20 separate scales, six validity scales and 14 sexual deviance scales. This tool is designed to measure the sexual characteristic of adolescent male sexual offenders aged 12 to 19 years or older, if they are ‘socially delayed’. Only the Cognitive Distortion and Immaturity Scale (CDI; 19 items) was used in this study. Although this scale is often described as measuring the degree to which the respondent takes a victim stance with regard to their offending, careful examination of the items reveals that the range of items is somewhat wider than this and that there are items relating to the degree of harm victims might experience, the perceived maturity of children and wanting to be with children. The adult version of this test was rigorously tested for reliability and validity, and provided good results. As the basic scales are the same in both tests, the authors propose that the adolescent version also demonstrates good reliability and validity. The

scores from the CDI scale were used to assess the convergent validity of the CASQ-AV measure.

The Risk Matrix 2000

The Risk Matrix 2000 (RM2000; Thornton et al. 2003) is a static risk algorithm designed to predict both sexual and non-sexual violent re-conviction for men convicted of at least one sexual offence. This can only be used with males who are aged 18 or older and have committed at least one sexual offence when they were aged 16 or older. This is an actuarial measure; that is, it uses factors identified as highly predictive of sexual reconviction from statistical analysis of relevant data. The scoring is based on simple facts about the individual's personal and criminal history (for example, age at release, not being in a stable long term relationship, number of criminal convictions, having a stranger victim). Based on these scores, individuals are classified into one of four risk groups: low, medium, high and very-high risk of sexual reconviction. These classifications can be used to make decisions about the likelihood of re-offending for a sexual crime (RM2000/S), a violent crime (RM2000/V) or either crime (RM2000/c), and the intensity of treatment required. Only the risk categorisation for sexual re-conviction was used in this study. Trained administrators calculated RM2000 assessments. Good predictive validity has been reported for the scale (Barnett, Wakeling, & Howard, 2010; Craig, Beech, & Cortoni, 2013). RM2000/S categories were used to examine the discriminant validity of the CASQ-AV scales and scores.

Procedure

The CSAQ-AV and other measures were administered to all young adult sexual offenders prior to commencing the SOTP, and again six weeks after completing the

programme. Measures are administered to groups in identical testing conditions pre- and post-treatment. Extra help was available from trained facilitators for individuals who required assistance. Pre- and post-treatment measures were then forwarded to the Intervention Services Group with the RM2000 risk categorisation, for pre-treatment only. Data from each measure was entered into separate SPSS databases.

This is an archival study of data from the CASQ-AV database. The data was extracted and examined to identify unusual and duplicate entries and incorrectly entered data. Data from the databases for the other measures was matched to CASQ-AV data using the designated research identification numbers. This data was also screened. Owing to the historical nature of the databases and changes to the way the data has been collected in young adult prisons and stored at ISG, not all participants with CASQ-AV data had completed other measures. Smaller subsamples were therefore created from the primary sample for some of the analyses.

Treatment of the data

The data was analysed using the Statistical Package for the Social Sciences (SPSS; Version 22). Table 3 presents a summary of analyses conducted, measures used, sample size and sample composition for each analysis:

Table 3: Sample size, sample composition and measure used for each analysis

Analysis	Measure	Sample size <i>n</i>	Sample composition
Principal components analysis.	CASQ-AV	356	Pre- and post-intervention
Internal consistency testing on retained components	CASQ-AV components	356	Pre- and post intervention
Test-retest reliability on retained components	CASQ-AV components	15	Not yet considered for treatment
Correlation - effects of socially desirable responding	PRI	160	Pre-intervention
Correlation - convergent validity	MSI	53	Pre-intervention
ANOVA/Kruskal-Wallis H - discriminant validity	RM2000	81	Pre-intervention
Paired sample t-test - sensitivity to treatment change	CASQ-AV	160	Matched pairs pre- and post intervention

Ethical considerations

As this was an archival study using anonymised data, there were few ethical considerations. Young adult sexual offenders completed a standard consent form before undertaking assessment and treatment on the SOTP. This covers permission to use anonymised data for evaluation of the SOTP (see Appendix 9). This study formed part of that evaluation. A participant information leaflet and consent form were developed for use with participants in the test-retest study (see Appendix 10).

As participants were not identifiable to the researcher, it was not possible to provide them with feedback about the study. A summary of the research, however, was made available to the treatment managers at the two YOIs that deliver SOTP, with permission to share this information with the young sexual offenders if they wished. Ethical approval to conduct this research was granted by NOMS National

Research Committee (Ref. 2013-276) and the University of Birmingham (Ref. ERN_13-1520).

Results

Factor structure

As highlighted in Chapter 3, although the full 84-item CASQ-AV is routinely administered to young sexual offenders, only 30 of the 84 items are scored in clinical practice. Chapter 3 also emphasised that the empirical research underpinning this specific factor structure has not been reported. Given that the items were derived from the same interview with sexual offenders, it was decided to subject all items apart from the ten 'lie items' (which were irrelevant to the factor structure as a measure of child abuse supportive beliefs) to Principal Components Analysis (PCA) to examine the underlying factor structure. A preliminary examination of the dataset identified that, as expected for such an extensive questionnaire, there was missing data. A missing values analysis revealed that the number of missing values was negligible (0.9% of values). The prescribed convention for both the adult and juvenile version of the CASQ was therefore adopted and missing data replaced with a '2' (Beckett et al., 2002; Rallings & Webster 2001).

The suitability of PCA as a data reduction technique was assessed prior to analysis. Inspection of an initial correlation matrix revealed that all items (except Q48: 'there is too much talk about children and sex') had at least one correlation co-efficient greater than 0.3. Q48 was therefore removed from subsequent analyses. The overall Kaiser-Meyer-Olkin measure of sampling adequacy (KMO; Kaiser, 1970) was 0.94, which is in the 'marvellous' range (Hutchenson & Sofroniou, 1999). All individual

items were greater than 0.8, well above the acceptable range of 0.5 (Field, 2013).

Bartlett's test of sphericity was statistically significant ($p < 0.0005$). The data was thus suitable for Principal Components Analysis (PCA).

A duplicate item was noted in the CASQ-AV: Q27 and Q54 are both, 'I think about children when I am alone'. It was not clear whether this was an error on the original measure, or a form of validity check. Both items were included initially so that the complete measure could be examined and this resulted in 73 items being subject to an initial PCA. Two methods were used to make a decision about how many components to extract: examination of the Scree Plot (Cattel, 1966) and examination of the eigenvalues, only extracting components with an eigenvalue greater than one (the Kaiser criterion; Kaiser, 1960). An oblique rotation was used, as the underlying components are considered theoretically related. The solutions deemed most parsimonious were retained.

14 components had an eigenvalue of more than one. These factors accounted for 63.31% of the total variance. Component 1 accounted for the vast majority of the variance (32.28%); all other components contributed less. This 14-factor solution was extracted and rotated loadings examined. As expected, many factors were not interpretable and had low loadings. With sample sizes larger than 200, the Scree Plot may better indicate the number of reliable factors than the Kaiser criterion (Stevens, 2009). In this case, however, the scree plot was ambiguous and showed inflexions that justified retaining either three or four components. Therefore, two separate PCAs were conducted that extracted three and four components, respectively. In order to interpret the different component solutions, factor loadings after rotation were examined using the pattern matrix (Tabachnick & Fidell, 2007).

The four-component solution explained 47% of the common variance and the three-component solution explained 44%. There was some factorial complexity in both solutions in that some items cross-loaded. There is no established rule for factor loading cut-off scores. Stevens (2009) suggests a cut-off of 0.4, irrespective of sample size – a commonly-adopted procedure. Tabachnick and Fidell (2007) suggest a more stringent range of cut-off points, from 0.32 (poor) to 0.71 (excellent). However, given the large sample size (300+), items with factors loadings as low as 0.3 could be retained (Hair, Tatham, Anderson, & Black, 1998). Choosing a cut-off point is a subjective task. A range of factors were considered here, including the number of items loading on each component, the size of the main loadings, the size of the cross loadings, the contribution each item made to the component and the meaning of that component. Another important consideration was the length of the test, and maintaining a balance between the number of items in the final solution and ensuring reliability of the measure overall (Kline, 2000). Retaining only those items with a factor loading of 0.45 or above, although stringent, appeared to offer a parsimonious solution, in which cross loading and excessive numbers of items per component were avoided and theoretically meaningful components were produced.

The first three components in the three and four component solutions were reasonably similar. The first component comprised items indicative of beliefs that being with, or thinking about, children will result in positive feelings (including sexual feelings) or a reduction in negative feelings (for example, ‘thinking about children makes me feel good’, ‘when a child smiles at me, it can stir me up’, ‘being with children stops me from being lonely’). Some of the items in this component appeared in Beckett’s (1987;1995) ‘emotional congruence with children’ sub-scale. The second

component comprised of items relating to beliefs that children are sexually provocative/enticing, they want to have sexual contact with adults and are not harmed by this (for example, ‘children can flirt with people of my age’, ‘children sometimes ask people my age for sex’, ‘not all sexual contact between people of my age and children causes harm’). This component was broadly similar to the ‘cognitive distortions’ subscale of the original CASQ-AV. The third component comprised of items that related to beliefs about having good skills with children or being ‘child orientated’ (for example, ‘I am good at making children laugh’, ‘I know what children like’). The fourth component was readily interpretable and related to beliefs about children being powerful and mature (for example, ‘children are powerful’, ‘children are not as innocent as most people think’). Although potentially representing a new type of child abuse supportive belief, this component had theoretical value in that it could be related to the implicit theory about the ‘uncontrollability of the world’, or the ‘nature of harm’. It was therefore retained and the four-component solution adopted for subsequent analyses. Adopting the four- rather than three-component solution had the added advantage of less cross-loading, as well as accounting for slightly more of the common variance.

Table 2 illustrates the rotated factor loadings (pattern matrix) for the four-component solution. The three-component solution is presented in Appendix 11. Component labels were assigned that captured the conceptual meaning of the items therein, consistent with the wider literature. Component 1 (Positive Affect) had 16 high loading items, Component 2 (Child Sexuality) the same, Component 3 (Child Orientation) had ten high loading items and Component 4 (‘Child Maturity’) five high loading items.

Table 4: Factor loadings from the CASQ-AV rotated four-factor solution

CASQ-AV item	Component			
	1 Positive Affect	2 Child Sexuality	3 Child Orientation	4 Child Maturity
Q27 I think about children when I am alone	.842	-.057	.020	.027
Q7 I sometimes get very strange feelings with children	.805	-.040	-.098	.084
Q54 I think about children when I am alone	.793	.026	.016	-.083
Q82 Newspapers and television stir up my feelings about children	.792	.158	-.042	-.111
Q55 when a child smiles at me, it can stir me up	.762	.106	-.045	.037
Q81 some children make me feel funny inside	.742	.123	.045	-.030
Q16 thinking about children makes me feel good	.721	-.149	.147	.147
Q77 children are very attractive	.659	.173	.123	-.025
Q84 some of my closest friends have been children	.629	.098	-.025	.044
Q58 I feel more comfortable with children than with people of my age	.617	.075	.036	.142
Q12 I prefer to spend my time with children	.587	-.068	.034	.273
Q28 being with children stops me from being lonely	.542	.034	.220	.097
Q37 children are special for me	.497	-.017	.311	.093
Q15 I have loved children at first sight	.489	-.110	.221	.104
Q13 I talk to children about my problems	.484	-.013	-.089	.261
Q83 I can talk about my feelings with children	.459	.203	.093	.045
Q22 children tease me	.420	.249	-.177	.240

CASQ-AV item	Component			
	1 Positive Affect	2 Child Sexuality	3 Child Orientation	4 Child Maturity
Q76 Some children prefer to be with me rather than their parents	.402	.281	.235	.087
Q30 I know more about children than their parents do	.377	.199	.186	.060
Q5 I envy children	.344	.117	.065	.093
Q78 girls are like women	.316	.293	.021	.135
Q61 children sometimes ask people my age for sex	-.024	.728	.147	-.012
Q33 children want sexual contact with people my age	.110	.727	.025	-.091
Q71 Some children could teach people my age about sex	-.103	.698	.037	.111
Q70 People don't realise how much children know about sex	-.191	.696	.159	.214
Q50 children can lead people on	-.041	.695	.064	.154
Q49 children can flirt with people of my age	.112	.680	.034	-.029
Q41 when adolescents & children have sexual relationships it is not the adolescents fault	-.024	.674	.124	-.007
Q66 children can lead people of my age astray	.015	.649	.105	.155
Q72 If children want they should have sexual contact with people of my age	.422	.611	.002	-.216
Q43 not all sexual contact between people of my age an children causes harm	.275	.573	.011	-.014
Q68 There is no harm in sexual contact between children and people of my age	.497	.547	.023	-.226
Q20 some children know more about sex than people of my own age	-.093	.525	-.141	.421
Q9 children like to talk about sex	.129	.515	.004	.060

CASQ-AV item	Component			
	1 Positive Affect	2 Child Sexuality	3 Child Orientation	4 Child Maturity
Q42 if children want they should be allowed to have sexual relationships with adolescents	.491	.502	-.085	-.029
Q53 some children find me attractive	.203	.489	.148	.063
Q45 children can blackmail people of my age	-.079	.484	.077	.331
Q34 there is nothing wrong with sexual contact between children and people of my age	.422	.441	-.110	-.223
Q19 children know a lot about sex	-.094	.433	-.114	.390
Q60 sometimes I meet a child who I know has special feelings about me	.372	.423	.123	-.013
Q80 children seem to seek me out	.260	.406	.137	.086
Q73 children can make me do things against my will	.344	.345	.002	-.069
Q51 children can look after themselves	.132	.338	.031	.182
Q1 children feel safe with me	-.225	.186	.764	-.183
Q4 children like to play with me	-.171	.159	.752	.009
Q2 children like my company	-.185	.193	.732	-.065
Q29 I am good at making children laugh	.050	-.012	.689	.118
Q24 I know how to talk to children	-.009	-.008	.655	.161
Q26 I love children	.241	-.064	.582	.008
Q6 some children make me feel good	.330	-.100	.578	.066
Q56 I know what children like	.334	-.104	.526	.133
Q38 children find me easy to make friends with	.174	.023	.489	.285

CASQ-AV item	Component			
	1 Positive Affect	2 Child Sexuality	3 Child Orientation	4 Child Maturity
Q74 When I feel low children cheer me up	.434	-.051	.465	.117
Q65 I have loved a child	.378	-.073	.424	-.132
Q79 I find it easy to talk to children	.327	-.028	.421	.076
Q8 I find it hard to resist children's' requests	.320	-.077	.376	.247
Q64 I am better than most people at getting along with children	.284	.227	.373	.206
Q57 children remind me of myself	.342	.063	.355	.183
Q63 I am better than most people at understanding children	.287	.241	.338	.065
Q17 I know when children are interested in me	.251	.089	.292	.231
Q11 children are powerful	.000	.051	.134	.547
Q21 society does not give children enough responsibility	.078	.203	.003	.512
Q25 people don't know what children are like	.099	.213	.023	.488
Q40 children are not as innocent as most people think	-.127	.441	.002	.476
Q32 children know what they like	-.027	.055	.250	.468
Q14 most people my age do not understand children	.095	.073	-.014	.446
Q23 sometimes children look at me in a special way	.246	.223	.017	.391
Q44 children tell lies about people of my age	-.030	.360	-.018	.368

CASQ-AV item	Component			
	1 Positive Affect	2 Child Sexuality	3 Child Orientation	4 Child Maturity
Q59 children speak with their eyes	.126	.059	.286	.352
Q18 children are more honest than adults	.025	-.064	.243	.327
Q36 children can be trusted	.250	-.114	.203	.307
Q46 adults cannot be trusted	.149	-.024	-.025	.303
Q47 many people have a sexual interest in children	.206	.139	.051	.239
Eigen values	23.57	4.94	3.85	1.99
% of variance	32.28	6.84	5.27	2.74

Reliability

Internal consistency

The internal consistency of each of the CASQ-AV components and the CASQ-AV total score were examined separately using Cronbach's alpha. Cattell and Kline (1977) have indicated that the internal consistency of a measure may be artificially inflated by having items that ask the same question in slightly different ways. It was found that Q27 and Q54 were worded identically and both loaded extremely highly on Component 1. The alpha coefficient for Component 1 was 0.939. If Q27 was removed, it would be .932. If Q54 were removed, it would be .934. It was clear that deleting either item would make very little difference to the internal consistency of the component, but it seemed nonsensical to ask exactly the same question twice; thus Q54 was removed from further analysis. The final list of items in each component is presented in Appendix 12.

It is generally accepted that when testing the reliability of a measure, a co-

efficient of 0.7 is adequate, 0.8 is good and 0.9 is excellent (Nunnally, 1978). Table 5 illustrates the Cronbach's coefficient of reliability for the CASQ-AV measure and component scores, (without Q54). All fall into the adequate range or above. Deleting further items did not increase the alpha coefficient for the CASQ-AV component or total scores.

Test-retest reliability

A sample of 15 young adult sexual offenders completed the CASQ-AV at the beginning of May 2015 and again one month later. The test-retest reliabilities for the CASQ-AV measure and the components were all in the excellent range ($r > 0.9$), except for Component 3, Child Orientation, which was in the very high range ($r = 0.82$). All correlations were significant ($p < .001$). The results of these analyses can be seen in Table 5.

Table 5: Mean scores, alpha and test-retest coefficients for the CASQ-AV measure and components

Component & range	Mean (SD)	Cronbach's α	Test-retest R
Positive Affect (0-60)	8.99 (12.11)	.93	0.92
Child Sexuality (0-64)	10.62 (11.76)	.93	0.91
Child Orientation (0-40)	18.15 (9.04)	.88	0.82
Child Maturity (0-20)	5.56 (4.03)	.71	0.93
CASQ-AV Total Score (0-184)	43.31 (30.17)	.96	0.92

Socially desirable responding as measured by the PRI

Not all participants completed a PRI assessment ($n = 160$; 88% of the pre-treatment sample). In order to test whether the CASQ-AV measure and components were susceptible to socially desirable responding, the total PRI score was correlated with

CASQ-AV total and component scores. There was a monotonic relationship between the PRI and the CASQ-AV scores (as demonstrated by examination of scatter and PP plots), but the CASQ-AV measure and component scores were not normally distributed and some components had outlying scores. Therefore the relationship between the PRI and CASQ-AV was assessed using Spearman's Rho. As can be seen from Table 4, there was no relationship between the PRI total score and any of the CASQ-AV scores, suggesting that CASQ-AV scores are not influenced by socially desirable responding.

Validity

Concurrent validity

Not all participants completed pre-treatment assessment using the MSI ($n = 52$; 29% of the pre-treatment sample). In order to test concurrent validity of the CASQ-AV, component and total scores were correlated with the CDI subscale of the MSI-J. Scores on the CDI and the CASQ-AV components were not normally distributed and some of the CASQ-AV components had outlying data. Thus the relationship between the MSI and CASQ-AV was assessed using Spearman's Rho. As seen illustrates in Table 6, the CASQ-AV component and total scores were all significantly correlated with the CDI scale of MSI, although correlations were not strong.

Table 6: Relationships of CASQ-AV with PRI (socially desirable responding) and CDI sub-scale (concurrent validity)

	CASQ component				
	Positive Affect	Child Sexuality	Child Orientation	Child Maturity	Total CASQ-AV score
PRI	-.09	-.02	-.09	-.04	-.12
CDI	.39**	.50**	.33*	.25*	.47**

** $p < 0.001$; * $p < 0.05$

Discriminant validity (RM2000)

Not all young people had completed an RM2000 pre-treatment ($n = 81$; 45% of pre-course sample). This was due to changes in assessments administered over time, and to assessment criteria of the RM2000: that is, it cannot be administered to sexual offenders who have not been convicted at least one sexual offence aged over 16. No sexual offenders were assessed as low risk on this measure¹². There were 25 medium-risk, 41 high-risk, and 15 very high-risk offenders. As seen in Table 7, with the exception of Component 3 (Child Orientation), all CASQ-AV scores increased as participants' risk level increased.

Table 7: CASQ-AV component and total scores for each risk category

	CASQ component				Total CASQ-AV score, range: 0-184
	Positive Affect range; 0-60	Child Sexuality range; 0-64	Child Orientation, range: 0-40	Child Maturity, range; 0-20	
Risk of reoffending	Mean (<i>SD</i>)	Mean (<i>SD</i>)	Mean (<i>SD</i>)	Mean (<i>SD</i>)	Mean (<i>SD</i>)
Medium risk ($n = 25$)	5.68 (8.08)	14.00 (10.84)	20.00 (8.21)	6.40 (3.83)	46.12 (22.80)
High risk ($n = 41$)	11.92 (14.39)	15.00 (13.18)	21.95 (8.62)	7.00 (3.77)	56.85 (33.72)
Very high risk ($n = 15$)	15.26 (14.01)	21.00 (12.86)	20.20 (8.95)	8.20 (2.91)	64.66 (33.61)

To examine whether these differences were significant, some basic assumptions required for the ANOVA were violated for the CASQ-AV measure and for Components

¹² This is an artefact of RM2000: being young and single raises risk on this measure

1 to 3. Component 4 (Child Maturity) did not violate any of these assumptions.

Therefore, a one-way ANOVA was conducted for Component 4 and a non-parametric test, Kruskal-Wallis H, for the other subscales and CASQ-AV total score.

The one-way ANOVA demonstrated that differences in scores between the risk groups on Component 4 were not statistically significant (Child Maturity: $F(2,78) = 1.59, p > 0.05$). The Kruskal-Wallis H test revealed significant differences between risk groups for Component 1 (Positive Affect; $H(2) = 6.39, p = 0.041$). No further significant differences were observed. Both corrected and uncorrected post-hoc pairwise comparisons for Component 1 were carried out using Dunn's procedure. Despite the significant overall effect, none of the individual comparisons between groups indicated significantly higher endorsement of positive affect beliefs owing to increased risk level.

Sensitivity to treatment change

One hundred and sixty young people were assessed with the CASQ-AV pre- and post-treatment. Differences between pre- and post-treatment scores were examined. These were not normally distributed. Outlying scores and a leptokurtic distribution were observed, but this was not extreme. In large datasets, paired sample t -tests are fairly robust to deviations to normality (Field, 2013). Because the violations were not serious, it was decided to proceed with this test. As seen in Table 8, there is statistically significant reduction in scores on the CASQ-AV, post-treatment. Cohen (1988) proposes that effect sizes of 0.20 should be considered 'small', 0.50 'moderate' and 0.80 'large'. Child Sexuality and Child Maturity components both showed a moderate degree of improvement post-treatment; the Child Orientation and CASV component approached this range. Pooled standard deviations were used.

Table 8: Mean pre- and post-treatment CASQ-AV component and total scores and effect sizes

Component	Mean (<i>SD</i>)	95% <i>CI</i>	<i>t</i> value	<i>Cohen's d</i>
Positive Affect				
Pre-treatment	9.95(12.50)	(0.90, 4.01)	3.21*	0.21
Post-treatment	7.49 (11.28)			
Child Sexuality				
Pre-treatment	13.92 (11.62)	(3.78, 7.56)	5.92**	0.53
Post-treatment	8.23 (9.90)			
Child Orientation				
Pre-treatment	19.76 (8.98)	(2.74, 5.23)	6.30**	0.45
Post-treatment	15.78 (8.65)			
Child Maturity				
Pre-treatment	7.11 (4.19)	(1.58, 2.79)	7.19**	0.53
Post-treatment	4.92 (4.09)			
CASQ-AV total				
Pre-treatment	50.76 (30.31)	(10.03, 18.59)	6.60**	0.49
Post-treatment	36.44 (28.02)			

** $p < 0.001$; * $p < 0.05$

Discussion

This study had four main aims. The first was to examine the underlying factor structure of the CASQ-AV and to compare this to the five implicit theories for child sexual offenders (Ward, 2000; Ward & Keenan, 1999). PCA of all items contained in the CASQ-AV did not produce a component solution that replicated the CD and EC subscales. Rather, PCA revealed four components, one of which was very similar to the original cognitive distortions sub-scale and all of which were readily interpretable and consistent with psychological theory on child abuse supportive beliefs generally and specifically with some of the five implicit theories for child sexual offenders (Ward & Keenan, 1999). The components relate to beliefs that: being with, or thinking about children, will result in positive feelings, including sexual feelings (Positive Affect); children are sexually provocative/enticing, they want to have sexual contact with adults

and are not harmed by this (Child Sexuality); the respondent has good skills with children or is 'child orientated' (Child Orientation); and children are mature and powerful (Child Maturity). This finding offers support for the construct validity of the measure. The second and third aims were to establish whether these components and the CASQ-AV measure overall were reliable and valid. The CASQ-AV measure had excellent internal consistency overall and alphas for the individual components ranged from adequate to excellent. Test-retest reliability over a month was good. Responses did not appear to be influenced by social desirability. The CASQ-AV demonstrated concurrent validity and correlated significantly with another established measure of child abuse supportive beliefs (Cognitive Distortions/Immaturity scale of the MSI). It also demonstrated a degree of discriminant validity. Higher-risk young adult sexual offenders generally endorsed more child abuse supportive beliefs than lower-risk individuals, although differences were only significant for one component one: Positive Affect. As the CASQ-AV appeared to be a reliable and valid measure of child abuse supportive beliefs, the fourth aim was to test the sensitivity of the measure to treatment change. Significant reductions in child abuse supportive beliefs were found for young adult sexual offenders following treatment on a cognitive-behavioural intervention.

Examination of Component 1 (Positive Affect) identified that all items in this component (except 'children are very attractive') are personal, using the word 'I' or 'me'. These relate to the positive personal gains that may come from thinking about or being with children, either by creating positive feelings (for example, 'when a child smiles at me it can stir me up', 'thinking about children makes me feel good') or eliminating negative feelings (such as, 'being with children stops me from being lonely', 'I can talk to children about my problems'). Six items in this component

appeared in Beckett's 'emotional congruence with children' sub-scale. There is certainly an emotional aspect to Component 1. Gannon et al. (2009) examined six measures of child abuse supportive beliefs to assess the extent to which they contained items related to Ward and Keenan's five implicit theories for child sexual offenders (Ward 2000; Ward & Keenan 1999). This study included the CASQ (adult version) but the EC subscale was excluded from analysis because this subscale was considered to measure the 'affective rather than beliefs based component of child molesters' attitudes towards children' (Gannon et al., 2009, p. 331) However, Gannon et al. (2009) also emphasised that having a strong emotional connection to children may be related to having a 'dangerous world' implicit theory. Although the specific belief reflected in Component 1 has not previously been identified, it is evident that this sort of belief could support the abuse of children. For example, a young adult sexual offender who believed positive things would happen or that they could eliminate negative feelings if they were in the company of a child, may seek children out, perhaps initially to meet their intimacy needs, but creating a situation in which they might offend. Furthermore, this belief could form part of a 'dangerous world' implicit theory, according to which children are seen as trustworthy, dependable and a safe haven in an otherwise hostile and rejecting world. The implicit theory conceptualisation of OSC proposes that maladaptive implicit theories develop in childhood and are modified during puberty and adolescence in response to sexual experiences, but the typical age at which such offence-related beliefs begin to develop has not been empirically tested. Gannon and Polaschek (2006) suggested that a longitudinal study design would be useful to test adolescent sexual offenders' cognition at various time points. Whilst this study is clearly not longitudinal, it offers some tentative support for the fact that some young

sexual offenders who are at the stage of transition between adolescence and adulthood have child abuse supportive beliefs that might form part of a ‘dangerous world’ implicit theory.

Examination of Component 2 (Child Sexuality) revealed that that there was a large overlap of items on this component and on Beckett’s (1987; 1995) ‘cognitive distortions’ subscale, with all but three items on the Beckett sub-scale being replicated in this component. The items contained in Component 2 reflect all the beliefs found in the Beckett subscale (children are sexually mature and motivated to have sex with older people, children are sexually provocative, and having sex with older people does not harm children). Ward and Keenan (1999) identified these beliefs as underpinning two important implicit theories: children as sexual beings and the nature of harm. This supports the proposal that young adults who abuse children have child abuse supportive beliefs connected to maladaptive implicit theories. Ward and Keenan (1999) also proposed that implicit theories may co-exist and be used in conjunction with one another, with ‘nature of harm’ identified as an implicit theory often associated with others. The fact that a range of child abuse supportive beliefs related to children as sexual beings and children not being harmed by sex have clustered together in this component, suggests that, perhaps for young adults, ‘children as sexual beings’ and the ‘nature of harm’ implicit theory do co-exist or that these two separate beliefs sets are generated by one larger implicit theory. This might be a useful area of further research.

Examination of the items in Component 3 (Child Orientation) revealed that with the exception of Q74 (‘when I feel low children cheer me up’), there was no overlap in items between this component and either of the Beckett (1987;1995) subscales. Q74 appears to be an anomalous item in this component, as all other items relate to beliefs

that the respondent has good skills with children or is child orientated (for example, ‘children like to play with me’, ‘children find me easy to make friends with’, ‘children feel safe with me’). These items appear to reflect a new, previously unidentified belief set, although it is similar to the risk factor, ‘feeling more comfortable with children than adults’, on the adult Structured Assessment of Risk and Need (SARN; Ministry of Justice, 2009) used in the HM Prison Service in the UK (also see, Tully, Browne, & Craig, 2014). Evidence for this risk factor includes having a child-orientated lifestyle, viewing children as more satisfying companions than adults and the individual seeing himself as a child. It is debatable whether having a belief that one has good skills with children might be considered offence supportive, but it is possible that if young adults believe they have such skills, they could use them to attract children and make them feel safe, although this is purely speculative at this stage. Equally, having good skills with children might be the reality for some young adult sexual offenders. Some may have a similar emotional maturity level to children. Therefore, self-proclaimed skills with children should not be assumed to be synonymous with pathology. Nonetheless, the fact that this belief has been identified for this age group (18 to 21 years) is of concern, as having good skills with children and wanting to engage in childlike activities might be viewed very differently for someone in this age group compared to an adult male.

Examination of items in Component 4 (Child Maturity) identified that these related to beliefs about children being powerful, mature and manipulative, but not in a sexual way. One item from this component (‘children are not as innocent as most people think’) appeared on Beckett’s (1987; 1995) ‘cognitive distortions’ subscale. Believing that children are powerful and mature has not previously been identified as a child abuse supportive belief for either young people or adult sexual offenders, although it

shares similarities with a group of items Gannon et al. (2009) could not relate to any existing implicit theory and that they referred to as 'children are unknowable'. In terms of how a belief that children are mature and powerful could support offending, it is likely that young adults who hold this belief would view children as responsible and mature enough to make their own decisions and choices, which could include whether to engage in sexual behaviour. These young people would therefore feel no responsibility towards protecting children and guilt would not act as a barrier to prevent them engaging in sex with children. This belief mirrors one of the generalised beliefs that underpins the 'children as sexual beings' implicit theory, whereby 'human beings are seen as capable of identifying their own needs and making their own decisions' and 'children are viewed as possessing the cognitive capabilities to identify their needs and preferences' (Ward & Keenan, 1999, p. 828). This finding offers further support for the proposal that young adult sexual offenders who abuse children may have child abuse supportive beliefs that are part of maladaptive implicit theories about children as sexual beings.

In summary, the four components arising from the PCA appear to reflect beliefs that fit with current psychological theory and conceptualisations of child abuse supportive beliefs generally, and Ward and Keenan's (1999) implicit theory hypothesis specifically. Kline (2000) argues that establishing the construct validity of a measure is a subjective and inferential task, and thus can never conclusively be proven (the best that can be expected is that strong support is found). These findings provide preliminary support for the construct validity of CASQ-AV in that it is able to measure child abuse supportive beliefs in young adults who have committed offences against children. These findings also offer support for the proposal that such beliefs may reflect the same

maladaptive implicit theories thought to exist with adults who sexually abuse children. Despite these positive results, this new factor structure would need to be replicated in a similar sample and a younger sample (under 18 years) before the four-factor CASQ-AV could be used for clinical purposes.

Many psychometricians (Guilford, 1956; Nunnally, 1978) argue that high internal consistency is a pre-requisite for validity. Cattell and Kline (1977), however, argue that most psychological tests measure variables have some breadth and that when alpha coefficients are high this indicates the items are highly correlated, and that the test might be too narrow and specific to have any meaning or utility. In practical terms, this might mean that test items simply ask the same question in slightly different ways. Furthermore, having a high number of items on a scale may also increase the alpha coefficient (Kline, 2000). Internal consistency of the CASQ-AV does not appear to have been artificially bolstered by repeat questioning (see Appendix 12) and the measure consists of a range of different questions that access different child abuse supportive beliefs. Additionally, one objective of setting the factor loading cut-off scores at a high level (0.45) was to reduce the number of items contained in each component, although it is accepted that this approach could reduce the breadth of the component. Overall, internal consistency of the CASQ-AV measure and individual components is supported.

The CASQ-AV measure and individual components also demonstrated good temporal stability. The correlation of scores over a month test-retest interval were all above 0.9 (excellent), except for Component 3, Child Orientation, which was still in the very high range (0.82). Kline (2000) suggests that reliability coefficients should be derived from a sample of at least 100 participants to minimise statistical error and that they should be measured over at least a three-month interval to reduce learning effects.

The practical problems encountered in securing a test-retest sample were described in the method section, and to leave a three-month gap between testing would inevitably have led to losing more participants owing to transfer to different prison establishments. It might also have resulted in higher rates of attrition owing to participants commencing treatment. Although neither the sample size or test-retest interval was ideal, they were generally consistent with similar published tests in which validation was conducted in secure conditions. For example, for the validation of the SWCH (Mann et al., 2007) the sample consisted of 30 sexual offenders and the test-retest interval was an average of 72 days. For validation of the Social Problem-Solving Inventory-Revised (Wakeling, 2007), the sample consisted of 30 offenders and the test-retest interval was an average of 28.7 days. Given concerns raised about both sample size and test-retest interval in this study, the findings presented here should be viewed as a preliminary indication of the temporal stability of the CASQ-AV, rather than as a definitive assessment.

There was no relationship between the PRI (a measure of socially desirable responding) and the CASQ-AV measure or any of its components. This suggests that the CASQ-AV is not susceptible to socially desirable responding and supports previous findings for this measure with young people (Edwards et al., 2012; van Outsem et al., 2006). As discussed, Gannon and Polaschek (2006) have questioned the validity of inferring child abuse supportive abuse measures as free from response style bias when no correlation is found between such measures and measures of socially desirable responding. They hypothesise that sexual offenders might depress their responses to items that are obviously about sexual contact with children and yet not provide evidence of elevated self-presentation bias. There was no control group in this study and so it was not possible to compare scores on the CASQ-AV from young adults who have

committed sexual offences against children, with either other types of sexual offenders, other types of non-sexual offenders or non-offenders. Using this method to examine socially desirable responding assumes that young adults who have committed sexual offences against children have child abuse supportive beliefs and that the other groups do not, an assumption not supported by the systematic review presented in Chapter 2.

As discussed, the scores of adult sexual offenders against children on child abuse supportive belief measures are often low. They do not typically ‘agree’ or ‘strongly agree’ with child abuse supportive statements. In fact, they tend to simply disagree less or provide neutral statements in comparison to other groups (Arkowitz & Vess, 2003). Similarly, raw scores, mean component and total CASQ-AV scores from this younger sample were low for all analyses. Respondents indicated that the child abuse supportive statements on the CASQ-AV were *very true* or *somewhat true* less frequently than *don’t know*, *somewhat untrue* or *very untrue*. However, it might be argued that any response apart from strong disagreement is tacit agreement on this sort of measure. Kolton, Boer and Boer (2001) used dichotomous scoring on an adaption of the ABCS, in which *strongly disagreeing* with items was given a score of ‘0’ and all other responses (*strongly agree*, *agree*, *neutral*, *disagree*) were coded ‘1’, although the ethics of re-categorising disagreement as agreement is questionable, particularly on measures of beliefs that support child abuse.

As part of the ongoing debate about the extent to which sexual offenders engage in socially desirable responding, Mathie and Wakeling (2011) examined the relationship between the Balanced Inventory of Desirable Responding-6 (BIDR; Paulhus, 1988) and a wide range of psychometric measures using a sample of 1,730 adult sexual offenders, concluding that the extent of socially desirable responding, with adults at least, is

actually less than often assumed. Self-report questionnaires continue to be a popular way of assessing individual treatment needs, progress in treatment, and the overall efficacy of treatment programmes (Beech et al, 2013; Beggs & Grace 2011; Wakeling & Barnett, 2014). More research is required to examine response bias in questionnaires in which the construct being measured is particularly distasteful. Piloting different ways of scoring the CASQ-AV may be beneficial. For example, separating respondents who agree with child abuse supportive statements to some extent (*somewhat true* or *very true*) from those who disagree or remain neutral (*somewhat untrue*, *very untrue*, *don't know*) may be useful for treatment planning purposes. However, for now, these findings suggest that when using an accepted methodology (correlation of the index measure with an established measure of socially desirable responding), the CASQ-AV does not appear affected by self-presentation bias.

There were significant correlations between the CDI scale of the MSI-J (a measure of cognitive distortions) and the CASQ-AV measure and component scores, although none of the correlations were particularly strong. Kline (2000) has stated that, to ensure that assessments of concurrent validity are meaningful, an instrument that measures the same construct as the one under development – with already-established validity and reliability (gold standard or ‘benchmark test’) – must be used in the correlation. This is often very difficult to achieve. Furthermore, there is no agreed standard for how high a correlation should be in order to indicate that a measure has concurrent validity. Kline (2000) suggests that, where a benchmark tests exists, a correlation of 0.75 or above indicates concurrent validity; where no such test exists (which is often the case), a correlation of 0.4 or 0.5 will suffice. When the lower correlation was used, results from this study indicated that only Component 2 (Child

Sexuality) and the total CASQ-AV score correlated highly with the CDI scale of the MSI although the Positive Affect component is approaching that value (0.39). Close examination of the items on the CDI scale of the MSI-J offers a partial explanation for these results. Many of the items on this scale are about children being sexually knowing and mature and not being harmed by sex, which are the same beliefs that are represented in Component 2 (Child Sexuality). Other items relate to loving children and wanting to be with children, themes captured in Component 1 (Positive Affect)¹³. There are no items on the CDI scale that relate to being skilled with children or to children being mature, which might explain the lower correlations with Components 3 and 4. The absence of valid and reliable measures of child abuse supportive beliefs for young people was a key finding of the systematic review presented in Chapter 2 and so finding a ‘benchmark test’ to use to establish the concurrent validity of the CASQ-AV illustrated the problems highlighted by Kline (2000), and utilising the CDI scale of the MSI-J was the best option available. It might be argued that if a measure of the same child abuse supportive beliefs as the CASQ-AV existed and it had already-established validity and reliability with young sexual offenders, why would it be necessary to develop and test the CASQ-AV? New measures are usually developed because they have different characteristics to existing tests. One of the strengths of the CASQ-AV lies in the range of beliefs that it measures. Overall, these results provide evidence that the measure has concurrent validity, as far as can be tested at the current time.

It was hypothesised that the CASQ-AV would be able to discriminate between high risk and lower risk individuals based on the number and strength of child abuse supportive beliefs held. Results indicate that scores on the CASQ-AV increase

¹³ The exact items are not reproduced here to comply with copyright legislation

incrementally as level of risk increased for all components (apart from Component 3, Child Orientation). Observed differences were only statistically significant for Component 1 (Positive Affect). Nonetheless, this finding supports the hypothesis to some extent and provides some evidence for discriminant validity. The relationship between risk levels (as measured by actuarial tools) and the endorsement of child abuse supportive beliefs has not previously been examined for young people. This is probably because no fully validated actuarial measure is available that may be used with young people under the age of 18¹⁴ (NCJA, 2014). However, these results support previous findings in the adult literature, in which significant differences in SWCH scores were found between risk groups, as measured by RM2000 (Mann et al., 2007). The reason that the incremental pattern did not hold for Component 3 (Child Orientation) is unclear, but examination of the mean scores for this component reveal that they are similar across risk levels. This may lend weight to the suggestion put forward earlier, that believing you have good skills with children and being child orientated might not be offence-supportive or pathological.

Significant differences were found between pre- and post-treatment scores on the CASQ-AV measure and all individual components. Moderate effect sizes were observed. Taken together, these results support the hypothesis that the CASQ-AV is sensitive to treatment change. This study did not have an untreated control group and so nothing can be inferred from these results about the effectiveness of the HM Prison Service treatment programme in modifying child abuse supportive beliefs. Establishing a meaningful control group for offending behaviour interventions is notoriously difficult (Friendship, Beech, & Browne, 2002). Innovative, ‘clinically significant change

¹⁴ Juvenile Sexual Offense Recidivism Risk Assessment Tool-II (JSORRAT-II; Epperson, Ralston, Fowers, Dewitt, & Gore, 2006).) is a validated actuarial risk assessment tool for juveniles but it is only approved for use in Utah and Iowa (National Criminal Justice Association, 2014)

methodology' has therefore been developed (Nunes et al., 2011) to assess whether individuals are in the 'functional range' post-treatment on dynamic risk variables, as measured by psychometric tests, and whether the change is statistically reliable and clinically observable. Individual changes can then be linked to recidivism data, allowing judgments to be made about the effectiveness of interventions overall (Barnett et al., 2013; Wakeling & Barnett, 2014). Assessing treatment change in this way relies on the availability of valid and reliable measures (Beggs, 2010). A standardised, normative sample was beyond the scope of this study, but given that the validity and reliability of the CASQ-AV is now largely supported, however, such research might follow logically.

The strengths of this study lie in its large sample size and the range of reliability and validity assessments undertaken. It is also unique in assessing whether child abuse supportive beliefs may be identified in a group of young people in the transition between adolescence and adulthood and in assessing whether these could be related to Ward's (1999) maladaptive implicit theories. The main limitation of this study is the absence of comparison/control groups. This meant that the ability of the CASQ-AV to discriminate young men who sexually abuse children from other groups could not be tested. It was encouraging child abuse supportive beliefs were reduced following treatment, but the lack of a control group meant that the reason for these changes could not be investigated. These changes could have been as a direct result of the intervention; alternatively, they could have been the result of passing time. Evidence presented in Chapters 2 and 3 suggests that child abuse supportive beliefs may be more relevant for some sub-groups of child sexual offenders; for example, extra-familial offenders. Psychometric data has not been collected in a systematic way over time for young adult sexual offenders engaging in sexual offender treatment in the HM Prison Service. Thus,

basic details about the nature of offence, such as the gender of the victim and the relationship to the victim, are unavailable for the majority of the sample. Sub-group analysis therefore could not be undertaken. The small test-retest sample was also a limitation. However, given the low numbers of untreated young adult sexual offenders, it would be difficult to recruit a sample of 100, as recommended by Kline (2000).

Future research should focus on replicating and confirming the component structure identified in this study both with a similarly aged sample and with a younger sample. It would also be helpful to administer the CASQ-AV to a sample of non-offenders (aged 18 to 21 years) to test the discriminant validity of the measure further, and to provide a potential normalisation group to assess the clinical significance of the statistical changes observed on the CASQ-AV following treatment. Increasing the test-retest sample size would enhance confidence in the reliability of the CASQ-AV.

In conclusion, Chapter 2 highlighted the need either for new, developmentally-sensitive measures of OSA&Bs for young sexual offenders or for psychometric testing of existing measures. Chapter 3 highlighted the urgent need to establish the psychometric properties of the CASQ-AV, an existing measure of child abuse supportive beliefs that is in widespread use. Whilst there may still be value in developing new measures based on direct clinical experience with young people, this study has demonstrated that it is possible to measure child abuse supportive beliefs in younger populations with existing measures. It could be argued that the young men aged 18 and 21 years who participated in this validation study were more similar to mature men than to adolescents. It would be prudent to replicate this study with a population aged under 18 years. The results nevertheless provide support for the reliability and validity of the four-component CASQ-AV as a measure of child abuse

supportive beliefs for young adult sexual offenders who have offended against children. These results are extremely encouraging. The four-component CASQ-AV could now be used for research, although it would require replication before it could be used to assess pre-treatment needs, treatment change and overall efficacy of sexual offender treatment programmes. This study is the first to provide evidence that younger sexual offenders have the types of child abuse supportive beliefs that Ward and Keenan (1999) identified as stemming from a range of maladaptive implicit theories in adults. As such, these findings offer tentative support for Ward's (2000) proposal that maladaptive implicit theories develop and are present before adulthood. It is hoped that this will stimulate further research into the development of child abuse supportive beliefs and implicit theories in younger sexual offender populations.

CHAPTER 5: GENERAL DISCUSSION

The aim of this thesis was to identify, assimilate and analyse data on OSC in younger sexual offenders in order to establishing the role it plays in their offending, its relevance as a criminogenic need for this group, and whether one particular form of OSC (attitudes and beliefs that support the sexual abuse of children) may be validly and reliably measured. The key findings of each chapter are now discussed in relation to these aims.

Key findings from Chapter 2: A Systematic Review

The systematic review revealed that OSC is under-researched in younger sexual offenders. A comprehensive search of five databases and 19 experts was conducted, yet only 13 relevant studies were identified. The review revealed that when adults describe their offending they make statements that justify, excuse and minimise the severity of their offending (Abel et al., 1984; 1989) clinical observation suggests that younger sexual offenders do the same. In the literature on adult sexual offenders, it has historically been assumed that these ‘cognitive distortions’ indicate the presence of underlying OSA&Bs and that these may be measured using self-report questionnaires (Keown et al., 2010). It appears as though this assumption has been extended to research with younger sexual offenders. The range of cognitive phenomena investigated and the methods used to assess these are limited. OSA&Bs were assessed in all studies and this construct was almost universally assessed using self-report psychometric measures.

If OSA&Bs play a causal role in sexual offending for young people, then it should be possible to discriminate young sexual offenders from community-based non-

offending controls and other types of offenders. A key finding of the review was that these groups cannot be discriminated from each other using self-report measures of OSA&Bs. A counterintuitive finding was that community non-sex offender comparison groups sometimes endorsed *more* child abuse supportive beliefs and rape myths than sexual offenders. Furthermore, child sexual abusers could not be discriminated from peer/adult abusers on offence-specific measures. There is a reluctance to consider that OSA&Bs may not play a causal role in or be relevant to adult sexual offending (Benbouriche et al., 2015). The same hesitancy was noted in studies reviewed here. Various explanations for the lack of discriminant validity of measures are proposed. A common explanation is that young sexual offenders engaged in socially desirable responding because they were assessed in a context in which the authorities knew them to be sexual offenders, whereas community controls responded anonymously (Becket 2006; van Outsem et al., 2006; Zgourides et al., 1997). However, in studies in which sex offenders were compared to other types of offenders and child abusers were compared to peer/adult abusers in the *same assessment setting* (that is, the demand characteristics were removed), the groups could still not be discriminated on OSC measures.

Another key finding was the lack of psychometrically-sound OSA&Bs measures for young sexual offenders. This could also account for the lack of discrimination between sexual offenders and other groups. A number of studies used the CASQ-AV (Beckett, 1995) to assess child abuse supportive beliefs, and the critique of the CASQ-AV (Chapter 3), showed that this measure has few reported psychometric properties. However, the results of the validation of this measure presented in Chapter 4 suggest that the CASQ-AV is psychometrically robust and *not* susceptible to social desirability,

although these findings were based on a different factor structure. Whilst the use of psychometric measures that have not been properly validated cannot be supported, it is possible that the lack of psychometric evidence has been overemphasised as a reason for the lack of discrimination between groups on OSA&B measures

Lack of discrimination between sexual offenders and other groups might mean that sexual offenders, other types of offenders, and non-offending young men in the community all have OSA&Bs, or that none of these groups do. Alternatively, OSA&Bs might be clinically significant in only a small minority of young people in offender or control groups making discrimination impossible. Research suggests that the responses of adult sexual offenders on OAS&B measures indicate that they are not actually agreeing with or endorsing offence-supportive statements; their scores are typically low, suggesting that they may hold OSA&Bs to a lesser extent than often assumed (Gannon & Polaschek, 2006). The results of the review studies lacked detail, making it impossible to compare the mean scores of respondents to the range of possible scores, to establish to what extent respondents were endorsing offence-supportive statements. Nonetheless, taken together, these results suggest that young sexual offenders do not typically hold OSA&Bs that differ either from other offenders or non-offenders and these findings present a challenge to the psychological theories that identify OSA&Bs as playing a causal role in sexual offending.

Some studies of OSC with adult sexual offenders suggest that OSA&Bs may play a role in offending only in certain types of child offender, namely those offend outside the family (Fisher et al., 1999; Seto et al., 2015). This finding was supported by one study in the current review, in which a significantly higher percentage of extra-familial offenders were found to have high scores on cognitive distortions compared with intra-

familial offenders (Beckett, 2006). However this trend requires further research before being confirmed as a significant finding for younger sexual offenders.

Few studies addressed the relevance of OSC as a treatment need for young sexual offenders. Only one assessed change in test scores pre- and post- intervention (Edwards et al., 2012). This study found that the majority of young sexual offenders were already in the 'functional range' as regards OSA&Bs pre-treatment (a result also observed with adults; Barnett et al., 2013). This calls into question the relevance of this construct as a treatment need. All the young people identified as having OSA&Bs as a treatment need were assessed as being in the functional range post- treatment. No attempt has been made to link these changes to re-offending, or to use psychometric test scores to predict recidivism for young sexual offenders. This area requires further research.

To some extent, these findings are not incompatible with the implicit theory hypothesis, as Ward and Keenan indicated that maladaptive implicit theories thought to contain offence-related beliefs are not universally held by sexual offenders: 'we are not suggesting that all sexual offenders hold these implicit theories' (Ward & Keenan, 1999, p. 822). It follows that, if only a minority of young sexual offenders hold OSA&Bs, this would not be identified when examining group level data. Ward and Keenan (1999) also suggested that offenders have OSA&Bs at three levels: general level beliefs about people and the world; middle level beliefs about categories of people, for example, women and children; and, at the most specific level, attitudes and beliefs about their own particular victim(s). It could be that when young sexual offenders describe their offences, the 'distorted' statements they make about their specific victims (such as, 'I didn't hurt her') for most of them – possibly because of their youth – may not generalise to all victims (for example, 'children are not hurt by sexual contact with older people').

The self-report questionnaires reviewed here focused almost exclusively on generalised child abuse supportive beliefs and rape myths. Evidence suggests that these cognitive structures do not play a causal role in sexual offending and are an irrelevant treatment need for many younger sexual offenders. Thus, a more holistic approach to OSC with this group might be useful. Although modifying post-hoc justifications for offending is unlikely to have much criminogenic value (Maruna & Mann, 2006), an examination of victim-specific beliefs present when the young person is contemplating offending (for example, 'she won't say anything because she is a slag') or those that are used to maintain offending ('he enjoyed it, he wants to have sex with me') has not been undertaken. Such research may provide new insights into the role and function of OSC for this group.

No single factor explains sexual offending, in either adults or young people (Ward et al., 2006). Sexual offending involves a complex interplay between biological, physical, emotional, cognitive, social and cultural factors (Ward 2009; Ward & Beech, 2006). Thus, attempting to isolate and examine OSC without consideration of other risk factors may be futile, perhaps particularly so for younger sexual offenders who, by nature of their age, are influenced more by the broader social ecology and systems in which they exist (Dennison & Leclerc, 2011).

Another important finding of this chapter is that that none of the measures of OSA&Bs in use with younger sexual offenders were developmentally sensitive: they were not constructed with or tested on young sexual offenders. Rather, adult measures were consistently used. Whilst there are likely similarities in OSC with adult and younger populations, it cannot be assumed that because a measure of OSA&Bs is proven valid and reliable in adult sexual offenders, the same is true for younger

populations. The lack of developmentally-sensitive measures of OSA&Bs formed the rationale for Chapter 3. Chapter 3 examined psychometric properties of the CASQ-AV, a measure of child abuse supportive beliefs designed for and tested on adult men who abuse children, but in wide use with younger populations in both the UK and Europe.

Key findings from Chapter 3: A Psychometric Critique

This chapter highlighted that the CASQ-AV forms part of a battery of tests used in a wide range of settings to assess individual needs and treatment progress, but also to assess the efficacy of sexual offender interventions. In particular, the CASQ-AV was administered to all young adult sexual offenders who had been convicted of abusing children and were undergoing treatment in two young offender establishments in England, with scores being used to assess pre-treatment need and treatment change, which in turn contributed to wider risk assessment. In these circumstances, this influenced parole and release decisions.

The reliability and validity of the measure for use with adults was shown to be reasonable. The reliability of the measure was adequate when tested on young sexual offenders younger than 18-years old, but the validity of the measure was not well established for this group. Further, problems that continue to affect the establishment of the psychometric properties of measures of OSA&Bs with adults (Benbouriche et al; 2015; Nunes et al., 2014) were also found with this group. One of the most pressing issues is how to develop a measure with face validity that is insusceptible to socially desirable responding, and can discriminate between sexual offenders and non-sexual offenders or community control groups. The most concerning finding was that the validity and reliability of the CASQ-AV for use with 18 to 21-year olds had not been

tested at all. Given the circumstances in which the CASQ-AV is used, this position is not defensible. This finding informed the rationale for Chapter 4, the content of which represents an attempt to establish the reliability and validity of the CASQ-AV.

Key Findings from Chapter 4: Research

The most important finding of Chapter 4 was that the factor structure resulting from the principal components analysis of the CASQ-AV did not replicate the two-factor structure that has been in use with this measure for the past 20 years. In the current solution, there was no distinction between items that measured ‘cognitive distortions’ and ‘emotional congruence with children’. It might be argued that the factor structure produced in this study is more conceptually sound, as all the components produced reflected distinct beliefs that could – theoretically at least – support offending.

Furthermore, three of the four beliefs identified reflected beliefs associated with Ward’s (1999) implicit theories, and this finding lends some support to the proposal that some young adults may have offence-related implicit theories, or at least the beliefs thought to originate from these. Responses to the CASQ-AV did not appear to be influenced by social desirability, and scores on all but one component increased incrementally as risk of re-offending increased. Ward and Keenan (1999) proposed that maladaptive implicit theories strengthen over time. It might thus be argued that OAS&Bs are more relevant for higher risk offenders who perhaps have offended for a longer time, more frequently or have more victims. There were significant differences between pre-and post-test scores following a cognitive-behavioural intervention and effect sizes were reasonable, although with adults it has been demonstrated that positive treatment change on the

CASQ is not associated with reduced re-offence (Barnett, et al., 2013; Wakeling et al., 2013).

Overall, the validity and reliability of the CASQ-AV was established for use with younger men who abuse children, but mean scores were low for all components, consistent with the adult OAS&B literature (Gannon & Polaschek, 2006). The implication is that this sample did not endorse child abuse supportive beliefs on the measure. Therefore, no conclusions can be drawn from this data about whether child abuse supportive beliefs as measured by the CASQ-AV are a relevant treatment need for this group. Furthermore, as a non-offending comparison group was not used in this study, nothing may be said about the ability of the CASQ-AV to discriminate between sexual offenders and non-offenders and therefore whether OSA&Bs are likely to play a causal role in sexual offending for this group.

Implications for Practice and Research

The central message for practitioners assessing and treating younger sexual offenders appears to be not to assume that ‘distorted’ offence accounts mean that all, or even any, of this population have entrenched OAS&Bs. Rather, it appears that the aim of assessment should be to identify which, if any, young people have relevant treatment needs in this area. Achieving this aim will not be possible whilst practitioners continue to use measures of OSC that have been developed for and tested on adults, and this practice should not be supported as it is unethical and indefensible. If the factor structure and corresponding sub-scales of the CASQ-AV as identified in this study are replicated on a similar and younger population, then this measure could be used with

confidence by practitioners as a valid and reliable measure of OSA&Bs for use with young people aged between the ages of ten and 21 years.

Although the CASQ-AV has demonstrated some psychometric properties with a younger sexual offender population, there is still a need for a developmentally-sensitive measure of OSC based on interviews with young sexual offenders. This might highlight important differences in the role and relevance of OSC for adults and younger sexual offenders and result in both assessment and treatment better suited to the developmental needs of younger populations. This thesis suggests that OSA&Bs are not significant treatment needs for young sexual offenders. By considering sexual offenders, in particular those who offend against children, as a homogenous group, the importance of this treatment need for some offenders may remain unidentified. It would be helpful for future research to take a lead from adult sexual offender research (Seto et al. 2015) and examine sub-groups of offenders; for example those who have offended within and outside of the family.

The extent to which young people respond in a socially desirable way when completing measures of OSA&Bs remains unresolved. The creative use of innovative methodology, such as the ‘bogus pipe line’ used with adult child abusers to examine the extent to which they depressed their scores on a measure of OSA&Bs (Gannon et al., 2007) would be welcomed, although the use of deception with younger populations would present some difficult ethical issues. In the meantime, other steps could be taken to minimise the extent to which respondents feel compelled to ‘fake good’, such as conducting research with young people in a research setting in which they are assured anonymity, rather than a treatment setting in which the consequences of appearing ‘deviant’ might be high (Wakeling & Barnett 2014). In addition, when reporting scores

on measures of OSA&Bs, it would be helpful if future research reported the range of scores that can be achieved so that the extent to which participants are agreeing with offence-supportive statements may be deduced.

Strengths and Limitations of the Thesis

This study addressed a previously neglected area of research and the results have challenged assumptions about the role and relevance of OSC for younger sexual offender populations. A further strength lies in its broad approach to the review of the literature on OSC with young people. Through the review, the lack of psychometrically and developmentally sound measures of OSC for use with young people was evident. This has resulted in both researchers and practitioners having to rely on adult measures, an ongoing and potentially unethical practice. This finding contributed directly to the critique of the CASQ-AV and subsequent psychometric validation of this measure on a large and relevant sample of young adult sexual offenders. Although the findings presented in Chapter 3 demonstrate a range of sound psychometric properties in the CASQ-AV, the main weakness of this thesis also lies here. The test-retest sample was small, there are issues with the scoring of the CASQ-AV that require resolution, and there is no benchmark test to measure the concurrent validity of the measure against. Furthermore the lack of non-offending and non-treated control groups meant that the discriminant properties of the CASQ-AV and the significance of the observed treatment change could not be further tested. Difficulty obtaining the approval of ethics committees to allow young people to complete questionnaires about the sexual abuse of children are well documented (Edwards et al., 2012; Whittaker, Brown, Beckett, & Gerhold, 2006), as are problems securing a non-treated control groups (Friendship et al.,

2002). This thesis was unable to answer which, if any, young sexual offenders have offence supportive attitudes and beliefs at a level that would render it a treatment need for them. However, this is perhaps a clinical rather than a research question. The current study has contributed towards providing a measure that may allow practitioners to answer this question in the future.

Summary and Conclusions

OSC and particularly OSA&Bs are identified as aetiological factors in many theories of sexual offending. Consequently, modifying cognitive phenomena is clinically important for both adult and younger sexual offenders. There is little research on OSC in young sexual offenders. The research that exists does not support OSA&Bs as having a causal role in sexual offending or as being a relevant treatment need for this group. However, this does not rule out the possibility that OSA&Bs play a causal role in the sexual offences of *some* young people and these cognitive structures could be a relevant treatment need for this group. The lack of valid and reliable measures of OSC for younger sexual offender populations has been identified as a major problem and one that may have contributed to the negative conclusions identified above. This thesis succeeded in its aim to validate an existing measure of child abuse supportive beliefs for use with younger sexual offenders who abuse children. It is hoped that this measure will be used in clinical practice and future research, and that this may help clarify which, if any young sexual offenders have beliefs that support child abuse and in whom these beliefs need to be addressed in order to reduce their risk of reoffending.

REFERENCES

* studies included in the systematic review

- Abbey, A., Wegner, R., Pierce, J., & Jacques-Tiura, A. (2012). Patterns of sexual aggression in a community sample of young men: Risk factors associated with persistence, desistance, and initiation over a 1-year interval. *Psychology Of Violence*, 2(1), 1-15. doi:10.1037/a0026346
- Abel, G. G., Becker, J. V., & Cunningham-Rathner, J. (1984). Complications, consent, and cognitions in sex between children and adults. *International Journal of Law and Psychiatry*, 7(1), 89-103. doi:10.1016/0160-2527(84)90008-6
- Abel, G., Becker, J., & Skinner, L. (1987). Behavioural approaches to treatment of the violent sex offender. In L. Roth (Ed.), *Clinical treatment of the violent person* (pp. 95-118). New York: Guilford Press.
- Abel, G. G., Gore, D., Holland, C., Camps, N., Becker, J., & Rathner, J. (1989). The measurement of the cognitive distortions of child molesters. *Annals Of Sex Research*, 2(2), 135-152. doi:10.1007/bf00851319
- Andrews, D. A., & Bonta, J. (2010). *The psychology of criminal conduct* (5th ed.). Newark, NJ: LexisNexis/Anderson. doi:10.1111/j.2044-8295.2011.02057.x
- Arkowitz, S., & Vess, J. (2003). An evaluation of the Bumby RAPE and MOLEST scales as measures of cognitive distortions with civilly committed sexual offenders. *Sexual Abuse: A Journal Of Research And Treatment*, 15(4), 237-249. doi:10.1177/107906320301500402

- Arnett, J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, 55(5), 469-480. doi:10.1037//0003-066x.55.5.469
- Barnett, G. D., Wakeling, H. C., & Howard, P. (2010). An examination of the predictive validity of the risk matrix 2000 in England and Wales. *Sexual Abuse: A Journal Of Research And Treatment*, 22(4), 443-470. doi:10.1177/1079063210384274
- Barnett, G. D., Wakeling, H. C., Mandeville-Norden, R., & Rakestrow, J. (2012). How useful are psychometric scores in predicting recidivism for treated sex offenders? *International Journal of Offender Therapy and Comparative Criminology*, 56(3), 420-446. doi:10.1177/0306624X11403125
- Barnett, G. D., Wakeling, H., Mandeville-Norden, R., & Rakestrow, J. (2013). Does change in psychometric test scores tell us anything about risk of reconviction in sexual offenders? *Psychology, Crime & Law*, 19(1), 85-110. doi: 10.1080/1068316x.2011.607820
- Beckett, R.C. (1987). *Children and Sex Questionnaire*. Unpublished.
- Beckett, R.C. (1994). Cognitive-behavioural treatment of sex offenders. In T. Morrisson, M. Erooga, & R.C. Beckett (Eds.), *Sexual offending against children* (pp. 80-101). London: Routledge. doi:10.4324/9780203411773
- Beckett, R.C. (1995). *Children and Sex Questionnaire – Adolescent Version*. Unpublished.
- *Beckett, R.C. (2006). Risk prediction, decision-making and evaluation of adolescent sexual abusers. In M. Erooga, & H. Masson (Eds.) *Children and young people who sexually abuse others: current developments and practice responses* (pp. 215-235).

Abingdon, UK: Routledge. doi:10.1093/bjsw/bcm001

Beckett, R.C., Gerhold, C. K. E., & Brown, S. (2002). *The Adolescent Sex Offender*

Assessment Pack: Background and description of measures and scoring guide.

Available c/o SWAAY, PO Box 292, Earley, Reading, Berks, RG6 7XQ.

Beech, A. R., Bartels, R. M., & Dixon, L. (2013). Assessment and treatment of distorted schemas in sexual offenders. *Trauma, Violence, & Abuse, 14*(1), 54-66.

doi:10.1177/1524838012463970

Beech, A. R., Fisher, D. & Beckett, R. C. (1999). *Step 3: An evaluation of the Prison Sex Offender Treatment Programme.* England: Home Office Publications Unit.

Retrieved January 15, 2014 from

<http://www.nationalarchives.gov.uk/ero/records/HO/421/2/P2/RDS/PDFS/OCC-STEP3.PDF>

Beech, A. R., & Hamilton-Giachritsis, C. E. (2005). Relationship between therapeutic climate and treatment outcome in group-based sexual offender treatment programs.

Sexual Abuse: A Journal of Research and Treatment, 17(2), 127-140.

doi:10.1177/107906320501700204

Beggs, S. (2010). Within-treatment outcome among sexual offenders: A review.

Aggression and Violent Behavior, 15(5), 369-379. doi:10.1016/j.avb.2010.06.005

Beggs, S. M., & Grace, R. C. (2011). Treatment gain for sexual offenders against

children predicts reduced recidivism: a comparative validity study. *Journal of consulting and clinical psychology, 79*(2), 182. doi: 10.1037/a0022900

Benbouriche, M., Longpré, N., Guay, J. P., & Proulx, J. (2015). Cognitive distortions in sexual aggressors against children: An examination of the Abel and Becker

- Cognition Scale (ABCS). *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology*, 65(1), 53-60. doi:10.1016/j.erap.2014.10.004
- Bumby, K. M. (1996). Assessing the cognitive distortions of child molesters and rapists: Development and validation of the MOLEST and RAPE scales. *Sexual Abuse: A Journal of Research and Treatment*, 8(1), 37-54.
- Burt, M. R. (1980). Cultural myths and supports for rape. *Journal of Personality and Social Psychology*, 38(2), 217. doi:10.1037/0022-3514.38.2.217
- *Butz, C., & Spaccarelli, S. (1999). Use of physical force as an offense characteristic in subtyping juvenile sexual offenders. *Sexual Abuse: Journal of Research & Treatment*, 11(3), 217-232. doi:10.1177/107906329901100305
- Caldwell, M. F. (2010). Study characteristics and recidivism base rates in juvenile sex offender recidivism. *International Journal of Offender Therapy and Comparative Criminology*, 54, 197-212. doi:10.1177/0306624x08330016
- Calhoun, K. S., Bernat, J. A., Clum, G. A., & Frame, C. L. (1997). Sexual coercion and attraction to sexual aggression in a community sample of young men. *Journal of Interpersonal Violence*, 12, 392-406. doi:10.1177/088626097012003005
- Calleja, N. G. (2013). Integrating research into practice: The forward-focused model of adolescent sexual behavior treatment. *Aggression and Violent Behavior*, 18(6), 686-694. doi:10.1016/j.avb.2013.07.02in pres
- Cann, S., Konopasky, R. J., & McGrath, M. L. (1995, June). *Chi-Mo: A new measure of cognitive distortions in sexual offenders*. Poster session presented at the annual convention of the Canadian Psychological Association, Prince Edward Island, Canada.

- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1(2), 245-276. doi:10.1207/s15327906mbr0102_10
- Cattell, R.B. (1973). *Personality and Mood by Questionnaire*. San Fransico, Jossey-Bass.
- Cattell, R. B., & Kline, P. E. (1977). *The scientific analysis of personality and motivation*. London: Academic Press.
- Cohen, J. (1998). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Cooke, A., Smith, D., & Booth, A. (2012). Beyond PICO: The SPIDER tool for qualitative evidence synthesis. *Qualitative Health Research*, 22(10), 1435–1443. doi: 10.1177/1049732312452938
- Craig, L. A., Beech, A. R., & Cortoni, F. (2013). What works in assessing risk in sexual and violent offenders. In L. A. Craig, L. Dixon, T. A. Gannon (Eds.), *What Works in Offender Rehabilitation: An evidence-based approach in assessment and treatment* (pp. 94-114). Chichester: Wiley. doi:10.1002/9781118320655.ch5
- Critical Appraisal Skills Programme (CASP), (2013). *Critical Appraisal Skills Programme: Making sense of evidence*. Retrieved from <http://www.casp-uk.net>
- Dalziel, K., Round, A., Stein, K., Garside, R., Castelnovo, E., Payne, L. (2005). *Do the findings of case series studies vary significantly according to methodological characteristics?* Retrieved from <http://www.hta.ac.uk/execsumm/summ902.htm>
- Dean, C., Mann, R. E., Milner, R., & Maruna, S. (2009). Changing child sexual abusers' cognition. In T. A. Gannon, T. Ward, A. R. Beech & D. Fisher (Eds.), *Aggressive offenders' cognition* (pp. 117-134). Chichester: John Wiley & Sons.

Dennison, S., & Leclerc, B. (2011). Developmental factors in adolescent child sexual offenders: A comparison of nonrepeat and repeat sexual offenders. *Criminal Justice and Behavior*, 38(11), 1089–1102. doi:10.1177/0093854811417076

Drake, C. R., Ward, T., Nathan, P., & Lee, J. K. (2001). Challenging the cognitive distortions of child molesters: An implicit theory approach. *Journal of Sexual Aggression*, 7(2), 25-40. doi:10.1080/13552600108416165

Dundar, Y., & Fleeman, N. (2014). Developing my search strategy and applying inclusion criteria. In A. Boland, M.G. Cherry, & R. Dickson (Eds.), *Doing a Systematic Review: A Student's Guide* (pp. 35 – 59). London: Sage

*Eastman, B. J. (2004). Assessing the efficacy of treatment for adolescent sex offenders: A cross-over longitudinal study. *The Prison Journal*, 84(4), 472-485. doi:10.1177/0032885504269346

*Eastman, B. J. (2005). Variables associated with treatment failure among adolescent sex offenders. *Journal of Offender Rehabilitation*, 42(3), 23-40. doi:10.1300/J076v42n03_02

*Edwards, R., Whittaker, M. K., Beckett, R., Bishopp, D., & Bates, A. (2012). Adolescents who have sexually harmed: An evaluation of a specialist treatment programme. *Journal of Sexual Aggression*, 18(1), 91-111. doi:10.1080/13552600.2011.635317

Epperson, D.L., Ralston, C.A., Fowers, D., Dewitt, J., & Gore, K.S. (2006). Actuarial risk assessment with juveniles who sexually offend: Development of the Juvenile Sexual Offense Recidivism Risk Assessment Tool-II (JSORRAT-II). In D.S. Prescott (Ed.), *Risk Assessment of Youth Who Have Sexually Abused: Theory,*

Controversy, and Emerging Strategies (pp. 118–169). Oklahoma City, OK: Wood & Barnes.

*Epps, K. J., Haworth, R., & Swaffer, T. (1993). Attitudes toward women and rape among male adolescents convicted of sexual versus nonsexual crimes. *Journal of Psychology*, 127(5), 501-506. doi:10.1080/00223980.1993.9914887

Falshaw, L., Bates, A., Patel, V., Corbett, C. & Friendship, C. (2003). Assessing reconviction, reoffending and recidivism in a sample of UK sexual offenders. *Legal and Criminological Psychology*, 8: 207–215.
doi:10.1348/135532503322362979

Fanniff, A. M., & Becker, J. V. (2006). Specialized assessment and treatment of adolescent sex offenders. *Aggression and Violent Behavior*, 11(3), 265–282.
doi:10.1016/j.avb.2005.08.003

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. London: Sage Publications.

Finkelhor, D. (1984). *Child sexual abuse: New theory and research*. New York: Free Press.

Fisher, D., & Beech, A.R. (2007). The implicit theories of rapists and sexual murderers. In T.A. Gannon, T. Ward, A.R. Beech, & D. Fisher (Eds.), *Aggressive offenders' cognition: Research, theory and practice* (pp31-52). Chichester, UK: Wiley.
doi:10.1002/9780470746295.ch2

Fisher, D., Beech, A. R., & Browne, K. (1999). Comparison of sex offenders to non-offenders on selected psychological measures. *International Journal of Offender*

- Therapy and Comparative Criminology*, 43(4), 473–491.
doi:10.1177/0306624X99434006
- Fleeman, N., & Dunder, Y. (2014). Data extraction: where do I begin? In A. Boland, M.G. Cherry, & R. Dickson (Eds.), *Doing a systematic review: A student's guide* (pp. 85–98). London: Sage
- Friendship, C., Beech, A.R. & Browne, K.D. (2002). Reconviction as an outcome in research: A methodological note. *British Journal of Criminology*, 42, 442–444.
doi:10.1093/bjc/42.2.442
- Gannon, T. (2009). Current cognitive distortion theory and research: An internalist approach to cognition. *Journal of Sexual Aggression*, 15(3), 225–246.
doi:10.1080/13552600903263079
- Gannon, T. A., & Alleyne, E. K. A. (2013). Female sexual abusers' cognition: A systematic review. *Trauma, Violence, & Abuse*, 14(1), 67–79.
doi:10.1177/1524838012462245
- Gannon, T. A., Collie, R. M., Ward, T., & Thakker, J. (2008). Rape: Psychopathology, theory and treatment. *Clinical Psychology Review*, 28(6), 982–1008.
doi:10.1016/j.cpr.2008.02.005
- Gannon, T. A., Keown, K., & Polaschek, D. L. L. (2007). Increasing honest responding on cognitive distortions in child molesters: The bogus pipeline revisited. *Sexual Abuse: a Journal of Research and Treatment*, 19(1), 5–22.
doi:10.1177/107906320701900103
- Gannon, T. A., Keown, K., & Rose, M. R. (2009). An examination of current psychometric assessments of child molesters' offense-supportive beliefs using

- Ward's implicit theories. *International Journal of Offender Therapy and Comparative Criminology*, 53(3), 316-333. doi:10.1177/0306624x07312791
- Gannon, T. A., & O'Connor, A. (2011). The development of the interest in child molestation scale. *Sexual Abuse: A Journal of Research and Treatment*, 23(4), 474-493. doi:10.1177/1079063211412390
- Gannon, T. A., & Polaschek, D. L. (2006). Cognitive distortions in child molesters: A re-examination of key theories and research. *Clinical psychology review*, 26(8), 1000-1019. doi:10.1016/j.cpr.2005.11.010
- Gannon, T., & Ward, T. (2009). Cognitive distortions 'debate': Introductory note. *Journal Of Sexual Aggression*, 15(3), 223-224. doi:10.1080/13552600903263020
- Gannon, T., Wright, D., Beech, A., & Williams, S. (2006). Do child molesters hold distorted beliefs? What does their memory recall tell us? *Journal Of Sexual Aggression*, 12(1), 5-18. doi: 10.1080/13552600500451655
- Gardner, M., & Steinberg, L. (2005). Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: An experimental study. *Developmental Psychology*, 41(4), 625. doi:10.1037/0012-1649.41.4.625
- Gerhold, C. K., Browne, K. D., & Beckett, R. (2007). Predicting recidivism in adolescent sexual offenders. *Aggression and Violent Behavior*, 12(4), 427-438. doi:10.1016/j.avb.2006.10.004
- Goddard, H. L. (2006). *Assessing the implicit theories of child molesters: Development and validation of the sex offenders' implicit theories scale* (Unpublished undergraduate thesis). University of Birmingham, UK.

- Grady, M. D., Brodersen, M., & Abramson, J. M. (2011). The state of psychological measures for adult sexual offenders. *Aggression and Violent Behavior, 16*(3), 227-240. doi:10.1016/j.avb.2011.03.001
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*(6), 1464. doi:10.1037//0022-3514.74.6.1464
- Greenwald, H. J., & Satow, Y. (1970). A short social desirability scale. *Psychological Reports, 27*(1), 131-131. doi:10.2466/pr0.1970.27.1.131
- Guilford, J.P. (1956). *Psychometric Methods*. New York: McGraw-Hill
- Hair, J.F. Tatham, R.L. Anderson, R.E. & Black, W. (1998). *Multivariate data analysis*. (7th ed.). London: Prentice-Hall.
- Hanson, R. K., Bourgon, G., Helmus, L., & Hodgson, S. (2009). The principles of effective correctional treatment also apply to sexual offenders: A meta-analysis. *Criminal Justice and Behavior, 36*(9), 865–891. doi:10.1177/0093854809338545
- Hanson, R. K., Gizzarelli, R., & Scott, H. (1994). The attitudes of incest offenders sexual entitlement and acceptance of sex with children. *Criminal Justice and Behavior, 21*(2), 187-202. doi:10.1177/0093854894021002001
- Hanson, R. K., & Morton-Bourgon, K. (2004). Predictors of sexual recidivism: An updated meta-analysis. *Corrections User Report No. 2004-02*. Ottawa, Ontario, Canada: Public Safety Canada.
- Hargreaves, C., & Francis, B. (2014). The long-term recidivism risk of young sexual offenders in England and Wales: Enduring risk or redemption? *Journal of*

Criminal Justice, 42(2), 164–172. doi:10.1016/j.jcrimjus.2013.06.017

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

Hoge, R. (2015). Risk/Need/Responsivity in Juveniles. In K. Heilbrun, D. DeMatteo, & N. Goldstein (Eds.), *APA Handbook of Psychology and Juvenile Justice*. (pp. 179-196). Washington, DC: American Psychological Association. doi:10.1037/14643-009

Hutchenson, G., & Sofroniou, N. (1999). *The multivariate social scientist*. London: Sage Publications. doi:10.4135/9780857028075

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20(1), 141-151. doi:10.1177/001316446002000116

Kaiser, H. (1970). A second generation little jiffy. *Psychometrika*, 35(4), 401-415. doi:10.1007/bf02291817

Keeling, J. A., Rose, J. L., & Beech, A. R. (2006). An investigation into the effectiveness of a custody-based cognitive-behavioural treatment for special needs sexual offenders. *The Journal of Forensic Psychiatry & Psychology*, 17(3), 372-392. doi:10.1080/14789940600658293

*Kenny, D. T., Keogh, T., & Seidler, K. (2001). Predictors of recidivism in Australian juvenile sex offenders: Implications for treatment. *Sexual Abuse: A Journal of Research and Treatment*, 13(2), 131-148. doi:10.1177/107906320101300206

- Keown, K., Gannon, T., & Ward, T. (2008). What were they thinking? An exploration of child sexual offenders' beliefs using a lexical decision task. *Psychology, Crime & Law*, 14(4), 317-337. doi./10.1080/10683160701770112
- Keown, K., Gannon, T. A., & Ward, T. (2010). What's in a measure? A multi-method study of child sexual offenders' beliefs. *Psychology, Crime & Law*, 16(1-2), 125–143. doi:10.1080/10683160802622022
- Kline, P. (1986). *A handbook of test construction: Introduction to psychometric design*. London: Methuen.
- Kline, P. (2000). *The handbook of psychological testing*. London: Psychology Press. doi:10.4324/9781315812274
- Kolton, D. J., Boer, A., & Boer, D. P. (2001). A revision of the Abel and Becker Cognition Scale for intellectually disabled sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 13(3), 217-219. doi:10.1177/107906320101300306
- Kroner, D. G., & Weekes, J. R. (1996). Balanced inventory of desirable responding: Factor structure, reliability, and validity with an offender sample. *Personality and Individual Differences*, 21(3), 323-333. doi:10.1016/0191-8869(96)00079-7
- Langevin, R. (1991). A note on the problem of response set in measuring cognitive distortions. *Annals of Sex Research*, 4(3-4), 287-292. doi:10.1177/107906329100400404
- Langevin, R., Curnoe, S., Feroroff, P., Bennett, R., Langevin, M., Peever, C. (2004). Lifetime sex offender recidivism: A 25-year follow up study. *Canadian Journal of*

Criminology and Criminal Justice, 46(5), 531–552.

doi:<http://dx.doi.org/10.3138/cjccj.46.5.531>

Letourneau, E. J., & Miner, M. H. (2005). Juvenile sex offenders: A case against the legal and clinical status quo. *Sexual Abuse: A Journal of Research and Treatment*, 17(3), 293-312. doi:10.1177/107906320501700304

Lindsay, W. R., Whitefield, E., Carson, D., Broxholme, S., & Steptoe, L. (2004).

Questionnaire on Attitudes Consistent With Sexual Offending. QACSO:

Administration and scoring manual. Unpublished manuscript. (Available from

Bill Lindsay: bill.lindsay@tpct.scot.nhs.uk)

Lösel, F., Bottoms, A., & Farrington, D. (2012). *Young adult offenders: Lost in transition*. Milton Park, Abingdon, Oxon: Routledge.

doi:10.4324/9780203128510

Maniglio, R. (2009). The impact of child sexual abuse on health: A systematic review of reviews. *Clinical Psychology Review*, 29(7), 647-657.

doi:10.1016/j.cpr.2009.08.003

Mann, R. E., & Beech, A. (2003). Cognitive distortions, schemas, and implicit theories.

In T. Ward, D. R. Laws, & S. M. Hudson (Eds.), *Sexual deviance: Issues and controversies* (pp. 135-153). London, England: Sage Publications.

doi:10.4135/9781483328751.n8

Mann, R., & Hollin, C. (2010). Self-reported schemas in sexual offenders. *Journal of Forensic Psychiatry & Psychology*, 21(6), 834-851.

doi:10.1080/14789949.2010.511240.

- Mann, R. E., & Shingler, J. (2006). Schema-driven cognition in sexual offenders: Theory, assessment and treatment. In W.L Marshall, Y.M. Fernandez, L.E., Marshall, & G.A., Seran (Eds.). *Sexual offender treatment: Controversial issues*, (pp. 173-185). Chichester, UK: Wiley. doi:10.1002/9780470713457
- Mann, R. E., & Thornton, D. (1998). The evolution of a multi-site sex offender treatment programme. In W. L. Marshall, Y. M. Fernandez, S. H. Hudson, & T. Ward (Eds.), *Sourcebook of treatment programs for sexual offenders* (pp. 47–58). New York, NY: Plenum Press. doi:10.1007/978-1-4899-1916-8_4
- Mann, R. E., Hanson, R. K., & Thornton, D. (2010). Assessing risk for sexual recidivism: Some proposals on the nature of psychologically meaningful risk factors. *Sexual Abuse: A Journal of Research and Treatment*, 22(2), 191-217. doi:10.1177/1079063210366039
- Mann, R. E., Webster, S., Wakeling, H., & Marshall, W. (2007). The measurement and influence of child sexual abuse supportive beliefs. *Psychology, Crime & Law*, 13(5), 443-458. doi:10.1080/10683160601061141
- Marshall, W. (1994). Treatment effects on denial and minimization in incarcerated sex offenders. *Behaviour Research And Therapy*, 32(5), 559-564. doi:10.1016/0005-7967(94)90145-7
- Marshall, W., & Barbaree, H. (1990). An integrated theory of the etiology of sexual offending. In W. Marshall, D. Laws & H. Barbaree, *Handbook of sexual assault: Issues, theories, and treatment of the offender* (pp. 257-275). New York: Plenum Press.

- Marsall, W. L., & Marshall, L. E. (2000). The origins of sexual offending. *Trauma, Violence, & Abuse*, 1(3), 250-263. doi:10.1177/15248380000001003003
- Marshall, W. L., Marshall, L. E., & Kingston, D. A. (2011). Are the cognitive distortions of child molesters in need of treatment? *Journal of Sexual Aggression*, 17(2), 118-129. doi:10.1080/13552600.2011.580572
- Maruna, S., & Mann, R. E. (2006). A fundamental attribution error? Rethinking cognitive distortions. *Legal and Criminological Psychology*, 11(2), 155-177. doi:10.1348/135532506x114608
- Marziano, V., Ward, T., Beech, A. R., & Pattison, P. (2006). Identification of five fundamental implicit theories underlying cognitive distortions in child abusers: A preliminary study. *Psychology, Crime & Law*, 12(1), 97-105. doi:10.1080/10683160500056887
- Mathie, N. L., & Wakeling, H. C. (2011). Assessing socially desirable responding and its impact on self-report measures among sexual offenders. *Psychology, Crime & Law*, 17(3), 215-237. doi:10.1080/10683160903113681
- McGrath, M., Cann, S., & Konopasky, R. (1998). New measures of defensiveness, empathy, and cognitive distortions for sexual offenders against children. *Sexual Abuse: A Journal of Research and Treatment*, 10(1), 25-36. doi:10.1177/107906329801000104
- McGrath, R. J., Cumming, G. F., Burchard, B. L., Zeoli, S., & Ellerby, L. (2010). *Current practices and emerging trends in sexual abuser management: The Safer Society 2009 North American survey*. Brandon, VT: Safer Society Foundation.

- Mills, J. and Kroner, D.G. (2006). Impression Management and Self-Report Among Violent Offenders. *Journal Of Interpersonal Violence*, 21(2), 178-192.
doi:10.1177/0886260505282288.
- Milner, J. S., Murphy, W. D., Valle, L. A., & Tolliver, R. M. (1998). Assessment issues in child abuse evaluations. In J. R. Lutzker (Ed.), *Handbook of child abuse research and treatment*. New York, NY: Plenum Press.
- Ministry of Justice, Home Office, Office for National Statistics (2013). *An Overview of Sexual Offending in England and Wales*. Retrieved from
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/214970/sexual-offending-overview-jan-2013.pdf
- Ministry of Justice, Office for National Statistics (2014). *Criminal Justice Statistics Quarterly Update to March 2014* (Published August 2014). Retrieved from
<https://www.gov.uk/government/statistics/criminal-justice-statistics-quarterly-april-2013-to-march-2014>
- Ministry of Justice. (2009). *Structured Assessment of Risk and Need (SARN) – Sexual Offending. Stage 2: Treatment Need Analysis Scoring Guide*. London: Ministry of Justice.
- Moga, C., Guo, B., Schopflocher, D., & Harstall, C. (2012). *Development of a quality appraisal tool for case series studies using a modified Delphi technique*. Edmonton, AB: Institute of Health Economics. Retrieved from
<http://www.ihe.ca/documents/Case%20series%20studies%20using%20a%20modified%20Delphi%20technique.pdf>
- Murphy, W. D. (1990). Assessment and modification of cognitive distortions in sex

- offenders. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.), *Handbook of sexual assault* (pp. 331-342). New York, NY: Plenum Press. doi:10.1007/978-1-4899-0915-2_19
- National Criminal Justice Association. (2014). Sex Offender Management Assessment and Planning Initiative. Washington, DC: National Criminal Justice Association, Office of Justice Programs (SMART), U.S. Department of Justice. Retrieved from http://www.smart.gov/SOMAPI/sec2/ch4_risk.html
- Nichols, H. R., & Molinder, I. (1984). *Multiphasic Sex Inventory manual: A test to assess the psychosexual characteristics of the sexual offender*. Tacoma, WA: Nichols & Molinder.
- Nunes, K. L., Babchishin, K. M., & Cortoni, F. (2011). Measuring treatment change in sex offenders: Clinical and statistical significance. *Criminal Justice and Behavior*, 38(2), 157-173. doi:10.1177/0093854810391054
- Nunes, K., Pettersen, C., Hermann, C., Looman, J., & Spape, J. (2014). Does Change on the MOLEST and RAPE Scales Predict Sexual Recidivism? *Sexual Abuse: A Journal Of Research And Treatment*. Advance online publication. doi:10.1177/1079063214540725
- Nunnally, J.O. (1978). *Psychometric Theory*. New York, NY: McGraw-Hill.
- Ó Ciardha, C., & Gannon, T. A. (2011). The cognitive distortions of child molesters are in need of treatment. *Journal of Sexual Aggression*, 17(2), 130-141. doi:10.1080/13552600.2011.580573
- Ó Ciardha, C., & Ward, T. (2013). Theories of cognitive distortions in sexual offending: What the current research tells Us. *Trauma, Violence, & Abuse*, 14(1), 5-21.

doi:10.1177/1524838012467856

O'Connor, D., Green, S., & Higgins, J. P. T. (2008). Defining the review question and developing criteria for including studies. In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions*. Retrieved from <http://www.cochrane-handbook.org>

Oneal, B. J., Burns, G. L., Kahn, T. J., Rich, P., & Worling, J. R. (2008). Initial psychometric properties of a treatment planning and progress inventory for adolescents who sexually abuse. *Sexual Abuse: a Journal of Research and Treatment*, 20(2), 161–187. doi:10.1177/1079063208317465

Paras, M. L., Murad, M. H., Chen, L. P., Goranson, E. N., Sattler, A. L., Colbenson, K. M., & Zirakzadeh, A. (2009). Sexual abuse and lifetime diagnosis of somatic disorders: A systematic review and meta-analysis. *The Journal of the American Medical Association*, 302(5), 550-561. doi:10.4065/mcp.2009.0583

Paulhus, D.L. (1988). *Manual for the Balanced Inventory of Desirable Responding (BIDR-6)*. Unpublished manual, University of British Columbia.

Polaschek, D. L. L., & Gannon, T. A. (2004). The implicit theories of rapists: What convicted offenders tell us. *Sexual Abuse: A Journal of Research and Treatment*, 16, 299–314. doi:10.1177/107906320401600404

Polaschek, D., & Ward, T. (2002). The implicit theories of potential rapists: What our questionnaires tell us. *Aggression and Violent Behavior*, 7(4), 385-406. doi:10.1016/s1359-1789(01)00063-5.

Prentky, R., & Righthand, S. (2003). *Juvenile Sex Offender Assessment Protocol-II (JSOAP-II) manual*. DC: Department of Justice. doi:10.1037/e377042004-001

Prentky, R.A., Righthand, S., & Lamade, R. (2015). Sexual offending: Assessment and intervention. In K. Heilbrun, D. DeMatteo, & N. Goldstein (Eds.), *APA handbook of psychology and juvenile justice* (pp. 641-672). Washington, DC: American Psychological Association. doi:10.1037/14643-029

Prior, D., Farrow, K., Hughes, N., Kelly, G., Manders, G., White, S., & Wilkinson, B. (2011). *Maturity, young adults and criminal justice*. Birmingham: Institute of Applied Social Studies, School of Social Policy, University of Birmingham.

*Racey, B., Lopez, N., & Schneider, H. (2000). Sexually assaultive adolescents: Cue perception, interpersonal competence and cognitive distortions. *International Journal of Adolescence and Youth*, 8(2-3), 229-239.
doi:10.1080/02673843.2000.9747852

Rallings, M., & Webster, S. D. (2001). *The psychometric properties of HM Prison Service SOTP assessment battery* (Unpublished document). Offending Behaviour Programmes Unit, HM Prison Service.

Rich, P. (2009). Understanding the complexities and needs of adolescent sex offenders. In A.R. Beech, L.A. Craig, & K.D. Browne (Eds.), *Assessment and treatment of sex offenders* (pp. 431-453). Chichester: Wiley-Blackwell.
doi:10.1002/9780470714362.ch23

Richardson, G., Bhate, S., & Graham, F. (1997). Cognitive-based practice with sexually abusive adolescents. In M.S. Hoghughi (Ed.), *Working with sexually abusive adolescents*, (pp. 128-143) London, UK: Sage.

Salter, A.C. (1998). *Treating child sex offenders and their victims: A practical guide*. London: Sage Publications. doi:10.1016/0145-2134(92)90114-7

- Seto, M. C., Babchishin, K. M., Pullman, L. E., & McPhail, I. V. (2015). The puzzle of intrafamilial child sexual abuse: A meta-analysis comparing intrafamilial and extrafamilial offenders with child victims. *Clinical Psychology Review, 39*(C), 42–57. doi: 10.1016/j.cpr.2015.04.001
- Seto, M., & Lalumière, M. (2010). What is so special about male adolescent sexual offending? A review and test of explanations through meta-analysis. *Psychological Bulletin, 136*(4), 526-575. doi: 10.1037/a0019700
- Snowden, R. J., Craig, R. L., & Gray, N. S. (2011). Indirect behavioral measures of cognition among sexual offenders. *Journal of Sex Research, 48*(2-3), 192-217. doi:10.1080/00224499.2011.557750
- Social Care Institute for Excellence (SCIE). (2010). *SCIE systematic research reviews*. Retrieved from <http://www.scie.org.uk/publications/researchresources/rr01.pdf>
- Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review, 28*(1), 78-106. doi:10.1016/j.dr.2007.08.002
- Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences* (5th ed.). Hove, East Sussex: Routledge.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson/Allyn & Bacon.
- Tierney, D. W., & McCabe, M. P. (2001). An evaluation of self-report measures of cognitive distortions and empathy among Australian sex offenders. *Archives of Sexual Behavior, 30*(5), 495-519. doi: 10.1023/A:1010239217517
- The Home Office. (1994). *Community-based treatment for sex offenders: An evaluation of seven treatment programmes*. London: The Home Office.

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

doi:10.1177/107906320201400205

Thornton, D., Mann, R. E., Webster, S., Blud, L., Travers, R., Friendship, C., & Erikson, M. (2003). Distinguishing and combining risks for sexual and violent recidivism.

Annals of the New York Academy of Sciences, 989(1), 225-235.

doi:10.1111/j.1749-6632.2003.tb07308.x

*Tidefors, I., Goulding, A., & Arvidsson, H. (2011). A Swedish sample of 45

adolescent males who have sexually offended: Background, individual

characteristics, and offending behavior. *Nordic Psychology, 63*(4), 18-34.

doi:10.1027/1901-2276/a000043

Tolan, P. H., Walker, T., & Reppucci, N. D. (2012). Applying Developmental

Criminology to Law. *Justice Research and Policy, 14*(1), 117-146.

doi:10.3818/jrp.14.1.2012.117

Tully, R. J., Browne, K. D., & Craig, L. A. (2014). An Examination of the Predictive Validity of the Structured Assessment of Risk and Need-Treatment Needs

Analysis (SARN-TNA) in England and Wales. *Criminal Justice and Behavior.*

doi:10.1177/0093854814553096

*van Outsem, R., Beckett, R., Bullens, R., Vermeiren, R., van Horn, J., & Doreleijers, T.

(2006). The Adolescent Sexual Abuser Project (ASAP) assessment measures -

Dutch revised version: A comparison of personality characteristics between

juvenile sex offenders, juvenile perpetrators of non-sexual violent offences and

non-delinquent youth in the Netherlands. *Journal of Sexual Aggression, 12*(2),

127–141. doi:10.1080/13552600600858304

*Van Vugt, E., Hendriks, J., Stams, G.J., van Exter, F., Bijleveld, C., van der Laan, P., & Asscher, J. (2011). Moral judgment, cognitive distortions and implicit theories in young sex offenders. *Journal of Forensic Psychiatry & Psychology*, 22(4), 603–619. doi:10.1080/14789949.2011.603189

Viljoen, J., Mordell, S., & Beneteau, J. (2012). Prediction of adolescent sexual reoffending: A meta-analysis of the J-SOAP-II, ERASOR, J-SORRAT-II, and Static-99. *Law and Human Behavior*, 36(5), 423-438. doi:10.1037/h0093938

Wakeling, H. C. (2007). The psychometric validation of the Social Problem-Solving Inventory – Revised with UK incarcerated sexual offenders. *Sexual abuse: A Journal of Research and Treatment*, 19(3), 217-236. doi:10.1177/107906320701900304

Wakeling, H. C., & Barnett, G. D. (2014). The relationship between psychometric test scores and reconviction in sexual offenders undertaking treatment. *Aggression and Violent Behavior*, 19(2), 138-145. doi:10.1016/j.avb.2014.01.002

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

Walton, J., Duff, S., & Chou, S. (2014). A brief discussion about measuring child molester cognition with the Sex With Children Scale. *Child Abuse Review*. Advance online publication. doi:10.1002/car.2361

Ward, T. (2000). Sexual offenders' cognitive distortions as implicit theories. *Aggression and Violent Behavior*, 5(5), 491-507. doi:10.1016/s1359-1789(98)00036-6

- Ward, T. (2009). The extended mind theory of cognitive distortions in sex offenders. *Journal of Sexual Aggression*, 15(3), 247–259. doi:10.1080/13552600903263087
- Ward, T., & Beech, A. (2006). An integrated theory of sexual offending. *Aggression and violent behavior*, 11(1), 44-63. doi:10.1016/j.avb.2005.05.002
- Ward, T., & Keenan, T. (1999). Child molesters' implicit theories. *Journal of interpersonal violence*, 14(8), 821-838. doi:10.1177/088626099014008003
- Ward, T., Hudson, S. M., Johnston, L., & Marshall, W. L. (1997). Cognitive distortions in sex offenders: An integrative review. *Clinical Psychology Review*, 17(5), 479-507. doi:10.1016/s0272-7358(97)81034-3
- Ward, T., Polaschek, D. L., & Beech, A. R (2006). Theories of cognitive distortions. In T. Ward, D. Polaschek, & A. Beech (Eds.), *Theories of Sexual Offending*, (pp.115-133). Chichester, UK: Wiley. doi:10.1002/9780470713648
- Webster, S. D., Mann, R. E., Carter, A. J., Long, J., Milner, R. J., O'Brien, M. D., & Ray, N. L. (2006). Inter-rater reliability of dynamic risk assessment with sexual offenders. *Psychology, Crime & Law*, 12, 439-452. doi:10.1080/10683160500036889
- Whittaker, M. K., Brown, J., Beckett, R., & Gerhold, C. (2006). Sexual knowledge and empathy: A comparison of adolescent child molesters and non-offending adolescents. *Journal of Sexual Aggression*, 12, 143-154.
- *Worling, J. R. (1995). Adolescent sex offenders against females: Differences based on the age of their victims. *International Journal of Offender Therapy and Comparative Criminology*, 39(3), 276–293. Z

- Worling, J. R., & Curwen, T. (2001). Estimate of Risk of Adolescent Sexual Offense Recidivism (Version 2.0: The 'ERASOR'). In M. C. Calder (Ed.), *Juveniles and Children who Sexually Abuse: Frameworks for Assessment* (pp. 372-397). Lyme Regis, Dorset, UK: Russell House Publishing. doi:10.1037/t04170-000
- Worling, J. R., & Långström, N. (2006). Risk of sexual recidivism in adolescents who offend sexually: Correlates and assessment. In H. E. Barbaree & W. L. Marshall (Eds.), *The juvenile sex offender* (2nd ed.), pp. 219–247. New York: Guilford Press.
- *Zgourides, G., Monto, M., & Harris, R. (1997). Correlates of adolescent male sexual offense: Prior adult sexual contact, sexual attitudes, and use of sexually explicit materials. *International Journal of Offender Therapy and Comparative Criminology*, 41(3), 272-283. doi:10.1177/0306624X97413006

APPENDICES

Appendix 1: Examples of search terms and syntax developed to access data on the OvidSP and Proquest platforms

OvidSP - PsycINFO terms (number of hits in brackets)

- 1 juvenile delinquency/ (13311)
- 2 juvenile*.mp. (24648)
- 3 young*.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (166967)
- 4 adolescen*.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (167760)
- 5 'young adult*'.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (28408)
- 6 child.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (216909)
- 7 children.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (348473)
- 8 youth*.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (58229)
- 9 teen*.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (15451)
- 10 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 (654043)
- 11 sex offenses/ or incest/ or paraphilias/ or pedophilia/ (11325)
- 12 rape/ (4265)
- 13 'sex* offen*'.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (8391)
- 14 'sex* abus*'.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (21707)
- 15 'child molest*'.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (952)
- 16 'child abus*'.mp. [mp=title, abstract, heading word, table of contents, key

concepts, original title, tests & measures] (24270)

17 'child offen*'.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (28)

18 'sex* harm*'.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (60)

19 rapist*.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (1066)

20 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 (47440)

21 ('offen* support*' adj3 (thought* or attitude* or belief* or cognit* or schema*)).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (30)

22 ('offen* relate*' adj3 (thought* or attitude* or belief* or cognit* or schema*)).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (17)

23 ('abuse* support*' adj3 (thought* or attitude* or belief* or cognit* or schema*)).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (4)

24 ('abuse* relate*' adj3 (thought* or attitude* or belief* or cognit* or schema*)).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (18)

25 ('sex* offen*' adj3 (thought* or attitude* or belief* or cognit* or schema*)).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (241)

26 ('child molest*' adj3 (thought* or attitude* or belief* or cognit* or schema*)).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (62)

27 (rape adj3 (thought* or attitude* or belief* or cognit* or myth*)).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (1008)

28 ('pro-offen*' adj2 (thought* or attitude* or belief* or cognit* or schema*)).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (11)

- 29 (distort* adj2 (thought* or attitude* or belief* or cognit* or schema*)).mp.
[mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (1587)
- 30 (implicit adj2 (theor* or schema*)).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures] (1691)
- 31 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 (4471)
- 32 10 and 20 and 31 (458)

Proquest terms

LINE ONE

juvenile* OR young OR adolescen* OR 'young adult*' OR child OR children OR youth OR teen*

LINE TWO

'sex* offen*' OR 'sex* abus*' OR 'child molest*' OR 'child abus*' OR 'child offen*' OR 'sex* harm*

LINE THREE

('offen* support*' NEAR/3 (thought* OR attitude* OR belief* OR cognit* OR schema*))

('offen* relate*' NEAR/3 (thought* OR attitude* OR belief* OR cognit* OR schema*))

('abuse* support*' NEAR/3 (thought* OR attitude* OR belief* OR cognit* OR schema*))

('abuse* relate*' NEAR/3 (thought* OR attitude* OR belief* OR cognit* OR schema*))

ab(('sex* offen*' NEAR/3 (thought* OR attitude* OR belief* OR cognit* OR schema*)))

ab(('child molest*' NEAR/3 (thought* OR attitude* OR belief* OR cognit* OR schema*)))

ab((rape NEAR/3 (thought* OR attitude* OR belief* OR cognit* OR myth*)))

ab(('pro-offen*' NEAR/3 (thought* OR attitude* OR belief* OR cognit* OR
schema*)))

ab((distort* NEAR/2 (thought* OR attitude* OR belief* OR cognit* OR schema*)))

ab((implicit NEAR/2 (theor* OR schema*)))

Appendix 2: Screening and selection tool (SST)

Reviewer name:		Date:
Reference:		
	INCLUDE if....	EXCLUDE if....
Population	<input type="checkbox"/> Male	<input type="checkbox"/> Female
	<input type="checkbox"/> Age 10-21 (or mean age under 18 if age range not specified)	<input type="checkbox"/> Participants are all aged 18 or over
	<input type="checkbox"/> Convicted of or charged with a sexual offence, or referred due to engaging in sexually harmful behaviour	<input type="checkbox"/> Members of the public or 'at risk' groups
Phenomenon of Interest	<input type="checkbox"/> OSC - cognitive structures and content including thoughts, beliefs, attitudes, schemes, implicit theories, (or distortions of these) that directly or indirectly support sexual offending	<ul style="list-style-type: none"> • Cognitive processes <input type="checkbox"/> • General pro-criminal or antisocial thoughts, beliefs, attitudes schemas and implicit theories <input type="checkbox"/> • Empathy/empathy deficits <input type="checkbox"/> • Absolute denial. <input type="checkbox"/>
Comparison group	<ul style="list-style-type: none"> • Non offender <input type="checkbox"/> • Non-sexual offender <input type="checkbox"/> • Sub-groups of sexual offender (e.g., child molester, peer aggressor) <input type="checkbox"/> • Pre- post-intervention <input type="checkbox"/> • Recidivist, non-recidivist <input type="checkbox"/> 	<ul style="list-style-type: none"> • No comparison <input type="checkbox"/> • Female comparison <input type="checkbox"/>
Outcome	<input type="checkbox"/> A comparison between the population of interest and at least one other group on a numerical measure of OSC	<input type="checkbox"/> No comparison group or no specific numerical measure of OSC
Research / type design	<input type="checkbox"/> Quantitative – MUST include a comparison group.	<input type="checkbox"/> Qualitative studies, Narrative reviews, editorials, commentaries
Publication type	<input type="checkbox"/> Published source	<input type="checkbox"/> Unpublished theses and dissertations
Language	<input type="checkbox"/> English	<input type="checkbox"/> Any other language
Decision		

Notes:

Appendix 3: List of excluded studies (n=24) and reasons for exclusion***Narrative review (9)***

- Beckett, R. (1999). Evaluation of adolescent sexual abusers. In M. Erooga & H. Masson (eds). *Children and young people who sexually abuse others: Challenges and responses* (pp. 204-224). Florence, KY: Taylor & Frances/Routledge; US.
- Epps, K. (1997). Developing a therapeutic environment for boys in secure settings who have sexually offended. *Journal of Sexual Aggression*, 3(2), 71-86.
doi:10.1080/13552609708413271
- Fanniff, A. M., & Becker, J. V. (2006). Specialized assessment and treatment of adolescent sex offenders. *Aggression and Violent Behavior*, 11(3), 265-282.
doi:10.1016/j.avb.2005.08.003
- Graham, F., Richardson, G., & Bhate, S. (1997). Assessment. In M.S. Hough, S.R. Bhate, & F.Graham. *Working with sexually abusive adolescents* (pp. 52-91). Thousand Oaks, CA: Sage. doi: 10.4135/9781446279687.n4
- Lakey, J. F. (1994). The profile and treatment of male adolescent sex offenders. *Adolescence*, 29(116), 755-761. doi:10.1080/02673843.1995.9747780
- Pierce, L. H., & Pierce, R. L. (1990). Adolescent/sibling incest perpetrators *The incest perpetrator: A family member no one wants to treat* (pp. 99-107). Thousand Oaks, CA: Sage
- Righthand, S., & Welch, C. P. D. (2002). Juveniles who have sexually offended: An introduction. *Prevention Researcher*, 9(4), 1-1, 3, 4. doi:10.1037/e313292005-001
- Whitford, R., & Parr, V. (1995). Uses of rational emotive behavior therapy with juvenile sex offenders. *Journal of Rational-Emotive & Cognitive-Behavior Therapy*, 13(4), 273-282. doi:10.1007/BF02354518

Will, D., Douglas, A., & Wood, C. Evolution of a group therapy programme for adolescent perpetrators of sexually abusive behaviour. *Journal of Sexual Aggression*, 1(2), 69-82. doi: 10.1080/13552609408413245

No specific measure of OSC (9)

Barroso, R., Manita, C., & Nobre, P. J. (2010). The influence of family violence in young sexual offenders behaviour. *Sexologies*, 19, S65.

Burton, D. L. (1999). An examination of social cognitive theory with differences among sexually aggressive, physically aggressive and nonaggressive children in state care. *Violence & Victims*, 14(2), 161-178.

DeLisi, M., Vaughn, M. G., Beaver, K. M., Wright, J. P., Hochstetler, A., Kosloski, A. E., & Ury, A. J. (2008). Juvenile sex offenders and institutional misconduct: The role of thought psychopathology. *Criminal Behaviour and Mental Health*, 18(5), 292-305. doi:10.1002/cbm.707

Grant, J., Indermaur, D., Thornton, J., Stevens, G., Chamarette, C., & Halse, A. (2009). Intrafamilial adolescent sex offenders: Psychological profile and treatment. *Australian Institute of Criminology, GPO Box 2944 Canberra ACT, 2601 SuppNotes - Trends & Issues in Crime and Criminal Justice, No. 375, June 2009.* doi:10.1080/13218719.2011.561763

Hastings, T., Anderson, S. J., & Hemphill, P. (1997). Comparisons of daily stress, coping, problem behavior, and cognitive distortions in adolescent sexual offenders and conduct-disordered youth. *Sexual Abuse: Journal of Research and Treatment*, 9(1), 29-42. doi: 10.1177/107906329700900103

- Ikomi, P. A., Harris-Wyatt, G., Doucet, G., & Rodney, H. E. (2009). Treatment for juveniles who sexually offend in a southwestern state. *Journal of Child Sexual Abuse, 18*(6), 594-610. doi:10.1080/10926770903307914
- Lambie, I., Hickling, L., Seymour, F., Simmonds, L., Robson, M., & Houlahan, C. (2000). Using wilderness therapy in treating adolescent sexual offenders. *Journal of Sexual Aggression, 5*(2), 99-117. doi:10.1080/13552600008413302
- Valliant, P. M., & Clark, L. M. (2009). An evaluation of non-assaultive, assaultive, and sexually assaultive adolescents at pretrial sentencing: a comparison on cognition, personality, aggression, and criminal sentiments. *Psychological Reports, 105*(3 Pt 2), 1077-1091. doi: 10.2466/pr0.105.3f.1077-1091
- Van Vugt, E., Stams, G. J., Dekovic, M., Brugman, D., Rutten, E., & Hendriks, J. (2008). Moral development of solo juvenile sex offenders. *Journal of Sexual Aggression, 14*(2), 99-109. doi:10.1080/13552600802248106¹⁵

No comparison group (2)

- Hunter, J. A. (2004). Developmental pathways in youth sexual aggression and delinquency: Risk factors and mediators. *Journal of Family Violence, 19*(4), 233-242. doi: 10.1023/B:JOFV.0000032633.37269.1d
- McCrary, F., Kaufman, K., Vasey, M. W., Barriga, A. Q., Devlin, R. S., & Gibbs, J. C. (2008). It's all about me: A brief report of incarcerated adolescent sex offenders' generic and sex-specific cognitive distortions. *Sexual Abuse: Journal of Research and Treatment, 20*(3), 261-271. doi:10.1177/1079063208320249

¹⁵ Did include a measure of OSC but this was not used with the comparison group

Meta-analytical studies (2)

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

Seto, M. C., & Lalumiere, M. L. (2010). What Is So Special About Male Adolescent Sexual Offending? A Review and Test of Explanations Through Meta-Analysis. *Psychological Bulletin, 136*(4), 526-575. doi:10.1037/a0019700

Non-convicted sample (1)

Fromuth, M. E., Burkhart, B. R., & Jones, C. W. (1991). Hidden child molestation: An investigation of adolescent perpetrators in a nonclinical sample. *Journal of Interpersonal Violence, 6*(3), 376-384. doi:10.1177/088626091006003009

All adult sample (1)

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

Appendix 4: Quality assessment form for case control studies

Name of study:

Criterion	Yes (2)	Partly (1)	No (0)	Can't tell (0)	Comments
<i>Initial Screening</i>					
1. Did the study address a clearly focused issue/aim/hypothesis?					
<i>Selection bias</i>					
2. Was there sufficient demographic info about the sample (cases and comparisons)					
3. Was the type(s) of sex offence(s) that defined the cases, and if part of the design, the comparison group, defined precisely					
4. Were the sample recruited in an appropriate way					
5. Are the cases representative of the defined population					
6. Were sufficient cases selected (power analysis /sufficient for analysis)					
7. Were the comparison group representative of a defined population					
8. Were sufficient comparisons selected (power analysis /sufficient for analysis)					

9. Are the cases and comparison group comparable in relation to demographic/confounding variables (e.g. age, ethnicity, SES, education, risk, location)					
<i>Measurement/classification bias</i>					
10. Was the measure of OSC appropriate given aims of the study					
11. Had the measure(s) of OSC been standardised, validated and reliability tested on an appropriate population (i.e. young people)					
12. Was the validity of the OSC measure good (2), adequate (1) poor (0), not reported (0)					
13. Was the reliability of the OSC measure good (2), adequate (1) poor (0) not reported (0)					
14. Was OSC measured in the same way for cases and controls					
15. Have the authors taken account of confounding factors (in design/analysis)					

Results					
16. Is the analysis appropriate to the design					
17. Have limitations been identified and discussed					
TOTAL					

Total Score (X) =

Max score: 34

Percentage: $X/34 \times 100 =$

Number of items missing / can't tell =

Any other comments:

Appendix 5: Quality assessment form for case series studies

Name of Study:

Criterion	Yes (2)	Partly (1)	No (0)	Can't tell (0)	Comments
<i>Initial Screening</i>					
1. Is the hypothesis / aims /objectives of the study clearly stated?					
<i>Selection bias</i>					
2. Are the characteristics of the participants included in the study described?					
3. Are the eligibility criteria (i.e. inclusion and exclusion criteria) for entry into the study clearly stated?					
4. Were participants recruited in an appropriate way?					
<i>Measurement bias</i>					
5. Was the intervention of interest clearly described?					
6. Was OSC measured with a relevant and appropriate objective and/or subjective method(s)					
7. Had the measure(s) of OSC been standardised, validated and reliability tested on the defined population					
8. Was the validity of the OSC measure good (2), adequate					

(1) poor (0), not reported (0)					
9. Was the reliability of the OSC measure good (2), adequate (1) poor (0), not reported (0)					
10. Were the statistical tests used to assess the relevant outcomes appropriate?					
<i>Attrition bias</i>					
11. Was the length of follow-up reported?					
12. Was the loss to follow-up reported?					
<i>Performance/detection bias</i>					
13. Were the providers of the intervention/assessment blinded?					
14. Was the study conducted prospectively?					
<i>Results</i>					
15. Are the conclusions of the study supported by results?					
16. Are both competing interests and sources of support for the study reported?					

Total Score (X) =

Max Score: 32

Percentage: $X/32 \times 100 =$

Number of items missing/can't tell =

Any other comments:

Appendix 6: Data extraction form

General Information	
Title of Study	
Author	
Year	
Source	
Country of study	
Study characteristics	
Design/type of study	
Aim/objective of study	
Measure of OSC - concept measured	

Standardisation, validity and reliability of measure	
Participant characteristics cases (case control) whole sample (case series)	
Type of offence / definition	
Number of cases/sample	
Age range and mean of cases/sample	
Ethnicity of cases/sample	
Location of cases /sample	
Recruitment procedure	
Comparison group characteristics (case control)	
Type of comparison group & definition	
Number in the comparison group(s)	

Age range and mean of comparison group(s)	
Ethnicity of the comparison group	
Location of comparison group	
Recruitment procedure	
Intervention information (CS only)	
Description of intervention including timing of post intervention assessment	
Attrition rate	
Study results	
Analysis used	
Results & Significance	
Conclusions	

Quality	
Quality assessment score	

Appendix 7: The CASQ-AV

ASAP Booklet Two (for child sexual offenders)

[removed to protect the integrity of the questionnaire]

Appendix 8: Items contained in the two sub-scales of the original CASQ-AV

[removed to protect the integrity of the questionnaire]

Appendix 9: HM Prison Service Consent Form**Consent Form for the following Offending Behaviour Programme:**

.....

I understand that I am consenting to:

- ☐ Being asked questions by members of the Programme Team that relates to my treatment.
- ☐ Filling in questionnaires for assessment.
- ☐ My assessments being used to inform my treatment needs, finding out about the risk I present and progress that I have made.
- ☐ Taking part in sessions.
- ☐ Being recorded in some assessments and all treatment sessions.
- ☐ Respecting other group members by keeping confidential any information I learn about them during the course of the programme. This includes not discussing other group members offences or private lives outside of the group room.
- ☐ Having my details entered into a database. I know that this data will then be used to carry out research into whether the programme is effective. I understand that this data will not be able to identify me.
- ☐ Details from my assessments and progress on the above programme will be shared with other members of staff and maybe used to inform decisions about my progress through sentence and release.

	Yes	No
I have a copy and been given time to read the 'Guide to Consent'.	<input type="checkbox"/>	<input type="checkbox"/>
I have a copy and been given time to read a booklet about the specific programme I am being assessed for.	<input type="checkbox"/>	<input type="checkbox"/>
I have been given the time to ask all the questions I wanted to. Those questions have been answered fully.	<input type="checkbox"/>	<input type="checkbox"/>
I understand that I can take back my consent at any time and someone will discuss with me what this means if it happens.	<input type="checkbox"/>	<input type="checkbox"/>
I agree to give my consent to take part in the assessment and treatment for the above programme.	<input type="checkbox"/>	<input type="checkbox"/>
Signed		
Date		

Staff Checklist

- ☐ I have checked the person concerned could read.
- ☐ I have given them a copy of the Guide to consent, and programme specific booklet.
- ☐ They have had at least one night to read these booklets.
- ☐ If the participant could not read I have explained the booklets to the best of my ability. I have also asked a mentor, listener or personal officer to go through the booklets again with the person.
- ☐ I have answered all their questions in a way that they could understand.
- ☐ I feel confident that they have given informed consent and have not be forced or misinformed into consenting.

Signed

Date

Appendix 10: Information sheet & consent form for the test-retest study**A study looking at the questionnaires young adults complete before starting the Sexual Offenders Treatment Programme (SOTP)
Information Sheet**

Who is doing the study? This study is being done by Louise Bowers. Louise is studying for a doctorate at the University of Birmingham. Louise is also working with researchers at Prison Service headquarters. Louise has permission from the University of Birmingham and the National Offender Management Service (NOMS) to do this study.

What is the study about? The study is looking at how good the questionnaires are that young adults fill out before and after they complete the SOTP. The questionnaire Louise is really interested in, is one about young adults' thoughts feelings and beliefs about children.

What will the study involve? You are being asked if you would mind filling out ONE questionnaire. The questionnaire will take about 5-10 MINUTES TO COMPLETE. You will just have to tick boxes to say how true you think things are. You will be asked to fill out the questionnaire again in one month's time

Why have you been asked to take part? You have been asked to take part because you have been convicted of a sexual offence.

Do you have to take part? The answer is NO! Taking part is voluntary – that means YOU CAN CHOOSE if you want to take part or not.

Can I change my mind after I have said yes? The answer is YES! If you change your mind you should contact the programme manager (XX) and she will contact Louise, and ask for your questionnaire to be taken out of the study. But, you can only change your mind up until XX because this is when Louise will start to analyse the data.

Will the information in the questionnaire be kept confidential? YES! When you have finished filling in the questionnaire, the treatment manager here will take your name off and put a CODE on it.

What will happen to the information I give? Louise will put the information from all of the questionnaires you and others have filled out onto a database. Then she will do some tests on it to try and work out how good the questionnaire is.

What will happen to the results? Louise will use the results as part of her qualification. She might also publish the results in a journal. But remember, no one will know that you took part in the study unless you want to tell them.

What might be bad about taking part? The only bad thing is that you will have to fill the questionnaire in now, and then again in one month's time. You may have to fill this questionnaire out again if you start treatment.

What might be good about taking part? You would be playing a part in making the assessment of young adult sexual offenders better.

If you need any further information please ask the member of staff who is with you

THANK YOU ☺

Consent Form

Your treatment manager will keep this form, and just tell Louise the code.

I confirm that.....

- Someone has talked to me about the research and I have read the information sheet about the study
- I know I can choose to take part or not
- I know I can change my mind about taking part and I can ask for my questionnaire to be taken out of the study up until XX 2015
- I know my information will not be shared with anyone else
- The results of the research may be published but I will not be identifiable

I..... want to take part in the research study.

☐

Signed.....

Date.....

Appendix 11: CASQ-AV Three Component Solution*Factor loadings from the CASQ-AV rotated three-factor solution*

	Component		
	1	2	3
RQ27 I think about children when I am alone	.861	-.108	.045
RQ82 Newspapers and television stir up my feelings about children	.828	.050	-.110
RQ54 I think about children when I am alone	.821	-.070	-.023
RQ7 I sometimes get very strange feelings with children	.818	-.058	-.044
RQ55 when a child smiles at me, it can stir me up	.787	.066	-.037
RQ81 some children make me feel funny inside	.772	.052	.016
RQ16 thinking about children makes me feel good	.727	-.143	.234
RQ77 children are very attractive	.691	.106	.084
RQ84 some of my closest friends have been children	.649	.070	-.014
RQ58 I feel more comfortable with children than with people of my age	.631	.089	.092
RQ12 I prefer to spend my time with children	.586	.007	.170
RQ28 being with children stops me from being lonely	.558	.027	.253
RQ68 There is no harm in sexual contact between children and people of my age	.555	.404	-.160
RQ42 if children want they should be allowed to have sexual relations with adults	.534	.450	-.166

RQ37 children are special for me	.510	-.025	.345
RQ15 I have loved children at first sight	.495	-.109	.277
RQ83 I can talk about my feelings with children	.482	.182	.082
RQ72 If children want they should have sexual contact with people of my age	.482	.478	-.185
RQ13 I talk to children about my problems	.481	.069	.039
RQ34 there is nothing wrong with sexual contact between children and people of my age	.469	.312	-.271
RQ22 children tease me	.430	.327	-.094
RQ76 Some children prefer to be with me rather than their parents	.429	.276	.225
RQ60 sometimes I meet a child who I know has special feelings about me	.410	.380	.050
RQ30 I know more about children than their parents do	.399	.187	.178
RQ65 I have loved a child	.399	-.176	.357
RQ73 children can make me do things against my will	.378	.285	-.079
RQ5 I envy children	.357	.128	.090
RQ78 girls are like women	.335	.325	.040
RQ46 adults cannot be trusted	.134	.099	.119
RQ70 People don't realise how much children know about sex	-.164	.790	.146
RQ50 children can lead people on	-.009	.756	.028

RQ71 Some children could teach people my age about sex	-.070	.745	-.018
RQ20 some children know more about sex than people of my own age	-.089	.717	-.021
RQ61 children sometimes ask people my age for sex	.021	.710	.027
RQ66 children can lead people of my age astray	.047	.705	.075
RQ33 children want sexual contact with people my age	.161	.670	-.126
RQ41 when adults and children have sexual relations it is not the adults fault	.017	.661	.015
RQ40 children are not as innocent as most people think	-.130	.654	.153
RQ49 children can flirt with people of my age	.157	.650	-.082
RQ45 children can blackmail people of my age	-.069	.627	.153
RQ19 children know a lot about sex	-.094	.612	.004
RQ43 not all sexual contact between people of my age an children causes harm	.317	.539	-.079
RQ9 children like to talk about sex	.159	.525	-.045
RQ44 children tell lies about people of my age	-.030	.520	.097
RQ53 some children find me attractive	.236	.490	.098
RQ21 society does not give children enough responsibility	.064	.419	.207
RQ25 people don't know what children are like	.089	.417	.214
RQ80 children seem to seek me out	.289	.414	.111
RQ51 children can look after themselves	.147	.403	.063

RQ23 sometimes children look at me in a special way	.245	.373	.163
RQ14 most people my age do not understand children	.078	.262	.181
RQ47 many people have a sexual interest in children	.208	.225	.138
RQ29 I am good at making children laugh	.057	.008	.715
RQ24 I know how to talk to children	-.007	.037	.700
RQ4 children like to play with me	-.153	.143	.698
RQ2 children like my company	-.162	.147	.640
RQ1 children feel safe with me	-.196	.089	.617
RQ6 some children make me feel good	.340	-.118	.599
RQ38 children find me easy to make friends with	.174	.115	.595
RQ56 I know what children like	.340	-.091	.581
RQ26 I love children	.254	-.102	.571
RQ74 When I feel low children cheer me up	.445	-.051	.508
RQ8 I find it hard to resist children's requests	.318	-.007	.486
RQ32 children know what they like	-.045	.252	.444
RQ79 I find it easy to talk to children	.338	-.036	.443
RQ59 children speak with their eyes	.120	.192	.426
RQ64 I am better than most people at getting along with children	.301	.279	.419
RQ57 children remind me of myself	.352	.103	.415
RQ18 children are more honest than adults	.009	.069	.392

RQ17 I know when children are interested in me	.256	.158	.373
RQ11 children are powerful	-.024	.285	.371
RQ36 children can be trusted	.237	-.004	.353
RQ63 I am better than most people at understanding children	.312	.231	.319

Appendix 12: Items Contained in the CASQ-AV Four Component Solution***Component 1, Positive Affect (15 items)***

- Q27 I think about children when I am alone
- Q7 I sometimes get very strange feelings with children
- Q82 Newspapers and television stir up my feelings about children
- Q55 When a child smiles at me, it can stir me up
- Q81 Some children make me feel funny inside
- Q16 Thinking about children makes me feel good
- Q77 Children are very attractive
- Q84 Some of my closest friends have been children
- Q58 I feel more comfortable with children than with people of my age
- Q12 I prefer to spend my time with children
- Q28 Being with children stops me from being lonely
- Q37 Children are special for me
- Q15 I have loved children at first sight
- Q13 I talk to children about my problems
- Q83 I can talk about my feelings with children

Component 2, Child Sexuality (16 items)

- Q61 Children sometimes ask people my age for sex
- Q33 Children want sexual contact with people my age
- Q71 Some children could teach people my age about sex
- Q70 People don't realise how much children know about sex
- Q50 Children can lead people on

Q49 Children can flirt with people of my age

Q41 When adolescents & children have sexual relationships it is not the adolescent's fault

Q66 Children can lead people of my age astray

Q72 If children want they should have sexual contact with people of my age

Q43 Not all sexual contact between people of my age and children causes harm

Q68 There is no harm in sexual contact between children and people of my age

Q20 Some children know more about sex than people of my own age

Q9 Children like to talk about sex

Q42 If children want they should be allowed to have sexual relationships with adolescents

Q53 Some children find me attractive

Q45 Children can blackmail people of my age

Component 3, Child Orientation (10 items)

Q1 Children feel safe with me

Q4 Children like to play with me

Q2 Children like my company

Q29 I am good at making children laugh

Q24 I know how to talk to children

Q26 I love children

Q6 Some children make me feel good

Q56 I know what children like

Q38 Children find me easy to make friends with

Q74 When I feel low children cheer me up

Component 4, Child Maturity (5 items)

Q11 Children are powerful

Q21 Society does not give children enough responsibility

Q25 People don't know what children are like

Q40 Children are not as innocent as most people think

Q32 Children know what they like