RAPE MYTH ACCEPTANCE: ITS ROLE, IMPORTANCE, AND PSYCHOMETRIC MEASUREMENT

by

LARISSA GABRIELLE JOHNSON

A thesis submitted to the University of Birmingham

for the degree of

Doctorate in Forensic Psychology Practice (ForenPsyD)

Centre for Forensic and Criminological Psychology

School of Psychology

College of Life and Environmental Science

University of Birmingham

June 2017

Abstract

Rape myth acceptance (RMA) is an important psychological concept in the research, assessment, education, and treatment of sexual violence. The current thesis presents the role, importance, and psychometric measurement of RMA in research and practice. This work aimed to review RMA as a criminogenic need in adult male rapists (i.e., whether RMA was affected by sex offender intervention, linked to recidivism, and/or presented differently across types of offenders and non-offenders) and found mixed results across the literature. Additionally, a focus was placed on the Illinois Rape Myth Acceptance (IRMA) Scale – a prominent measure of endorsement of rape myths. Its psychometric properties were evaluated, and it was found to be a reliable measure of RMA based on previous studies. However, it is highlighted that, due to cultural changes over time, it is necessary to continue measuring the reliability and validity of this measure. The current thesis employed the IRMA to measure acceptance of rape myths amongst university students to establish the factor structure, dimensionality, and reliability of the IRMA. A four-component factor structure was found across two dimensions as well as high internal consistency of the scale. Implications for practice and future research are discussed.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my academic supervisor, Professor Anthony Beech for the continuous support across academic and professional realms. I am grateful for your guidance, motivation, and immense knowledge.

Besides my academic supervisor, I would like to thank the rest of the ForenPsyD faculty, past and present: Dr Caroline Oliver and Dr Myfanwy Ball for your warmth, insightful comments, and encouragement over the years; Dr Darren Charles Frances Bishopp, for this collaboration, your dark sense of humour, and your never-ending list of qualms with statistical analysis. You took something so daunting and made it approachable; I am also grateful for the work of Sue Hanson, once the backbone of this course, for your endless support and amazing tenacity. You are missed.

My sincere thanks also go to the Cayman Islands Government Scholarship Committee, for without your award, I might never have been able to pursue this dream.

I thank my fellow colleagues for the stimulating discussions, shared appreciation for anything edible, and endless laughter and fun over the last three years. To Dr Philip Howard, thank you for your companionship and your perspicacious editing eye. Also, I thank my colleagues at Her Majesty's Cayman Islands Prison Service, Aduke Joseph-Caesar, Michael Chester, Dr Elma Augustine, and Helen Reynolds. I would also like to thank Kathryn Dinspel-Powell of the Cayman Islands' Ministry of Home Affairs.

Last but not the least, I would like to thank my family – my parents, Lascelles and Marjorie, brother, LJ, sister-in-law, Cybel, my best friends, and my partner for supporting me spiritually throughout writing this thesis and throughout my life in general.

Contents

Table of Appendices

List of Tables

List of Figures

CHAPTER 1: INTRODUCTION TO THE THESIS

Introduction

Rape is a widespread and significant societal issue which has a devastating impact on those whom it affects (Baldwin-White, Thompson, & Gray, 2016; Branscombe, Wohl, Owen, Allison, & N'gbala, 2003; Paul, Gray, Elhai, & Davis, 2009; Sleath & Bull, 2015). However, despite its severity, cases of rape remain underreported in the United Kingdom (Kelly, Lovett, & Regan, 2005; Rape Crisis England and Wales, 2017; Xue et al., 2016). In England and Wales, approximately 85,000 adult women and 12,000 adult men are raped or are victims of attempted rape every year, and nearly half a million adults are sexually assaulted (Ministry of Justice (MoJ), Home Office, & Office of National Statistics (ONS), 2013). Of these, however, only 15% of victims of sexual violence report to the police (MoJ, Home Office, & ONS, 2013). Furthermore, approximately just 5.7% of rape cases end in conviction of the perpetrator (Kelly et al., 2005; Rape Crisis England and Wales, 2017). Such low reporting and conviction rates may be indicative of an endemic societal issue. It could be that the occurrence of rape, how society views victims, and how it is handled in the criminal justice system are all influenced by shared societal beliefs and attitudes. Indeed, psychological and sociological researchers as well as social justice advocates highlight that rape supportive attitudes and beliefs contribute to not only the extensive prevalence of sexual assault, but also the underreporting of sexual assault in society (Xue et al., 2016). For example, feelings of entitlement to women's bodies can act as a disinhibitor for men forcing sexual intercourse on women they deem are "asking for it" (Bohner et al., 1998; Chapleau & Oswald, 2010). Furthermore, victims of rape are likely to be blamed for the attack and automatically associated with negative stereotypes assigned to victims of rape (e.g. promiscuity and irresponsibility; Baldwin-White et al., 2016; Buddie & Miller, 2001; Chapleau & Oswald, 2010; Gray, 2006). Individuals, offenders, and non-offenders alike, tend to hold particular

perceptions, prejudices, and sometimes false beliefs about rapists, rape and victims of rape. These are referred to as rape myths.

Rape Myths and Rape Myth Acceptance

Expanding on the work of Brownmiller (1975), Martha Burt first defined the concept of rape myths in 1980 as "prejudicial, stereotyped, or false beliefs about rape, rape victims, and rapists" (p. 217). In later years, Lonsway and Fitzgerald (1994) went on to expand on the definition, stating that rape myths are "attitudes and beliefs that are generally false but are widely and persistently held, and that serve to deny and justify male sexual aggression against women" (p. 134). Some examples of rape myths in the literature include: "in the majority of rapes, the victim is promiscuous or had a bad reputation," "any healthy woman can successfully resist a rapist if she really wants to," (Burt, 1980) and "If a woman goes home with a man she doesn't know, it is her own fault if she is raped" (Payne, Lonsway, & Fitzgerald, 1999). Rape myths influence attitudes towards victims on a social level. Endorsement of such myths allows an individual to shift the blame for the crime towards the victim (Chapleau & Oswald, 2010; Gray, 2006).

Burt (1980) reported that rape myths serve the function of lowering a man's inhibitions so that he may proceed to offend. Chapleau and Oswald (2010) added to this explanation, purporting that endorsement of rape myths – or rape myth acceptance (RMA) – may reduce the expectation of negative outcomes or consequences in sexual offenders. Rape myth acceptance has been a major topic in rape literature and research has identified the negative impact of RMA across a variety of settings.

The ongoing endorsement of rape myths poses many challenges for victims of rape as well as the criminal justice system (Sleath & Bull 2015). The literature has shown that rape survivors who endorse and internalise rape myths, blaming themselves for the rape, are

negatively impacted psychologically (Baldwin-White et al., 2016; Branscombe et al., 2003), with some exhibiting posttraumatic stress disorder (PTSD) symptoms (Baldwin-White et al., 2016; Paulet al., 2009). Additionally, there are survivors who will not report an incident of rape because they are unsure whether what took place was a punishable offense or if they were to blame (Baldwin-White et al., 2016; Buddie & Miller, 2001; Durán, Moya, Megías, & Viki, 2010; Franiuk, Seefelt, & Vandello, 2008; Hayes-Smith & Levett, 2010; King & Roberts, 2011; Peterson & Muehlenhard, 2004). These beliefs are not far-fetched when one considers that, in general, rape myth acceptance is a widespread, societal issue. For example, Amnesty (2005) reported that a third of people believe women who flirt are partially responsible for being raped (Amnesty, 2005; Rape Crisis England and Wales, 2017). Lonsway and Fitzgerald (1994) found using various measures of RMA that between 25 to 35 per cent of both male and female participants endorsed the majority of rape myths. Further to this, it has been shown that men are more likely to accept rape myths than women (Edwards, Turchik, Dardis, Reynolds, & Gidycz, 2011; Suarez & Gadalla 2010).

Rape myth acceptance has also been identified in the literature as a criminogenic need – a risk factor – for sexual offenders. Past research revealed that high levels of RMA are strongly associated with rape proclivity – one's likelihood or tendency to choose to rape – and, more significantly, highlighted RMA as a causal antecedent for rape proclivity in men (Bohner et al., 1998; Chapleau & Oswald, 2010; Chiroro, Bohner, Viki, & Jarvis, 2004; Edwards et al., 2011; Gray, 2006; Malamuth & Check, 1985). There is evidence of RMA amongst convicted rapists, using myths to rationalise their behaviours (Chiroro et al., 2004). Baldwin-White et al. (2016) note that, on an individual level, RMA can act as an 'enabler' for sexual assault. They write that while contemplating engaging in sexual assault, RMA allows men to turn off dissenting thoughts, and these men are more likely to perpetrate sexual assault than men who do not accept rape myths (Baldwin-White et al., 2016). Overall, RMA

influences how individuals infer social information and serves the function of increasing victim blaming (Jones & Aronson, 1973), decreasing perpetrator blame (Eyssel & Bohner, 2011), decreasing women's anxiety around becoming potential victims (Bohner & Lampridis, 2004), and lowering men's inhibitions to be sexually aggressive (Bohner et al., 1998; Chapleau & Oswald, 2010).

The Importance of RMA Research

Rape myths can be 'protective' for both men and women and the reasons behind why men and women endorse these myths may differ (Bratcher, 2011; Payne et al., 1999). It is likely that the influx of information about rapists is perceived as an attack on men and rape myths function to shift the blame and responsibility away from men onto women; thus, this would be beneficial for men to believe in rape myths. Women, however, may endorse rape myths as a guide of sorts. If a woman believes that if she acts a certain way or refrains from engaging in certain behaviours, then she does not have to face the frightening possibility that rape is out of her hands; in this sense, rape myths offer a sense of safety. These beliefs have a dangerous impact on society. Eyssel, Bohner, and Siebler's (2006) research supported the notion that societal norms – or our perceptions of social norms – influence our beliefs and, consequently, our behaviour. Paul and colleagues (2009) found that individuals believe their levels of RMA are lower than the social norm. These appraisal comparisons are problematic, as they leave the individual with a sense of complacency that serves to ignore the societal issue that is sexual assault. The omnipresence of rape myths and inherent individual comparison of RMA social norms within society is likely to lead potential perpetrators to believe these rationalisations are reasonable. These perpetrators are more likely to use the rationalisations to excuse and justify their behaviour (Wegner, Abbey, Pierce, Pegram, & Woerner, 2015).

This social endorsement of rape myths leaves people with the belief that this is an issue that does not affect them and, should the conversation arise, they may feel uncomfortable and disconnect. Rich, Utley, Janke, and Moldoveanu (2010) carried out a survey with college men to gauge their responses to sexual assault education. Not only did they find the men to be disinterested in the programme, they found that half (51%) thought it was irrelevant to their lives, while 11% were offended at the thought of having to attend a sexual assault prevention programme. Thirty-two percent of the men felt that a mandatory or voluntary sexual assault programme would be beneficial, however, they voiced that it would be unlikely they would attend. This highlights a need to continue to allow for open discussion and further education around sexual assault. Sexual assault prevention programmes and bystander intervention programmes have been found to be effective in reducing rates of victimisation and increase bystander intervention, respectively, despite resistance (Bratcher, 2011; Lee et al., 2003; McMahon, 2010; Rich et al., 2010; Rothman & Silverman, 2007). However, McMahon (2010) found that an individual's willingness to intervene as a bystander was negatively correlated with acceptance of rape myths and encouraged future programmes to include content on rape myths.

It has been suggested that RMA rates have declined over time (Edward & McLeod, 1999), however, Edwards et al. (2011) noted that methodologically, it is quite difficult to draw a valid comparison between RMA rates across time. They also highlighted literature which showed that when students took part in rape education programmes, it was linked to a decrease in scores on RMA scales (Edwards et al., 2011; Hinck & Thomas, 1999) but noted that, in actuality, this apparent decrease in RMA is reflective of greater unwillingness to acknowledge rape myths due to increased awareness of sexual assault being socially unacceptable (Edwards et al., 2011). Thus, it is pertinent that research is able to produce valid and reliable measures, both explicit and implicit.

Wegner and colleagues (2015) posited that attitudes supportive of rape are likely to be "activated" in contexts which associate strongly with an individual's pre-existing attitudes. These rape-supportive attitudes, which include belief in rape myths, may bias a perpetrator's perception of their victims' actions as their RMA sets an expectation (Wegner et al., 2015; Snyder & Higgins, 1988). For example, when a perpetrator interprets situational cues consistent with their pre-existing attitudes about rape, such as a victim's alcohol consumption or "provocative" attire, they are likely to feel justified in their use of sexual aggression (Wegner et al., 2015). The authors also noted that this likely affects criminal justice proceedings as well.

Research in the field of RMA is necessary for the purpose of educating the general public. It is important that studies are able to identify appropriate implicit and explicit measures of rape myths and also whether different myth subtypes are present within the population, if any. This would inform awareness, prevention, and intervention programmes as well as play a possible role in the criminal justice system.

Aim of the thesis

Despite increasing knowledge in the area of rape myth acceptance, there remains debate as to its utility as a dynamic risk factor (i.e., whether endorsement of RMA can be lowered by sex offender treatment programmes and whether it is related to risk of recidivism), its importance in societal issues, and how to appropriately measure RMA as a construct. The aim of this thesis is the draw together the research on RMA and its clinical utility, explore its role in forensic practice and in society, and to evaluate whether RMA can be validly and reliably measured with the possibility of it being separated into categories based in psychological theory.

Summary of chapters

Chapter 2 presents a systematic review of the literature on rape myth acceptance in convicted rapists. The review sought to establish whether rape myth acceptance is an important treatment need for adult male rapists. To address this, data from a myriad of research endeavours were amalgamated. This included literature which investigated changes in RMA due to intervention as well as research offering comparisons between sexual offenders and non-offenders. Studies making comparisons within the offending population and studies observing relationships between RMA and other criminogenic factors were also included in the review. The findings of the review are discussed in terms of the role RMA has in sexual offending, in particular within typologies of rapists. Implications for treatment and for future assessment of rape supportive attitudes are discussed. The amended version of this review has been published in Aggression and Violent Behavior (Johnson & Beech, 2017).

Chapter 3 offers a critique of the Illinois Rape Myth Acceptance Scale (IRMA; Payne et al., 1999), a measure created for the purpose of evaluating individuals' endorsement of rape myths. The IRMA was selected as it is a measure that has been found to be both reliable and valid and is, arguably, one of the most widely used measures of rape myth acceptance. The findings of the critique are discussed in relation to the IRMA's clinical use and utility in future research efforts. The results unveil that although there is some evidence of the reliability and validity of the IRMA as a measure, there appears to be little use of its updated version (McMahon & Farmer, 2011) in current research, which is concerning given that it is a tool which is temporally and culturally bound.

Chapter 4 presents an empirical research study which aims to establish reliability and validity of the IRMA using data from a sample of undergraduate students. A principal components analysis was conducted to assess the underlying factor structure of the IRMA. Further to this, an alternating least-squares algorithm (ALSCAL) was used to perform multidimensional scaling on the data to give a representation of the items in two dimensional

space. These results were interpreted considering current conceptualisations of rape supportive attitudes. The findings of the study are discussed in terms of future research utility of the IRMA and its use in clinical settings.

Chapter 5 draws together the reviews, research, and findings from the previous chapters to provide an overview of the role, importance, and psychometric measurement of RMA. Conclusions are drawn and implications for future research, social intervention, and clinical practice are discussed.

CHAPTER 2: RAPE MYTH ACCEPTANCE IN CONVICTED RAPISTS: A SYSTEMATIC REVIEW OF THE LITERATURE

Abstract

Aim: There is evidence to suggest that addressing rape myth acceptance is a relevant criminogenic need (i.e., a need directly related to an offender's likelihood of re-offending) for adult male rapists, but the research is mixed on this matter and a systematic review looking specifically at convicted offenders has yet to be carried out. This review examines studies on rape myth acceptance within populations of convicted sexual offenders and will review literature around changes in RMA due to interventions, comparisons made between sexual offenders and community controls, comparisons made within the offending population and relationships found between RMA and other psychological constructs linked to criminogenic need.

Method: General and specific searching strategies were carried out to gauge the need for the current review. The search strategy utilised three major search platforms, OvidSP, Web of Science, and Proquest; hand searching the reference lists of included studies; and contacting 35 experts in the field. Specific inclusion/exclusion and quality appraisal criteria were applied to each study.

Results: Eight studies met the inclusion criteria. Narrative data analyses highlighted that differences in subgroups of rapists were evident for different aspects of RMA; while rapists can be distinguished from non-offenders and non-sexual offenders on measures of RMA, they cannot be significantly discriminated from child molesters by relying on these measures; in regards to rapists and sexual murderers, the two groups could not be distinguished using

RMA scores; RMA was not found to be a significant predictor of sexual or violence recidivism; and significant positive change in RMA was reported after sex offenders completed treatment programmes.

Conclusions: Due to the nature of the research being sought in this review, a completely "randomised controlled trial" would be impossible to attain. So, it is worth noting that for future reviews, studies examining this construct would best not be marked as high risk based on the fact that they are conceptually different from randomised control studies. Differences in scores on RMA subscales among rapists' typologies were discovered, which may be indicative of the differences in beliefs of each of the typologies. If this is the case, then it is important that these differences be identified to develop specific treatment programmes to target these beliefs. Studying the power-sex dynamic as it is related to RMA may be beneficial in helping to understand the cognitive associations that sexually aggressive men have.

Keywords: rape myth acceptance, rapist typology, rapists, sex offending, offence-supportive attitudes.

Introduction

Sexual offending research is often heavily weighted towards the topic of child sexual abuse. Rape is underrepresented in the literature, resulting in limited knowledge and inefficient treatment. Often, sexual offenders will receive a generic treatment programme despite it being important to offer separate treatment for those that differ in their criminogenic needs (Reid, Wilson, & Boer, 2011). Helmus, Hanson, Babchishin and Mann (2013) note that holding cognitive distortions, specifically "attitudes supportive of sexual offending", is a risk factor that has predictive validity for sexual recidivism. Rape myth acceptance has been identified as one of these cognitive distortions and will be the topic of this review.

Rape myths and rape myth acceptance

Martha Burt first introduced and subsequently defined the concept of rape myths in 1980 as "prejudicial, stereotyped, or false beliefs about rape, rape victims, and rapists" (p. 217). In later years, Lonsway and Fitzgerald (1994) went on to expand on the definition, stating that rape myths are "attitudes and beliefs that are generally false but are widely and persistently held, and that serve to deny and justify male sexual aggression against women" (p. 134). For example, women "ask for rape" and rape is a result of the "uncontrollable" male sex drive (Payne, Lonsway, & Fitzgerald, 1999), shifting the blame for the crime towards the victim (Chapleau & Oswald, 2010; Gray, 2006). Rape myths influence attitudes towards victims on a social level. High levels of rape myth acceptance (RMA) are strongly associated with rape proclivity – one's likelihood or tendency to choose to rape (Chapleau & Oswald, 2010; Chiroro, Bohner, Viki, & Jarvis, 2004; Edwards, Turchik, Dardis, Reynolds, & Gidycz, 2011; Gray, 2006). Rape myths are thought to reduce the expectation of negative outcomes or consequences in sexual offenders (Chapleau & Oswald, 2010). There is evidence of RMA amongst convicted rapists, using myths to rationalise their behaviours

(Chiroro et al., 2004). Chiroro et al.'s findings suggest desire to exert power, but not the desire to obtain sex, mediates the relationship between RMA and rape proclivity. Rape myth acceptance has been a major topic in rape literature and research has identified the devastating impact of RMA across a variety of settings.

Measures of rape myth acceptance

There are a wide range of instruments designed to assess constructs related to rape myths. However, it should be noted that within the literature what defines a "rape myth" will vary across authors. Some experts state that the term "rape myth acceptance" is now interchangeable with "offence supportive attitudes" or "rape supportive attitudes" (C. Hermann, personal communication, May 4th, 2015; J. W. Van den Berg, personal communication, April 28th, 2015). Alternatively, these terms could be viewed, arguably more appropriately, as overarching terminology under which "rape myth acceptance" falls as a subcategory. The varied literature on the topic looks at rape attitudes, knowledge on rape, empathy towards rape and rape aversion (Lonsway & Fitzgerald, 1994).

Before the official introduction of the term "rape myths" by Burt in 1980, Feild (1978) developed the Attitudes Toward Rape Scale (ATR). In this study, Feild sought to investigate dimensionality of attitudes about rape, whether background characteristics of participants were related to their perceptions of rape, and if these participants differed in their attitudes toward rape. The research involved a diverse group of individuals who would have "different points of contact with rape" (p. 157). The subject group consisted of the following: 1. crisis counsellors who were likely to interact with rape victims and might have an understanding of victims' psychological response to rape; 2. police officers who were highlighted as often being the first people to meet with the victim following the offence and whose judgment is likely to affect whether a case is pursued; 3. citizens – Field noted that victims would need to interact with members of the community following an attack and these

citizens' perceptions of rape are likely to affect how a victim is able to readjust into their social surroundings after a rape; and 4. rapists who were included as their attitudes towards rape is likely to affect the decision to commit rape. The researcher found that counsellors differed from police, citizens, and rapists in the beliefs about rape, with citizens and the police being most similar. However, the scale failed to discriminate between rapists and police on approximately half of the attitudinal dimensions. As a result, many studies after this have chosen to utilise other tools for measuring rape myth acceptance or to pull aspects from the ATR and combine these with items that better discriminate rapists from non-offenders.

Arguably, the most widely used measure of rape myths is the Rape Myth Acceptance Scale developed by Burt (1980). The Rape Myth Acceptance Scale measures distorted beliefs around the sexual assault of adult women. This was the introductory measure for rape myth terminology. Research with the scale has found that men who are sexually aggressive toward adult women endorse more of these distorted beliefs about rape than do non-sexually aggressive men (Burt, 1980; Muehlenhard & Linton, 1987). Bumby (1996) noted that approximately a third of the scale's items do not specifically measure rape myths. Rather, he explained, the Rape Myth Acceptance Scale appeared to reveal how people's biases regarding age, race, and gender affect their likelihood of believing an allegation of rape.

Bumby (1996) felt that Burt's scale was highly susceptible to socially desirable responding and that there was weak evidence of its ability to discriminate between offenders and non-offenders. In response, he created the Bumby RAPE scale and found that it was able to discriminate between sex offenders and controls, but could not discriminate amongst sex offenders (i.e., separate rapists from child sex offenders). However, the RAPE scale has been discounted as well as a measure of rape myth acceptance and seen as a measure overall of sexual-assault-supportive attitudes (W. Murphy, personal communication, April 27th, 2015).

Also, building on Burt's scale, and attempting to enhance it, Payne, Lonsway, and Fitzgerald (1999) created the Illinois Rape Myth Acceptance Scale to assess myths about female victims of rape, male perpetrators, and rape as a violent crime by examining gender-role stereotyping, adversarial sexual and heterosexual beliefs, hostility towards women, and acceptance of interpersonal violence.

Many researchers have developed extended or modified versions of Burt's RMAS and others have developed scales that are conceptually similar (see Lonsway and Fitzgerald (1994) for a comprehensive list of measures relating to rape myth acceptance and rape-supportive attitudes).

The current review

There is evidence to suggest that addressing rape myth acceptance is a relevant treatment need for adult male rapists but the research is mixed on this matter and a systematic review looking specifically at convicted offenders has yet to be carried out. To understand sexually aggressive behaviours, it is critical to understand the cognitive associations of sexually aggressive men and it is important that this research be done with the criminally convicted. Studying rape proclivity, though beneficial, may lose the cognition inherent in a criminal that may not be present in members of the general population. Also, from a rehabilitative and reintegration standpoint, it is more appropriate to target those needing rehabilitation. Helmus et al. (2013) carried out a meta-analysis on offence-supportive attitudes as a risk factor in sexual offending as an update to Hanson and Morton-Bourgon's (2004) analysis. They looked at the role of cognition in sexual offending, however, they did not look into offence-specific justifications (e.g., rape myth acceptance in rapists) which is a gap in the literature that this review will attempt to fill. The review examines studies on rape myth acceptance within populations of convicted sexual offenders and will review literature

around changes in RMA due to interventions, comparisons made between sexual offenders and community controls, comparisons made within the offending population and relationships found between RMA and other psychological constructs linked to criminogenic need.

Aims and objectives

The aim of this systematic review was to explore rape myth acceptance as a criminogenic need for adult males who have committed sexual assaults against adult women.

The specific objectives of the review were:

- To determine if adult, male rapists can be distinguished from adult, male child molesters, non-sexual offenders, or non-offenders on measures on RMA
- To determine if there are differences in levels of RMA between different sub-groups of rapists, for example, those motivated by sex vs those motivated by anger
- To determine if difference in levels of RMA can discriminate between rapists who reoffend (recidivists) and those who do not
- To establish if RMA is responsive to sex offender treatment programmes

Method

Scoping exercise

Prior to the commencement of this review, several databases were searched to establish whether previous reviews of a similar or identical nature had already been carried out or planned. Searches were conducted on the 6th of April 2017 using the Centre for Reviews and Dissemination (DARE), the Campbell Corporation, the Cochrane Database of Systematic Reviews, and the International Prospective Register of Systematic Reviews

(PROSPERO). Through this search, no existing or planned reviews were identified, which acts as confirmation for the necessity of this review. Afterwards, a brief, preliminary scoping search was conducted in Embase, Ovid MEDLINE, PsycARTICLES, and PsycIFNO through the OvidSP database to establish the quantity and scope of information that was potentially available for this review and to check the practicality of the review question. Basic free text search terms were used to capture some of the ways RMA is referred to in the literature:

"sex* offen*" AND (rape myth* OR rape myth* accept*" OR rape support* attitud* OR rape support* belief*)

This basic search retrieved literature that was relevant to the review question, however, results were very limited around offenders convicted of rape (as compared to studies on rape proclivity). Still, the data that were found were sufficient enough to allow the continuation of a more thorough and comprehensive search.

Overview of search strategy

The search required for this review on the relevant research occurred in three stages. First, major electronic databases were searched. These included the OvidSP platform (within this the following databases were searched: Books@Ovid, CAB Abstracts, Embase, Embase Classic,
HMIC Health Management Information Consortium, Journals@Ovid Full Text,
Ovid MEDLINE® In-Process & Other Non-Indexed Citations and Ovid MEDLINE ®, PsycARTICLES Full Text, PsycINFO, and Social Policy and Practice); Web of Science; and Proquest. OvidSP, Web of Science and Proquest were searched on the 6th of April 2017. The search syntax used for each database can be found below. Second, the reference lists of the

full text articles – those which met the inclusion and exclusion criteria – and Helmus' (2013) meta-analysis were hand searched for relevant articles which could potentially be included in the review. Third, 35 recognised experts in the field of sex offender research and rape-supportive attitudes were contacted and queried about any relevant and pertinent studies (published or unpublished) that could be included. Twenty-one experts responded.

Search terms

Ten bibliographic databases searched (Books@Ovid, CAB Abstracts, Embase, Embase Classic,
HMIC Health Management Information Consortium, Journals@Ovid Full Text,
Ovid MEDLINE® In-Process & Other Non-Indexed Citations and Ovid MEDLINE ®,
PsycARTICLES Full Text, PsycINFO, and Social Policy and Practice) were accessed through the OvidSP platform. This search platform uses subject headings to index the databases' contents. Free text words were searched on OvidSP to be found in the title, abstract or main body of the articles available on the databases.

Rape myth acceptance is a concept that is described using varied terminology. In a similar sense, rapists are referred to in the literature in many different ways and these variations have developed and altered over time. To capture the different terms, several papers relating to rape myth acceptance and the assessment and treatment of rapists were examined. Additionally, experts were contacted for any alternative search terms. A list of key terms were drawn up from the experts and papers. Although "rapist" could be defined and alternative search terms generated fairly easily, an adjacency searching strategy was employed for phrases or terms related to the concept of RMA.

The following search terms were used along with the Boolean operators 'AND' (to combine the search concepts), 'OR' (to combine synonyms) and 'NOT' (to eliminate

particular terms) where necessary. Search terms and operators were modified to accommodate the different search conventions requisite for different databases and platforms.

rape myth* OR rape myth accept* OR cognit* distort* OR attitud* OR attitud* adj/3 towards women OR rape adj/3 support* attitude* OR victim* adj/2 blam* OR attribut* adj/3 blam*
AND
rapist* OR sex* offend* OR Convict* rapist* OR convict* sex* offend* OR incarcerate* sex* offend* OR incarcerate* rapist* OR sex* aggress* OR sex* molest* OR sex* assault*
NOT
rape propensity OR rape proclivity

Search syntax used in the OvidSP, Web of Science and Proquest platforms are attached in Appendix A. After the completion of each search, references were exported to RefWorks – a reference management system.

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

Of the above searches, 3,111 hits were returned from OvidSP, 438 from Web of Science, and 123 from Proquest. First, duplicate references were removed from OvidSP (*n*=1,062). Second, all titles, abstracts, and sources of the remaining articles in OvidSP (*n*=2,049), Web of Science and Proquest were screened. Those which were not relevant to the review in that they did not meet the inclusion/exclusion criteria outlined in the screening and selection tool (outlined below; the complete tool can be found in Appendix B) were removed. Third, full text copies were obtained for all citations that remained *(n*=15) and the inclusion and exclusion criteria in the screening and selection tool was applied to each paper with reasoning behind inclusion/exclusion recorded on the form. A list of the excluded studies

(*n*=13) and the reasons for their exclusion can be found in Appendix C. Fourth, the screening and selection tool was applied to the papers obtained from hand searching references, this returned 4 articles. Lastly, the screening and selection tool was applied to articles acquired from experts in the field; from this, four articles were obtained. Figure 1 shows a diagrammatic representation of this process.

The screening and selection tool

The Population, Intervention, Comparators, Outcome, Study design (PICOS) framework is arguably the most widely used tool adopted for use for defining systematic review questions, creating search terms and establishing inclusion and exclusion criteria in quantitative research (Higgins & Green, 2011). This review is predominantly exploratory and is not making a specific attempt to evaluate intervention efficacy and as such, some components of the PICOS framework (e.g., intervention or comparators) may not be relevant to this review. Cooke, Smith and Booth (2012) developed an alternative framework used for qualitative and mixed methods studies referred to as SPIDER (Sample, Phenomenon of Interest, Design, Evaluation, Research type). This review utilised aspects of both frameworks to best capture all angles of the review questions and these were incorporated into the screening and selection tool in Appendix B. A summary of the major inclusion and exclusion criteria used for the screening and selection tool is as follows:

Population:	Male, adult rapists (older than 18)
Phenomenon of Interest:	Rape myth acceptance as operationalised by Burt (1980) and extended by Lonsway and Fitzgerald (1999)

Comparison Group:	• Non-offenders (community controls) • Non-sexual offenders (e.g. violent offenders) • Other categories of sexual offender (e.g. child molester) • Recidivists and non-recidivists • Pre- and post-intervention
Outcome:	Comparison between male adult rapists and one of the aforementioned comparison groups on a specific, quantitative (numerical) measure of RMA
Research type/study design:	Any that is not solely qualitative in design, needs comparison group
Language:	English

Exclusion criteria consisted of the following:

1. Studies looking only at rape proclivity (i.e. utilising a non-clinical sample in which no one has been convicted of a rape)
2. Studies that did not utilise measures that were specifically used for the measurement of RMA
3. Studies that relied solely on qualitative measures

When applying inclusion/exclusion criteria, there were studies that utilised the Bumby RAPE scale as a measure of RMA. However, when discussed with experts, the decision was made to exclude these studies as experts could not come to an agreement on whether the Bumby RAPE scale *specifically* measures rape myth acceptance versus general rape-supportive attitudes (J. Abracen, personal communication, April 24th, 2015; K. Nunes,

personal communication, April 24th, 2015; W. Murphy, personal communication, April 27th, 2015).

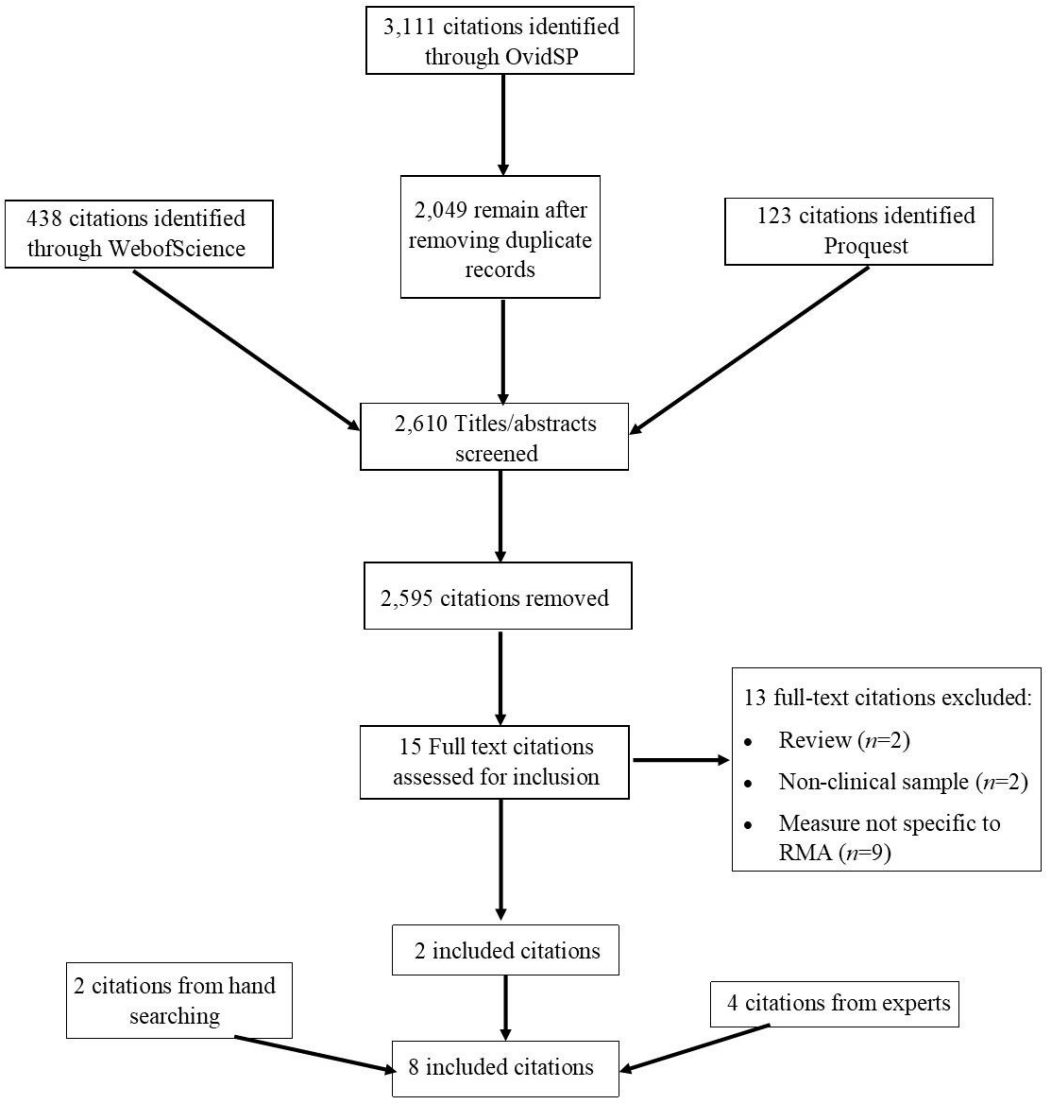

Figure 1. Flowchart of the study selection process.

Quality assessment

After application of the screening and selection tool, eight studies remained. The research design did vary quite a bit amongst the studies: three studies were of the before-and-after observational design; one was a case control (controlled observational) study; two were

cross-sectional (observational) studies; one study was a case series (observational); and the final study used a quasi-experimental design.

Due to the large variability in research study designs, quality assessment tools that cater to specific study designs were deemed inappropriate and would not provide the flexibility required to assess the methodological rigour of the studies in this review. Additionally, it would be impractical to utilise four separate quality appraisal tools and even though these tools offer great simplicity, the Cochrane Handbook explicitly discourages the use of these scales for assessing research quality. Often these scales give a summary score which involves having to allot 'weight' to different items on the scale, however, it is quite difficult to justify the weight distributions. Additionally, Higgins and Green (2011) mention that such scales or checklists have been shown to be unreliable in assessing validity and are not very transparent to assessors. As such, the Cochrane Collaborate recommends using a domain-based evaluation – which is neither a scale nor a checklist – that is used to make critical assessment separately for different domains (of bias; Higgins & Green, 2011). Consequently, the Cochrane Collaboration's tool for assessing risk of bias was used for this review. This tool can be found in Appendix D and supplementary information (criteria for rating and examples used as guides for raters) in Appendices E, F, G and H. Assessment of overall risk of bias for each study was agreed on through deliberation with each of the two quality assessors and were informed by the empirical evidence of bias, likely direction of bias, and likely magnitude of bias as guided by the Cochrane Collaboration's Handbook (Higgins & Green, 2011). The two quality assessors were both past researchers in the field of forensic psychology, with one currently in medical school and the other in law school. Both individuals have carried out systematic literature reviews previously. No "cut-off point" for exclusion was applied, due to the small number of studies (n=8) that were included in the review and the sparse amount of literature in this topic area in general.

Data extraction

A data extraction form was used for the purpose of extricating relevant information from each included study in a consistent manner (see Appendix I). General as well as specific information was extracted to include characteristics of the papers and outcomes relevant to the aims of this systematic literature review. The data extraction form focused on these key variables:

- General information (date form completed, person completing form, reference citation, study author contact details, and publication type)
- Study characteristics (methods, participants and other pertinent info, e.g. any conflicts of interest)
- Key conclusions

Study characteristics such as the type of study, type of intervention, comparison groups and outcome measures are all recorded in the screening and selection tool and are not repeated again in the data extraction tool.

Results

Overview of studies

Table 1 depicts a summary of the synthesised data for the 13 studies, allowing for evaluation of how rape myth acceptance is measured amongst rapists. The comprehensive information on each study along with their risk of bias forms can be found in Appendix J.

Methodological and study characteristics

There was some variability in the studies with regard to countries. Four countries were represented in the data: the United Kingdom, the United States of America, Canada and Israel. There was also variability in the dates of the studies with the oldest study being published 29 years ago (Overholser & Beck, 1986) and the most recent having been published in 2015 (Stefanska, Carter, Higgs, Bishopp, & Beech, 2015). Study design was also quite varied with three studies using a before-and-after observational design; one study was a case control (controlled observational) study; two utilised a cross-sectional (observational) design; one study was a case series (observational); and the final study used a quasi-experimental design.

Table 1. Characteristics and overall risk of bias scored for the 8 included studies

Author & year [Study ID]	Aims of study & design	Population	Intervention	Comparison/Control	Measure of RMA	Findings (in relation to RMA)	Overall risk of bias
Beech, Oliver, Fisher, and Beckett (2006) [Beech 2006]	Commissioned by Home Office and Her Majesty's Prison Service to evaluate prison and probation treatment services for sexual offenders; specifically, the CORE Sex Offender Treatment Programme (SOTP) and its appropriateness for use with rapists and sexual murderers. Used psychometrics to highlight criminogenic needs of sexual offenders as well as investigate the immediate or short-term effects of the treatment programmes. Observational (before-and-after).	112 rapists and 58 sexual murderers involved in the CORE SOTP from 1998-2002.	CORE SOTP	Pre- and post-intervention results	Burt's Rape Myth Acceptance Scale (RMAS; 1980) Cronbach's alpha: .88	Pre-treatment: no significant differences found between rapists and sexual murders re: RMA. Post-treatment: no differences found between rapists and sexual murderers and no overall change in RMA in rapists or in sexual murderers; within rapists typologies: groups differed significantly on their scores for the 'adversarial sexual beliefs' subscale of Burt's RMAS and the 'sexually motivated' offenders were found to score higher on the 'sex role stereoptyping' subscale than non-sexual violent offenders and community non-offending males.	High

Cohen (2012) [Cohen 2012]	Investigated the existence of the "uncontrollability" and "entitlement" schemata rapists and child molesters purportedly hold, as well as the schema of "sexy children" in child molesters using an emotional stroop task (ES) and lexical decision task (LD). Additionally, the author measures cognitive distortions using the Burt RMAS and the Hanson Sex Attitude Questionnaire and social desirability using the Marlowe-Crowne Social Desirability Scale (MCSDS). Controlled observational (case control).	44 sex offenders in community-based treatment programme in Israel	N/A	44 undergraduate students	Burt's RMAS	No difference in cognitive distortion levels between sex offenders and students, but social desirability was not a factor in this finding. Overall, found that cognitive distortions are present in both sex-offenders and non-offenders. However, in sex offenders they interact with lack of sex-role satisfaction, high trait anger and trait anxiety, lending them an emotional salience not present in non-offenders. Cognitive distortions manifest in high risk situations for offenders but not for non-offenders	High
Marshall and Hambley (1996) [Marshall 1996]	Examined the relationship among rapists of their responses to measures of loneliness, intimacy, rape myth acceptance, and hostility toward women. Observational (cross-sectional).	27 incarcerated male rapists	N/A	N/A	Burt's RMAS	All expected relationships between variables confirmed through correlational analyses. Results suggest rape is a function of hostility toward women combined with the acceptance of rape myths, which are also related to intimacy and loneliness deficits among sex offenders. It was found that the link with intimacy was stronger than the link with loneliness in this group of offenders.	Low

Olver, Nicholaichuk, and Wong (2014) [Olver 2014]	Examined sex offenders' risk and treatment change based on a battery of psychometric assessment measures followed up an average 18 years post-release. Observational (case series).	276 federal sex offenders	Clearwater Programme (High intensity Sex Offender Treatment Programme)	Pre- and post-intervention; with follow up	Burt's RMAS	The mean Rape Myth Acceptance (RMA) score was approximately one full standard deviation below the normative mean for both offenders and non-offenders (Burt, 1980) at pre-treatment, and approximately two-thirds of a standard deviation lower at post treatment. There was a significant decrease in rape myths endorsed within the sample from pre- to post-treatment.	High
Overholser and Beck (1986) [Olverholser 1986]	Assessed rapists, child molesters, and three control groups on five potentially relevant variables: heterosocial skills, social anxiety, hostility, impulsivity, and attitudinal variables. Quasi-experimental.	12 rapists 12 child molesters	N/A	1. 12 prisoners who were non-sex offenders 2. group of 12 community-based low socioeconomic status (SES) men 3. group of 12 "minimal-dater" college students	Burt's RMAS	No significant effect was found on the Rape Myth Acceptance Scale. A significant main effect for group was found for the Sex Role Stereotyping scale, $F(4, 55) = 4.00$, $p < .01$. A Newman-Keuls analysis indicated that child molesters displayed significantly higher levels (more conservative) of sex role stereotyping than did both the community-based low-SES men and the minimal-dater college students.	High

Pithers (1994) [Pithers 1994]	Carried out a process evaluation to assess the extent to which a specialised treatment group might enhance the offenders' empathy for sexual abuse survivors. Observational (before-and-after)	20 convicted males: 10 paedophiles, 10 rapists	Survivor empathy group (as part of Vermont Treatment Programme for Sexual Aggressors; Pithers, Martin & Cumming, 1989)	Pre- and post-intervention	Burt's RMAS	Results suggest the intervention results in decreased endorsement of cognitive distortions predisposing rape. Paedophiles and rapists did not differ in pre-treatment or post-treatment endorsement of cognitive distortions hypothetically related to or rape. Scores on Burt's Rape Myth Acceptance Scale, which would be expected to reveal deficits in the rapists, did not discriminate these samples of child abusers and rapists.	High
Stefanska, Carter, Higgs, Bishopp, and Beech (2015) [Stefanska 2015]	Aimed to examine pathways to sexual killing while also separating sexual killers on the basis of whether or not they had a previous conviction for rape. Consideration was then given to whether the pathways to offending differ based on this distinction. Observational (cross-sectional).	150 sexual murderers	N/A	N/A	The Rape Myths Scale (Offending Behaviour Programmes Unit, 1995)	Rape myths were not analysed in isolation. However, upon reading the results tables, it is shown that of the men in the *high problem* group (offenders who were likely to report high levels of sexual entitlement beliefs, rape myths, have problems with being open to others and tend to believe that women are deceitful) 35% ($p<.001$) were found to endorse rape myths, whereas in the *low problem* group (those who did not report problems in the aforementioned areas), 13% ($p<.001$) endorsed rape myths.	Unclear

Webster, Akhtar, Bowers, Mann, Rallings, Marshall (2004) [Webster 2004]	Explored the impact of the Prison Service CORE Sex Offender Treatment Programme (SOTP) upon Black sexual offenders as compared with White sexual offenders. Observational (before-and-after).	52 Black sexual offenders	CORE SOTP	52 White sexual offenders Pre- and post-intervention	Burt's RMAS	Groups significantly improved on the RMAS post-treatment. There were no other significant within-subjects main effects or interactions. *It should be noted that when reviewing the statistics it would appear that child molesters had significant change in their RMAS post-treatment as compared to the rapists that seem to exhibit very little/no change at all. Additionally, an interaction seems to be apparent here where the White rapists experienced a positive change in RMAS post treatment. As this was not the main focus of the study, the in-depth statistics for these were not reported.	Low

Participants and recruitment

Participant groups ranged in mean age from 31.22 to 35.38 years. However, it should be noted that this range in mean ages was from within the same study (Webster et al., 2004) and does not include the mean age of 25.87 in Stefanska et al.'s (2015) study as this mean was an average of ages taken at the time of the offenders' index offence whereas the other studies reported the mean age at the time of assessment. Sample sizes were quite reasonable for the majority of studies with the smallest being reported in Pithers (1994) where process evaluation of a specialised treatment programme was carried out with 10 rapists and 10 paedophiles. The study with the largest sample size was of a case series design (Olver, Nicholaichuk, & Wong, 2014) and examined risk and treatment change in 276 federal sex offenders. The locations from where participants were recruited varied with individuals drawn from prisons, psychiatric facilities and treatment centres, community-based programmes and probation departments. Control groups, where applicable, were recruited from universities and community-based organisations. Due to the design of a few studies (e.g. case series) an active "recruitment" process was not required. In these cases, researchers gathered information on participants from national databases and criminal justice records.

Study focus and aims

There were no studies included in this review that had rape myth acceptance as the only variable to be examined; however, one study did have RMA as one of its main variables under investigation and explored the relationship RMA had to loneliness, intimacy and hostility toward women among rapists (Marshall & Hambley, 1996). Four studies included an RMA measure in a battery of psychometric assessment measures to evaluate risk and change in offenders as a means

of some form of process or treatment programme evaluation (Beech, Oliver, Fisher, & Beckett, 2006; Olver et al., 2014; Pithers, 1994; Webster et al., 2004). In the remaining three studies, rape myth acceptance fell under the category of some all-encompassing attitudinal variable that was just one factor amongst many under investigation.

The main aim of half of the studies included in this review (*n*=4) was to evaluate the impact of treatment on sexual offenders utilising a set of psychometrics as pre- and post-intervention measures of risk and change. Two of these studies were evaluating the CORE Sex Offender Treatment Programme in the UK (Beech et al., 2006; Webster et al., 2004), however, in addition to pre- and post-treatment scores, Webster and colleagues were focusing particularly on any differences that occurred across ethnic backgrounds (i.e. Black sexual offenders versus White sexual offenders); one study evaluated the "Clearwater Programme" a high intensity sex offender treatment programme run in a maximum-security forensic psychiatric facility in Saskatoon, Saskatchewan, Canada (Olver et al., 2014); and Pithers (1994) carried out a process evaluation of the Vermont Treatment Programme for Sexual Aggressors, focusing specifically on the survivor empathy group (a group of offenders exposed to the experiences of sexual abuse survivors). Two studies utilised control groups – a non-offending population (Cohen, 2012; Overholser and Beck, 1986). Cohen (2012) compared 44 sex offenders in a community-based treatment programme to 44 undergraduate student to explore, firstly, whether the "uncontrollability" and "entitlement" schemata rapists and child molesters purportedly hold existed, as well as the schema of "sexy children" in child molesters; and secondly, whether these cognitive distortions were absent in the control group or no difference existed between sex offenders and controls. Overholser & Beck (1986), on the other hand, wanted to investigate whether heterosocial skills, social anxiety, hostility, impulsivity and attitudinal variables differed

between sex offenders and non-sex-offenders as well as community controls. It should be noted, however, that Beech et al. (2006) carried out a post-hoc comparison between the main study sample and a sample of non-violent offenders and community non-offending males (data for the comparator sample was gathered from an older study: Beazley Richards, 2000). This data will be discussed in the narrative data synthesis. The final two studies carried out cross-sectional investigations. Marshall and Hambley (1996) looked at a single group of incarcerated rapists exploring their responses to measures of loneliness, intimacy, rape myth acceptance and hostility toward women and the relationship between these. Stefanska and her colleagues (2015) explored the pathways to offending in sexual killers and whether there was a distinction between those who had a previous conviction of rape and those who did not.

Measures of RMA

There was virtually no variability in terms of the type of measures used to evaluate RMA. In fact, all studies used Burt's Rape Myth Acceptance Scale (1980) to measure RMA, with the exception of one study (Stefanksa et al., 2015). Stefanska and colleagues utilised the Rape Myths Scale developed by the Offending Behaviour Programmes Unit (1995) in the National Offender Management Service (now known as Her Majesty's Prison and Probation Service).

Burt's Rape Myths Acceptance Scale is a 19-item self-report measure that assesses the extent to which respondents endorse false beliefs about the rape of adult women that tend to externalise blame. The scale includes 11 items which are related to victim blaming and justification for rape and eight additional items that relate to false accusations and the likelihood that the respondent believes individuals' claims of rape. Each item is scored on a 7-point Likert scale ranging from 1 = "strongly disagree" to 7 = "strongly agree". A higher score is an

indication of a greater acceptance of rape myths. Burt (1980) reported initial validation studies in a sample of 598 US adults as a Cronbach's alpha of .88 and item-to-item correlations of between .27 and .62. However, research on the scale's discriminant validity or the effect of social desirability on the scale is still mixed (Bumby, 1996; Lonsway & Fitzgerald, 1995). It should be mentioned that Beech et al. (2006) report using a 23-item version of this scale. It is also worth noting that in Cohen's (2012) study, an extended version of Burt's Rape Myth Acceptance Scale was utilised which was not used in any of the other studies (or not specified). This version contains 55 items scored on the same 7-point Likert scale as the 19-item and 23-item versions. Cohen provides details of this extended scale, stating that the inventory includes six subscales relating to sexual behaviour: Adversarial Sexual Beliefs, Sex Role Satisfaction, Rape Myth Acceptance, Sex Role Stereotype, Sexual Conservatism and Acceptance of Interpersonal Violence. The Adversarial Sexual Beliefs subscale refers to the core belief that sexual relationships are fundamentally exploitative and contains nine items. Sex Role Satisfaction refers to "familial, work, and interpersonal role elements relevant to sex role stereotyping" (Burt, 1980, p. 219). The scale contains 10 items, asking the respondent to rate how satisfied they are with their "competence and skillfulness", "amount of socializing", etc. The Rape Myth Acceptance scale contains the first 11 items as in the 19-item scale which target false beliefs about rape, rapists, and rape victims. The Sex Role Stereotype scale contains nine items which reflect the respondent's endorsement of stereotyped sex roles for men and women. The Sexual Conservatism scale contains 10 items which reflect "restrictions on the appropriateness of sexual partners, sexual acts, conditions or circumstances under which sex should occur" (Burt, 1980, p. 219). The final scale, Acceptance of Interpersonal Violence, contains six items which refer to the

notion that force and coercion are legitimate ways to gain compliance, especially in sexual relationships (Cohen, 2012, p.50).

Not much could be uncovered for this review on the properties of the Operational Services and Intervention Group's Rape Myths Scale as it appears to be integrated into Her Majesty's Prison Service's Sex Offender Treatment Programme psychometric battery which remains unpublished and has now been withdrawn. Stefanska et al. (2015) offer a brief outline. The Rape Myths Scale is a 17-item measuring externalisation around rape and a higher score indicates a greater acceptance of justifications for rape. Good internal consistency ($\alpha = .83$) and test re-test reliability ($r = .85$) were reported.

Risk of bias ratings of included studies

The risk of bias ratings across studies included in the review varied extensively. Two studies had what was deemed the least amount of bias, which can be judged to equate to a higher quality, with overall ratings of "low risk" of bias (Marshall & Hambley, 1996; Webster et al., 2004). On the other hand, five studies were rated as being at an overall "high risk" of bias (Beech et al., 2006; Cohen, 2012; Olver et al., 2014; Overholser & Beck, 1986; Pithers, 1994) for the following reasons: incomplete outcome data (Beech et al., 2006; Cohen, 2012; Olver et al., 2014; Overholser & Beck, 1986; Pithers, 1994); no blinding of the participants and/or personnel (Beech et al., 2006; Cohen, 2012; Olver et al., 2014; Overholser & Beck, 1986; Pithers, 1994); no blinding of the outcome assessment (Beech et al., 2006; Cohen, 2012; Olver et al., 2014; Pithers, 1994); lack of random sequence generation (Cohen, 2012; Overholser & Beck, 1986); lack of allocation concealment (Overholser & Beck, 1986); selective reporting (Pithers, 1994); and other biases (detailed in Appendix J; Cohen et al., 2014, Olver et al., 2014; Pithers, 1994)

Narrative data synthesis and key findings

The data extracted from the studies varied in aims, methodology, and participant group and consequently, the results of the studies are not homogenous. As such, it was deemed inappropriate to combine the results and evaluate them as part of a meta-analysis. Alternatively, a narrative data synthesis will be carried out for this review, highlighting key findings of the studies in relation to the aims outlined in the introduction.

Can adult, male rapists be distinguished from adult, male child molesters, non-sexual offenders, or non-offenders on measures on RMA?

Two studies in this review explicitly compared rapists with non-rapists. Both studies employed Burt's Rape Myth Acceptance Scale as a measure of RMA with Cohen (2012) utilising the 55-item scale. Overholser and Beck (1986) did not specify the number of items in the scale used, however, the results reported an effect on the "Sex Role Stereotyping Scale" (p. 686) and it may be reasonable to assume that the researchers used a similar scale to Cohen.

Cohen (2012) reported that for most measures of offence-related cognition non-offenders scored similarly to sexual offenders. However, on the measure of rape myth acceptance, sex offenders reported a significantly *lower* level of rape myth acceptance than non-offenders. They found that this response could not be attributed to social desirability, at least not in that particular study, any more than could be attributed to social desirability in non-offenders. Overall, Cohen (2012) found that cognitive distortions were present in both sex offenders and non-offenders. However, in sex offenders they found that these distortions interact with lack of sex-role satisfaction (a subscale of the Burt Rape Myth Acceptance Scale), high trait anger and trait

anxiety, lending them an emotional salience not present in non-offenders. The study found that cognitive distortions manifest in high risk situations for offenders but not for non-offenders.

Overholser and Beck (1986) did not find any significant effect on Burt's Rape Myth Acceptance Scale, however they did find a significant main effect for group on the Sex Role Stereotyping subscale. Further analysis showed that child molesters displayed significantly higher levels of sex role stereotyping (i.e. more conservative) than did both the community control groups, however, the rapists did not appear to be distinguishable from the child molesters.

Although a comparison group was not included in the main study, Beech et al. (2006) compared their results to those of a similar study by Beazley Richards (2000) whose sample consisted of UK male non-offender employees of a civil engineering company and UK male non-sexual violent offenders. Beech and colleagues found that *sexually motivated* offenders (those who were primarily motivated to have sex and have used some form of force or violence against the victim to achieve this aim) scored higher on the Sex Role Stereotyping subscale than non-sexual violent offenders and community non-offending males.

Even though comparison between rapists and non-rapists was not the focus of the study, Pithers (1994) reported that Burt's Rape Myth Acceptance Scale did not discriminate between child abusers and rapists.

Overall, the results of these studies suggest that rapists may be distinguished from other non-sexual offenders and from community non-offending males on measures of rape myth acceptance, particularly on the sex-role subscales of the Burt Rape Myth Acceptance Scale. It could be that these subscales are more sensitive to differences that separate sexual offenders from non-sexual offenders and non-offenders. However, whether these results are reliable or

consistent may be called into question. Furthermore, there still appears to be difficulty in discriminating between child molesters and rapists when relying on rape myth acceptance as the distinguishing factor. Also, an unexpected result was noted in the Cohen study in which sex offenders reported *lower* acceptance of rape myths than non-offenders.

Are there differences in levels of RMA between different sub-groups of rapists?

Two studies examined sub-groups of rapists. Beech et al. (2006) carried out typological comparisons by categorising rapists into one of five main types as described by Knight and Prentky (1990): opportunistic, pervasively angry, vindictive, sexual non-sadistic and sexual sadistic rapists. The researchers found that groups differed significantly on their scores for the Adversarial Sexual Beliefs subscale of Burt's Rape Myth Acceptance Scale. Post-hoc analyses revealed that opportunistic rapists scored significantly higher than both the sexual sadistic and sexual non-sadistic types. Difference in scores between the sexual sadistic and non-sadistic types were not significant. Overall, the sexual non-sadistic rapists had the lowest scores on the scale. This finding may likely reflect that the sexual non-sadistic rapists hold less negative views about sexual relationships when compared to the other groups as Beech and colleagues found no correlation between measures of socially desirable responding and this subscale. Conversely, according to Burt (1980) the opportunistic rapists' scores revealed beliefs that sexual partners are manipulative, cunning and not to be trusted. Vindictive rapists had the second largest mean scores on the Adversarial Sexual Beliefs subscale. Beech and colleagues (2006) do note, however, that despite these findings, the highest mean score overall ('opportunistic' sub-types: Mean = 22.5) was only 0.4 of a standard deviation above the mean of 20.6 of a non-offender sample (from Beazley Richards, 2000). On the Acceptance of Interpersonal Violence subscale,

the opportunistic rapists scored significantly higher than both the sexual sadistic and non-sadistic types. This time, the opportunistic rapists' mean score was over two-and-a-half standard deviations above the non-offending sample's mean score and 1.3 of a standard deviation above the mean of a sample of incarcerated non-sexual violent offenders. The researchers offer an explanation for this finding, that this scale reflects the notion that coercion and force are legitimate modes through which to gain compliance in intimate and sexual relationships, versus relationships in general.

Although not explicitly defined as a "subgroup" of rapists, sexual murderers will be included in this section of the review. In regards to the sexual murderers, Beech et al. (2006) found no difference in RMA between rapists and sexual murderers.

Stefanska et al.'s (2015) study on sexual murderers' pathways to offending explored potential differences between rapists who were also sexual killers and sexual killers who had never solely committed rape (i.e. without killing the victim). It was difficult to gather data from this study as rape myths were not analysed in isolation. However, upon reading the results tables, it is shown that of the men in the "high problem" group (offenders who were likely to report high levels of sexual entitlement beliefs, rape myths, have problems with being open to others and tend to believe that women are deceitful) 35% ($p<.001$) were found to endorse rape myths, whereas in the low problem group (those who did not report problems in the aforementioned areas), 13% ($p<.001$) endorsed rape myths. However, this data combines both groups of sexual murderers so it is impossible to extract and separate the data to allow for comparison between the two groups.

Can differences in levels of RMA discriminate between rapists who reoffend (recidivists) and those who do not?

Two studies examined recidivism outcomes (Beech et al., 2006; Olver et al., 2014), however, in the Beech study, recidivism was not explored in terms of RMA. In the Olver study, RMA was not found to be a significant predictor of sexual or violent recidivism nor did the study compare recidivists with non-recidivists.

Is RMA amenable to sex offender treatment programmes?

Four studies examined the effect of treatment on rape myth acceptance. Beech et al. (2006) found no main effect of treatment in sexual murderers and no effect in rapists overall. However, when looking at the impact of treatment on typologies, Beech et al. (2006) grouped the Knight and Prentky (1990) typologies into three groups: sexually motivated (opportunistic and sexual non-sadistic rapists), anger motivated (vindictive and pervasively angry rapists) and sexual sadistic rapists remained a standalone "sadistic" type. Analysis revealed a significant change in scores overall on the Sex Role Stereotyping scale showing scoring actually increasing post-treatment. This is indicative of greater endorsement of stereotypical beliefs. There were no differences found between typologies. A result like this could indicate something inherent in the programme that would change these scores for the worse, however it should be noted that Beech et al. (2006) mentioned that quite a few offenders were removed from this sample as they could not be grouped into the categories.

Olver and colleagues (2014) reported that the mean RMA score was approximately one full standard deviation below the normative mean for both offenders and non-offenders (as reported in Burt, 1980) at pre-treatment, and approximately two-thirds of a standard deviation

lower at post-treatment. There was a significant decrease in rape myths endorsed within the sample from pre- to post-treatment which would suggest that the Clearwater programme has the capabilities to effect positive change in cognitive distortions around rape. Similarly, Pithers (1994) reported a significant treatment effect and found that there was a reduction in acceptance of rape myths after completion of the programme. This, perhaps, points to some effectiveness of victim empathy programmes for use with rapists.

Although the major outcome for Webster et al.'s (2004) study was differences in psychometric data between Black versus White sexual offenders, the researchers did investigate changes in RMA and found that both groups significantly improved on rape myths. However, these data look at child molesters and rapists combined. Also, it should be noted that when reviewing the statistics it would appear that child molesters had significant, positive change in their rape myths post-treatment as compared to the rapists that seem to exhibit very little/no change at all. Additionally, an interaction seems to be apparent here where the White rapists experienced a positive change in rape myths post treatment. As this was not the main focus of the study, the detailed statistics for these were not reported.

Discussion and Conclusions

Main findings of the review

This systematic review explored the relevance of rape myth acceptance as a treatment need for rapists. Overall, the literature on rape myth acceptance is quite vast, however, it was suprising how few studies were found that addressed this in rapists (n=8) versus the general public. Only studies which identified specific measures of RMA were included in this review, which led to the exclusion of studies solely using the Bumby Rape Scale as there was

disagreement amongst experts as to its use as a measure of RMA. Also included were studies which had convicted rapists as participants. Studies focusing on rape proclivity with non-offending samples only were excluded. Being quite strict with measures of rape myth acceptance may have biased this review in a way as seven out of the eight studies included utilised the Burt Rape Myth Acceptance scale. Perhaps broadening the definition could allow for the inclusion of more studies. It may be worth mentioning that Lonsway and Fitzgerald's (1999) Illinois Rape Myth Acceptance Scale would have been suitable for this review, however, the author could not locate or gain access to studies which examined its use that met the inclusion and exclusion criteria. However, the search carried out was extensive. The search was conducted across three platforms, OvidSP, Web of Science, and Proquest and also included hand searching of the reference lists of included studies and a meta-analysis, and contacting a large number of experts in the field for any published or grey (unpublished) literature. The response from experts was very good and four additional papers were garnered from this. Overall, there is confidence that most relevant research has been included in this review and that the consequent conclusions are from the synthesis of a solid evidence base.

The results indicate that while rapists can be distinguished from non-offenders and non-sexual offenders on measures of RMA, they cannot be significantly discriminated from child molesters by relying on these measures. Also, in regards to rapists and sexual murders, the two groups could not be distinguished using RMA scores. In analysing differences that were found, Cohen (2012) had results in the opposite direction from what is to be expected and reported that sexual offenders scored lower on rape myth acceptance than non-offenders. In terms of differentiating between sub-groups of rapists, Beech et al.'s (2006) findings were quite enlightening. The opportunistic rapists scored higher on the Adversarial Sexual Beliefs subscale

with the sexual non-sadistic rapists scoring the lowest. Additionally, the opportunistic rapists scored higher on the Acceptance of Interpersonal Violence subscale. Knight and Prentky (1990) posited that the opportunistic rapist views violence as an instrument to be used if needed in order to succeed in a sexual attack. They are described as taking advantage of an opportune situation and are indifferent to any impact this may have on the victim. Beech et al.'s (2006) study is congruent with this assertion. Studies did not compare recidivists and nonrecidivists, nor was RMA found to be a significant predictor of sexual or violence recidivism (Stefanska et al., 2015). RMA did, however, appear to be affected by sex offender treatment programmes in a positive manner. Studies reported significant positive change in RMA after sex offenders completed a treatment programme.

Strengths and weaknesses of the review

The major weakness of this review has been mentioned above, namely the limitation on measures included in the review. Additionally, the assessment of quality guidelines used to judge the literature was quite strict and as such some of the studies reviewed may be deemed 'poor quality'; this is discussed further below. However, this review employed a comprehensive research strategy guided by the advice of experts within the field. Additionally, new information around the applicability of RMA in rapist typologies has been introduced and long held assumptions about RMA in sexual offenders versus in the general public have been challenged in this review. The publication for this review can be found in Appendix K. The implications for future directions are outlined in the following section.

Implications for practice and future direction

The review adopts the Cochrane principles of systematic review, however, the randomised control trial is championed by the Cochrane Collaboration as the belief is held that other types of trial evidence are likely to inflate the positive findings for the intervention (Sackett, Rosenberg, Gray, Haynes & Richardson, 1996). Due to the nature of the research being sought in this review, a completely "randomised controlled trial" would be impossible to attain. Even far more robust reviews struggle with adhering to Cochrane principles for judging sex offender treatment. The Hanson, Bourgon, Helmus and Hodgson (2009) review of 129 sex offender treatment studies could rate none as 'strong' according to Collaborative Outcome Data Committee guidelines. For this reason multiple quality assessors (two in addition to the author) would come to an informed agreement about the risk of bias present in studies. As this review is interested in the relevance of an attitudinal variable rather than treatment efficacy or effectiveness as the primary outcome, attaining the "gold standard" as determined by Cochrane is unnecessary. So, it is worth noting that for future reviews, studies examining this construct would best not be marked as high risk of bias solely because they are methodologically different from randomised control studies.

Out of the eight studies in this review, the most comprehensive was Beech et al.'s (2006) as it was a part of a large-scale project with the Prison Service. They found differences in scores on RMA subscales among rapists' typologies, which may be indicative of the differences in beliefs of each of the typologies. If this is the case, then it is important that these differences be identified to develop specific treatment programmes to target these beliefs. Sex Role Stereotyping was a subscale on which sexually motivated offenders scored highly. It could be possible that this stereotyping is linked to the feeling of one's entitlement to male dominance or power (Eyssel & Bohner, 2011). Additionally, it could possibly be that power is a necessary

component for sexual attraction in sexual offenders, particularly rapists (Chapleau & Oswald, 2010). It may follow, then, that maintaining and strengthening such an association could lead to misinterpreting innocent social interactions and dire consequences. As such, studying the power-sex dynamic as it is related to RMA may be beneficial in helping to understand the cognitive associations that sexually aggressive men have. It is evident that future work needs to target differentiating amongst sexual offenders and utilising implicit measures to measure these associations may be a means of overcoming the transparency of using explicit measures alone.

The advantages and disadvantages of explicit measures will be addressed throughout this report, with specific focus on Payne et al.'s (1999) IRMA.

CHAPTER 3: THE ILLINOIS RAPE MYTH ACCEPTANCE SCALE (IRMA): A PSYCHOMETRIC CRITIQUE

Introduction

Sexual violence has devastating effects on society and as a result, researchers have endeavoured to understand what underlies sexually aggressive behaviour. One construct which influences the perpetration of sexual assault is the acceptance of rape myths (Chiroro, Bohner, Viki, & Jarvis, 2004). Researchers in the fields of sociology and feminism first took note of the concept of rape myths in the seventies (e.g., Brownmiller, 1975; Schwendinger & Schwendinger, 1974). They highlighted that one could often find male myths about rape appearing in literature, being quoted by individuals who were deemed "experts", and in comedy. These researchers noted that these "rape myths" served to perpetuate male sexual aggression towards women and blame the victim. In 1980, Martha Burt first scientifically examined the construct of rape myth acceptance (RMA) and, with this, offered the definition of rape myths as "prejudicial, stereotyped, or false beliefs about rape, rape victims, and rapists" (p. 217). Burt's definition emphasised that endorsement of rape myths allows for minimisation and justification of sexual crimes and victim blaming. Some years later, Lonsway and Fitzgerald (1994) expanded upon this definition, defining rape myths as "attitudes and beliefs that are generally false but are widely and persistently held, and that serve to deny and justify male sexual aggression against women" (p. 134). Examples of commonly endorsed myths include: most women lie about rape, women "ask for rape", and rape is a result of the "uncontrollable" male sex drive (Edwards, Turchik, Dardis, Reynolds, & Gidycz, 2011; Payne, Lonsway, & Fitzgerald, 1999). A critical discovery in the research on RMA revealed that higher levels of rape myth acceptance in men is strongly associated with rape proclivity, that is, one's likelihood or tendency to choose to rape (Chapleau

& Oswald, 2010; Chiroro et al., 2004; Edwards et al., 2011; Gray, 2006). More importantly, research findings highlight that a high endorsement of rape myths is a causal antecedent for the actual perpetration of sexual violence or rape proclivity (Bohner, Jarvis, Eyssel, & Siebler, 2005; Hinck & Thomas, 1999). RMA has been a significant topic in rape literature, and research has identified the devastating impact of RMA across a variety of settings and social contexts.

Historical and Current Measures of RMA

There are a wide range of instruments designed to assess constructs related to rape myths. Before Burt first introduced the term "rape myths" in 1980, Feild (1978) developed the Attitudes Toward Rape Scale (ATR). This was the first instrument developed to measure attitudes about the act of rape, victims of rape, and perpetrators of rape. Feild proposed that rape attitudes were multidimensional and identified eight factors in his scale: *women's responsibility in rape prevention*; *sex as motivation for rape*; *severe punishment for the rape*; *victim precipitation of rape*; *normality of rapists*; *power as motivation for rape*; *favorable perception of a woman after rape*; and *resistance as woman's role during rape*. In his research, Feild found that counsellors differed from police, citizens, and rapists in their beliefs about rape, with citizens and the police being most similar. However, the scale failed to discriminate between rapists and police on approximately half of the attitudinal factors. Following this, researchers investigating rape myths have chosen to implement other tools for measuring rape myth acceptance or to pull aspects from the ATR and combine these with items that better discriminate rapists from non-offenders.

Arguably, the most widely used measure of rape myths is the Rape Myth Acceptance Scale (RMAS) developed by Burt (1980). The RMAS measures distorted beliefs around the sexual assault of adult women, i.e. individual levels of rape myth endorsement. This was the

introductory measure for rape myth terminology. Research with the scale has found that men who are sexually aggressive toward adult women endorse more of these distorted beliefs about rape than do non-sexually aggressive men (Burt, 1980; Muehlenhard & Linton, 1987). However, Bumby (1996) noted that approximately a third of the scale's items do not specifically measure rape myths. Rather, he explained, the RMAS was highly susceptible to socially desirable responding and that there was weak evidence of its ability to discriminate between offenders and non-offenders. Consequently, he created the Bumby RAPE scale and found that it was able to discriminate between sex offenders and controls, but could not discriminate amongst sex offenders (i.e., separate rapists from child sex offenders). Furthermore, the RAPE scale itself has been discounted as a measure of rape myth acceptance and seen as a measure overall of sexual-assault-supportive attitudes (W. Murphy, personal communication, April 27th, 2015) while the RMAS currently remains in circulation.

One issue in the research of RMA is the tendency for authors to vary in their terminology and phrasing for slightly different aspects of these attitudes. This complicates discussion in the research as some address these beliefs as myths, others as cognitive distortions (Abel, Becker, & Cunningham-Rathner, 1984; Abel, Becker, & Skinner, 1987) and still others as implicit theories (Polaschek & Ward, 2002). This makes it difficult to draw comparisons between tests as no two tests are actually measuring the same constructs. Furthermore, over time, criticism about existing measures of rape myth acceptance began to increase. In 1994, Lonsway and Fitzgerald reviewed 24 RMA measures, including Burt's RMAS and found that the measures lacked psychometric precision. They noted problematic aspects such as use of colloquial phrases (for example, "fair game" and "necking"), lack of clarity, problems with item format (for example, questions assessing knowledge or rape statistics rather than information related to rape myths), a high

likelihood of response bias due to impression management, and the fact that the RMAS failed to discriminate between offenders and non-offenders. Because of this, Lonsway and Fitgerald (1994, 1995) set out to reconceptualise and redefine rape myth as a construct. They offered the definition as: “attitudes and beliefs that are generally false, but are widely and persistently held, and that serve to deny and justify male sexual aggression against women” (Lonsway & Fitzgerald, 1994, p. 134). In an attempt to enhance Burt’s RMAS, Payne, Lonsway, and Fitzgerald (1999) created the Illinois Rape Myth Acceptance Scale (IRMA) to assess myths about female victims of rape, male perpetrators, and rape as a violent crime by examining gender-role stereotyping, adversarial sexual and heterosexual beliefs, hostility towards women, and acceptance of interpersonal violence.

Recently, a new measure has been developed to explore a broader definition of rape myth acceptance, the Acceptance of Modern Myths about Sexual Aggression (AMMSA) scale (Gerger, Kley, Bohner, & Siebler, 2007). Gerger and colleagues noted that not only were most RMA measures over a decade old, they noted that traditional measures used explicit item wording to assess attitudes about sexual assault. Most often, individuals’ views around sexism and topics related to sexism (such as RMA) tend to be more covert than in the past. Thus, the researchers endeavoured to create a scale that reflected these more subtle attitudes around rape, victims, and perpetrators: the AMMSA. One of the criticisms that Gerger et al. (2007) made about the IRMA and its predecessors was that participants are able to recognise explicit items regarding attitudes toward sexual assault and engage in socially desirable responding. Additionally, the researchers wanted to offer a broader scope than the IRMA in their RMA measure. As a result, the AMMSA scale examines denial of rape as a problem in society, views that blame the victim and exonerate the perpetrator, attitudes about male coercion as a natural

part of sexual encounters, and beliefs about public policies designed to address sexual assault (Gerger et al., 2007). The AMMSA scale incorporates items both covertly and overtly related to RMA. However, the AMMSA is not intended for use with the offending population which makes it difficult to assess whether it is actually measuring a construct related to rape. Another criticism of the AMMSA is that most individuals are not agreeing with many of the statements in the scale, thus it mainly measures disagreement with the items (D. C. F. Bishopp, personal communication, February 2nd, 2017).

McMahon and Farmer (2011) tackled one of the major criticisms raised by Gerger et al. (2007): the explicit nature of the items in the original IRMA. They recognised the high likelihood of response bias and sought to update the IRMA to offer an examination of subtle beliefs about rape victims, and perpetrators. Additionally, they updated the language used in the tool so that it was relevant to the cohort of participants. As the updated Illinois Rape Myth Acceptance Scale only explores attitudes and beliefs related to RMA, it fits within the research endeavours of the current author as a psychological tool. This is opposed to the AMMSA which is comprised of broader content areas that operate on a more sociological level. Furthermore, the AMMSA does not have the same range of established factors as those demonstrated within the IRMA (Sleath & Bull, 2015). Specifically, this review will examine the psychometric assessment by Payne, Lonsway, and Fitzgerald (1999), the Illinois Rape Myth Acceptance Scale, as updated by McMahon and Farmer (2011), in terms of its psychometric properties and its research uses.

Overview of the IRMA

Background to the Measure

The Illinois Rape Myth Acceptance scale is a measure that was developed to understand prominent rape myths and how rape myths are endorsed in society. The updated IRMA is a 19-item, self-report measure designed to assess subtle rape myths (McMahon & Farmer, 2011). It is a modified and current version of the original IRMA containing more modern language and offering a more accurate reflection of the subtleties in rape myths.

Scale Development

The original version of the IRMA was developed by Payne, Lonsway, and Fitzgerald (1999) in an attempt to expand and improve upon Burt's (1980) original RMAS. While crediting Burt for her immense contributions to the field of rape myth acceptance, Lonsway and Fitzgerald (1995) underlined the shortcomings of the RMAS (e.g. overlap in subscales, lack of criterion validity, use of colloquial phrasing, lack of discriminatory ability between offenders and non-offenders). When setting out to develop their own scale, Lonsway and Fitzgerald (1994) conducted stringent investigations and carefully considered clarity of items, colloquial terms, representation of the construct "rape myths", and psychometric properties.

In 1999, Payne, Lonsway, and Fitzgerald carried out a series of six studies to investigate the culture of rape myths and to develop a measure of rape myth acceptance. The first study involved a set of over 600 undergraduate students to evaluate 95 statements identified as rape myths. These 95 statements fit into 19 rape myth categories. Following multivariate analyses, the researchers found that the concept of RMA had a general component and seven subcomponents. The subcomponents of rape myth acceptance were: She asked for it; It wasn't really rape; He didn't mean to; She wanted it; She lied; Rape is a trivial event; and Rape is a deviant event.

In the second study, the researchers looked at individual perceptions of similarities in cultural rape myths using individual differences scaling (INDSCAL) as developed by Carroll and Chang (1970). In this study, participants consisted of both students and university employees. There were 24 men and 23 women in the study. The researchers took 19 rape myth statement pairs from their original 95 items and asked participants to rate similarity on a 9-point scale from not at all similar to very similar. Through INDSCAL and cluster analysis, the researchers found nine clusters ("women exaggerate about the effect of rape; it's not rape if…; rape only occurs in the bad part of town; she wanted it or enjoyed it; she led him on; the woman is responsible for preventing the rape; she was a tease/promiscuous; male absolution; women lie about rape") and two dimensions ("deny versus justify rape" and "victim versus perpetrator focus"). This structure mimicked that of study one (Payne et al., 1999, p. 43-44).

Study three explored the development and psychometric properties of the Illinois Rape Myth Acceptance Scale's structure and constructs. The researchers' criteria for item selection were structural integrity (items that highly correlated to the general and specific subcomponents outlined in study one), clarity (straightforward and clearly worded items), content coverage (range of content covered in subscales), reliability (each subscale possesses an alpha greater than .75), content weighting (items represent the depth of the component), and employment of some colloquial language. From these criteria, a 45-item assessment was produced containing 40 rape myths and five filler items to aid in controlling response sets.

To allow for wide-ranging use of the scale, the authors developed the short form of the IRMA (Illinois Rape Myth Acceptance Scale – Short Form; IRMA-SF). Payne and colleagues (1999) chose 17 items from each of the seven subscales out of the original 45 items along with three filler items.

Study four examined the IRMA's relationship to measures of constructs similar to RMA. The researchers found that individuals with higher scores on the IRMA and IRMA-SF endorsed more traditional sex-role stereotypes, accepted the idea that the relationship between the sexes is inherently adversarial, were more hostile toward women, and were comparatively accepting of interpersonal violence and general violence. Study five explored groups shown in the past to differ in RMA scores (i.e. police officers and advocates for rape victims) and contrasted their IRMA scores. The mean IRMA scores were statistically different between the rape advocates and police officers.

In study six, the researchers employed the use of rape narratives and examined IRMA scores in relation to the content of these narratives. Eighty-one university undergraduate students volunteered for the study. Of these, forty-five were included in the final sample due to availability. These individuals were asked to write two different stories. The first story was to detail what they thought occurred during a rape scenario presented to them. The second story was to be about a person from a Native American or African American descent. This story was used as a control and was not examined for rape myths. Analysis of story content revealed that inclusion of rape myths in the narratives was positively correlated to IRMA scores. Additionally, there was a moderate negative correlation between mean victim empathy ratings and IRMA scores supporting the scale's validity. In sum, both the IRMA and IRMA-SF demonstrated strong psychometric properties.

Updated IRMA. A little over a decade after the publishing of the original IRMA, McMahon and Farmer (2011) offered an update to the scale to reflect covert rape myths and contemporary language. They operated from the understanding that expression of rape myths had actually become subtler over time (Gerger et al., 2007). Their study had two stages. The first

stage involved the use of focus groups. The researchers conducted three focus groups: two with undergraduate sexual violence peer educators and one with professionals who have worked with sexual assault victims. The groups were asked to consider victim blaming remarks they had heard from students. Subsequently, they were presented with the 45-item IRMA and were asked to comment on the language and how relevant they thought differing items were. Four of the seven subscales were deemed relevant: She asked for it; It wasn't really rape; He didn't mean to; and She lied. Additionally, the researchers updated language in the measure to be better suited to their university sample and to capture more subtle rape myths. Furthermore, they added three items to the "He didn't mean it" subscale and one to the "It wasn't really rape" subscale. At this point, the researchers were down to 27 items and four subscales. Following this, McMahon and Farmer (2011) held interviews with 100 undergraduate and 40 graduate students and a panel of experts to have them review the scale. After several updates and modifications, the measure had an item total of 22.

The second stage of their research was a psychometric study of the properties of the updated IRMA. McMahon and Farmer (2011) administered their 22-item measure to 951 undergraduate students. To assess construct validity, the authors used Exploratory Structural Equation Modeling (ESEM) and found that there were actually five subscales instead of four. The researchers found that three of the items loaded onto a factor that centred around rape myths about alcohol. Additionally, three items did not load on to any of the factors and thus, they were removed from the item pool, leaving a 19-item scale. They evaluated criterion validity, exploring factors that might predict RMA (prior exposure to sexual assault education and knowing a victim of sexual assault) and their relatedness to gender using a Multivariate Analysis of Variance

(MANOVA). The MANOVA revealed gender as a significant predictor variable, however, this did not hold true for the other variables under investigation.

The psychometric properties of the IRMA will be discussed in detail later.

Administration and Scoring

The updated IRMA is a self-report scale comprising 19 items. Respondents are asked to rate items using a 5-point Likert scale ranging from 5 (strongly disagree) to 1 (strongly agree). The scale takes approximately 15 minutes to complete. On completion, the IRMA is scored by taking a sum of the ratings. Total scores can range from 19 to 95. Higher scores are considered to represent greater rejection of rape myths (McMahon & Farmer, 2011).

Psychometric Properties of the IRMA

Level of Measurement

According to Kline (2000), ideally one would want to employ a ratio scale when scientifically measuring a construct. However, psychological tests are most often based on interval level data. All the same, Kline (2000) states that these tests are still deemed acceptable as it is assumed that they have been constructed in a manner that allows for a wide coverage of item content and subsequent analysis. The IRMA uses a five-point Likert-type scale. Although not a true interval scale, it is treated as one for the purpose of statistical analyses. That is, the five points are viewed as equally spaced.

Reliability

Reliability refers to the quality of a measurement. More specifically, how consistent a test is both internally and over time (Kline, 2013).

Internal consistency. Internal consistency refers to how well the items on a measure relate to each other and whether, as representations of the same construct, they yield similar results. Internal consistency reliability is most often measured using Cronbach's alpha (α) and is based on the average correlation between items on a particular scale (Kline, 1986; Pallant, 2016). Nunnally (1978) recommends a minimum correlation of 0.7. The IRMA subscales had alphas ranging from α = .74 to .84, with the overall α of the final scale being .93. The overall internal consistency for the IRMA-SF was α = .87. The updated IRMA's subscales had alphas ranging from α = .64 to .80, and the overall Cronbach's alpha for the measure was .87.

Test-retest reliability. Test-retest reliability denotes that on administration of a scale on two separate occasions to the same participants, the scores obtained are replicated and consistent across administrations. These scores are correlated, and Pearson's *r* is the statistic used to represent the correlation between the two scores. Kline (2013) notes that a good measure should have a test-retest correlation of 0.7 or above. The IRMA demonstrated good test-retest stability at $r(495) = .90$, $p < .001$ (Payne et al., 1999). McMahon and Farmer (2011) did not report test-retest reliability for the updated IRMA.

Validity

A measure is valid if the scale measures what it claims to measure (Kline, 2013). There are number of ways in which the validity of a test can be demonstrated.

Face validity. A measure is said to have face validity if it *appears* to measure what it purports to measure (Kline, 2013). Face validity is not in actuality related to the true validity of a test. In fact, there are cases in which high face validity may prove to be disadvantageous (Cattell & Warburton, 1967, as cited in Kline, 2013). For example, in particular tests, participants can detect what the test is measuring, due to its high face validity and change their responses so that they are viewed in a more favourable light (Kline, 2000). This is a problem when attempting to measure attitudes that may be viewed negatively, as is the case with the IRMA. However, McMahon and Farmer addressed this issue when they updated the IRMA in 2011 to capture subtler rape myths. Effectively this weakened the face validity but in an effort to elicit attitudes on a topic of contention, this might be preferable. Lowering the face validity of such a measure would make the measure less likely to fall prey to response bias due to social desirability of participants' answers.

Concurrent validity (criterion-related). A measure demonstrates concurrent validity when it is shown to correlate highly with other tests claiming to measure the same construct (Kline, 2000). Therefore, if the IRMA claims to measure the same construct as existing RMA scales, it should follow that the IRMA will correlate with said scales. It should be noted, however, that difficulties may arise in choosing a second test as this test itself must be valid and reliable (Kline, 2000). In cases where it is difficult to find other tests for correlation, for example when both tests are measuring a construct the new test aims to improve upon, Kline suggests that significant but moderate correlations of 0.4 or 0.5 are acceptable.

Payne, Lonsway, and Fitzgerald (1999) did not report attempts to correlate their measure with other complete measures of rape myths (only subscales to assess for construct validity as

detailed below) nor did McMahon and Farmer (2011). Instead, the IRMA has since been used to validate newer measures, namely, the AMMSA. Gerger and colleagues (2007) correlated the AMMSA with the IRMA-SF across four studies. Sample sizes ranged from 40 to 848 participants (N = 1,279) and the researchers found significant, positive correlations ranging from .80 to .88.

Predictive validity (criterion-related). Predictive validity refers to a measure's ability to predict future performance by examining correlations between administration of the test on one occasion and another criterion at a later occasion (Kline, 1986). While there are many studies that have investigated the relationship between rape proclivity and rape myth acceptance, many of these used Burt's (1980) RMAS as a measure of RMA and are egregiously outdated (see Briere & Malamuth, 1983; Check & Malamuth, 1985; Greedlinger & Byrne, 1987; Malamuth, 1989a; Malamuth, 1989b; and Malamuth & Ceniti, 1986). A few studies have used newer measures to investigate the relationship between rape proclivity and RMA. That said, McMahon and Farmer (2011) reported that in 2009, Stephens and George demonstrated that the IRMA had predictive validity though its correlation with men's rape proclivity and sexual aggression. However, the current author could not find figures for this correlation. Additionally, Chapleau, Oswald, and Russell (2007) found a positive correlation between hostile sexism toward women and RMA (β = .24, t = 7.29, p < .001).

In using the MANOVA to evaluate the criterion validity of their updated IRMA, McMahon and Farmer (2011) assessed gender, knowledge of a victim of sexual assault, and prior experience in sexual assault education for differences in levels of RMA. They found an effect

with gender. Males and females differed significantly on the updated IRMA (Wilks' λ = .88 [F(5, 912) = 24.36, p < .01]) with females being more rejecting of rape myths.

Content validity. Content validity is the degree to which a measure includes all the elements necessary to represent the construct under investigation. If a measure is lacking in content validity, then it is not offering a complete representation of the concept it claims to assess.

As mentioned previously in the discussion on the IRMA's scale development, Payne and colleagues (1999) had five criteria for item selection: structural integrity, clarity, reliability, content weighting, and the use of colloquial language. Through adherence to their criteria and use of item-to-total correlations, the researchers were able to select 40 scale items and 5 filler items. McMahon and Farmer (2011) acknowledged that culture changes would threaten the IRMA's content validity and adjusted the language to reflect change in culture, less ambiguity in items, and a cultural shift from overt to covert sexism. Baldwin-White, Thompson, and Gray (2016) argue, however, that the IRMA does not address victim blaming, thereby lacking a grasp of all aspects of RMA. The current author counters that items falling under the subscales "She wanted it," "She lied," and "She asked for it" all serve to blame the victim. Overall, the IRMA appears to have good content validity.

Construct Validity. Construct validity involves a test's ability to measure the concept of interest to the researcher or researchers involved. That is, construct validity concerns whether results obtained on a test measuring a specific concept are in line with theoretically derived hypotheses about, or psychological nature of, that concept (Kline, 2000). Construct validity is

explored by investigating whether the measure is related to another measure with related constructs (i.e., convergent validity) and unrelated to a measure with dissimilar constructs (i.e., discriminant validity; Pallant, 2016).

To measure convergent validity, Payne and colleagues (1999) examined the IRMA and IRMA-SF in relation to scales that measured constructs that were theoretically and/or empirically linked to RMA. The constructs of comparison were: Sex-role stereotyping, Adversarial sexual beliefs, Hostility toward women, and Attitudes toward violence. The researchers correlated the IRMA and IRMA-SF with the following scales: Burt's (1980) *Sex-Role Stereotyping Scale*; Rombough and Ventimighlia's (1981) *Sexism Scale*; Burt's (1980) *Adversarial Sexual Beliefs Scale*; Lonsway and Fitzgerald's (1995) *Adversarial Heterosexual Beliefs Scale*; Lonsway and Fitzgerald's (1995) *Hostility Toward Women Scale*; Burt's (1980) *Acceptance of Interpersonal Violence Scale*; and Lonsway and Fitzgerald's (1995) *Attitudes Toward Violence Scale*. Correlations between the IRMA, IRMA-SF, and these related measures ranged from $r(174) = .47$, $p < .001$, to $r(174) = .74$, $p < .001$, that is, ranging from moderate to strong (Cohen, 1988).

Additionally, Payne et al. (1999) were able to demonstrate the discriminant validity of their scale when they recorded a moderate negative correlation between IRMA scores and mean victim empathy ratings, $r(43) = -.51$, $p < .01$).

Furthermore, in the updated IRMA, McMahon and Farmer (2011) used ESEM to assess the construct validity of the tool. They found that the fit for a five-factor model was acceptable (comparative fit index [CFI] = .90, Tucker-Lewis index [TLI] = .97, Root mean square error of approximation [RMSEA] = .07) versus the fit for the four-factor model which was not as good (CFI = .87, TLI = .91, RMSEA = .09).

Normative Samples

In order to carry out interpretation of a test at an individual or group level, a good test must have appropriate group norms (Kline, 2000). Establishing norms involves standardisation of a test, allowing mean scores from the sample under investigation to be compared to the "normed" population. Without a normative comparison, scores may be less meaningful and of less value to both the examiner and the participant (Kline, 2000). Interestingly, however, Kline (1986) stated that "for the use of psychological tests in the scientific study of human attributes – the psychometrics of individual differences – norms are not as useful. For this the direct, raw test-scores are satisfactory" (p. 159). The IRMA was not developed as a diagnostic tool, thus, standardised scores were not developed so that individual scores could be compared to a norming population. Because the IRMA arguably falls into the category of "psychometrics of individual differences", standardisation would not be pertinent for studying the construct of RMA. Consequently, raw scores may be sufficient to compare differences between individual participants and groups.

Limitations

As demonstrated above, there is evidence to support the IRMA's validity, reliability, and integrity as an interval scale. Despite this, there are limitations that merit consideration. Firstly, Payne, Lonsway, and Fitzgerald (1999) made note of the fact that the items on their scale were not discriminating between issues concerning victims of stranger versus acquaintance rape. They highlighted that individual participants may each visualise different types of rape when completing the questionnaire, subsequently producing different patterns of endorsement. This limits the scale's construct validity as important patterns in the makeup of rape myth statements

may be missed. Further research could include priming stimuli to place either stranger or acquaintance rape in the forefront of participants' minds.

Furthermore, in terms of administration and scoring, the scoring of the IRMA seems counterintuitive. Usually, in measurement, a higher score indicates a greater presence rather than an absence. However, the reverse is the case for the IRMA. This may be indicative of poor construction of the Likert items with a score of '5' as strongly disagree. Most tests would range from a score of '1', strongly disagree, to a score of '5', strongly agree.

Another limitation, which has been mentioned extensively in the literature, is that the IRMA is bound by culture and time. This is actually a necessary element of rape myth acceptance scales as the items need to be relevant to the cohort undertaking the survey. However, colloquial terms and phrases may be ambiguous or completely unknown when presented cross-culturally. Furthermore, these phrases can quickly become antiquated. This is a difficult problem to avoid as slang terminology is at the core of sexual communication (McMahon & Farmer, 2011).

A limitation of the research field of rape myth acceptance in general is the difficulty in the fact that what defines a "rape myth" will vary across authors. There are experts who will deem the term "rape myth acceptance" as interchangeable with "offense supportive attitudes" or "rape supportive attitudes" (C. Hermann, personal communication, May 4th, 2015; J. W. Van den Berg, personal communication, April 28th 2015). Alternatively, these terms could be viewed, possibly more appropriately, as overarching terminology under which "rape myth acceptance" falls as a subcategory. Nonetheless, this divergence on the operationalisation of rape myth acceptance will only serve to hinder standardisation, reliability, and concurrent validity.

Conclusion

The IRMA appears to demonstrate a reasonable degree of reliability and validity. However, when conducting the search for measure bias and/or distortion, it seems as though many research studies have accepted that the IRMA is reliable and valid without any further examination. Baldwin-White and colleagues (2016) had a similar finding. Another odd finding was that although the IRMA had been updated in 2011, there is very little evidence of the updated version being used in current research, even when attempting to create new RMA tools. While the initial reports and the updated report on the IRMA were respectable in their validity and reliability results, it is on the onus of researchers to use the most up to date resources. It is also their responsibility in utilising the tool to check for any kind of biases within the context of their experiments. This will allow for future researchers to update the tool accordingly. While there are limitations to the IRMA, the scales have been used extensively over the last 15 years and have directed the sexism and sexual assault literature.

As expressions of rape myths become more covert (Gerger et al., 2007), it may be worth comparing the IRMA with a measure of implicit association as a means of gauging the validity of the inferences we make from the IRMA scores. For example, the Implicit Association Test has demonstrated predictive validity in the prediction of behaviour and reportedly has greater validity than self-report in topics that are socially sensitive (Greenwald, Poehlman, Uhlmann, & Banaji, 2009), such as sexism and endorsement of rape myths.

In summary, the IRMA satisfies the minimum level of measurement and there is some evidence of its reliability and validity as a measure of rape myth acceptance. The IRMA is not without its shortcomings, and researchers are urged to report their reliability and validity statistics when utilising the tool for their purposes. There is most certainly scope for more

comprehensive research with this tool to further our understanding of its psychometric integrity. Thus, the structure of the IRMA was investigated further to explore its factor structure and dimensionality as an RMA measure.

CHAPTER 4: EXPLORING THE FACTOR STRUCTURE AND RELIABILITY OF THE ILLINOIS RAPE MYTH ACCEPTANCE SCALE

Abstract

The Illinois Rape Myth Acceptance Scale (IRMA) is one of the most widely used measures of rape myth acceptance (RMA). The current study had four aims: 1. To explore the factor structure of the IRMA using Principal Component Analysis (PCA) to test its construct validity; 2. To compare factor structure to the five implicit theories identified by Polaschek and Ward (2002); 3. To explore multivariate structure of the IRMA using an alternating least-squares algorithm (ALSCAL) to perform multidimensional scaling; and 4. To test the reliability of the IRMA, specifically its internal consistency. The IRMA was found to have an underlying four-component factor structure which was supported by the spatial structure produced by ALSCAL. The four components included: Shift of Responsibility/Minimisation of Harm, Male Sex Drive is Uncontrollable, She Lied, and Victim Devaluation/Promiscuity. These components are discussed in relation to the five implicit theories. Additionally, the four components of the IRMA were found to have high internal consistency (reliability). Implications for future research and practice are discussed.

Introduction

Sexual assault awareness and prevention programmes aim to educate individuals about the impact of sexual violence and sometimes, as mentioned previously, target attitudes towards rape which are believed to influence the perpetration and maintenance of sexual aggression. Beyond prevention programmes, rape supportive attitudes are targeted within sexual offender treatment programmes (SOTPs) as well (Johnson & Beech, 2017). Thus, it is of great importance that practitioners and researchers have a reliable and valid way of measuring rape myth acceptance.

The role of RMA

Although the relevance and impact of RMA was discussed in Chapter 1, it would be helpful to briefly review its role in society. RMA serves to blame the victim and exculpate the perpetrator. With this being a far-reaching theme in society, it is no wonder why victims struggle with deciding whether it is worth the effort to seek justice. Juries will vary in their opinions about rape, rapists, and victims of rape. Chapleau and Oswald's (2013) study found that rape myths may be more difficult to quash despite the apparent decrease in RMA; additionally, the researchers questioned the effectiveness of rape education programmes. They argue that in order to truly dissect, understand, and eradicate rape myths, we must venture into exploring the underlying mechanisms of RMA.

Maruna and Mann (2006) suggested that the examination of cognitive content alone, as assessed by explicit measures, may provide results that appear to demonstrate treatment change but are really superficial in nature. They posit that this overt change was actually a representation of individuals' increased capacity to modify how they communicate propensity to

rape or endorsement of rape myths. Maruna and Mann (2006) argued for the importance of assessing cognitive structure, such as schemas and implicit theories.

RMA as a cognitive structure

Ó Ciardha and Gannon (2011) provide a helpful distinction between cognitive structure, cognitive processes, and cognitive products. Cognitive structures are described as the schemas, beliefs, attitudes, and implicit theories individuals might hold. Cognitive processes are influenced by cognitive structures and relate to how an individual processes information. Cognitive products are the thoughts and/or statements that are produced as a result of this information processing through cognitive structures (Ó Ciardha & Gannon, 2011). There was a shift from evaluating cognitive content to classifying, understanding, and adjusting cognitive structures in treatment. This shift was aided by the introduction of the implicit theories hypothesis (Ward 2000; Ward & Keenan, 1999).

Implicit theories

In 1999, Ward and Keenan posited that sexual offenders hold underlying maladaptive (implicit) theories about themselves and the world and these implicit theories create and are maintained by a network of correlated beliefs and expectations.

As it relates to rape myths, Ward (2000) brought forward an additional explanation behind the rape myth phenomenon, suggesting that people hold implicit theories of how they understand the social world. These implicit theories influence an individual's interpretation of events (e.g. victim's actions) allowing for rationalization of sexually aggressive behavior (Polaschek & Ward, 2002).

Polaschek and Ward (2002) highlight four major assumptions about implicit theories which emphasise the different forms of knowledge within implicit theories and how they can be considered parallel to scientific theories (Fletcher, 1995; Polaschek & Ward, 2002):

1. Implicit theories may involve an ontology, involving a set of concepts that shows the psychological structures and processes within human beings;
2. These constructs, their properties and the relationships between them can explain behavior across different contexts, and are linked to underlying "psychological states";
3. Implicitly theories contain a number of interrelated beliefs and concepts and remain relatively comprehensible;
4. Implicit theories dictate how evidence is perceived and interpreted; evidence considered relevant is dependent on the cognitive structures or constructs within a theory. Polaschek and Ward (2002) capture this in the following example:

> "a rapist who believes women to be in a constant state of sexual awareness will interpret a woman's friendly behavior towards him as sexual interest. Alternatively, rapists who believe women to be inherently deceptive in their interactions with men, will interpret the same behavior as indicating the woman's hostile intent" (p. 391).

Polaschek and Ward (2002) went on to posit that sexual offenders create implicit theories about their offending and their victims which help them form perceptions of their victim and their own intentions and predict future behavior. The researchers proposed that sexual offenders hold the following rape-supportive implicit theories: Women are Unknowable (the theory that women are inherently different from men and it cannot be expected that a woman's mind can be understood by men; an example myth would be ''a lot of times when women say no, they are just playing hard to get and really mean yes''; p. 395); Women as Sex Objects (the belief that

women exist to fulfil the sexual needs of men. This includes the theory of sexual precedence and an example of a statement which supports this theory is “a woman who goes to the home or apartment of a man on their first date implies that she is willing to have sex”; p.397); Male Sex Drive is Uncontrollable (beliefs that men are unable to control their sexual urges and must be satisfied when aroused. An example of a thought generated by this belief is “when women go around braless or wearing short skirts and tight tops, they are just asking for trouble”; p. 398)); Entitlement (the belief that men’s sexual needs are to be met on demand; they are entitled to sex. An example of this is “men rape because women reject them”; p. 398); and Dangerous World (beliefs that the world is inherently dangerous and hostile, where others cause harm for their own benefit. Cognition supportive of this theory includes, “she would have done the same to me, if I hadn’t got to her first”; p.398).

Many researchers, as well as practitioners, agree with the shift from exploration of cognitive content to gaining an understanding of cognitive structures within those who offend and those with the propensity to offend sexually (Beech, Bartels, & Dixon, 2013; Dean, Mann, Milner, & Maruna, 2009; Mann & Beech, 2003; Maruna & Mann 2006; Ó Ciardha & Gannon, 2011). These underlying structures are often measured using implicit measures (see Nunes & Hermann (2013) for further detail), however, explicit measures are most often accessible. Researchers have begun to look at subtypes of RMA instead of RMA as an overarching construct in an attempt to tap into possible schemas.

Types of RMA

Lonsway and Fitzgerald (1994) noted that overall measurement of RMA masks the heterogeneity of rape myths (Johnson, Kuck, & Schander, 1997). They emphasised the

importance of looking at subscales of measures and categorisation of rape myths. Categorisations of RMA vary throughout the literature, including victims lying about rape (Cuklanz, 2000), victims are to be blamed and to be held responsible for their rape (Scully, 1990), and acquaintance rape justifications (Johnson et al., 1997).

Arguably, one of the most widely used RMA measures is the Illinois Rape Myth Acceptance Scale (IRMA; Payne et al., 1999). IRMA and its subscale categories of RMA have been cited throughout the literature. As previously mentioned in Chapter 3, the main categories identified in the updated IRMA are: 1. He didn't mean it (the man has uncontrollable sexual urges or he was drunk); 2. It wasn't really rape (no visible signs of an attack, i.e. bruising, no physical fight); 3. She asked for it (the victim went into a room alone with the perpetrator; she wore provocative clothing; she was drunk); and 4. She lied (the victim does not want to be viewed poorly so pretends to be raped, or is trying to obtain some legal or financial benefit; McMahon & Farmer, 2011; Pauna & Pleszewski 2012; Payne et al., 1999**)**. This typology will be utilised in the interpretation of results within this report.

The current study

As described previously, the IRMA is a widely implemented tool for measuring RMA; however, it appears as though the updated version is not utilised as often in research as the original version. Also, the updated IRMA's factor structure has not been repeatedly tested in different individuals and, at last search, it appears little has been done in the way of exploring possible underlying dimensions of the tool. It is important to establish the IRMA as psychometrically sound across multiple contexts.

As this study hopes to examine rape myths within a wider social context, beyond offenders, the population will consist of undergraduate students, who are at high risk for perpetration and victimisation of sexual assault (Baldwin-White et al., 2016; Crawford, Wright, & Birchmeier, 2008; Fisher, Daigle, & Cullen, 2010). This study intends to test the reliability, factor structure, and underlying dimensions of the IRMA within a group of undergraduate students.

Specific aims

Through secondary data analysis this study will attempt to:

1. Explore the factor structure of the IRMA using Principal Component Analysis (PCA) to test its construct validity;
2. Compare factor structure to the five implicit theories identified by Polaschek and Ward (2002);
3. Explore multivariate structure of the IRMA using an alternating least-squares algorithm (ALSCAL) to perform multidimensional scaling;
4. Test the reliability of the IRMA, specifically its internal consistency.

Method

Participants

The sampling data were collected via Lime Survey, the University of Birmingham's online survey tool. Participants were sampled from undergraduates, mainly Psychology students. There was also purposive sampling of males through the University’s social and sports clubs. This was to compensate for the small number of males in the Psychology course. Students were

awarded course credits for participation. Consent was given online prior to the survey being presented.

The dataset contained IRMA scores on 346 undergraduate students aged 18 to 33 years ($\mu = 19.90$, $SD = 1.69$). The sample was predominantly White British/Other white background (61.85%), with 10.69% describing themselves as Asian/Other Asian background. The remainder described themselves as Black, of Mixed/Multiple background, or of another ethnic group. Of this group, 218 were women, 51 were men. The remainder identified as other types of gender or chose not to answer.

Measure

The measure used in this study was the Illinois Rape Myth Acceptance scale (IRMA; Payne et al., 1999) as updated by McMahon and Farmer (2011). Although, McMahon and Farmer (2011) finalised a 19-item scale, the scale utilised in this dataset was the 22-item measure McMahon and Farmer had in place prior to removal of three items which did not fit their five-factor model. The 22 items found in the IRMA can be found in Table 2.

The IRMA is used to evaluate beliefs about rape myths and attitudes towards rape, rapists, and victims of rape. The IRMA uses a Likert-type scale, which requires the respondent to rate each item on a 5-point scale from 1, Strongly Agree, to 5, Strongly disagree. Items were reverse scored as appropriate. Higher scores represent a greater degree of rejection of rape myths.

Treatment of the data

The data was analysed using the Statistical Package for the Social Sciences (SPSS; IBM, 2016). All subjects with missing data were removed for the analysis, leaving a total of 265. A

principal component analysis (PCA) was run on 22 items. For improved interpretability, the items were rotated using direct oblimin (oblique rotation) as the expectation was that the items would be correlated. For confirmation, ALSCAL (Young, Takane, & Lewyckyj, 1978), a form of multidimensional scaling, was applied to the data. This analysis allows for the data to be represented in multidimensional Euclidean space. Following this, a reliability analysis was conducted on each of the subscales. Cronbach's alpha is reported.

Ethical considerations

Participants were provided with information relevant to the study for which the data was obtained. Participants also provided consent for their data to be used in research conducted at the University of Birmingham. However, as participants were not identifiable to the current researcher, it was not possible to provide them with information or feedback about the current study.

Results

Descriptive Analysis

Descriptive statistics for each of the items are provided in Table 2. Examination of the table indicates some skewness in the distribution of items.

Table 2

Item Descriptive Statistics (N = 265)

Item	M (SD)	Skewness	Scale Responses (%) 1	2	3	4	5
1. If a girl is raped while she is drunk, she is at least somewhat responsible for letting things get out of hand.	1.42 (0.88)	2.55	75.1	16.6	2.3	3.8	2.3
2. When girls go to parties wearing slutty clothes, they are asking for trouble.	1.43 (0.87)	2.20	74.3	14.7	5.3	4.5	1.1
3. If a girl goes to a room alone with a guy at a party, it is her own fault if she is raped	1.26 (0.76)	3.55	85.3	8.7	3.0	0.8	2.3
4. If a girl acts like a slut, eventually she is going to get into trouble.	1.88 (1.07)	1.04	49.4	26.4	12.8	9.8	1.5
5. When girls get raped, it's often because the way they said "no" was unclear.	1.40 (0.80)	2.40	73.2	17.7	6.0	1.5	1.5
6. If a girl initiates kissing or hooking up, she should not be surprised if a guy assumes she wants to have sex.	2.17 (1.22)	0.60	41.9	21.9	16.2	17.4	2.6
7. When guys rape, it is usually because of their strong desire for sex.	2.65 (1.20)	-0.05	23.8	20.8	24.5	28.3	2.6
8. Guys don't usually intend to force sex on a girl, but sometimes they get too sexually carried away.	2.29 (1.13)	0.49	29.8	31.3	21.1	15.1	2.6

9. Rape happens when a guy's sex drive goes out of control.	2.29 (1.14)	0.40	32.1	26	24.5	15.1	2.3
10. If a guy is drunk, he might rape someone unintentionally.	2.10 (1.13)	0.67	40.0	27.5	16.2	14.7	1.5
11. It shouldn't be considered rape if a guy is drunk and didn't realize what he was doing.	1.35 (0.72)	2.90	74	21.5	1.9	1.1	1.5
12. If both people are drunk, it can't be rape.	1.54 (0.90)	2.12	64.2	25.7	5.3	2.3	2.6
13. If a girl doesn't physically resist sex, even if protesting verbally, it can't be considered rape.	1.32 (0.76)	2.99	79.2	14.7	2.3	2.3	1.5
14. If a girl doesn't physically fight back, you can't really say it was rape.	1.23 (0.69)	4.13	85.3	11.3	1.1	0.0	2.3
15. A rape probably doesn't happen if a girl doesn't have any bruises or marks.	1.15 (0.63)	5.27	91.7	6.0	0.0	0.0	2.3
16.If the accused "rapist" doesn't have a weapon, you really can't call it rape.	1.14 (0.63)	5.38	93.2	4.2	0.4	0	2.3
17. If a girl doesn't say "no" she can't claim rape.	1.79 (1.04)	1.27	53.2	24.5	14.7	4.9	2.6
18. A lot of times, girls who say they were raped agreed to have sex and then regret it.	1.95 (1.00)	0.87	41.5	30.9	20.8	4.9	1.9
19. Rape accusations are often used as a way of getting back at guys.	1.93 (1.03)	0.95	44.2	28.7	19.6	5.3	2.3
20. A lot of times, girls who say they were raped often led the guy on and then had regrets.	1.76 (0.95)	1.12	52.1	25.3	18.5	2.6	1.5
21. A lot of times, girls who claim they were raped have emotional problems.	1.77 (1.03)	1.25	55.1	22.3	15.5	4.9	2.3
22. Girls who are caught cheating on their boyfriends sometimes claim it was rape.	1.93 (1.05)	0.80	47.2	23.8	18.9	9.4	0.8

Factor Structure

As highlighted previously, although the 45-item IRMA scale is routinely administered in research, the scale used in data collection was the 22-item version as presented in McMahon and Farmer's (2011) report. A PCA was conducted on 22 items with oblique rotation (direct oblimin). The Kaiser-Meyer-Olkin (KMO) measure verified the sampling adequacy for the analysis, KMO = .92 ('superb' according to Field, 2009), and all KMO values for individual items were >.84, above the acceptable limit of .5 (Field, 2009). Bartlett's test of sphericity χ^2 (231) = 4179.16, $p < .001$, indicated that correlations between items were sufficiently large for PCA. Determining the number of components is one of the complexities of PCA and authors differ on the most appropriate way to proceed. Two of the more traditional approaches include the Kaiser (1960) criterion (Eigenvalues greater than 1) and Cattell's (1966) scree test. There is no fool-proof method for determining the number of factors. It is suggested that the best way to determine parsimony is to explore a range of solutions which best explain the data (D. Bishopp, personal communication, July 17th, 2017). The scree plot was slightly ambiguous and showed inflexions that would justify retaining both four and six components. As such, a number of solutions were explored from four to six, with the four-component model explaining the most variance. Four components had eigenvalues over Kaiser's criterion of 1 and in combination explained 66.74% of the variance[1]. Given the large sample size, and the convergence of the scree plot and Kaiser's criterion on four components, this is the number of components that were retained in the final analysis. The pattern matrix of loadings (after rotation) for the four factor

[1] A principal component analysis was conducted on McMahon and Farmer's (2011) 19-item scale. A 3-factor model was produced that, 1. Explained 63.34% of the variance and 2. Was less theoretically sound than the 4-factor model reported here.

PCA is provided below in Table 3. The structure matrix is also provided in Table 4 along with the correlation matrix between factors (Table 5).

The items that clustered on component 1 are indicative of beliefs that the victim should be responsible for preventing or stopping a rape (e.g. “If a girl goes to a room alone with a guy at a party, it is her own fault if she is raped,” or “If a girl doesn’t physically fight back, you can’t really say it was rape”) and beliefs around what “real” rape “looks” like (e.g. “A rape probably doesn’t happen if a girl doesn’t have any bruises or marks,” or “If the accused “rapist” doesn’t have a weapon, you really can’t call it rape.”) This component has been labelled Shift of Responsibility/Minimisation of Harm. Items in this component serve to belittle the experience of the victim and minimises the harm (physical or psychological) endured. The second component was related to the belief that men are not in control when perpetrating a rape as their desire for sex is overwhelming and uncontainable. Items in this component include “When guys rape, it is usually because of their strong desire for sex,” and “Guys don’t usually intend to force sex on a girl, but sometimes they get too sexually carried away.” Items in this component relate to the implicit theory “Male sex drive is uncontrollable” (as suggested in Polaschek and Ward, 2002) and has been labelled as such. The items clustered on component 3 represent beliefs that the victim lied about the rape and includes “Girls who are caught cheating on their boyfriends sometimes claim it was rape,” and “A lot of times, girls who say they were raped agreed to have sex and then regret it” as items. This component is very similar to the She Lied subscale offered in McMahon and Farmer (2011) and, thus, has been given this label. The fourth component is comprised of items which relate to beliefs about ‘promiscuity’ of the victim, highlighting conservative ideology regarding women’s gender role and appearance. This component is entitled “Victim Devaluation/Promiscuity”.

As shown in Table 3, Component 1 (Shift of Responsibility/Minimisation of Harm; SR) had nine main high loading items, Component 2 (Male Sex Drive is Uncontrollable; US) had four high loading items, Component 3 (She Lied; SL) had six high loading items and Component 4 (Victim Devaluation/Promiscuity; VD) had three main high loading items.

Table 3

Pattern Matrix for four-component solution

Item		Component 1 SR	2 US	3 SL	4 VD
16.	If the accused "rapist" doesn't have a weapon, you really can't call it rape.	**0.99**	-0.02	-0.01	-0.13
15.	A rape probably doesn't happen if a girl doesn't have any bruises or marks.	**0.96**	-0.10	-0.04	-0.09
14.	If a girl doesn't physically fight back, you can't really say it was rape.	**0.93**	-0.06	-0.04	0.00
13.	If a girl doesn't physically resist sex, even if protesting verbally, it can't be considered rape.	**0.79**	-0.05	0.01	0.11
11.	It shouldn't be considered rape if a guy is drunk and didn't realize what he was doing.	**0.75**	0.15	-0.04	-0.08
3.	If a girl goes to a room alone with a guy at a party, it is her own fault if she is raped.	**0.75**	0.00	-0.01	0.23
5.	When girls get raped, it's often because the way they said "no" was unclear.	**0.68**	0.12	-0.08	-0.04
12.	If both people are drunk, it can't be rape.	**0.63**	0.06	-0.07	0.06
1.	If a girl is raped while she is drunk, she is at least somewhat responsible for letting things get out of hand.	**0.54**	0.03	0.04	**0.40**
7.	When guys rape, it is usually because of their strong desire for sex.	-0.08	**0.83**	-0.02	-0.05
9.	Rape happens when a guy's sex drive goes out of control.	-0.08	**0.81**	0.08	0.13
8.	Guys don't usually intend to force sex on a girl, but sometimes they get too sexually carried away.	0.13	**0.68**	-0.08	-0.05
10.	If a guy is drunk, he might rape someone unintentionally.	0.09	**0.63**	-0.11	0.00

19.	Rape accusations are often used as a way of getting back at guys.	-0.05	-0.11	**-0.91**	0.06
22.	Girls who are caught cheating on their boyfriends sometimes claim it was rape.	-0.05	0.03	**-0.85**	-0.04
20.	A lot of times, girls who say they were raped often led the guy on and then had regrets.	0.05	0.11	**-0.84**	-0.02
18.	A lot of times, girls who say they were raped agreed to have sex and then regret it.	0.05	0.04	**-0.80**	0.02
21.	A lot of times, girls who claim they were raped have emotional problems.	0.11	0.06	**-0.62**	0.03
17.	If a girl doesn't say "no" she can't claim rape.	0.18	0.02	**-0.39**	0.20
4.	If a girl acts like a slut, eventually she is going to get into trouble.	0.03	0.13	-0.08	**0.74**
6.	If a girl initiates kissing or hooking up, she should not be surprised if a guy assumes she wants to have sex.	-0.09	-0.02	-0.18	**0.69**
2.	When girls go to parties wearing slutty clothes, they are asking for trouble.	**0.40**	0.13	0.05	**0.60**
Cronbach's alpha (α)		0.93	0.76	0.88	0.72

Table 4

Structure Matrix

Item		Component 1 SR	2 US	3 SL	4 VD
16.	If the accused "rapist" doesn't have a weapon, you really can't call it rape.	**0.94**	0.24	-0.45	0.25
15.	A rape probably doesn't happen if a girl doesn't have any bruises or marks.	**0.92**	0.18	-0.45	0.27
14.	If a girl doesn't physically fight back, you can't really say it was rape.	**0.94**	0.24	-0.49	0.35
13.	If a girl doesn't physically resist sex, even if protesting verbally, it can't be considered rape.	**0.82**	0.22	-0.42	0.39
11.	It shouldn't be considered rape if a guy is drunk and didn't realize what he was doing.	**0.79**	0.37	-0.45	0.27
3.	If a girl goes to a room alone with a guy at a party, it is her own fault if she is raped.	**0.84**	0.29	-0.48	0.52
5.	When girls get raped, it's often because the way they said "no" was unclear.	**0.74**	0.35	-0.46	0.29
12.	If both people are drunk, it can't be rape.	**0.70**	0.30	-0.44	0.35
1.	If a girl is raped while she is drunk, she is at least somewhat responsible for letting things get out of hand.	**0.69**	0.30	-0.42	0.60
7.	When guys rape, it is usually because of their strong desire for sex.	0.16	**0.80**	-0.32	0.18
9.	Rape happens when a guy's sex drive goes out of control.	0.17	**0.79**	-0.28	0.31
8.	Guys don't usually intend to force sex on a girl, but sometimes they get too sexually carried away.	0.36	**0.74**	-0.42	0.24
10.	If a guy is drunk, he might rape someone unintentionally.	0.34	**0.71**	-0.43	0.27
19.	Rape accusations are often used as a way of getting back at guys.	0.40	0.29	**-0.86**	0.39

22.	Girls who are caught cheating on their boyfriends sometimes claim it was rape.	0.38	0.37	**-0.82**	0.31
20.	A lot of times, girls who say they were raped often led the guy on and then had regrets.	0.50	0.48	**-0.90**	0.38
18.	A lot of times, girls who say they were raped agreed to have sex and then regret it.	0.47	0.41	**-0.85**	0.38
21.	A lot of times, girls who claim they were raped have emotional problems.	0.46	0.37	**-0.71**	0.35
17.	If a girl doesn't say "no" she can't claim rape.	0.46	0.30	**-0.57**	0.44
4.	If a girl acts like a slut, eventually she is going to get into trouble.	0.39	0.40	-0.46	**0.83**
6.	If a girl initiates kissing or hooking up, she should not be surprised if a guy assumes she wants to have sex.	0.25	0.23	-0.41	**0.72**
2.	When girls go to parties wearing slutty clothes, they are asking for trouble.	**0.64**	0.40	-0.46	**0.77**

Table 5

Component Correlation Matrix for four-component solution

Component	1 SR	2 US	3 SL	4 VD
1 SR	1.000	0.300	-0.509	0.382
2 US	0.300	1.000	-0.433	0.297
3 SL	-0.509	-0.433	1.000	-0.416
4 VD	0.382	0.297	-0.416	1.000

Multivariate structure

To confirm the factor structure, ALSCAL was applied to the data. Inter-variable correlations are represented by a matrix of proximities, which are then represented as coordinates unfolded in dimensional space (Bishopp, 2003). The distances between coordinates are representative of the association between variables; so, the variables which are closest together, are more strongly related to each other (Bishopp, 2003; Borg & Groenen, 1997). The ALSCAL

output can be seen in Figure 2. Points are labelled with item numbers, i.e., Q01 corresponds to the first item on the IRMA, “If a girl is raped while she is drunk, she is at least somewhat responsible for letting things get out of hand,” while Q02 corresponds to the second item on the IRMA, “When girls go to parties wearing slutty clothes, they are asking for trouble,” and so on. Four clusters of items were identified and the plot has been sectioned and labelled to aid interpretation.

Multidimensional scaling attempts to minimise the stress associated with obtaining a multidimensional solution. Stress is a goodness-of-fit statistic and the aim is to keep this as low as possible (Wilkinson, 2000). The statistic varies between 0 and 1, with values closer to 0 representative of better fit. As to what constitutes an ‘acceptable’ amount of stress, there is difference in scientific opinion within the literature. Kruskal’s (1964) rule of thumb is quite conservative, requiring a stress limit below .15. However, Bell (1997) suggests that a more appropriate guide for acceptable levels of stress is found in Spence and Ogilvie (1973), which is the guide used for this report. That said, it is argued that of highest importance is the configuration of the points and interpretability of the multidimensional scaling solution (Bishopp, 2003; Bell 1997; Shye, 1978). The stress for this solution gives a value of 0.26, which is satisfactory (below 0.309) according to Spence and Ogilvie (1973). The corresponding correlation is the RSQ (r-squared) value, the variance accounted for in the MDS solution. The RSQ for this solution is 0.82.

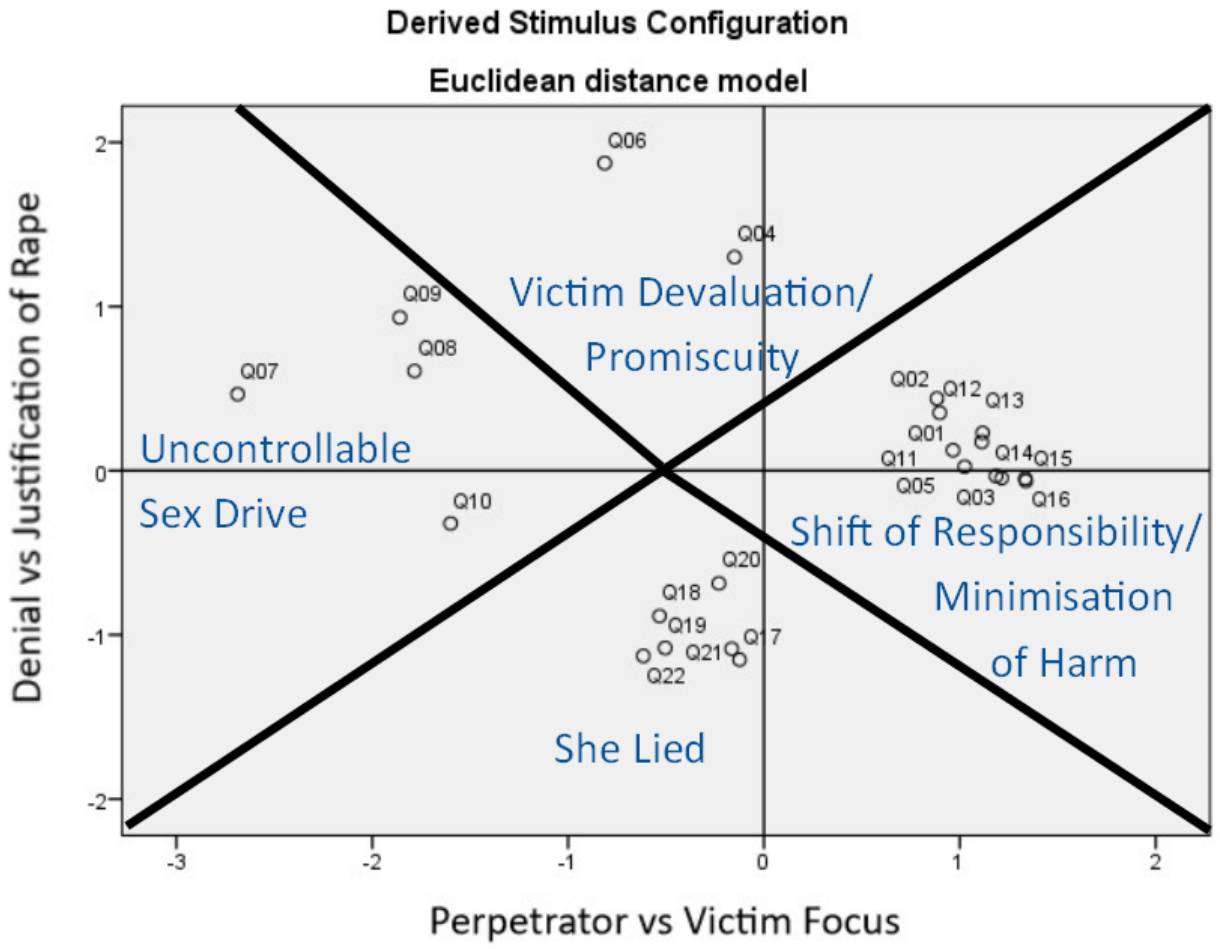

Figure 2. ALSCAL spatial structure for IRMA

Reliability

The internal consistency of the IRMA was examined using Cronbach's alpha. Cronbach's alpha will not be used here as a measure of unidimensionality, as it has been shown that a high alpha can be achieved when items are moderately correlated as well as when some factors are correlated (Cortina, 1993; Grayson, 2004). Cronbach (1951) suggested that in cases where several factors exist, the reliability formula should be applied separately to the items related to each factor. In this case, Cronbach's alpha was calculated separately for each of the IRMA components.

The alpha values for each component is reported under the pattern matrix in Table 3. Overall, all items within each subscale are positively contributing to the overall reliability. However, the item "If a girl initiates kissing or hooking up, she should not be surprised if a guy assumes she wants to have sex" within the Victim Devaluation/Promiscuity subscale had a corrected item-total correlation of 0.46 which is low relative to other items. Also, if this item were to be removed, the alpha of the subscale would increase from 0.72 to 0.76. While this is not substantial, it is worth noting as this is an item which McMahon and Farmer chose to remove from their final scale. This is similar to the item "If a girl doesn't say "no" she can't claim rape," which falls in the She Lied subscale. This item had a corrected item-total correlation of .50 and if this item were to be deleted the alpha would increase to 0.88 from 0.89, which may be considered negligible. These were the only noteworthy statistics to report from the reliability analysis, which overall may be a testament to fairly good internal consistency of the subscales. The Item-Total Statistics and Reliability Statistics for each component can be found in Table 6.

Table 6

Item-Total Statistics and Reliability Statistics for IRMA's Four Components

Component 1: Shift of Responsibility/Minimisation of Harm	
Reliability Statistics	
Cronbach's Alpha	N of Items
0.93	9
Item-Total Statistics	

Item	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
1. If a girl is raped while she is drunk, she is at least somewhat responsible for letting things get out of hand.	10.38	24.00	0.65	0.93
3. If a girl goes to a room alone with a guy at a party, it is her own fault if she is raped	10.54	23.97	0.80	0.92
5. When girls get raped, it's often because the way they said "no" was unclear.	10.40	24.48	0.68	0.93
11. It shouldn't be considered rape if a guy is drunk and didn't realize what he was doing.	10.45	24.64	0.74	0.93
12. If both people are drunk, it can't be rape.	10.26	23.89	0.65	0.93
13. If a girl doesn't physically resist sex, even if protesting verbally, it can't be considered rape.	10.48	24.25	0.75	0.93
14. If a girl doesn't physically fight back, you can't really say it was rape.	10.57	23.99	0.89	0.92
15. A rape probably doesn't happen if a girl doesn't have any bruises or marks.	10.65	24.64	0.86	0.92
16.If the accused "rapist" doesn't have a weapon, you really can't call it rape.	10.66	24.51	0.89	0.92

Component 2: Male Sex Drive Is Uncontrollable

Reliability Statistics

Cronbach's Alpha	N of Items
0.76	4

Item-Total Statistics

Item	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
7. When guys rape, it is usually because of their strong desire for sex.	6.69	7.15	0.58	0.69
8. Guys don't usually intend to force sex on a girl, but sometimes they get too sexually carried away.	7.05	7.54	0.56	0.70
9. Rape happens when a guy's sex drive goes out of control.	7.05	7.42	0.58	0.70
10. If a guy is drunk, he might rape someone unintentionally.	7.24	7.74	0.52	0.73

Component 3: She Lied

Reliability Statistics

Cronbach's Alpha	N of Items
0.88	6

Item-Total Statistics

Item	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
17. If a girl doesn't say "no" she can't claim rape.	9.34	17.80	0.50	0.89
18. A lot of times, girls who say they were raped agreed to have sex and then regret it.	9.18	16.07	0.78	0.85
19. Rape accusations are often used as a way of getting back at guys.	9.20	16.08	0.74	0.85
20. A lot of times, girls who say they were raped often led the guy on and then had regrets.	9.37	15.98	0.84	0.84
21. A lot of times, girls who claim they were raped have emotional problems.	9.36	17.03	0.61	0.88

22. Girls who are caught cheating on their boyfriends sometimes claim it was rape.	9.20	16.21	0.70	0.86

Component 4: Victim Devaluation/Promiscuity	
Reliability Statistics	
Cronbach's Alpha	N of Items
0.72	3

Item-Total Statistics

Item	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
2. When girls go to parties wearing slutty clothes, they are asking for trouble.	4.05	3.76	0.59	0.60
4. If a girl acts like a slut, eventually she is going to get into trouble.	3.60	3.07	0.62	0.53
6. If a girl initiates kissing or hooking up, she should not be surprised if a guy assumes she wants to have sex.	3.31	3.07	0.46	0.76

Discussion

This study had four main objectives, which will be discussed in detail here.

Exploring the factor structure of the IRMA using Principal Component Analysis (PCA) to test its construct validity

The PCA produced four components. Component 1 (Shift of Responsibility/Minimisation of Harm; SR) related to the belief that no real "harm" was inflicted on the victim as there are no

visible signs of "real rape" as well as ideas that the responsibility for prevention of rape is the responsibility of the victim. Component 2 (Male Sex Drive is Uncontrollable; US) is representative of the belief that a man is victim to his sexual urges and thus any devious sexual behaviour is the corollary of his overpowering and irrepressible sexual desires and should be excused. Component 3 (She Lied; SL) corresponds to beliefs that victims fabricate incidents of sexual assault in order to preserve their dignity if caught in an unsavoury act or because they are either emotional or devious. Lastly, component 4 (Victim Devaluation/Promiscuity; VD represents the idea that victims who are perceived as presenting provocatively are lowly and are "asking to be raped." The four components explained 67% of the variance in the model, which is indicative of a good model, i.e. it identifies the complex associations inherent in the data and emergent structures are likely to be robust and theoretically defensible (Bishopp, 2003).

The PCA did not completely replicate McMahon and Farmer's (2011) four-factor model (which later became a five-factor model). The items on the She Lied scale for the current study mapped almost completely onto the She Lied scale in McMahon and Farmer's 2011 study with the exception of an extra item "If a girl doesn't say "no" she can't claim rape," added into the She Lied scale of the current study.

It is difficult to make a direct comparison as different data analysis strategies were used and it appeared that the researchers carried out their structural equation modelling in a less 'exploratory' manner than the current study as the parameters they set were more restrictive. It is worth considering other explanations for the observed differences. These differences might indicate that the factor structure of the measure is not stable. However, this explanation does not account for the effect of time and how the difference of six years between McMahon and Farmer's study and the current study could affect outcomes. There may be a shift in culture or a

shift in how individuals respond to specific types of myths. Future research should involve using IRMA as a measure over time, across different populations, and over different treatment programmes, while keeping the treatment of data and its analysis constant. Overall, the findings offer support for construct validity of the IRMA.

Factor Loadings. Using a correlation cut-off of 0.39, we see that the only two cross-loading items are 1, "If a girl is raped while she is drunk, she is at least somewhat responsible for letting things get out of hand," and item 2, "When girls go to parties wearing slutty clothes, they are asking for trouble," which both load on to components 1 and 4, with item 1 having the higher pattern coefficient for factor 1 and item 2 having the higher pattern coefficient for factor 4. For both items, the corresponding structure coefficients are not that dissimilar from their pattern coefficients. This may be indicative that these are items which could explain either component, Shift of Responsibility or Victim Devaluation. With regard to item 1, although it fits better with the Shift of Responsibility component, it could, theoretically, be linked to the Victim Devaluation component. If one were to extrapolate, it could be argued that being intoxicated meant that you have willingly opened yourself up to all things dreadful. The idea here is that one's worth is lowered when drunk, as though drinking is an abhorrent and unsavoury decision and is associated with being flagrant, thus devaluing the victim and justifying any harm perpetrated upon said victim. This line of thinking is not too unbelievable when considering the implicit theories of sexual offenders (Polaschek & Ward, 2002) which serve to justify rape; these theories are expanded upon below. Similarly, with item 2, although it contributes more to Victim Devaluation than it does Shift of Responsibility, it is easy to understand how this could cross load onto Shift of Responsibility. In placing the onus on to the victim, there is a shifting of

responsibility, implying that the victim sought trouble of their own volition, the moment they chose a particular set of clothing.

Factor Correlations. When reviewing the Component Correlation Matrix (Table 5), it is apparent that components 1, 2, and 4, are most highly correlated with component 3 as opposed to with each other. In looking at the structure matrix, we see that almost all items, save for 7 and 9, have a structure coefficient above 0.40 indicating that all items are in part representative of this factor. This is where it is useful to examine both the pattern and structure matrices as the pattern matrix helps to identify which items best fit into this component.

Shared Variance. It is apparent from the structure matrix that the majority of items in the scale appear to contribute to components Shift of Responsibility and She Lied. With She Lied, all items, save for 7 and 9, have structure coefficients above .40, which suggests that all these items contribute to the component structure; however, for those items representative of factors Shift of Responsibility, Male Sex Drive is Uncontrollable, and Victim Devaluation, their structure coefficients in She Lied are not comparable or as meaningful as the coefficients with the factors those items actually represent. As stated prior, items which factored into Shift of Responsibility are also contributing to the variance in She Lied and vice versa. This could be a testament to the similarity of these two components, Shift of Responsibility and She Lied. Theoretically, the two components are linked as assuming the victim in lying is shifting the blame toward the victim. Also, this assumption not only minimises any possible harm, but completely negates that any harm is done. Both these components can serve to vilify the victim. Further explanations on the similarities of these components as they relate to implicit theories can be found below.

There are cases in which the structure coefficients are higher than the pattern coefficient. Looking only at items corresponding to their main component (i.e. items 1 and 2 corresponding to factors 1 and 4 respectively), items 2 and 17 stand out from the rest. Their structure coefficients are 0.77 and -0.57 respectively. This may indicate that these variables' association with their respective factors is not unique (Graham, Guthine, & Thompson, 2003). Thus, a greater portion – than could be determined by viewing the pattern matrix alone – of these items' contribution to their factors is shared with the other items loading on to this factor. Thereby the pattern matrix, in a sense, downplays the importance of these items to their corresponding factors. There were no issues regarding suppressor effects (large pattern coefficient with relatively small structure coefficient which causes suppression of error in other items; Graham et al., 2003) present in the data.

Comparing factor structure to the five implicit theories identified by Polaschek and Ward (2002)

The four components produced by the PCA were readily interpretable and consistent with general rape supportive attitudes as well as specifically with a few of Polaschek and Ward's (2002) five implicit theories of sexual offending.

Women are unknowable. This theory focuses on the belief that men are inherently different from women and men are unable to develop a realistic theory of mind for women (Polaschek &Ward, 2002). The thinking is very black-and-white, in that it fosters and maintains stereotypes that women are either "nice girls or whores" (p. 394). Polaschek and Ward (2002) offer a modification to this theory in that women are viewed as deceptive individuals who are fully aware of the innate differences between themselves and men and understand that their

needs are incompatible, but refuse to communicate their needs explicitly. The component which is theoretically best related to this theory is the third component (She Lied) of the IRMA, which involves women's deception as a means to capitalise on the competing goals they have with men. The component was not labelled as Women are unknowable as this was considered too broad a label. The title "She Lied" appropriately and succinctly captures the items within it.

Women as sex objects. This theory posits that women are in a "constant state of sexual reception" (Polaschek & Ward, 2002, p. 395) and exist to fulfil men's sexual desires. The belief is that women will always want sex, regardless of whether consent is given, and should always be available for sex. In 1984, Scully and Marolla noted that this ideology was a ''cultural view of women as sexual commodities, dehumanized and devoid of autonomy and dignity'' (p.542). This aspect of the theory may be most closely linked to the Victim devaluation component. However, it is unlikely that such unsavoury attitudes would be captured in a modern self-report measure. It is even less likely that researchers would consider including items representative of this ideology due to the high likelihood of socially desirable responding. For example, in the original IRMA (Payne et al., 1999), the 23rd item on the scale stated, "Some women prefer to have sex forced on them so they don't have to feel guilty about it," (p. 49). However, over time, researchers have become cognizant of the cultural shift toward political correctness and are creating scales that reflect more "subtle" myths (Gerger et al., 2007; McMahon & Farmer, 2011; Sleath & Bull, 2015).

Another aspect of this theory suggests that women do not know about their sexual needs and, thus, are unaware of their role as sexual objects. The belief is that even if a woman is saying "no", unconsciously she desires sex and men can pick up on their body language which alludes to this. It is suggested that a woman might respond to a sexual attack "with ready acceptance"

unable to resist the sexual assault (Freud, 1901, p. 143). This aspect of the theory most closely links with the Shift of responsibility/Minimisation of Harm component as this component represents beliefs that a woman should "fight hard enough" if she really does not want non-consensual sex.

The authors offer yet another, more specific, variant of this theory: women are "gatekeepers" to sexual gratification (see Koss & Cleveland, 1997). Women are viewed as unpredictable in whether they will grant access to men's sexual gratification or not; thus, in this version of the theory they are viewed as hostile. Polaschek and Ward (2002) describe the gatekeeper theory as it relates to consent:

> In the gatekeeper theory, women's consent functions like a gate that when open, provides a man or men with the opportunity to have sex with her. Because, in this theory, a woman is still viewed as existing in a continuously sexually receptive state, the gate's functioning is asymmetrical. It is assumed to be open, or easily opened, by default. Only extreme evidence to the contrary (e.g., very aggressive resistance) serves to disconfirm this theory, and then only temporarily. (p. 396)

This theory then results in men overgeneralising consent, assuming that consent on one occasion accounts for all and consent for one man is consent for all. Once the gate has been "opened" (i.e. consent has been given to a man) for sexual access, it is viewed as unreasonable for a woman to withdraw her consent anytime afterwards. This aspect of the theory can be matched to the Shift of Responsibility scale as it represents such ideas as "a woman should feel guilty following a rape''; ''a raped woman is a responsible victim, not an innocent one''; and ''any healthy woman can resist rape." (p. 396). Also, through stereotyping, some men will apply this ideology to

specific women, e.g. those ‘known’ to engaged in promiscuous behaviours. This would, then, more closely link to the Victim Devaluation/Promiscuity component.

Male sex drive is uncontrollable. The third implicit theory proposed by Polaschek and Ward (2002) is based on the belief that men’s libido is impossible to control and women are at the centre of this lack of control. Here, men’s sex drive is described as a “hydraulic mechanism” (p. 397) in that it is an unstoppable, overwhelming, and overpowering urge. The second component of the IRMA PCA maps directly on to this implicit theory, hence its labelling. However, when a man holds both this implicit theory as well as the gatekeeper theory as presented above, women are given the responsibility for not only provoking but also preventing rape. Women are believed to hold the key to men’s sexuality and the perpetrator is viewed as the true victim. In this way, the theory could be connected to the shift of responsibility/minimisation of harm component as there is minimisation of any psychological or emotional harm caused by sexual assault when overt, physical evidence is lacking.

Entitlement. Another implicit theory of rapists is that of Entitlement, i.e. that men should have all needs met upon demand. This includes sexual needs. All items under the Victim Devaluation/Promiscuity component fall under this theory as men make the call as to what is acceptable and unacceptable behaviour for women. Unsuitable conduct (e.g. dressing provocatively) may be met with punishment (e.g. rape). This is also another theory which researchers would have difficulty with measuring explicitly. For example, in the original IRMA, item 36 read “A woman who ‘‘teases’’ men deserves anything that might happen.” Subtle changes in wording rather than complete elimination may prove beneficial in current day studies (see McMahon & Farmer, 2011).

Dangerous World. The final implicit theory holds that the world is an inherently dangerous and hostile place. The authors note that this theory is often combined with Entitlement and together, they serve to justify and maintain harmful behaviours. In the realm of sexual offending, men who hold this implicit theory view everyone as an adversary regardless of age or gender. The authors also note that these beliefs are not often measured in scales which examine rape myths as the examples of this ideology are often present in attitudes towards violence scales, general cognitive measures, and antisocial personality tools. Concurrent with this, this theory is not present in the updated IRMA and is unlikely to be present in any current RMA scale. The updated IRMA shows a high victim focus as shown in the MDS solution but would need to have items representative of high perpetrator focus and high justification of rape elements to map onto the Dangerous World implicit theory. The Uncontrollable Sex Drive component is probably the most comparable, with the man's sex drive being the "danger"; this is represented in the MDS spatial structure.

As noted, these implicit theories are not independent of each other. Many, in fact, can be found to occur in tandem to form beliefs and attitudes about women, victims, rape, gender roles and other constructs. Similar to the current study's components, the implicit theories are all very much interlinked. This is to be expected as these beliefs are thought to all be part of the same cognitive processing.

Exploring multivariate structure of the IRMA using an alternating least-squares algorithm (ALSCAL) to perform multidimensional scaling

Factor analysis assumes that components can be fit to a solitary underlying dimension. Multidimensional scaling overcomes this assumption and allows for true exploration of the data

structure (Bishopp, 2003). Payne et al. (1999) discuss the importance of investigating the data structure, highlighting that rape myths are heterogenous and function differently across individuals and group. The ALSCAL output provided a spatial structure of the IRMA items.

Spatial structure. The scale items mapped on to the MDS solution almost exactly as they did in the four-component model, save for item 2 being clustered with the Shift of Responsibility items instead of the Victim Devaluation items. The structure matrix did highlight that item 2 contributed much more to factor 1 than the pattern matrix gave credit for. That said it is the highest of the cluster in the plot, i.e. closer to the justification of rape end of the dimension than the other items. Examining the item's content, it theoretically is more justifiable to include it in the Victim Devaluation cluster/component.

Looking at the plot, overall, items' coordinates in the spatial structure make theoretical sense. In considering just the end-points, we see that item 6 falls in the most extreme end of the Victim Devaluation cluster, while item 17 falls at the opposite end, the farthest end of the SL cluster. Additionally, items 7 and 13 lie at the end-points of their clusters, Male Sex Drive is Uncontrollable and Shift of Responsibility, respectively.

Looking at the plot without the sectioning, one could imagine item 6 as an outlier; however, the same could be argued for items in US. Theoretically, it makes sense to have 4 and 6 as a cluster, nearer the justification of rape end of the dimension rather than to label either an outlier and remove them from the scale. Read in conjunction with the pattern and structure matrices, we see that each item strongly contributes to their factor (Victim Devaluation) and can justifiably remain in the model. Dimensionality is described in further detail below.

Dimensionality. Dimensions in the ALSCAL solution were not rotated for interpretation, nor was this necessary. The resulting dimensions are psychologically meaningful and serve the

purposes of this study. The results paralleled the structure found in the PCA. The resulting solution was a 2-dimensional solution accounting for 82% of the variance among the data. Similar to Payne and colleagues' study (1999), the first dimension appeared to separate items by perpetrator versus victim focus and the second dimension separated items according to whether they deny vs justify rape. The Perpetrator vs Victim dimension ranges from items focused on characterisation of the perpetrator (left) to victim culpability (right), with item 7 being at the perpetrator end-point of the dimension and item 13 being at the victim end-point of the dimension. This dimension is focused on where the responsibility lies, although even when it is acknowledged that it is the perpetrator's fault, his behaviour is excused in some way. Hence, the majority of the scale clusters toward the victim end with fewer items having a perpetrator focus. The second dimension is labelled denial versus justification of rape with items nearer the bottom representing denial of rape, e.g. item 17, and items toward the top of the plot signifying justification of rape, with item 6 being at the very end-point of this dimension.

Testing the reliability of the IRMA, specifically its internal consistency

As explained above, the reliability of each component was tested and all were found to have high alpha values, indicative of internal consistency. As McMahon and Farmer (2011) removed items 5, 6, and 15 from their final scale, these alpha values were closely scrutinised to observe whether removal would have greatly affected the model. For item 5, "When girls get raped, it's often because the way they said "no" was unclear" The corrected item-total correlation was $\alpha = 0.68$ and the alpha value if the item was removed would have been $\alpha = 0.931$ as compared to component SR's alpha value of $\alpha = 0.934$. Removal of this item would not make a difference to the model; however, it does offer a decent contribution to the model. Item number

6, “If a girl initiates kissing or hooking up, she should not be surprised if a guy assumes she wants to have sex,” has a corrected item-total correlation of $\alpha = 0.46$. The alpha value of the component Victim Devaluation would actually increase (from 0.71 to 0.76), should this item be deleted. This would indicate that it might be beneficial to remove this item as this corresponds to the idea that item 6 is an “outlier” as may be interpreted from the MDS solution. However, in reviewing the structure matrix, it is shown that this item is a good estimate of the component Victim Devaluation. Also, its theoretical link to the component, as discussed previously, is psychologically meaningful. Lastly, item 15, “A rape probably doesn’t happen if a girl doesn’t have any bruises or marks,” had a corrected item-total correlation of $\alpha = 0.86$, while the Cronbach’s alpha if it were to be deleted would be $\alpha = 0.921$ (as compared to an alpha value of $\alpha = .934$ for the component Shift of Responsibility). In this case, the reliability of the component would decrease should this item be removed. This corresponds with the item’s structure coefficient, 0.92, which indicates that item 15 is one of the best estimates for the component, Shift of Responsibility, out of nine items. While it statistically made sense for McMahon and Farmer (2011) to remove those items, it appears they would still be of benefit. Also, when removed from the scale, as mentioned previously, the variance explained decreased.

Limitations

The study is not with its limitations. Firstly, the sample of undergraduate students may make generalisation difficult, and one could argue that conducting the research with individuals convicted of sex offending would give a truer representation of RMA and increase clinical utility. Additionally, the data from sexual offenders might provide a model more parsimonious with Polaschek and Ward’s (2002) Implicit Theories. However, university students are

overrepresented in the literature of sexual violence perpetration and victimisation (e.g., Fitzgerald et al., 1996; Koss, Gidycz, & Wisniewski, 1987; Muehlenhard & Linton, 1987), thus this is a population of high interest. Payne et al. (1999) also noted that individuals within this developmental phase are highly exposed and sensitive to shifts in culture, including violence, sexuality, and gender roles and that research should not look to eliminate this as a study population, but should attempt to focus on the cultural issues within this group and determine the limits of generalisability to other populations. Furthermore, the fact that within the current population produced a dataset which could be applied to implicit theories of sexual offending may be indicative that some implicit theories are present in the non-offending population. Implicit theories are posited to be implemented in childhood and thus, further research should be carried out to explore the factor structure within this population to assess replicability.

Another limitation of this study involved the nature of the IRMA and is a common factor of all RMA scales. As mentioned in chapter 3, these measures are, necessarily, bound by culture and time and are at risk of being outdated within a few years. As Payne et al. (1999) note, this is virtually unavoidable due to the nature of sexual language which relies heavily on colloquialisms. Moreover, these colloquialisms may not translate cross-culturally. It would be helpful for future research to explore culturally appropriate ideas and phraseology in a way that maintains the integrity of the measure.

A further limitation with regard to data collection, is that a measure for response bias was not included in the participants' survey, thus it is unclear whether there was elevated self-presentation bias within the sample. Self-report questionnaires, especially those focused around contentious topics such as rape myths, are liable to socially desirable responding. Future studies might benefit from regularly including deception scales in addition to the IRMA as populations

will have various levels of socially desirable responses for different reasons (e.g. sexual offenders versus police officers).

In regards to the data analysis, it is possible that ALSCAL could lead to unreliable results due to applying a least squares method to non-linear structures (D. Bishopp, personal communication, July 17th, 2017). Alternatively, a smallest space analysis (Guttman, 1968) may be useful to explore in the future.

Moreover, it is a tenet of factor analysis and multidimensional scaling that there is some flexibility in deciding what a factor and spatial structure represent. The interpretation of structure falls within the hands of the researchers conducting the analyses. Therefore, even though these analytic procedures offer a guide for structure and researchers are expected to link their interpretations to an evidence base, the process is very subjective. Features of the factor structure and MDS output could produce the same structures but across different populations, researchers could ascribe different interpretations to the data (Bishopp, 2003). For example, if sex offenders were participating in the study, the interpretation might focus more on their motivation for offending. Additionally, MDS solutions can be rotated to suit the researcher's needs, which can bring the technique's evidentiary value and accuracy into question. It may be best to use MDS as a supplementary technique to other multidimensional unfolding solutions and include the use of unidimensional analyses.

Implications

Despite the above limitations, the current study has implications for research and clinical practice.

Dimensionality

As noted in this research study, most items on the IRMA scale cluster toward the victim end of dimension 1 in the MDS solution with fewer items having a perpetrator focus. Future updates of the IRMA should look to include items with more of a perpetrator focus as, currently, most scales are heavily focused on the act of rape or victims of rape and do not offer in-depth assessment of the range of myths about perpetrators of rape. Niemi and Young (2016) note that changing the focus of such statements from victims to perpetrators are likely to result in lower rates of victim blaming. Thus, it would be informative to explore various levels of endorsement across populations and cultures using items with more of a perpetrator focus (such as, men deserve sex from women, testosterone levels are linked to sexual violence, men are predisposed to violence, and men are likely to misinterpret sexual signals from women because women are unknowable; Polaschek & Ward, 2002). It could be argued that of the individuals in the population who highly endorse rape myths, the difference between those who go on to offend and those who do not may lie in how strongly they believe a perpetrator is justified in their actions. Adding items with more of an emphasis on the perpetrator may highlight these differences. Further exploration is needed to determine what it is about a sexual offender that sets them apart from non-offenders who may hold the same myths but never commit rape. This will then shape more appropriate intervention responses.

When analyzing RMA data, although Payne et al. (1999) advised against treating RMA as a completely unidimensional construct, it is continually treated as such within the literature. Previous research has failed to find differences in RMA between rapists and child sexual offenders and to predict recidivism of sexual offending using RMA (Johnson & Beech, 2017) and this may be due in part to the limited view and application of RMA as multidimensional. RMA as a unidimensional construct may not be effective in measuring subtle differences

amongst offenders or estimating recidivism rates; thus, use of RMA subscales may be more appropriate. Past research (as reported in Chapter 3) and the current study highlight its multidimensional nature. It is most beneficial to consider both uni- and multidimensionality of the RMA construct (Payne et al., 1999).

Factor Stability

As mentioned previously, the differences in factor structure found between McMahon and Farmer's (2011) study and the current research could indicate that the IRMA is not a stable measure. Additionally, research has found a shift in IRMA factor structure pre- and post-intervention (Baldwin-White et al., 2016). Keeping in mind its multidimensional nature, one could argue that a change in RMA factor structure is expected with a decrease in RMA endorsement. However, future research should explore alterations in factor structure pre- and post-treatment to observe the different structural changes that occur in response to specific treatments. If there is evidence of how a tool's factor structure might change in response to treatment, it may be possible to use this to predict outcome (i.e., changes in RMA or predictions of recidivism).

Implicit and explicit measurement of RMA

With regard to measuring recidivism rates, the current research highlights that RMA has not been found to be a significant predictor of reoffending (Johnson & Beech, 2017; Olver et al., 2014); however, it has been found to be linked to rape proclivity (Baldwin-White et al., 2016; Chapleau & Oswald, 2010; Edwards et al., 2011). It would be logical, then, to conclude that RMA is mediated by another factor or other factors which allow for the strong correlation with rape proclivity but not with sexual reoffending. One explanation could be that the use of explicit measures allows for high levels of socially desirable responding (Gerger et al.,2007). It may be

of benefit to use implicit measures to uncover mediators for the effect of RMA on sexual offending and reoffending and to avoid impression management.

Furthermore, research has failed to find differences between child molesters and adult rapists and between rapists and sexual murders in their endorsement of RMA (Beech et al., 2006; Pithers, 1994). Arguably, this could be attributed to the poor ability of explicit measures to elicit subtle differences between the groups of offenders (Gerger et al., 2007). If one were to compare the implicit theories of child sexual offenders (such as those proposed in Marziano, Ward, Beech, Pattison, 2006) and those of adult sexual offenders as detailed by Polaschek and Ward (2002), there would be few theoretical differences observed. Thus, it might be difficult to pinpoint exactly where the cognitive differences lie between adult sexual offenders and child molesters when relying on explicit measures, especially when, in theory, the two types of offenders' cognitive schemas and behavioural justifications are quite similar.

As mentioned in chapter 3, combining an explicit measure of RMA with an implicit measure may add to the validity and reliability of the measure. Ideally, this would countermand socially desirable responding and possible allow for truer representations of not only individuals' RMA endorsement but also denote whether RMA is a predictor of recidivism and treatment efficacy and if subtle differences are present amongst offenders.

Risk, Need, and Responsivity

The Risk-Need-Reponsivity model of offender rehabilitation (Andrews & Bonta, 2006) highlights essentially why more research on the heterogeneity of sexual offenders and assessment of their needs is warranted. The risk principle emphasises that more intensive treatment be applied to higher risk offenders. The need principle states that rehabilitation should target the needs of a particular offender; targeting the criminogenic needs of each rapist will

enhance intervention efficacy (Reid, Wilson, & Boer, 2011). Sexual offenders who endorse the ideology that their sex drive is uncontrollable as compared to those who minimise harm done to the victim would require specialised programming to address each of these needs. Furthermore, there may be subgroups of rapists which present similarly to child sexual offender. Therefore, when all types of rapists are referenced as a solitary group to be compared to child sexual offenders, this may mask any possible differences between specific categories of rapists and child sexual offenders.

Moreover, Chapter 2 demonstrated that the literature is mixed on whether RMA is a treatment need for sexual offenders. This calls for expansion of our understanding of RMA and its utility (i.e., used as a justification for behavior or a disinhibitor) within the offending population. Acknowledging this gap in knowledge, researchers should refrain from relying solely on RMA to determine a person's risk of reoffending. Rapists with different motives (e.g., sexually motivated versus motivated by anger) might endorse different rape myths and likely use different myths to justify their actions. Further research might look to extend Beech et al.'s (2006) research by linking rape myth subscales to types of rapists in an attempt to explore whether higher endorsement in certain subscales is correlated to rapist type.

Additionally, in the general population, rape awareness education should look to not only address rape myths generally but also address their heterogeneity. It may be that within particular cultures or in different populations of individuals (e.g. police officers versus social workers versus university students), the elevations in RMA differ. This discrepancy between groups would need to be addressed in education programmes, most easily by addressing all identified components of the RMA construct.

Intervention

The widespread epidemic of rape myth acceptance has been discussed and due to its prevalence throughout the general population – regardless of social standing, i.e., offenders and non-offenders – it would be beneficial to implement programmes which address stereotypes and misunderstandings about sex, gender roles, and rape at the early stages of development – namely, within primary and secondary education. For example, Malo-Juvera (2014) researched the use of young adult literature to observe the effects this would have on adolescent (13 to 17 years of age) children's RMA. Using a young adult novel in which the adolescent, female protagonist is a victim of rape, students were asked to engage in open dialogue and to write essays about how they might explain to the protagonist that she was raped, convince her to seek help, and how they would attempt to convince the perpetrator that he raped the victim. The researcher found that through discussion-based instruction, such literary devices could lower adolescents' endorsement of rape myths. This teaching could be generalised to primary school curricula, with less controversial material. It might be beneficial to introduce social issues around gender roles and expectations within primary education curricula.

In tertiary level education, intervention most often cited in the literature is bystander intervention programming which has been found to have a positive effect on RMA (Bratcher, 2011; Lee et al., 2003; McMahon, 2010; Rich et al., 2010; Rothman & Silverman, 2007). Bystander intervention programmes are likely effective in reducing RMA because they take into account the "continuum of social contexts and practices" (Lievore, 2003, p. 112) which justify sexual assault in ways that are seen as "normal." These intervention programmes require individuals to challenge societal constructions of gender roles.

In research with convicted offenders, RMA has been found to be responsive to treatment (Olver et al., 2014; Pithers, 1994; Webster et al., 2004); however, it is difficult to know whether

this is linked to lower risk of reoffending. Past research indicates that this is not the case (Beech et al., 2006; Johnson & Beech, 2017; Olver et al., 2014). Endorsement of rape myths is so deeply intertwined with society and societal values, it would be presumptuous to assume that RMA on its own might predict recidivism. As such, research should refrain from treating RMA as limited to the "forensic" sphere.

Lievore (2003) notes that sex offender treatment must go beyond specific cognitive distortions and should involve understanding the "causes, prevalence, and consequences of sexual violence" (p. 114). Lievore presents a similar suggestion to the current author in that social measures will be required to reduce sexual offending due to the interaction between sexual offending, violence, and gender roles. Lievore (2003) suggests the following social interventions:

- Public campaigns which raise awareness of sexual violence and how it is supported in society – from its presence the media to interpersonal relationships;
- On a societal level, including through educational instruction, challenging rape myths and the ideology of femininity as passive and masculinity as aggressive;
- Building healthy attitudes about sexuality, gender, and interpersonal relationships and providing positive models of these;
- Educating children around and protecting them from hypermasculinity and witnessing or being subject to sexual violence;
- Taking a wider, contextual approach to interventions with offenders considering the societal impact of their actions.

The mixed research on RMA as a criminogenic need and its links to sexual recidivism highlight the need for further investigation into whether there is need for specialised treatment of offenders and, if so, identification of which group of sexual offenders require specialised

treatment. Additionally, further clarification is needed on which treatment programmes contribute most to changes in RMA and whether it is necessary that these programmes are run within institutional settings or whether a more societal approach would be better suited for reduction in rape myths and consequently, sexual recidivism.

Conclusions

The purpose of this study was to explore the factor structure and dimensionality of the IRMA. A four-component solution was obtained through principal component analysis producing the following four factors: Shift of Responsibility/Minimisation of Harm, Male Sex Drive is Uncontrollable, She Lied, and Victim Devaluation/Promiscuity. These four components appear to reflect psychological theory which offers conceptualisations of rape supportive beliefs and attitudes, specifically Polaschek and Ward's (2002) implicit theories for sexual offending. Examining the construct validity of a measure is a somewhat subjective task, however, Kline (2000) notes that this cannot be conclusively proven. The current study's findings offer support for a multidimensional structure. The findings also provide support for the internal consistency of the four components. Of benefit would be to have future studies test whether the factor structure and spatial structure found in this study can be replicated with a similar population and with different populations. Also, development of items with a perpetrator focus might offer another dimension to the RMA construct which is understudied in the field.

CHAPTER 5: GENERAL DISCUSSION

The aim of this thesis was to disseminate research on RMA and its clinical utility, establish its role and relevance within society and clinical practice, and to evaluate the factor structure and dimensionality of a well-established measure or RMA, the Illinois Rape Myth Acceptance Scale (IRMA). The key findings of each chapter in relation to these aims are discussed below.

Key findings from Chapter 2: A Systematic Review of the Literature

The systematic review set out to explore the literature on rape myth acceptance in adult males who have committed sexual assaults against women. The review sought to determine if measures of RMA could accurately discriminate between rapists, child molesters, non-sexual offenders and offenders and if elevations on RMA measures varied within rapist typologies. The review also explored whether there were differences in RMA endorsement between rapists who reoffended and those who did not. The final aim of the review was to evaluate the responsiveness of RMA to sex offender treatment programmes.

The review found that rapists may be distinguished from other non-sexual offenders and from community non-offending males on measures of rape myth acceptance, especially on the sex-role subscales of the Burt Rape Myth Acceptance Scale (Cohen, 2012). However, the reliability and consistency of these results was questionable. It appeared that RMA could not offer clear distinction between rapists and child molesters in the data and, unexpectedly, in Cohen's (2012) study, sex offenders endorsed fewer rape myths than non-offenders and found that this was not attributed to socially desirable responding.

Another key finding was that in Beech et al.'s (2006) study, the researchers found that opportunistic rapists scored significantly higher than both the sexual sadistic and sexual non-sadistic types on the Adversarial Sexual Beliefs subscale of Burt's Rape Myth Acceptance Scale. They also found that, overall, sexual non-sadistic rapists had the lowest scores on the scale. Socially desirable responding was assessed in this study and this was found to be a non-factor, suggesting that sexual non-sadistic rapists hold less negative beliefs about sexual relationships when compared to other groups of rapists. Furthermore, a study on sexual murderers (Stefanska et al., 2015) reported that offenders who identified problems with being open to others and believing that women are deceitful, were more likely to endorse sexual entitlement beliefs and rape myths than offenders who did not report these problems.

With regard to whether differences in levels of RMA could distinguish between sexual recidivists and sexual non-recidivists, there were no studies included in the review which found RMA to be a significant predictor of recidivism (violent or sexual), nor did any study directly compare sexual reoffenders with those who did not reoffend.

In exploring whether RMA treatment was responsive to sex offender treatment programmes, four studies were found to investigate this. Key findings showed that, overall, as a homogenous group, rapists showed no effect of treatment on levels of RMA. However, when RMA was examined at the typological level, analysis revealed that scores on the Sex Role Stereotyping scale increased (i.e. more items were endorsed) post-treatment. Additionally, Pithers (1994) found a significant decrease in RMA after individuals partook in a victim empathy programme. Similarly, Olver et al. (2014) found that RMA levels decreased significantly in response to the Clearwater programme.

It was highlighted in the review that, due to conservative exclusion criteria, the range of RMA measures was severely limited, thus leading to the exclusion of arguable the most widely used tool in RMA measurement, the Illinois Rape Myth Acceptance Scale (IRMA; Payne et al., 1999). The tool was critiqued on its psychometric properties in the third chapter.

Key findings from Chapter 3: A Psychometric Critique

This critique examined the IRMA in terms of its psychometric properties and utility in research and practice. The IRMA appears to demonstrate a reasonable degree of reliability and validity but is not without its shortcomings, e.g. poor construction of the scoring of Likert items and lack of discrimination between victims of stranger versus acquaintance rape.

Also of note was that when conducting the search for measure bias and/or distortion, many research studies appeared to have accepted that the IRMA is reliable and valid without carrying out any checks for validity and reliability. Another odd finding was that although the IRMA had been updated in 2011, there is very little evidence of the updated version being used in current research, even when attempting to create new RMA tools.

The critique concluded that the IRMA satisfies the minimum level of measurement and there is some evidence of its reliability and validity as a measure of rape myth acceptance. Overall, it is a good measure of RMA.

Key findings from Chapter 4: Empirical Research Study

The most notable finding of Chapter 4 was that a four-component solution was produced from the principal components analysis of the IRMA, which was confirmed in a multidimensional spatial solution. These four components mapped on to implicit theories

proposed by Polaschek and Ward (2002), however multiple components were found to fit under multiple implicit theories, as expected.

The components showed high internal consistency, indicating that the IRMA in this study has good reliability. The chapter highlights that future research will reap the most benefits from the IRMA by examining RMA as both a general unidimensional construct and as multidimensional.

Implications for Practice and Research

The central message for researchers and practitioners using RMA measures to guide studies, assessment, and treatment is to assume heterogeneity of RMA as a construct and try to examine its underlying structure before using it for assessment. Also, researchers and practitioners must be cognizant of the fact that factor structure and dimensionality of RMA measures will change over time and will present differently across different sub-groups of the population and different cultures. Understanding its heterogeneity will help to guide tailored education and treatment programmes for the general public and offenders.

Although the IRMA has demonstrated some reliable and valid psychometric properties, there is still a need for current research to examine its construct validity and its reliability as a measure. The updated IRMA is underrepresented in current literature and is not reflective of the shifts in culture, potentially skewing results obtained.

Conclusions

Rape myths are damaging to the fabric of society as it is represented in the beliefs and attitudes of not only those who sexually offend but victims, members of the criminal justice

system, and the lay public. Repeated measurement, validation, reliability testing, and application of RMA measures is necessary to ensure that RMA is being examined against the ever-changing culture and social climate. Additionally, researchers should aim to further develop indirect measures that assess levels of RMA to complement the explicit measures already in use and further inform research and practice.

References

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,* 89-103.

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.

Amar, A. F., Sutherland, M., & Kesler, E. (2012). Evaluation of a bystander education program. *Issues in Mental Health Nursing, 33,* 851-857.

Amnesty International. (2005). *Amnesty International Report 2005 – United Kingdom.* UK: Amnesty International.

Andrews, D. A., & Bonta, J. (2006). *The Psychology of Criminal Conduct* (4th ed.). Cincinnati: Anderson Publishing.

Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology, 28,* 369-404.

Avais, M. A., Chandio, R. A., & Brohi, A. A. (2014). Rape myths: An overview of perceptions and beliefs. *Academic Research International, 5,* 356-360.

Baldwin-White, A., Thompson, M. S., & Gray, A. (2016). Pre-and postintervention factor analysis of the Illinois Rape Myth Acceptance Scale. *Journal of Aggression, Maltreatment & Trauma*, *25*, 636-651.

Beazley Richards, J. (2000). *A comparison of attitudes of convicted violent sexual offenders, convicted violent non-sexual offenders and non-offenders* (Unpublished Masters dissertation). University of Leicester, UK.

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

Beech, A., Oliver, C., Fisher, D. & Beckett, R.C. (2006). *STEP 4: The Sex Offender Treatment Programme in prison: Addressing the needs of rapists and sexual murderers.* Birmingham: University of Birmingham. Available electronically from www.hmprisonservice.gov.uk/assets/documents/100013DBStep_4_SOTP_report_2005.pdf.

Bell, R. C. (1997). *Using SPSS to analyse repertory grid data*. Melbourne: University of Melbourne.

Bishopp, D. C. F. (2003). *Dimensions of sexual aggression* (Doctoral dissertation). Retrieved from Surrey Research Insight Open Access.

Blake, B., Schulze, S., & Hughes, J. M. (2003). *Perceptual Mapping by Multidimensional Scaling: A Step by Step Primer*. Research Reports in Consumer Behavior. Cleveland, OH: Cleveland State University.

Blake, E., & Gannon, T. A. (2012). The implicit theories of rape-prone men: An information-processing investigation. *International Journal of Offender Therapy and Comparative Criminology, 54*, 895.

Bohner, G., Jarvis, C., Eyssel, F., & Siebler, F. (2005). The causal impact of rape myth acceptance on men's rape proclivity: Comparing sexually coercive and noncoercive men. *European Journal of Social Psychology, 35*, 819-828.

Bohner, G., Reihnard, M., Rutz, S., Sturm, S., Kerschbaum, B., Effler, D. (1998). Rape myths as neutralizing cognitions: Evidence for a causal impact of anti-victim attitudes on men's self-reported likelihood of raping. *European Journal of Social Psychology, 28,* 257-268.

Bohner, L., & Lampridis, E. (2004). Expecting to meet a rape victim affects women's self-esteem: The moderating role of rape myth acceptance. *Group Process and Intergroup Relations, 7*, 77–88.

Borg, I., & Groenen, P. (1997). Multitrait-multimethod by multidimensional scaling. *Softstat*, *97*, 59-66.

Bratcher, A. (2011). The effect of sexual assault awareness program on rape myth acceptances. In *Academic Creativity And Excellence (Ace) Day Proceedings* (pp. 48–56). Eugene, Oregon. Retrieved from http://www.nwcu.edu/wp-content/uploads/2014/10/ACE-Day-Proceedings-2010-2011.pdf#page=52

Branscombe, N., Wohl, M., Owen, S., Allison, J., & N'gbala, A. (2003). Counterfactual thinking, blame assignment, and well-being in rape victims. *Basic and Applied Social Psychology, 25*, 265–273.

Briere, J., & Malamuth, N. M. (1983). Self-reported likelihood of sexually aggressive behavior: Attitudinal versus sexual explanations. *Journal of Research in Personality, 17*, 315-323.

Brownmiller, S. (1975). *Against our will.* New York, NY: Simon & Schuster.

Buddie, A., & Miller, A. (2001). Beyond rape myths: A more complex view of perceptions of rape victims. *Sex Roles, 45*, 139–160.

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*, 37-54.

Burt, M. R. (1980). Cultural myths and supports for rape. *Journal of Personality and Social Psychology, 38*, 217.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*, 629-637.

Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an N-way generalization of ‘‘Eckart-Young’’ decomposition. *Psychometrika, 35*, 283–319.

Chapleau, K. M., & Oswald, D. L. (2010). Power, sex, and rape myth acceptance: Testing two models of rape proclivity. *Journal of Sex Research, 47*, 66–78.

Chapleau, K. M., & Oswald, D. L. (2013). Status, threat, and stereotypes: Understanding the function of rape myth acceptance. *Social justice research, 26*, 18–41.

Chapleau, K. M., Oswald, D. L., & Russell, B. L. (2007). How ambivalent sexism toward women and men support rape myth acceptance. *Sex Roles, 57,* 131–136.

Check, J. V. P., & Malamuth, N. (1985). An empirical assessment of some feminist hypotheses about rape. *International Journal of Women's Studies, 8*, 414-423.

Chiroro, P., Bohner, G., Viki, G. T., & Jarvis, C. I. (2004). Rape myth acceptance and rape proclivity: Expected dominance versus expected arousal as mediators in acquaintance-rape situations. *Journal of Interpersonal Violence, 19,* 427–442.

Cohen, D. A. (2012). *Attention bias to offense related stimuli in rapists and child molesters* (Unpublished doctoral dissertation). Bar-Ilan University, Israel.

Cohen, W. (1988). *Statistical power analysis for the behavioural sciences: Second edition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cooke, A., Smith, D., & Booth, A. (2012). Beyond PICO: The SPIDER Tool for qualitative evidence synthesis. *Qualitative Health Research, 22*, 1435-1443.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of applied psychology, 78*, 98.

Crawford, A., Green, S. B., Levy, R., Lo, W.-J., Scott, L., Svetina, D. S., & Thompson, M. S. (2010). Evaluation of parallel analysis methods for determining the number of factors. *Educational and Psychological Measurement, 70*, 885–901.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297-334.

Cuklanz, L. M. (2000). *Rape on prime time: Television, masculinity, and sexual violence*. Philadelphia, PA: University of Pennsylvania Press.

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.

Durán, M., Moya, M., Megías, J., & Viki, G. (2010). Social perception of rape victims in dating and married relationships: The role of perpetrator's benevolent sexism. *Sex Roles, 62,* 505–519.

Edward, K. E., & McLeod, M. D. (1999). The reality and myth of rape: Implications for the criminal justice system. *Expert Evidence, 7,* 37 –58.

Edwards, K. M., Turchik, J. A., Dardis, C. M., Reynolds, N., & Gidycz, C. A. (2011). Rape myths: History, individual and institutional-level presence, and implications for Change. *Sex Roles, 65,* 761–773.

Eyssel, F., & Bohner, G. (2011). Schema effects of rape myth acceptance on judgments of guilt and blame in rape cases: The role of perceived entitlement to judge. *Journal of Interpersonal Violence, 26,* 1579–1605.

Eyssel, F., Bohner, G., & Siebler, F. (2006). Perceived rape myth acceptance of others predicts rape proclivity: Social norm or judgmental anchoring?. *Swiss Journal of Psychology/Schweizerische Zeitschrift für Psychologie/Revue Suisse de Psychologie, 65*, 93-99.

Feild, H.S. (1978). Attitudes toward rape: A comparative analysis of police, rapists, crisis counselors, and citizens. *Journal of Personality and Social Psychology*, *36*(2), 156-179.

Field, A. (2009). Logistic regression. *Discovering statistics using SPSS*. London, UK: Sage Publications.

Fisher, B. S., Daigle, L. E., & Cullen, F. T. (2010). What distinguishes single from recurrent sexual victims? The role of lifestyle-routine activities and first-incident characteristics. *Justice Quarterly, 27*, 102–129.

Fitzgerald, L. F., Waldo, C. R., Cortina, L. M., Radhakrishnan, P., Reid, L. D., & Swan, S. (1996). *Campus climate at the University of Illinois at Urbana-Champaign: The effects of gender, race/ethnicity and sexual orientation.* Technical report submitted to the University of Illinois at Urbana-Champaign.

Fletcher, G. J. O. (1995). *The scientific credibility of folk psychology.* Mahwah, NJ: Erlbaum.

Franiuk, R., Seefelt, J., & Vandello, J. (2008). Prevalence of rape myths in headlines and their effects on attitudes toward rape. *Sex Roles, 58*, 790–801.

Freud, S. (1901). *Psychopathology of everyday life*. London, UK: Ernest Benn.

Gerger, H., Kley, H., Bohner, G., & Siebler, F. (2007). The acceptance of modern myths about sexual aggression scale: Development and validation in German and English. *Aggressive Behavior, 33*, 422-440.

Graham, J. M., Guthrie, A. C., & Thompson, B. (2003). Consequences of not interpreting structure coefficients in published CFA research: A reminder. *Structural equation modeling: A multidisciplinary journal, 10*, 142–153.

Gray, J. M. (2006). Rape myth beliefs and prejudiced instructions: Effects on decisions of guilt in a case of date rape. *Legal and Criminological Psychology, 11*, 75–80.

Grayson, D. (2004). Some myths and legends in quantitative psychology. *Understanding statistics*, *3*, 101-134.

Greedlinger, V., & Byrne, D. (1987). Coercive sexual fantasies of college men as predictors of self-reported likelihood to rape and overt sexual aggression. *Journal of Sex Research*, *23*, 1-11.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of personality and social psychology*, *97*, 17.

Grubb, A., & Turner, E. (2012). Attribution of blame in rape cases: A review of the impact of rape myth acceptance, gender role conformity and substance use on victim blaming. *Aggression and Violent Behavior*, *17*(5), 443-452.

Guttman, L. (1968). A general nonmetric technique for finding the smallest coordinate space for a configuration of points. *Psychometrika*, *33*, 469-506.

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.

Hare, R. D. (1991). *The Hare psychopathy checklist-revised: Manual*. Multi-Health Systems, Incorporated.

Hayes, R. M., Lorenz, K., & Bell, K. A. (2013). Victim blaming others. *Feminist criminology*, *8*, 202–220.

Hayes-Smith, R., & Levett, L. (2010). Student perceptions of sexual assault resources and prevalence of rape myth attitudes. *Feminist Criminology, 5*, 335–354.

Heath, N. M., Lynch, S. M., Fritch, A. M., & Wong, M. M. (2013). Rape myth acceptance impacts the reporting of rape to the police. *Violence Against Women, 19*, 1065-1078.

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*, 34-53.

Higgins, J. P. T. & Green, S. (2011). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration. Available from www.cochrane-handbook.org.

Hinck, S. S., & Thomas, R. W. (1999). Rape myth acceptance in college students: How far have we come? *Sex Roles, 40,* 815– 832.

IBM Corp. (2016). *IBM SPSS Statistics for Windows, Version 24.0*. Armonk, NY: IBM Corp.

Johnson, B. E., Kuck, D. L., & Schander, P. R. (1997). Rape Myth Acceptance and Sociodemographic Characteristics: A Multidimensional Analysis. *Sex Roles, 36*, 693–707.

Johnson, L. G., & Beech, A. (2017). Rape myth acceptance in convicted rapists: A systematic review of the literature. *Aggression and Violent Behavior, 34*, 20-34.

Jones, C., & Aronson, E. (1973). Attribution of fault to a rape victim as a function of respectability of the victim. *Journal of Personality and Social Psychology, 26*, 415–419.

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

Kelly, L., Lovett, J., & Regan, L. (2005). A gap or a chasm. Attrition in reported rape cases. Retrieved from https://www.researchgate.net/profile/Jo_Lovett/publication/238713283_Home_Office_Research_Study_293_A_gap_or_a_chasm_Attrition_in_reported_rape_cases/links/00b7d52a09b4935e0e000000/Home-Office-Research-Study-293-A-gap-or-a-chasm-Attrition-in-reported-rape-cases.pdf

King, L., & Roberts, J. (2011). Traditional gender role and rape myth acceptance: From the countryside to the big city. *Women & Criminal Justice, 21*, 1–20.

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

Kline, P. (2000). *A psychometrics primer*. London, UK: Free Association Books.

Kline, P. (2013). *Handbook of psychological testing*. London, UK: Routledge.

Knight, R. A., & Prentky, R. A. (1990). Classifying sexual offenders: The development and corroboration of taxonomic models. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.), *Handbook of sexual assault: Issues, theories, and treatment of the offender* (pp. 23–52). New York: Plenum.

Koss, M. P., & Cleveland, H. H. (1997). Stepping on toes: social roots of date rape lead to intractability and politicization. In M. D. Schwartz (Ed.), *Researching sexual violence*

against women: methodological and personal perspectives (pp. 4–21). Thousand Oaks, CA: Sage.

Koss, M. P., Gidycz, C. A., & Wisniewski, N. (1987). The scope of rape: Incidence and prevalence of sexual aggression and victimization in a national sample of higher education students. Journal of *Consulting and Clinical Psychology, 55*, 162–170.

Kruskal, J. B. (1964) Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, *29*, 1-27.

Lee, R., Caruso, M. E., Goins, S. E., & Southerland, J. P. (2003). Addressing Sexual Assault on College Campuses: Guidelines for a Prevention / Awareness Week. *Journal of College Counseling, 6*, 14.

Lievore, D. (2004). *Recidivism of sexual assault offenders: Rates, risk factors and treatment efficacy*. Canberra: Australian Institute of Criminology.

Lonsway, K. A., & Fitzgerald, L. F. (1994). Rape myths: In review. *Psychology of Women Quarterly, 18*, 133-164.

Lonsway, K. A., & Fitzgerald, L. F. (1995). Attitudinal antecedents of rape myth acceptance: A theoretical and empirical reexamination. *Journal of Personality & Social Psychology, 68,* 704–711.

Lutz-Zois, C., Moler, K. A., & Brown, M. J. (2015). Mechanisms for the relationship between traditional masculine ideologies and rape myth acceptance among college men. *Journal of Aggression, Maltreatment & Trauma*, *24*, 84-101.

Malamuth, N. M. (1989a). The attraction to sexual aggression scale: Part one. *Journal of Sex Research, 26,* 26-49.

Malamuth, N. M. (1989b). The attraction to sexual aggression scale: Part two. *Journal of Sex Research, 26,* 324-354.

Malamuth, N. M., & Ceniti, J. (1986). Repeated exposure to violent pornography: likelihood of raping ratings and laboratory aggression against women. *Aggressive Behavior, 12,* 129-137.

Malamuth, N., & Check, J. (1985). The effects of aggressive pornography on rape myths: Individual differences. *Journal of Research in Personality, 19,* 299-320.

Malo-Juvera, V. (2014). Speak: The effect of literary instruction on adolescents' rape myth acceptance. *Research in the Teaching of English, 48,* 407-427.

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.

Marshall, W. L., & Hambley, L. S. (1996). Intimacy and loneliness, and their relationship to rape myth acceptance and hostility toward women among rapists. *Journal of Interpersonal Violence, 11,* 586-592.

Maruna, S., & Mann, R. E. (2006). A fundamental attribution error? Rethinking cognitive distortions. *Legal and Criminological Psychology, 11,* 155-177.

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,* 97-105.

McMahon, S. (2010). Rape myth beliefs and bystander attitudes among incoming college students. *Journal of American College Health, 59*, 3-11.

McMahon, S., & Farmer, G. (2011). An updated measure for assessing subtle rape myths. *Social Work Research, 35*, 71–81.

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

Muehlenhard, C. L., & Linton, M. A. (1987). Date rape and sexual aggression in dating situations: Incidence and risk factors. *Journal of Counseling Psychology, 34*, 186–196.

Niemi, L., & Young, L. (2016). When and why we see victims as responsible. *Personality and Social Psychology Bulletin, 42*, 1227-1242.

Nunes, K. L., Hermann, C. A., & Ratcliffe, K. (2013). Implicit and explicit attitudes toward rape are associated with sexual aggression. *Journal of Interpersonal Violence, 28*, 2657-2675.

Nunnally, J. (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*, 130-141.

Offending Behaviour Programmes Unit. (1995). *Rape Myths scale.* Unpublished. National Offender Management Service: Ministry of Justice, London, England.

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

Overholser, J. C., & Beck, S. (1986). Multimethod assessment of rapists, child molesters, and three control groups on behavioral and psychological measures. *Journal of Consulting and Clinical Psychology*, *54*, 682.

Pallant, J. (2016). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS. 6th edition.* New South Wales, Australia: Allen & Unwin.

Parratt, K. A., & Pina, A. (2017). From "real rape" to real justice: A systematic review of police officers' rape myth beliefs. *Aggression and Violent Behavior, 34*, 68-83.

Paul, L., Gray, M., Elhai, J., & Davis, J. (2009). Perceptions of peer rape myth acceptance and disclosure in a sample of college sexual assault survivors. *Psychological Trauma: Theory, Research, Practice and Policy, 1*, 231–241.

Pauna, A., & Pleszewski, Z. (2012). Clinical and Methodological Issues in the Research on the Rape Myth. *International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering, 6*, 1088–1095.

Payne, D. L., Lonsway, K. A., & Fitzgerald, L. F. (1999). Rape myth acceptance: Exploration of its structure and its measurement using the Illinois Rape Myth Acceptance Scale. *Journal of Research in Personality, 33*, 27–68.

Peterson, Z., & Muehlenhard, C. (2004). Was it rape? The function of women's rape myth acceptance and definitions of sex in labeling their own experiences. *Sex Roles, 51,* 129–144.

Pithers, W. D. (1994). Process evaluation of a group therapy component designed to enhance sex offenders' empathy for sexual abuse survivors. *Behavioral Research and Therapy, 32,* 565–570.

Pithers, W. D., Martin, G. & Cimming, G. F. (1989). The Vermont Treatment Program-for sexual aggressors. In D. R. Laws, (Ed.), *Relapse prevention with sex offenders* (pp. 292-310). New York: Guilford Press.

Polaschek, D. L. L., & Ward, T. (2002). The implicit theories of potential rapists: What our questionnaires tell us. *Aggression and Violent Behavior, 7,* 385-406.

Polaschek, D., & Gannon, T. A. (2004). The implicit theories of tapists : What convicted offenders tell us. Sexual abuse: *A Journal of Research and Treatment, 16,* 299–314.

Rape Crisis England and Wales. (2017). *Rape Crisis England & Wales headline statistics 2015-16.* Retrieved from https://rapecrisis.org.uk/statistics.php

Reid, S., Wilson, N. J., & Boer, D. P. (2011). Risk, needs, and responsivity principles in action: Tailoring rapist's treatment to rapist typologies. In D. P. Boer, R. Eher, & M. H. Miner, (Eds.), *International perspectives on the assessment and treatment of sexual offenders* (pp. 287-297). Chichester, UK: Wiley-Blackwell.

Rich, M.D., Utley, E.A., Janke, K., & Moldoveanu, M. (2010). "I'd rather be doing something else:" Male resistance to rape prevention programs. *Journal of Men's Studies, 18,* 268-288.

Rombough, S., & Ventimighlia, J. C. (1981). Sexism: A tri-dimensional phenomenon. *Sex Roles, 7*, 747–755.

Rothman, E., & Silverman, J. (2007). The effect of a college sexual assault prevention program on first-year students' victimization rates. *Journal of American College Health, 55*, 283-290.

Sackett, D. L., Rosenberg, W., Gray, J. A., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: what it is and what it isn't. *British Medical Journal, 312*, 71-72

Scully, D. (1990). *Understanding sexual violence: A study of convicted rapists*. London, United Kingdom: Harper Collins.

Scully, D., & Marolla, J. (1984). Convicted rapists' vocabulary of motive: excuses and justifications. *Social Problems, 31*, 530–544.

Shye, S. (1978). Achievement motive: A faceted definition and structural analysis. *Multivariate behavioral research, 13*, 327-346.

Sleath, E., & Bull, R. (2015). A brief report on rape myth acceptance: Differences between police officers, law students, and psychology students in the United Kingdom. *Violence and Victims, 30*, 136-147.

Snyder, C. R., & Higgins, R. L. (1988). Excuses: Their effective role in the negotiation of reality. *Psychological Bulletin, 104*, 23-35.

Spence, I., & Ogilvie, J. G. (1973). A table of expected stress values for random rankings in nonmetric multidimensional scaling. Multivariate behavioral research, 8(4), 511–517.

Stefanska, E. B., Carter, A. J., Higgs, T., Bishopp, D., & Beech, A. R. (2015). Offense pathways of non-serial sexual killers. *Journal of Criminal Justice, 43*, 99-107.

Stephens, K. A., & George, W. H. (2009). Rape prevention with college men: Evaluating risk status. *Journal of Interpersonal Violence, 24,* 996–1013.

Suarez, E. B., & Gadalla, T. (2010, January). *Stop blaming the victim: A meta-analysis on rape myths*. Paper presented at the Annual Meeting of the Society for Social Work and Research, Tampa, FL.

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

Ward, T. (2000). Sexual offenders' cognitive distortions as implicit theories. *Aggression and Violent Behavior, 5,* 491-507.

Ward, T., & Keenan, T. (1999). Child molesters' implicit theories. *Journal of interpersonal violence, 14,* 821-838.

Webster, S. D., Akhtar, S., Bowers, L. E., Mann, R. E., Rallings, M., & Marshall, W. L. (2004). The impact of the prison service sex offender treatment programme on minority ethnic offenders: a preliminary study. *Psychology, Crime and Law*, *10*, 113-124.

Wegner, R., Abbey, A., Pierce, J., Pegram, S. E., & Woerner, J. (2015). Sexual assault perpetrators' justifications for their Actions. *Violence against women, 21,* 1018–1037.

Wilkinson, L. (2000). Multidimensional Scaling. *Systat 10 Statistics II,* 185–214. Retrieved from http://www.springerlink.com/index/v6j14w058h7n4407.pdf

Xue, J., Fang, G., Huang, H., Cui, N., Rhodes, K. V., & Gelles, R. (2016). Rape myths and the cross-cultural adaptation of the Illinois Rape Myth Acceptance Scale in China. *Journal of Interpersonal Violence, 31,* 1-33.

Young, F. W., Takane, Y., & Lewyckyj, R. (1978). ALSCAL: A nonmetric multidimensional scaling program with several individual-differences options. *Behavior Research Methods, 10*, 451-453.

APPENDICES

Appendix A

Databases and search strategies utilised to access data on the Ovid, Web of Science, and Proquest platforms.

OvidSP

(Databases: *Books@Ovid, CAB Abstracts, Embase, Embase Classic, HMIC Health Management Information Consortium, Journals@Ovid Full Text, Ovid MEDLINE® In-Process & Other Non-Indexed Citations and Ovid MEDLINE ®, PsycARTICLES Full Text, PsycINFO,* and *Social Policy and Practice*) (number of hits in brackets)

1 rape myth*.mp. [mp=tx, bt, ti, ab, ct, hw, tn, ot, dm, mf, dv, kw, id, cc, nm, kf, px, rx, ui, tc, tm, pt, an] *(1569)*
2 rape myth* accept*.mp. [mp=tx, bt, ti, ab, ct, hw, tn, ot, dm, mf, dv, kw, id, cc, nm, kf, px, rx, ui, tc, tm, pt, an] *(925)*
3 cognit* distort*.mp. [mp=tx, bt, ti, ab, ct, hw, tn, ot, dm, mf, dv, kw, id, cc, nm, kf, px, rx, ui, tc, tm, pt, an] *(5942)*
4 attitud* to* women.mp. [mp=tx, bt, ti, ab, ct, hw, tn, ot, dm, mf, dv, kw, id, cc, nm, kf, px, rx, ui, tc, tm, pt, an] (3356)
5 rape support* attitud*.mp. [mp=tx, bt, ti, ab, ct, hw, tn, ot, dm, mf, dv, kw, id, cc, nm, kf, px, rx, ui, tc, tm, pt, an] *(240)*
6 victim* blam*.mp. [mp=tx, bt, ti, ab, ct, hw, tn, ot, dm, mf, dv, kw, id, cc, nm, kf, px, rx, ui, tc, tm, pt, an] (2086)
7 attribut* blam*.mp. [mp=tx, bt, ti, ab, ct, hw, tn, ot, dm, mf, dv, kw, id, cc, nm, kf, px, rx, ui, tc, tm, pt, an] *(1385)*
8 1 or 2 or 3 or 4 or 5 or 6 or 7 *(13679)*
9 rapist*.mp. [mp=tx, bt, ti, ab, ct, hw, tn, ot, dm, mf, dv, kw, id, cc, nm, kf, px, rx, ui, tc, tm, pt, an] (3971)
10 sex* offend*.mp. [mp=tx, bt, ti, ab, ct, hw, tn, ot, dm, mf, dv, kw, id, cc, nm, kf, px, rx, ui, tc, tm, pt, an] *(18293)*
11 convict* rapist*.mp. [mp=tx, bt, ti, ab, ct, hw, tn, ot, dm, mf, dv, kw, id, cc, nm, kf, px, rx, ui, tc, tm, pt, an] (269)
12 convict* sex* offend*.mp. [mp=tx, bt, ti, ab, ct, hw, tn, ot, dm, mf, dv, kw, id, cc, nm, kf, px, rx, ui, tc, tm, pt, an] (663)
13 incarcerate* rapist*.mp. [mp=tx, bt, ti, ab, ct, hw, tn, ot, dm, mf, dv, kw, id, cc, nm, kf, px, rx, ui, tc, tm, pt, an] *(127)*
14 incarcerate* sex* offend*.mp. [mp=tx, bt, ti, ab, ct, hw, tn, ot, dm, mf, dv, kw, id, cc, nm, kf, px, rx, ui, tc, tm, pt, an] *(365)*
15 sex* aggress*.mp. [mp=tx, bt, ti, ab, ct, hw, tn, ot, dm, mf, dv, kw, id, cc, nm, kf, px, rx, ui, tc, tm, pt, an] (6635)
16 sex* molest*.mp. [mp=tx, bt, ti, ab, ct, hw, tn, ot, dm, mf, dv, kw, id, cc, nm, kf, px, rx, ui, tc, tm, pt, an] (1312)

17 sex* assault*.mp. [mp=tx, bt, ti, ab, ct, hw, tn, ot, dm, mf, dv, kw, id, cc, nm, kf, px, rx, ui, tc, tm, pt, an] *(24196)*
18 sex* assault*.mp. [mp=tx, bt, ti, ab, ct, hw, tn, ot, dm, mf, dv, kw, id, cc, nm, kf, px, rx, ui, tc, tm, pt, an] *(48727)*
19 8 and 18 *(2684)*
20 remove duplicates from 19 *(1794)*

Web of Science *(1747 results)*
Query: (TS=(Rape myth* OR rape myth accept* OR cognit* distort* OR attitud* OR attitud* adj/3 towards women + rape adj/3 support* attitude* OR victim* adj/2 blam* OR attribut* adj/3 blam*) AND TS=(Rapist* OR sex* offend* OR Convict* rapist* OR convict* sex* offend* OR incarcerate* sex* offend* OR incarcerate* rapist* OR sex* aggress* OR sex* molest* OR sex* assault*) NOT TS=(Rape propensity OR rape proclivity)) *AND* **LANGUAGE:** (English)

Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH Timespan=All years

Proquest *(39 results)*
Query: **SU(Rape myth* OR rape myth accept* OR cognit* distort* OR attitude* OR attitude* NEAR/3 towards women OR rape NEAR/3 support* attitude* OR victim* NEAR/2 blame* OR attribut* NEAR/3 blame*) AND SU(Rapist* OR sex* offend* OR Convict* rapist* OR convict* sex* offend* OR incarcerate* sex* offend* OR incarcerate* rapist* OR sex* aggress* OR sex* molest* OR sex* assault*) NOT SU(Rape propensity OR rape proclivity)**

Appendix B

Screening and Selection Tool

Study reference:

Study Characteristics	Eligibility criteria	Eligibility criteria met? Yes No Unclear	Location in text or source *(pg & ¶/fig/table)*
Type of study	Experimental study including randomized controlled trials (RCTs) or cluster-randomized trials (CRTs).	☐ ☐ ☐	
	Quasi-experimental studies including quasi-randomized trials, controlled before-after studies (CBAs) and interrupted time series studies (ITSs).	☐ ☐ ☐	
	Observational studies including cohort, case-control and cross-sectional studies.	☐ ☐ ☐	
Participants and setting	Participants must include *convicted* rapists – not necessarily incarcerated. Men who have been convicted of sexually assaulting an adult female.	☐ ☐ ☐	

Types of intervention	Not necessary, but e.g. Sex Offender Treatment Programme	☐ ☐ ☐	
Types of comparison	Not Necessary	☐ ☐ ☐	
Types of outcome measures	Measure of Rape Myth Acceptance	☐ ☐ ☐	
Results	Quantitative only	☐ ☐ ☐	
INCLUDE ☐		EXCLUDE ☐	
Reason for exclusion			
Notes:			

DO NOT PROCEED TO QUALITY ASSESSMENT IF STUDY EXCLUDED FROM REVIEW

Appendix C

Excluded Studies and Reasons for Their Exclusion

Measure did not look specifically at rape myths (9)

Blumenthal, S., Gudjonsson, G., & Burns, J. (1999). Cognitive distortions and blame attribution in sex offenders against adults and children. *Child Abuse and Neglect, 23*(2), 129-143.

DeGue, S., DiLillo, D., & Scalora, M. (2010). Are all perpetrators alike? Comparing risk factors for sexual coercion and aggression. *Sexual abuse: A Journal of Research and Treatment, 22*, 402-426.

Feild, H. S. (1978). Attitudes toward rape: A comparative analysis of police, rapists, crisis counselors, and citizens. *Journal of Personality and Social Psychology, 36*(2), 156.

Howells, K., & Wright, E. (1978). The sexual attitudes of aggressive sexual offenders. *The British Journal of Criminology*, 170-174.

Kroner, D. G., Boer, D. P., & Mills, J. F. (2004). Explaining rape-supportive attitudes among rapists. *American Journal of Forensic Psychology, 22*(3), 65-76.

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

O'Reilly, G., Carr, A., Murphy, P., & Cotter, A. (2010). A controlled evaluation of a prison-based sexual offender intervention program. *Sexual Abuse: A Journal of Research and Treatment, 22*(1), 95-111.

Scott, R. L., & Tetreault, L. A. (1987). Attitudes of rapists and other violent offenders toward women. *The Journal of Social Psychology, 127*(4), 375-380.

Segal, Z. V., & Stermac, L. (1984). A measure of rapists' attitudes towards women. *International Journal of Law and Psychiatry, 7*(3), 437-440.

Non-clinical sample (2)

Loh, C., Gidycz, C. A., Lobo, T. R., & Luthra, R. (2005). A prospective analysis of sexual assault perpetration: Risk factors related to perpetrator characteristics. *Journal of Interpersonal Violence, 20*, 1325–1348.

Thompson, M. P., Swartout, K. M., & Koss, M. P. (2013). Trajectories and predictors of sexually aggressive behaviors during emerging adulthood. *Psychology of Violence*, *3*(3), 247.

Review (2)

DeGue, S., Massetti, G. M., Holt, M. K., Tharp, A. T., Valle, L. A., Matjasko, J. L., & Lippy, C. (2013). Identifying links between sexual violence and youth violence perpetration. *Psychology of Violence*, *3*(2), 140-156.

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

Appendix D

Risk of Bias Tool (Quality Assessment)

Entry	Judgment	Support for Judgment
Random sequence generation (selection bias)		
Allocation concealment (selection bias)		
Blinding of participants and personnel (performance bias)		
Blinding of outcome assessment (detection bias)		
Incomplete outcome data addressed (attrition bias)		
Selective reporting (reporting bias)		
Other bias		

Appendix E

The Cochrane Collaboration's Tool for Assessing Risk of Bias

Domain	Support for judgement	Review authors' judgement
Selection bias.		
Random sequence generation.	Describe the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups.	Selection bias (biased allocation to interventions) due to inadequate generation of a randomised sequence.
Allocation concealment.	Describe the method used to conceal the allocation sequence in sufficient detail to determine whether intervention allocations could have been foreseen in advance of, or during, enrolment.	Selection bias (biased allocation to interventions) due to inadequate concealment of allocations prior to assignment.
Performance bias.		
Blinding of participants and personnel *Assessments should be made for each main outcome (or class of outcomes).*	Describe all measures used, if any, to blind study participants and personnel from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.	Performance bias due to knowledge of the allocated interventions by participants and personnel during the study.
Detection bias.		
Blinding of outcome assessment *Assessments should be made for each main outcome (or class of outcomes).*	Describe all measures used, if any, to blind outcome assessors from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.	Detection bias due to knowledge of the allocated interventions by outcome assessors.
Attrition bias.		
Incomplete outcome data *Assessments should be made for each main outcome (or class of outcomes).*	Describe the completeness of outcome data for each main outcome, including attrition and exclusions from the analysis. State whether attrition and exclusions were reported, the numbers in each intervention group (compared with total randomized participants), reasons for attrition/exclusions where reported, and any re-inclusions in analyses performed by the review authors.	Attrition bias due to amount, nature or handling of incomplete outcome data.
Reporting bias.		
Selective reporting.	State how the possibility of selective outcome reporting was examined by the review authors, and what was found.	Reporting bias due to selective outcome reporting.
Other bias.		

Other sources of bias.	State any important concerns about bias not addressed in the other domains in the tool. If particular questions/entries were pre-specified in the review's protocol, responses should be provided for each question/entry.	Bias due to problems not covered elsewhere in the table.

Appendix F

Criteria for Judging Risk of Bias in the 'Risk of Bias' Assessment Tool

RANDOM SEQUENCE GENERATION **Selection bias (biased allocation to interventions) due to inadequate generation of a randomised sequence.**	
Criteria for a judgement of 'Low risk' of bias.	The investigators describe a random component in the sequence generation process such as: • Referring to a random number table; • Using a computer random number generator; • Coin tossing; • Shuffling cards or envelopes; • Throwing dice; • Drawing of lots; • Minimization*. *Minimization may be implemented without a random element, and this is considered to be equivalent to being random.
Criteria for the judgement of 'High risk' of bias.	The investigators describe a non-random component in the sequence generation process. Usually, the description would involve some systematic, non-random approach, for example: • Sequence generated by odd or even date of birth; • Sequence generated by some rule based on date (or day) of admission; • Sequence generated by some rule based on hospital or clinic record number. Other non-random approaches happen much less frequently than the systematic approaches mentioned above and tend to be obvious. They usually involve judgement or some method of non-random categorization of participants, for example: • Allocation by judgement of the clinician; • Allocation by preference of the participant; • Allocation based on the results of a laboratory test or a series of tests; • Allocation by availability of the intervention.
Criteria for the judgement of 'Unclear risk' of bias.	Insufficient information about the sequence generation process to permit judgement of 'Low risk' or 'High risk'.
ALLOCATION CONCEALMENT **Selection bias (biased allocation to interventions) due to inadequate concealment of allocations prior to assignment.**	
Criteria for a judgement of 'Low risk' of bias.	Participants and investigators enrolling participants could not foresee assignment because one of the following, or an equivalent method, was used to conceal allocation: • Central allocation (including telephone, web-based and pharmacy-controlled randomization); • Sequentially numbered drug containers of identical appearance; • Sequentially numbered, opaque, sealed envelopes.

Criteria for the judgement of 'High risk' of bias.	Participants or investigators enrolling participants could possibly foresee assignments and thus introduce selection bias, such as allocation based on: • Using an open random allocation schedule (e.g. a list of random numbers); • Assignment envelopes were used without appropriate safeguards (e.g. if envelopes were unsealed or nonopaque or not sequentially numbered); • Alternation or rotation; • Date of birth; • Case record number; • Any other explicitly unconcealed procedure.
Criteria for the judgement of 'Unclear risk' of bias.	Insufficient information to permit judgement of 'Low risk' or 'High risk'. This is usually the case if the method of concealment is not described or not described in sufficient detail to allow a definite judgement – for example if the use of assignment envelopes is described, but it remains unclear whether envelopes were sequentially numbered, opaque and sealed.

BLINDING OF PARTICIPANTS AND PERSONNEL

Performance bias due to knowledge of the allocated interventions by participants and personnel during the study.

Criteria for a judgement of 'Low risk' of bias.	Any one of the following: • No blinding or incomplete blinding, but the review authors judge that the outcome is not likely to be influenced by lack of blinding; • Blinding of participants and key study personnel ensured, and unlikely that the blinding could have been broken.
Criteria for the judgement of 'High risk' of bias.	Any one of the following: • No blinding or incomplete blinding, and the outcome is likely to be influenced by lack of blinding; • Blinding of key study participants and personnel attempted, but likely that the blinding could have been broken, and the outcome is likely to be influenced by lack of blinding.
Criteria for the judgement of 'Unclear risk' of bias.	Any one of the following: • Insufficient information to permit judgement of 'Low risk' or 'High risk'; • The study did not address this outcome.

BLINDING OF OUTCOME ASSESSMENT

Detection bias due to knowledge of the allocated interventions by outcome assessors.

Criteria for a judgement of 'Low risk' of bias.	Any one of the following: • No blinding of outcome assessment, but the review authors judge that the outcome measurement is not likely to be influenced by lack of blinding; • Blinding of outcome assessment ensured, and unlikely that the blinding could have been broken.
Criteria for the judgement of 'High risk' of bias.	Any one of the following: • No blinding of outcome assessment, and the outcome measurement is likely to be influenced by lack of blinding;

	• Blinding of outcome assessment, but likely that the blinding could have been broken, and the outcome measurement is likely to be influenced by lack of blinding.
Criteria for the judgement of 'Unclear risk' of bias.	Any one of the following: • Insufficient information to permit judgement of 'Low risk' or 'High risk'; • The study did not address this outcome.

INCOMPLETE OUTCOME DATA

Attrition bias due to amount, nature or handling of incomplete outcome data.

Criteria for a judgement of 'Low risk' of bias.	Any one of the following: • No missing outcome data; • Reasons for missing outcome data unlikely to be related to true outcome (for survival data, censoring unlikely to be introducing bias); • Missing outcome data balanced in numbers across intervention groups, with similar reasons for missing data across groups; • For dichotomous outcome data, the proportion of missing outcomes compared with observed event risk not enough to have a clinically relevant impact on the intervention effect estimate; • For continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes not enough to have a clinically relevant impact on observed effect size; • Missing data have been imputed using appropriate methods.
Criteria for the judgement of 'High risk' of bias.	Any one of the following: • Reason for missing outcome data likely to be related to true outcome, with either imbalance in numbers or reasons for missing data across intervention groups; • For dichotomous outcome data, the proportion of missing outcomes compared with observed event risk enough to induce clinically relevant bias in intervention effect estimate; • For continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes enough to induce clinically relevant bias in observed effect size; • 'As-treated' analysis done with substantial departure of the intervention received from that assigned at randomization; • Potentially inappropriate application of simple imputation.
Criteria for the judgement of 'Unclear risk' of bias.	Any one of the following: • Insufficient reporting of attrition/exclusions to permit judgement of 'Low risk' or 'High risk' (e.g. number randomized not stated, no reasons for missing data provided); • The study did not address this outcome.

SELECTIVE REPORTING

Reporting bias due to selective outcome reporting.

Criteria for a judgement of 'Low risk' of bias.	Any of the following:

	• The study protocol is available and all of the study's pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way; • The study protocol is not available but it is clear that the published reports include all expected outcomes, including those that were pre-specified (convincing text of this nature may be uncommon).
Criteria for the judgement of 'High risk' of bias.	Any one of the following: • Not all of the study's pre-specified primary outcomes have been reported; • One or more primary outcomes is reported using measurements, analysis methods or subsets of the data (e.g. subscales) that were not pre-specified; • One or more reported primary outcomes were not pre-specified (unless clear justification for their reporting is provided, such as an unexpected adverse effect); • One or more outcomes of interest in the review are reported incompletely so that they cannot be entered in a meta-analysis; • The study report fails to include results for a key outcome that would be expected to have been reported for such a study.
Criteria for the judgement of 'Unclear risk' of bias.	Insufficient information to permit judgement of 'Low risk' or 'High risk'. It is likely that the majority of studies will fall into this category.
OTHER BIAS **Bias due to problems not covered elsewhere in the table.**	
Criteria for a judgement of 'Low risk' of bias.	The study appears to be free of other sources of bias.
Criteria for the judgement of 'High risk' of bias.	There is at least one important risk of bias. For example, the study: • Had a potential source of bias related to the specific study design used; or • Has been claimed to have been fraudulent; or • Had some other problem.
Criteria for the judgement of 'Unclear risk' of bias.	There may be a risk of bias, but there is either: • Insufficient information to assess whether an important risk of bias exists; or • Insufficient rationale or evidence that an identified problem will introduce bias.

Appendix G

Examples of Supports for Judgement for Sequence Generation Entry

Sequence generation.	Comment: No information provided.
Sequence generation.	Quote: “patients were randomly allocated”.
Sequence generation.	Quote: “patients were randomly allocated”. Comment: Probably done, since earlier reports from the same investigators clearly describe use of random sequences (Cartwright 1980).
Sequence generation.	Quote: “patients were randomly allocated”. Comment: Probably not done, as a similar trial by these investigators included the same phrase yet used alternate allocation (Winrow 1983).
Sequence generation.	Quote (from report): “patients were randomly allocated”. Quote (from correspondence): “randomization was performed according to day of treatment”. Comment: Not randomized.

Appendix H

Example of a 'Risk of Bias' Table for a Single Study (fictional)

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	Low risk.	Quote: "patients were randomly allocated." Comment: Probably done, since earlier reports from the same investigators clearly describe use of random sequences (Cartwright 1980).
Allocation concealment (selection bias)	High risk.	Quote: "...using a table of random numbers." Comment: Probably not done.
Blinding of participants and personnel (performance bias)	Low risk.	Quote: "double blind, double dummy"; "High and low dose tablets or capsules were indistinguishable in all aspects of their outward appearance. For each drug an identically matched placebo was available (the success of blinding was evaluated by examining the drugs before distribution)." Comment: Probably done.
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	Low risk.	Quote: "double blind". Comment: Probably done.
Blinding of outcome assessment (detection bias) (Mortality)	Low risk.	Obtained from medical records; review authors do not believe this will introduce bias.
Incomplete outcome data addressed (attrition bias) (Short-term outcomes (2-6 weeks))	High risk.	4 weeks: 17/110 missing from intervention group (9 due to 'lack of efficacy'); 7/113 missing from control group (2 due to 'lack of efficacy').
Incomplete outcome data addressed (attrition bias) (Longer-term outcomes (>6 weeks))	High risk.	12 weeks: 31/110 missing from intervention group; 18/113 missing from control group. Reasons differ across groups.
Selective reporting (reporting bias)	High risk.	Three rating scales for cognition listed in Methods, but only one (with statistically significant results) is reported.

Appendix I

Data Extraction Form

Notes:

- Be consistent in the order and style you use to describe the information for each report.
- Record any missing information as unclear or not described, to make it clear that the information was not found in the study report(s), not that you forgot to extract it.

Review title	Systematic Review on Rape Myth Acceptance in Rapists
Study ID *(surname of first author and year first full report of study was published e.g. Smith 2001)*	
Notes	

General Information

Date form completed *(dd/mm/yyyy)*	
Name/ID of person extracting data	
Reference citation (e.g. Medline)	

Study author contact details	
Publication type *(e.g. full report, abstract, letter)*	
Notes:	

Characteristics of included studies

Methods

	Descriptions as stated in report/paper	**Location in text or source** *(pg & ¶/fig/table)*
Aim of study		
Design		
Unit of observation		
Start date		

End date			
Duration of participation *(from recruitment to last follow-up)*			
Ethical approval needed/ obtained for study	☐ Yes ☐ No ☐ Unclear		
Notes:			

Participants

	Description	Location in text or source *(pg & ¶/fig/table)*
Population description *(from which study participants are drawn)*		

Setting and context *(including but not limited to healthcare system characteristics and health financing - e.g. user fees or financial coverage of PNC services - as well as social context, location).*		
Inclusion criteria		
Exclusion criteria		
Method of recruitment of participants *(e.g. phone, mail, clinic patients)*		

Informed consent obtained	☐ Yes ☐ No ☐ Unclear		
Total no. of subjects			
Clusters *(if applicable, no., type, no. people per cluster)*			
Withdrawals and exclusions			
Number of total person-years (if applicable)			
Missing data			
Outcome(s)	As they relate to Rape Myth Acceptance		
Definition, measure & classification			

Confounding factors/ effect modifiers accounted for		
Results	Crude	
	Adjusted	
Authors' reported limitations of study's methods/results		
Scientific quality		

Notes:

Other information

Study funding sources *(including role of funders)*		
Possible conflicts of interest *(for study authors)*		
	Description as stated in report/paper	**Location in text or source**
Key conclusions of study authors		
References to other relevant studies		

Correspondence required for further study information *(from whom, what and when)*	
Notes:	

Appendix J

Characteristics of Included Studies (ordered by study ID)

Beech 2006

Methods	Design: observational (before-and-after) Measure (of interest): Burt's Rape Myth Acceptance Scale (RMAS; 1980)
Participants	Participants: male prisoners incarcerate for rape (n=112) or sexual murder (n=58) Sex: all male Age: rapists - mean 34.9 (SD 8.4) years; sexual murderers - mean 39.3 (SD 10.5) years Setting: 7 prisons Inclusion criteria: prisoner with conviction for sexual offence apart from convictions related to consensual sexual behaviour; prisoner falls into medium- or high-risk group as determined by Risk Matrix 2000 (RM2000; Thornton et al., 2003); prisoner with a homicide conviction with a clear sexual element to the homicide Exclusion criteria: Psychopathy Checklist - Revised (PCL-R; Hare, 1991) score > 30; IQ < 80; men suffering from current mental illness; men suffering from mental illness/brain damage at time of offence; men deemed 'not ready' for treament (treatment not suitable for him at this time); total denial of the offence; refusal of treatment; does not speak English; physical disability incl. deafness or blindness; poor ltieracy; suicidal or self-harming
Interventions	CORE Sex Offender Treatment Programme (SOTP) -average treatment dose in study: 188 hours (94 sessions); two to five sessions per week
Outcomes	Difference in RMA: •none found between rapists and sexual murderers •no within-group change in pre-treatment to post-treatment for both rapists and sexual murderers •rapists typologies differed on 'Adversarial Sexual Beliefs' subcales scores on RMAS •*sexually motivated* offenders scored higher of 'Sex Role Stereotyping' subscale of RMAS than non-sexual violent offenders and community non-offending males
Notes	N/A

Risk of Bias

Bias	Author's judgment	Support for judgment
Random sequence generation	Low risk.	There is a clear risk of selection bias when the person recruiting participants knows in advance the clinical characteristics of a participant and which intervention they will receive. However, due to the study design, this is not very feasible to randomise. *Decision made to override 'High risk' rating.*
Allocation concealment	Low risk.	Lack of concealed randomised allocation increases the risk of selection bias. However, due to the study design, this is not possible to conceal. *Decision made to override 'High risk' rating.*
Blinding of participants and personnel	High risk.	Participants were not blinded. Personnel were not blinded. All interviews and the treatment programme itself were conducted by the investigators.
Blinding of outcome assessment	High risk.	Outcome assessment was not blinded. All interviews and treatment programme carried out by the investigators.
Incomplete outcome data addressed	High risk.	Rapists (pre-treatment): 86/112 completed pre-treatment assessments; of the remaining 16, 14 refused at interview to complete the questionnaires and 12 simply did not fill them out. Rapists (post-treatment): 65/86 that completed the questionnaires before treatment also completed them afterwards; the remaining 21 failed to complete and return the questionnaires. Sexual murderers (pre-treatment): 45/58 completed pre-treatment assessments; the remaining 13 either refused or failed to complete and/or return the questionnaires. Sexual murderers (post-treatment): 40/45 that completed the questionnaires before treatment, completed them after treatment; of the remaining 5, 2 refused to complete them and 3 failed to complete/return them.

Selective reporting	Low risk.	The published report included all expected outcomes, including those that were pre-specified
Other bias	Low risk.	The study appeared to be free of other sources of bias.

Cohen 2012

Methods	Design: controlled observational (case control) Measure (of interest): Burt's Rape Myth Acceptance Scale (RMAS; 1980)
Participants	Participants: Sex offenders (experimental group; *N*=46 intially, *N*=44 final); Non-offender students (control; *N*=50 initally, *N*=44 final) Sex: all male Age: Offenders - mean 36.3 (SD 14.78) years; Students - mean 27.8 (SD 3.5) years Setting: Treatment center in central-Israel and the Probation Department. Inclusion criteria: These were not explicitly outlined. However, the following information could be gathered from the article: "sex offender" in Israeli law (someone who has committed an offence according to Chapter 6 of the 1977 Criminal Code. These offences include rape, sodomy, sexual assault without penetration and 'flashing'. Also included are 'consensual' sexual conduct between an adult and a teenager under 16, between a therapist and a patient, or between an employer and an employee.); physical ability to complete questionnaires; appropriate reading ability; ability to read and comprehend Hebrew Exclusion criteria: These were not explicitly outlined. However, the following information could be gathered from the article: student's disclosure of any undetected sexual coercion; does not speak Hebrew; physical disability; poor literacy; colour-blindness
Interventions	N/A

Outcomes	Sex offenders and students had similar scores on the attribution subscale and the whole-form (both combined subscales) of the MCSDS. Sex offenders had significantly lower scores on the denial subscale of the MCSDS. For most inventories dealing with offence-related cognitions, the scores of non-offenders and sex offenders were similar, but sex offenders reported significantly lower levels of rape myth acceptance than did non-offenders. The sex offenders' self-reported levels of sex role satisfaction and sexual entitlement were correlated with measures of social desirability. Participants in both groups showed slower reaction times on the ES and the LD, but no difference in accuracy on the ES in response to "general threat" words compared to neutral words.Non-offenders showed slower reaction times on the ES and LD, but no difference in levels of accuracy on the ES in response to "general threat" words compared to neutral words and "uncontrollability", "entitlement", and "sexy children" words". Sex offenders reacted similarly to "general threat" and to "uncontrollability" and "entitlement", on the ES and LD. Reaction times to these words were slower on the ES and the LD, and but levels of accuracy on the ES did not differ significantly to these words than to neutral words. The Emotional Stroop showed significant differences in response for sex offender-specific words but not "general threat" words and the LD showed a difference for "uncontrollability" words only. Sex offenders who victimized children showed slower reaction times on the LD in response to "sexy children" words than offenders with older victims, but this was not the case for the emotional stroop.
Notes	N/A

Risk of Bias

Bias	**Author's judgment**	**Support for judgment**
Random sequence generation	High risk.	There is a clear risk of selection bias when the person recruiting participants knows in advance the clinical characteristics of a participant and which intervention they will receive. However, due to the study design, random sequence generation for the offenders was not very feasible. Although, there was some bias in the selection of students as they were only recruited from a particular department.

Allocation concealment	Low risk.	Lack of concealed randomised allocation increases the risk of selection bias. However, due to the study design, this is not possible to conceal.
Blinding of participants and personnel	High risk.	Participants were not blinded. Personnel were not blinded. All interviews and the treatment programme itself were conducted by the investigators. Quote: "Almost all of the sex offenders who participated in the study were in under some form of judicial impetus to participate in this treatment." "Students were solicited in the usual way, through advertisements on bulletin boards in the Criminology departments. The students participated in the study in return for "signatures" that they partook in an experiment (a requirement for undergraduate students). As not enough respondents were recruited in this way, the author approached a colleague who taught a summer course in Criminology at Bar-Ilan University, and requested that she enlist her students to participate."
Blinding of outcome assessment	High risk.	Outcome assessment was not blinded. All interviews carried out by the investigators.
Incomplete outcome data addressed	High risk.	Sex offenders: 2/46 removed from study as they were unable to complete both the ES and LD tasks. 44 male sex offenders in the final experimental group. 1/44 refused to provide full demographic information on himself. Students: 6/50 rejected due to physical problem which affected their performance (3/6), recent immigration to Israel resulting in imperfect command of Hebrew (2/6), and one (1/6) admitted to having physically coerced a woman to have sex.
Selective reporting	Low risk.	The published report included all expected outcomes, including those that were pre-specified.

Other bias	Unclear risk.	Decision to use two implicit tests but in methodologically dissimilar ways. "The present study used a 'fast' ES task and an untimed LD task, and then compared the results. The decision to use the untimed LD rather than an untimed ES, was based on the necessity to minimize as much as possible the possibility of 'cheating' (e.g., conscious efforts to avoid focusing on word content). Despite the aforementioned difficulty of 'cheating' on the ES, it was decided that an untimed version of that task, which does not necessitate the actual reading of the target word, was more open to manipulation than an untimed LD, in which words must be read. Therefore, an untimed LD was used. While the decision to use the different tests is, I believe, the correct one, the fact that ES and LD do not work in exactly the same way cannot be ignored (for example, LD is more susceptible to variations in word frequency)."

Marshall 1996

	Design: observational (cross-sectional) Measure (of interest): Burt's Rape Myth Acceptance Scale (1980)
Participants	Participants: male prisoners who had been convicted of sexually assaulting an adult female (*N*=27) Sex: all male Age: Range: 21-58 years. Mean age: 33.89 (SD 8.31) Setting: Canadian penitentiary unit or Canadian medium-security penitentiary. Inclusion criteria: male; incarcerated; convicted of sexually assaulting an adult female. Exclusion criteria: These were not explicitly outlined.
Interventions	N/A
Outcomes	Loneliness and intimacy scores were negatively correlated with and were found to share more than 60% variance in common. Similarly, hostility toward women and rape myth acceptance were significantly related, sharing 67% of their variance in common. Regarding the relationship between scores on the loneliness and intimacy measures on one hand and scores on the hostility and rape myth acceptance scales on the other: intimacy appeared to be more strongly related to hostility toward women (*r* = .79) than does loneliness (*r* = .53), and intimacy is also more strongly related to the acceptance of rape myths (*r* = .68) than is loneliness (*r* = .39).
Notes	N/A

Risk of Bias

Bias	Author's judgment	Support for judgment
Random sequence generation	Low risk.	Study design did not allow for random sequence generation. *Decision made to override 'High risk' rating.*
Allocation concealment	Low risk.	Lack of concealed randomised allocation increases the risk of selection bias. However, due to the study design, this is not possible to conceal. *Decision made to override 'High risk' rating.*
Blinding of participants and personnel	High risk.	Neither participants nor personnel could be blinded.

Blinding of outcome assessment	High risk.	Outcome assessment was not blinded. All interviews carried out by the investigators.
Incomplete outcome data addressed	Low risk.	There were no missing outcome data.
Selective reporting	Low risk.	All outcomes reported as pre-specified.
Other bias	Low risk.	The study appeared to be free of other sources of bias.

Olver 2014

Methods	Design: observational (case series) Measure (of interest): Burt's Rape Myth Acceptance Scale (1980)
Participants	Participants: male federal sex offenders who attended treatment services at a high intensity sex offender treatment programme (*N*=267) Sex: all male Age (at programme admission): Range: 18-66 years. Mean age: 32.22 (SD 8.99) Setting: Maximum-security forensic psychiatric facility in Saskatoon, Saskatchewan, Canada: The Regional Psychiatric Centre (RPC) Inclusion criteria: male; one or more index convictions for contact sexual offences Exclusion criteria: These were not explicitly outlined.
Interventions	The Clearwater Sex Offender Programme: a cognitive-behaviourally based treatment programme, approximately 6-8 months in duration, mandated to target moderate to high risk sex offenders.

Outcomes	The 257 offenders included in outcome analyses were followed up an average of 18.2 years (SD 4.7) post release. Employing a 20-year cap on follow-up time, 73 (27.3%) men were convicted for a new sexual offence and 135 (50.6%) were convicted for any new violent (including sexual) offence. Secondary Outcomes: The mean Rape Myth Acceptance (RMA) score was approximately one full standard deviation below the normative mean for both offenders and non-offenders (Burt, 1980) at pre-treatment, and approximately two-thirds of a standard deviation lower at posttreatment. There was a significant decrease in rape myths endorsed within the sample from pre- to post-treatment. Additionally, there was a significant convergent validity correlation pre-treatment between RMA and two of the three VRS-SO factors - criminality (r = .16, p<.05) and treatment responsivity (r = .22, p<.01). Post-treatment, RMA had no correlation with criminality (r = .00) and still, a significant positive correlation with treatment responsivity (r = .21, p<.01)
Notes	N/A

Risk of Bias

Bias	Author's judgment	Support for judgment
Random sequence generation	Low risk.	Study design did not allow for random sequence generation. *Decision made to override 'High risk' rating.*
Allocation concealment	Low risk.	Lack of concealed randomised allocation increases the risk of selection bias. However, due to the study design, this is not possible to conceal. *Decision made to override 'High risk' rating.*
Blinding of participants and personnel	High risk.	Neither participants nor personnel could be blinded.
Blinding of outcome assessment	High risk.	Outcome assessment was not blinded.
Incomplete outcome data addressed	High risk.	Missing outcome data apparent in results table, however, no explanation offered.

Selective reporting	Low risk.	All outcomes reported as pre-specified.
Other bias	High risk.	Quote:"Sexual recidivism was defined as any conviction for a new sexual offence following first release to the community after programme participation. Violent recidivism was defined as any conviction for a personinvolved offence (including sexual offences) with the potential for physical or psychological harm (e.g. non-sexual assault and robbery). Both violent and sexual recidivism were coded in a binary manner (i.e., 0*no recidivism; 1*recidivism)." Investigators chose to include sexual offences in violent recidivism results which may lead to skewed finding. For example, in the same offender it would be impractical compare sexual recidvism to their violent recidivism if there is cross-over of offences. Quote: "content of the risk need domains was constrained by the availability of measures used in the sex offender treatment programme at the RPC. For instance, a measure of child molester cognitive distortions was not introduced until some years later into the programme (Attitudes towards Sex with Children) and could not be included owing to large amounts missing data. As a result, the attitudinal domain in the present study did not contain a measure of child molester cognitive distortions in contrast to other related studies (e.g. Allan et al., 2007; Craig et al., 2007; Thornton, 2002;Wakeling et al., 2013)."

Overholser 1986

Methods	Design: quasi-experimental Measure (of interest): Burt's Rape Myth Acceptance Scale (1980)
Participants	Participants: male inmates at a medium-secure prison Experimental: 12 rapists; 12 child molesters Controls: 12 non-sex-offender inmates; 12 community-based low SES men; 12 "minimal-dater" college students Sex: all male Age: Rapists - *M* = 34.5 (SD 12.2); Child molesters - *M*= 38.8 (SD 6.1); Non-sex-offender prisoners - *M*=37.8 *(*SD 9.5); Low-SES volunteers - *M* = 33.8 (SD 8.5); College students - *M* = 20.4 (SD 1.3)

	Setting: Not specified. Inclusion criteria: Rapists: male; committed sexual offence involving nonconsensual sexual contact with a female nonrelative who was over the age of 17. Child molesters: male; committed sexual offence against a female, nonrelative who was under the age of 12; the offender was 18 years of age or older Non-sex-offender prisoners: male; no prior record of sexual offences; denied ever participating in coercive sexual activity Community-based men: low SES; (matched to prison participants) College-students: male; adult; single, no girlfriend; dated less than twice in the past month and less than four times in the past six months; reported feeling at least moderately anxious when in social settings with women Exclusion criteria: These were not explicitly outlined.
Interventions	N/A
Outcomes	No significant effect was found on the Rape Myth Acceptance Scale. A significant main effect for group was found for the Sex Role Stereotyping scale, $F(4, 55) = 4.00$, $p < .01$. A Newman-Keuls analysis indicated that child molesters displayed significantly higher levels (more conservative) of sex role stereotyping than did both the community-based low-SES men and the minimal-dater college students. Heterosocial skills deficits were observed in both child molesters and rapists, in comparison with the nonincarcerated control groups, while they participated in the naturalistic controlled interaction and in the role-play scenes. Rapists in the study displayed higher levels of physiological arousal in the assertive role-play scenes than did the other groups. College students who were minimal daters appeared more behaviorally anxious in the role-play scenes than did the other groups. Additionally, behavioural and physiological differences were found among the groups in interactions with the confederate, which suggests that the controlled interaction scene and the role-play scenes still appeared to provide assessments of all subjects' general ability to satisfactorily interact with a woman. Child molesters displayed significantly higher levels of fear of negative evaluations. Hostility and impulsivity as measured in this study did not differentiate child molesters and rapists from the control groups. In general, child molesters and rapists did not appear all that dissimilar on several diverse measures.
Notes	N/A

Risk of Bias

Bias	Author's judgment	Support for judgment
Random sequence generation	High risk.	Person recruiting participants knew in advance the clinical characteristics of]participants and which intervention they receive.
Allocation concealment	High risk.	Lack of concealed randomised allocation increases the risk o selection bias.
Blinding of participants and personnel	High risk.	Participants blinded. Personnel only partially. Quote: "The female confederate was a 22-year-old undergraduate student who knew the purpose of the study did not know the status of the prisoner subjects. Data from college students and community men were collected at a university laboratory, and thus the confederate was aware the status of the two nonprisoner groups."
Blinding of outcome assessment	Low risk.	Quote: "Three undergraduate students, who were naive as the purpose and types of men in the study, were trained to observe and score the controlled interaction and role-play scenes on the molecular and global ratings of social skills ar social anxiety. The observers were trained with practice tap until their agreement was at least 80%."
Incomplete outcome data addressed	High risk.	There were no missing outcome data.
Selective reporting	Low risk.	All outcomes reported as pre-specified.
Other bias	Low risk.	The study appeared free of other sources of bias.

Pithers 1994

Methods	Design: observational (before-and-after) Measure (of interest): Burt's Rape Myth Acceptance Scale (1980)
Participants	Participants: convicted males (*N = 20)*: paedophiles (*n* = 10); rapists (*n* = 10) Paedophiles that abused prepubescent males exclusively (*n*=4) Paedophiles that abused prepubescent females exclusively (*n*=4) Paedophiles that abused children of both genders (*n*=2) All rapists had abused adult females (*n*=10) Sex: all male Age: Rapists - *M* = 32.2 (SD 7.53); Paedophiles - *M*= 36.3 (SD 9.79) Setting: Northwest State Correctional Facility

	Inclusion criteria: These were not explicitly outlined. Exclusion criteria: These were not explicitly outlined.
Interventions	Survivor empathy group (as part of Vermonst Treatment Programme for Sexual Aggressors; Pithers, Martin & Cumming, 1989)
Outcomes	A univariate repeated-measures ANOVA did not identify a significant Group effect, F (1,18) <1, although the Treatment effect was significant, F (1,18) = 117.47, p<0.001. The Group x Treatment interaction was not significant, F (1,18) < 1. Thus both groups displayed reduced acceptance of rape myths. Paedophiles and rapists did not differ in pre-treatment or post-treatment endorsement of cognitive distortions hypothetically related to or rape. Scores on Burt's Rape Myth Acceptance Scale, which would be expected to reveal deficits in the rapists, did not discriminate these samples of child abusers and rapists.
Notes	N/A

Risk of Bias

Bias	**Author's judgment**	**Support for judgment**
Random sequence generation	Unclear risk.	Insufficient information about the sequence generation process to permit judgment of 'Low risk' or 'High risk'
Allocation concealment	Low risk.	Lack of concealed randomised allocation increases the risk of selection bias. However, due to the study design, this is not possible to conceal. *Decision made to override 'High risk' rating.*
Blinding of participants and personnel	High risk.	Participants were not blinded. Personnel were not blinded
Blinding of outcome assessment	High risk.	Investigators were not blind to outcome assessment.
Incomplete outcome data addressed	High risk.	There were no missing outcome data.
Selective reporting	High risk.	Did not report relationship between RMAS and established measures of empathy change.

Other bias	High risk.	Used Rape Myth Acceptance Scale as an "indirect measure of empathy" but its use in this manner had not been validated (or there is no mention of its use in this way in the article). But states an alternative use: "or…assess[es] a construct that may be expected to vary as empathy varies".

Stefanska 2015

Methods	Design: observational (cross-sectional) Measure (of interest): The Rape Myths Scale (Offending Behaviour Programmes Unit, 1995)
Participants	Participants: Sexual killers (*N = 150);* sexual killers with previous rape or attempted rape offence (*n=44)* Sex: all male Age: Range: 18-45 years. Mean age at the time of the offence: 25.87 (SD 7.23) Setting: Data retrieved from National Offender Management Service, OASys research database; Sex Offender Treatment Programme (SOTP) database; and Public Protection Unit Database (PPUD) Inclusion criteria: These were not explicitly outlined. However, the following information could be gathered from the article: male; sex offender (rapist or child molester); non-serial sexual killers (those convicted of killing one or two victims without an emotional cool-off period, e.g. two victims killed at the same time or within a period indicative of a single event) have been convicted and served or are serving a custodial sentence within HM Prison Service; victims are females aged 14, or older; a sexual killing includes murders and manslaughters where a sexual element and/or a sexual motivation was evidenced, suspected or admitted; completed the SOTP ; appropriate reading ability; ability to read and comprehend English Exclusion criteria: These were not explicitly outlined. However, the following information could be gathered from the article: serial sexual killers; sexual murderers of men; sexual murderers of children; does not speak English; physical disability; poor literacy; negative attitude to treatment/low motivation
Interventions	N/A

Outcomes	Rape myths were not analysed in isolation. However, upon reading the results tables, it is shown that of the men in the high problem group (offenders who were likely to report high levels of sexual entitlement beliefs, rape myths, have problems with being open to others and tend to believe that women are deceitful) 35% (p<.001) were found to endorse rape myths, whereas in the low problem group (those who did not report problems in the aforementined areas), 13% (p<.001) endorsed rape myths.
Notes	N/A

Risk of Bias

Bias	Author's judgment	Support for judgment
Random sequence generation	Unclear risk.	Insufficient information about the sequence generation process to permit judgement of 'Low risk' or 'High risk'
Allocation concealment	Low risk.	Lack of concealed randomised allocation increases the risk of selection bias. However, due to the study design, this is not possible to conceal. *Decision made to override 'High risk' rating.*
Blinding of participants and personnel	High risk.	Neither participants nor personnel could be blinded.
Blinding of outcome assessment	High risk.	Outcome assessment was not blinded. All interviews carried out by the investigators.
Incomplete outcome data addressed	Low risk.	There were no missing outcome data.
Selective reporting	Unclear risk.	As the study was carried out for exploratory purposes, there were no predetermined hypotheses.
Other bias	Low risk.	The study appeared to be free of other sources of bias.

Webster 2004

Methods	Design: observational (before-and-after) Measure (of interest): Burt's Rape Myth Acceptance Scale (RMAS; 1980)
Participants	Participants: Black sex offenders (experimental group; *N*=52); White sex offenders (comparator; *N*=52) Sex: all male

	Age: Black SOs - mean 31.22 (SD 12.14) years; White SOs - mean 35.38 (SD 10.54) years Setting: Her Majesty's Prison Service (data retrieved from the national database) Inclusion criteria: These were not explicitly outlined. However, the following information could be gathered from the article: male; sex offender (rapist or child molester); appropriate reading ability; ability to read and comprehend English Exclusion criteria: These were not explicitly outlined. However, the following information could be gathered from the article: does not speak English; physical disability; poor literacy; negative attitude to treatment/low motivation
Interventions	CORE SOTP
Outcomes	A main effect was not found for either ethnicity or type of victim. There was also no interaction effect between ethnic group and victim type. Within-subjects analysis showed that the groups significantly improved on rape myths, $F(1,69) = 20.71$, $p<0.001$. There were no other significant within-subjects main effects or interactions re: RMA.
Notes	N/A

Risk of Bias

Bias	**Author's judgment**	**Support for judgment**
Random sequence generation	Unclear risk.	Insufficient information about the sequence generation process to permit judgement of 'Low risk' or 'High risk'
Allocation concealment	Low risk.	Lack of concealed randomised allocation increases the risk of selection bias. However, due to the study design, this is not possible to conceal. *Decision made to override 'High risk' rating.*
Blinding of participants and personnel	Low risk.	Participants were not blinded. Personnel were not blinded. However, the outcome is not likely to be influenced by lack of blinding.
Blinding of outcome assessment	Low risk.	Investigators were blind to outcome assessment.

Incomplete outcome data addressed	Low risk.	There were no missing outcome data.
Selective reporting	Low risk.	The published report included all expected outcomes, including those that were pre-specified.
Other bias	Low risk.	The study appeared to be free of other sources of bias.

Appendix K

Rape Myth Acceptance in Convicted Rapists: A Systematic Review of the Literature as published in Aggression and Violent Behavior

Publication DOI: 10.1016/j.avb.2017.03.004

Publication title: "Rape Myth Acceptance in Convicted Rapists: A Systematic Review of the Literature"

Publication reference: Johnson, L. G., & Beech, A. (2017). Rape myth acceptance in convicted rapists: A systematic review of the literature. *Aggression and Violent Behavior, 34*, 20-34.

Aggression and Violent Behavior 34 (2017) 20–34

Contents lists available at ScienceDirect

Aggression and Violent Behavior

Rape myth acceptance in convicted rapists: A systematic review of the literature

Larissa Gabrielle Johnson *, Anthony Beech

Centre for Forensic and Criminological Psychology, School of Psychology, University of Birmingham, Birmingham, West Midlands B15 2TT, United Kingdom

ARTICLE INFO

Article history:
Received 17 October 2016
Received in revised form 24 February 2017
Accepted 13 March 2017
Available online 18 March 2017

Keywords:
Rape myth acceptance
Rapist typology
Rapists
Sex offending
Offence-supportive attitudes

ABSTRACT

Aim: The review examines studies on rape myth acceptance (RMA) within populations of convicted sexual offenders, changes in RMA due to interventions, comparisons between sexual offenders and community controls, comparisons within the offending population, and relationships between RMA and other psychological constructs linked to criminogenic need.
Method: The search employed electronic databases, OvidSP, Web of Science, and Proquest; hand searching reference lists; and contacting 35 experts in the field. Inclusion/exclusion and quality appraisal criteria were applied to each study.
Results: Eight studies met the inclusion criteria. Results highlighted differences in subgroups of rapists for different aspects of RMA; while rapists can be distinguished from non-offenders and non-sexual offenders on measures of RMA, they cannot be significantly discriminated from child molesters; rapists and sexual murders cannot be distinguished using RMA scores; RMA was not found to be a significant predictor of sexual or violence recidivism; and significant positive change in RMA was reported after sex offenders completed treatment programs.
Conclusions: Differences in scores on RMA subscales amongst rapists' typologies were discovered, which may indicate differences in beliefs within each type. Implications for practice are discussed.

Contents

* Corresponding author.
E-mail address: lgj489@bham.ac.uk (L.G. Johnson).

http://dx.doi.org/10.1016/j.avb.2017.03.004
1359-1789/

1. Introduction

Sexual offending research is often heavily weighted in the topic of child sexual abuse. Rape is underrepresented in the literature, resulting in limited knowledge and inefficient treatment. Often, sexual offenders will receive a generic treatment program despite it being important to separate treatment needs for those that differ in their criminogenic needs (Reid, Wilson, & Boer, 2011). Helmus, Hanson, Babchishin, and Mann (2013) note that cognitive distortions, specifically "attitudes supportive of sexual offending", are a risk factor that have predictive validity for sexual recidivism. Rape myth acceptance has been identified as one of these cognitive distortions and will be the topic of this review.

1.1. Rape myths and rape myth acceptance

Martha Burt first introduced and subsequently defined the concept of rape myths in 1980 as "prejudicial, stereotyped, or false beliefs about rape, rape victims, and rapists" (p. 217). In later years, Lonsway and Fitzgerald (1994) went on to expand on the definition, stating that rape myths are "attitudes and beliefs that are generally false but are widely and persistently held, and that serve to deny and justify male sexual aggression against women" (p. 134). For example, women "ask for rape" and rape is a result of the "uncontrollable" male sex drive (Payne, Lonsway, & Fitzgerald, 1999), shifting the blame for the crime toward the victim (Chapleau & Oswald, 2010; Gray, 2006). Rape myths influence attitudes toward victims on a social level. High levels of rape myth acceptance (RMA) are strongly associated with rape proclivity – one's likelihood or tendency to choose to rape (Chapleau & Oswald, 2010; Chiroro, Bohner, Viki, & Jarvis, 2004; Edwards, Turchik, Dardis, Reynolds, & Gidycz, 2011; Gray, 2006). Rape myths are thought to reduce the expectation of negative outcomes or consequences in sexual offenders (Chapleau & Oswald, 2010). There is evidence of RMA amongst convicted rapists, using myths to rationalize their behaviours (Chiroro et al., 2004). Rape myth acceptance has been a major topic in rape literature and research has identified the devastating impact of RMA across a variety of settings.

1.2. Measures of rape myth acceptance

There are a wide range of instruments designed to assess constructs related to rape myths. However, it should be noted that within the literature what defines a "rape myth" will vary across authors. Some experts state that the term "rape myth acceptance" is now interchangeable with "offence supportive attitudes" or "rape supportive attitudes" (C. Hermann, personal communication, May 4th, 2015; J. W. Van den Berg, personal communication, April 28th 2015). Alternatively, these terms could be viewed, arguably more appropriately, as overarching terminology under which "rape myth acceptance" falls as a subcategory. The varied literature on the topic looks at rape attitudes, knowledge on rape, empathy toward rape, and rape aversion (Lonsway & Fitzgerald, 1994).

Before the official introduction of the term "rape myths" by Burt in 1980, Feild (1978) developed the Attitudes Toward Rape Scale (ATR). The researcher found that counsellors differed from police, citizens, and rapists in their beliefs about rape, with citizens and the police being most similar. However, the scale failed to discriminate between rapists and police on approximately half of the attitudinal dimensions. As a result, many studies after this have chosen to utilize other tools for measuring rape myth acceptance or to pull aspects from the ATR and combine these with items that better discriminate rapists from non-offenders.

Arguably, the most widely used measure of rape myths is the Rape Myth Acceptance Scale developed by Burt (1980). The Rape Myth Acceptance Scale measures distorted beliefs around the sexual assault of adult women. This was the introductory measure for rape myth terminology. Research with the scale has found that men who are sexually aggressive toward adult women endorse more of these distorted beliefs about rape than do non-sexually aggressive men (Burt, 1980; Muehlenhard & Linton, 1987). Bumby (1996) noted that approximately a third of the scale's items do not specifically measure rape myths. Rather, he explained, the Rape Myth Acceptance Scale appeared to reveal how peoples' biases regarding age, race, and gender affect their likelihood of believing an allegation of rape.

Bumby (1996) felt that Burt's scale was highly susceptible to socially desirable responding and that there was weak evidence of its ability to discriminate between offenders and non-offenders. In response, he created the Bumby RAPE scale and found that it could discriminate between sex offenders and controls, but could not discriminate amongst sex offenders (i.e. separate rapists from child sex offenders). However, the RAPE scale has been discounted as well as a measure of rape myth acceptance and seen as a measure overall of sexual-assault-supportive attitudes (W. Murphy, personal communication, April 27th, 2015).

Also, building on Burt's scale, and attempting to enhance it, Payne et al. (1999) created the Illinois Rape Myth Acceptance Scale to assess myths about female victims of rape, male perpetrators, and rape as a violent crime by examining gender-role stereotyping, adversarial sexual and heterosexual beliefs, hostility toward women, and acceptance of interpersonal violence.

Many researchers have developed extended or modified versions of Burt's RMAS and others have developed scales that are conceptually similar (see Lonsway and Fitzgerald (1994) for a comprehensive list of measures relating to rape myth acceptance and rape-supportive attitudes).

1.3. The current review

There is evidence to suggest that addressing rape myth acceptance is a relevant treatment need for adult male rapists but the research is

mixed on this matter and a systematic review looking specifically at convicted offenders has yet to be carried out. To understand sexually aggressive behaviours, it is critical to understand the cognitive associations of sexually aggressive men and it is important that this research be done with the criminally convicted. Studying rape proclivity, though beneficial, may lose the cognition inherent in a criminal that may not be present in members of the general population. Also, from a rehabilitative and reintegration standpoint, it is more appropriate to target those needing rehabilitation. Helmus et al. (2013) carried out a meta-analysis on offence-supportive attitudes as a risk fact in sexual offending as an update to Hanson and Morton-Bourgon's (2004) analysis. They looked at the role of cognition in sexual offending, however, they do not consider offence-specific justifications (e.g. rape myth acceptance in rapists) which is a gap in the literature that this review will attempt to fill. The review examines studies on rape myth acceptance within populations of convicted sexual offenders and will review literature around changes in RMA due to interventions, comparisons made between sexual offenders and community controls, comparisons made within the offending population and relationships found between RMA and other psychological constructs linked to criminogenic need.

1.4. Aims and objectives

The aim of this systematic review was to explore rape myth acceptance as a criminogenic need for adult males who have committed sexual assaults against adult women.

The specific objectives of the review were:

- To determine if adult, male rapists can be distinguished from adult, male child molesters, non-sexual offenders, or non-offenders on measures on RMA
- To determine if there are differences in levels of RMA between different sub-groups of rapists, for example, those motivated by sex versus those motivated by anger
- To determine if difference in levels of RMA can discriminate between rapists who reoffend (recidivists) and those who do not
- To establish if RMA is responsive to sex offender treatment programs

2. Method

2.1. Scoping exercise

An electronic search of the Centre for Reviews and Dissemination (DARE), the Campbell Corporation, the Cochrane Database of Systematic Reviews, and the International Prospective Register of Systematic Reviews (PROSPERO) was performed to establish whether reviews of a similar or identical nature had been completed or planned. No existing or planned reviews were identified.

2.2. Overview of search strategy

The search for this review occurred in three stages. First, three electronic platforms (OvidSP, Web of Science, and Proquest) were searched for articles published before May 9th 2015 – the final day of the literature search. Within the OvidSP platform the following databases were searched: Books@Ovid, CAB Abstracts, Embase, Embase Classic, HMIC Health Management Information Consortium, Journals@Ovid Full Text, Ovid MEDLINE® In-Process & Other Non-Indexed Citations and Ovid MEDLINE®, PsycARTICLES Full Text, PsycINFO, and Social Policy and Practice. Second, the reference list of the full text articles – those which met the inclusion and exclusion criteria – and Helmus et al.' (2013) meta-analysis were hand searched for relevant articles which could potentially be included in the review. Third, 35 recognized experts in the field of sex offender research and rape-supportive attitudes were contacted and queried about any relevant and pertinent studies (published or unpublished) that could be included. Twenty-one experts responded.

2.3. Search terms

The following search terms were used along with the Boolean operators 'AND' (to combine the search concepts), 'OR' (to combine synonyms) and 'NOT' (to eliminate particular terms) where necessary. Search terms and operators were modified to accommodate the different search conventions requisite for different databases and platforms. The terms and operators were as follows: [rape myth* OR rape myth accept* OR cognit* distort* OR attitud* OR attitud* adj/3 toward women OR rape adj/3 support* attitude* OR victim* adj/2 blam* OR attribut* adj/3 blam*] AND [rapist* OR sex* offend* OR Convict* rapist* OR convict* sex* offend* OR incarcerate* sex* offend* OR incarcerate* rapist* OR sex* aggress* OR sex* molest* OR sex* assault*] NOT [rape propensity OR rape proclivity].

2.4. Inclusion/exclusion criteria

This review is predominantly exploratory and is not making a specific attempt to evaluate intervention efficacy and as such, some components of the Population, Intervention, Comparators, Outcome, Study design (PICOS) framework were not deemed relevant to this review. Cooke, Smith, and Booth (2012) developed an alternative framework used for qualitative and mixed methods studies referred to as SPIDER (Sample, Phenomenon of Interest, Design, Evaluation, Research type). This review utilized aspects of both frameworks to best capture all angles of the review questions and these were incorporated into the screening and selection process. Studies were eligible for the systematic review if: (a) participants were adult (older than 18), male rapists; (b) the phenomenon of interest was rape myth acceptance (as operationalized by Burt in 1980 and extended by Lonsway and colleagues in 1999); and (c) the study used any of the following comparators: non-offenders (community controls), non-sexual offenders (e.g. violent offenders), other categories of sexual offender (e.g. child molester, recidivists and non-recidivists, or participants pre- and post-intervention. Furthermore, only quantitative studies were included in this review. Exclusion criteria consisted of the following: (a) studies looking only at rape proclivity (i.e. utilizing a non-clinical sample in which no one has been convicted of a rape); (b) studies that did not utilize measures that were specifically used for the measurement of RMA; (c) studies that relied solely on qualitative measures. Additionally, studies that were not in English were not included in this review. It is noteworthy that in applying the inclusion/exclusion criteria, there were studies that utilized the Bumby RAPE scale as a measure of RMA. However, when discussed with experts, the decision was made to exclude these studies as experts could not come to an agreement on whether the Bumby RAPE scale *specifically* measure rape myth acceptance versus general rape-supportive attitudes (J. Abracen, personal communication, April 24th, 2015; K. Nunes, personal communication, April 24th, 2015; W. Murphy, personal communication, April 27th, 2015). Eight studies in total were included in this systematic review. These were: Beech, Oliver, Fisher, and Beckett (2006), Cohen (2012), Marshall and Hambley (1996), Olver, Nicholaichuk, and Wong (2014), Overholser and Beck (1986), Pithers (1994), Stefanska, Carter, Higgs, Bishopp, and Beech (2015), and Webster et al. (2004).

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

Of the above searches, 2686 hits were returned from OvidSP, 1747 from Web of Science, and 24 from Proquest. First, duplicate references were removed from OvidSP ($n = 892$). Second, all titles, abstracts, and sources of the remaining articles in OvidSP ($n = 1794$), Web of Science and Proquest were screened. Those which did not meet the inclusion

criteria were removed. Third, full text copies were obtained for all citations that remained ($n = 15$). The inclusion and exclusion criteria were applied to each paper. Fourth, the criteria were applied to the papers obtained from hand searching references, this returned 4 articles. Lastly, the articles acquired from experts in the field were also scrutinized using the inclusion/exclusion criteria; from this, four articles were obtained. Fig. 1 shows a diagrammatic representation of this process.

2.6. Quality assessment

Due to the large variance in study design across the eight articles, quality assessment tools that cater to specific study designs were deemed inappropriate and would not provide the flexibility required to assess the methodological rigor of the studies in this review. The Cochrane Collaboration recommends using a domain-based evaluation which is neither a scale nor a checklist that is used to make critical assessment separately for different domains (of bias; Higgins & Green, 2011). Consequently, the Cochrane Collaboration's tool for assessing risk of bias was used for this review. Its application to the eight studies can be seen in Appendix A. Assessment of overall risk of bias for each study was informed by the empirical evidence of bias, likely direction of bias, and likely magnitude of bias as guided by the Cochrane Collaboration's Handbook (Higgins & Green, 2011). This was carried out by two independent assessors. No "cut-off point" for exclusion was applied, due to the small number of studies ($n = 8$) that were included in the review and the sparse amount of literature in this specific topic area.

2.7. Data extraction

The following data were extracted: general information (date of extraction, reference citation, study author contact details, publication type); study characteristics (methods, participants, and other pertinent information, e.g. any conflicts of interest); and key conclusions. Study characteristics such as the type of study, type of intervention, comparison groups, and outcome measures were all recorded during screening and selection of the studies and thus were not repeated in data extraction.

3. Results

3.1. Overview of studies

Table 1 depicts a summary of the synthesized data for the 13 studies, allowing for evaluation of how rape myth acceptance is measured amongst rapists. The comprehensive information on each study along with their risk of bias forms can be found in Appendix A.

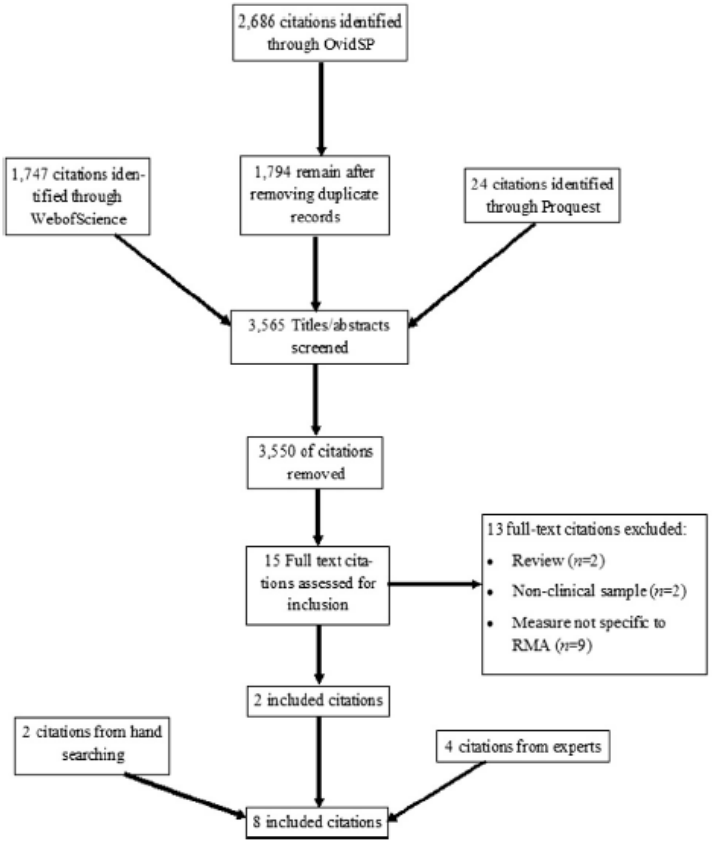

Fig. 1. Flowchart of the study selection process.

3.2. Methodological and study characteristics

There was some variability in the studies regarding countries. Four countries were represented in the data: The United Kingdom, the United States of America, Canada, and Israel. There was also variability in the dates of the studies with the oldest study being published 29 years ago (Overholser & Beck, 1986), and the most recent having been published within the last two years (Stefanska et al., 2015). Study design was also quite varied with three studies using a before-and-after observational design; one study was a case control (controlled observational) study; two utilized a cross-sectional (observational) design; one study was a case series (observational); and the final study used a quasi-experimental design.

3.3. Participants and recruitment

Participants ranged in age from 31.22 to 35.38 years. However, it should be noted that this range in mean ages were from within the same study (Webster et al., 2004) and does not include the mean age of 25.87 in Stefanska et al.'s (2015) study as this mean was an average of ages taken at the time of the offenders' index offence whereas the other studies reported the mean age at the time of assessment. Sample sizes were quite reasonable for most studies with the smallest being reported in Pithers (1994) where process evaluation of a specialized treatment program was carried out with 10 rapists and 10 paedophiles. The study with the largest sample size was of a case series design (Olver et al., 2014) and examined risk and treatment change in 276 federal sex offenders. The locations from where participants were recruited varied with individuals drawn from prisons, psychiatric facilities and treatment centers, community-based programs, and probation departments. Control groups, where applicable, were recruited from universities and community-based organizations. Due to the design of a few studies (e.g. case series) an active "recruitment" process was not required. In these cases, researchers gathered information on participants from national databases and criminal justice records.

3.4. Study focus and aims

There were no studies included in this review that had rape myth acceptance as the only variable to be examined; however, one study did have RMA as one of its main variables under investigation and explored the relationship RMA had to loneliness, intimacy, and hostility toward women amongst rapists (Marshall & Hambley, 1996). Four studies included an RMA measure in a battery of psychometric assessment measures to evaluate risk and change in offenders as a means of some form of process or treatment program evaluation (Beech et al., 2006; Olver et al., 2014; Pithers, 1994; Webster et al., 2004). In the remaining three studies, rape myth acceptance fell under the category of some all-encompassing attitudinal variable that was just one factor amongst many under investigation.

The main aim of half of the studies included in this review ($n = 4$) was to evaluate the impact of treatment on sexual offenders utilizing a set of psychometrics as pre- and post-intervention measures of risk and change. Two of these studies were evaluating the CORE Sex Offender Treatment Programme in the UK (Beech et al., 2006; Webster et al., 2004), however, in addition to pre- and post-treatment scores, Webster and colleagues were focusing particularly on any differences that occurred across ethnic backgrounds (i.e. Black sexual offenders versus White sexual offenders); one study evaluated the "Clearwater Programme" a high intensity sex offender treatment program run in a maximum-security forensic psychiatric facility in Saskatoon, Saskatchewan, Canada (Olver et al., 2014); and Pithers (1994) carried out a process evaluation of the Vermont Treatment Program for Sexual Aggressors, focusing specifically on the survivor empathy group. Two studies utilized control groups – a non-offending population (Cohen, 2012; Overholser & Beck, 1986). Cohen (2012) compared 44 sex offenders in a community-based treatment program to 44 undergraduate students to explore, firstly, whether the "uncontrollability" and "entitlement" schemata rapists and child molesters purportedly hold existed, as well as the schema of "sexy children" in child molesters; and secondly, whether these cognitive distortions were absent in the control group or no difference existed between sex offenders and controls. Overholser and Beck (1986), on the other hand, wanted to investigate whether heterosocial skills, social anxiety, hostility, impulsivity, and attitudinal variables differed between sex offenders and non-sex-offenders as well as community controls. It should be noted, however, that Beech et al. (2006) carried out a post-hoc comparison between the main study sample and a sample of non-violent offenders and community non-offending males (data for the comparator sample was gathered from an older study: Beazley Richards, 2000). This data will be discussed in Section 3.7. The final two studies carried out cross-sectional investigations. Marshall and Hambley (1996), looked at a single group of incarcerated rapists exploring their responses to measures of loneliness, intimacy, rape myth acceptance, and hostility toward women and the relationship between these. Stefanska et al. (2015) explored the pathways to offending in sexual killers and whether there was a distinction between those who had a previous conviction of rape versus those who did not.

3.5. Measures of RMA

There was virtually no variability in terms of the type of measures used to evaluate RMA. In fact, all studies used Burt's Rape Myth Acceptance Scale (1980) to measure RMA, with the exception of one study (Stefanska et al., 2015). Stefanska and colleagues utilized the Rape Myths Scale developed by the Offending Behaviour Programmes Unit (now known as the Operational Services and Intervention group; 1995) in the National Offender Management Service.

Burt's Rape Myths Acceptance Scale is a 19-item self-report measure that assesses the extent to which respondents endorse false beliefs about the rape of adult women that tend to externalize blame. The scale includes 11 items which are related to victim blaming and justification for rape and eight additional items that relate to false accusations and the likelihood that the respondent believes individuals' claims of rape. Each item is scored on a 7-point Likert scale ranging from 1 = "strongly disagree" to 7 = "strongly agree". A higher score is an indication of a greater acceptance of rape myths. Burt (1980) reported initial validation studies in a sample of 598 US adults as a Cronbach's alpha of 0.88 and item-to-item correlations of between 0.27 and 0.62. However, research on the scale's discriminant validity or the effect of social desirability on the scale is still mixed (Bumby, 1996; Lonsway & Fitzgerald, 1995). It should be mentioned that Beech et al. (2006) report using a 23-item version of this scale. It is also worth noting that in Cohen's (2012) study, an extended version of Burt's Rape Myth Acceptance Scale was utilized which was not used in any of the other studies (or not specified). This version contains 55 items scored on the same 7-point Likert scale as the 19-item and 23-item versions. Cohen provides details of this extended scale, stating that the inventory includes 6 subscales relating to sexual behaviour: Adversarial Sexual Beliefs, Sex Role Satisfaction, Rape Myth Acceptance, Sex Role Stereotype, Sexual Conservatism, and Acceptance of Interpersonal Violence. The Adversarial Sexual Beliefs subscale refers to the core belief that sexual relationships are fundamentally exploitative and contains 9 items. Sex Role Satisfaction refers to "familial, work, and interpersonal role elements relevant to sex role stereotyping" (Burt, 1980, p. 219). The scale contains 10 items, asking the respondent to rate how satisfied they are with their "competence and skillfulness", "amount of socializing", etc. The Rape Myth Acceptance scale contains the first 11 items as in the 19-item scale which target false beliefs about rape, rapists, and rape victims. The Sex Role Stereotype scale contains 9 items which reflect the respondent's endorsement of stereotyped sex roles for men and women. The Sexual Conservatism scale contains 10 items which reflect "restrictions on the

Table 1
Characteristics and overall risk of bias scored for the eight included studies.

Author & year [study ID]	Aims of study & design	Population	Intervention	Comparison/Control	Measure of RMA	Findings (in relation to RMA)	Overall risk of bias
Beech et al. (2006) [Beech 2006]	Commissioned by Home Office and Her Majesty's Prison Service to evaluate prison and probation treatment services for sexual offenders; specifically, the CORE Sex Offender Treatment Programme (SOTP) and its appropriateness for use with rapists and sexual murderers. Used psychometrics to highlight criminogenic needs of sexual offenders as well as investigate the immediate or short-term effects of the treatment programs. Observational (before-and-after).	112 rapists and 58 sexual murderers involved in the CORE SOTP from 1998 to 2002.	CORE SOTP	Pre- and post-intervention results	Burt's Rape Myth Acceptance Scale (RMAS; 1980) Cronbach's alpha: 0.88	Pre-treatment: no significant differences found between rapists and sexual murders re; RMA. Post-treatment: no differences found between rapists and sexual murderers and no overall change in RMA in rapists or in sexual murderers; within rapists typologies: groups differed significantly on their scores for the 'adversarial sexual beliefs' subscale of Burt's RMAS and the 'sexually motivated' offenders were found to score higher on the 'sex role stereotyping' subscale than non-sexual violent offenders and community non-offending males.	High
Cohen (2012) [Cohen 2012]	Investigated the existence of the "uncontrollability" and "entitlement" schemata rapists and child molesters purportedly hold, as well as the schema of "sexy children" in child molesters using an Emotional Stroop task (ES) and lexical decision task (LD). Additionally, the author measures cognitive distortions using the Burt RMAS and the Hanson Sex Attitude Questionnaire and social desirability using the Marlowe-Crowne Social Desirability Scale (MCSDS). Controlled observational (case control).	44 sex offenders in community-based treatment program in Israel	N/A	44 undergraduate students	Burt's RMAS	No difference in cognitive distortion levels between sex offenders and students, but social desirability was not a factor in this finding. Overall, found that cognitive distortions are present in both sex-offenders and non-offenders. However, in sex offenders they interact with lack of sex-role satisfaction, high trait anger and trait anxiety, lending them an emotional salience not present in non-offenders. Cognitive distortions manifest in high risk situations for offenders but not for non-offenders	High
Marshall and Hambley (1996) [Marshall 1996]	Examined the relationship amongst rapists of their responses to measures of loneliness, intimacy, rape myth acceptance, and hostility toward women. Observational (cross-sectional).	27 incarcerated male rapists	N/A	N/A	Burt's RMAS	All expected relationships between variables confirmed through correlational analyses. Results suggest rape is a function of hostility toward women combined with the acceptance of rape myths, which are also related to intimacy and loneliness deficits amongst sex offenders. It was found that the link with intimacy was stronger that the link with loneliness in this group of offenders.	Low
Olver et al. (2014) [Olver 2014]	Examined sex offenders' risk and treatment change based on a battery of psychometric assessment measures followed up an average 18 years post-release. Observational (case series).	276 federal sex offenders	Clearwater Programme (High intensity Sex Offender Treatment Programme)	Pre- and post-intervention; with follow up	Burt's RMAS	The mean Rape Myth Acceptance (RMA) score was approximately one full standard deviation below the normative mean for both offenders and non-offenders (Burt, 1980) at pre-treatment, and approximately two-thirds of a standard deviation lower at post treatment. There was a significant decrease in rape myths endorsed within the sample from pre- to post-treatment.	High

(continued on next page)

Table 1 (*continued*)

Author & year [study ID]	Aims of study & design	Population	Intervention	Comparison/Control	Measure of RMA	Findings (in relation to RMA)	Overall risk of bias
Overholser and Beck (1986) [Olverholser 1986]	Assessed rapists, child molesters, and three control groups on five potentially relevant variables: heterosocial skills, social anxiety, hostility, impulsivity, and attitudinal variables. Quasi-experimental.	12 rapists 12 child molesters	N/A	1. 12 prisoners who were non-sex offenders 2. group of 12 community-based low socioeconomic status (SES) men 3. group of 12 "minimal-dater" college students	Burt's RMAS	No significant effect was found on the Rape Myth Acceptance Scale. A significant main effect for group was found for the Sex Role Stereotyping scale, $F(4, 55) = 4.00$, $p < 0.01$. A Newman-Keuls analysis indicated that child molesters displayed significantly higher levels (more conservative) of sex role stereotyping than did both the community-based low-SES men and the minimal-dater college students.	High
Pithers (1994) [Pithers 1994]	Carried out a process evaluation to assess the extent to which a specialized treatment group might enhance the offenders' empathy for sexual abuse survivors. Observational (before-and-after)	20 convicted males: 10 paedophiles, 10 rapists	Survivor empathy group (as part of Vermont Treatment Program for Sexual Aggressors; Pithers, Martin, & Cimming, 1989)	Pre- and post-intervention	Burt's RMAS	Results suggest the intervention results in decreased endorsement of cognitive distortions predisposing rape. Paedophiles and rapists did not differ in pre-treatment or post-treatment endorsement of cognitive distortions hypothetically related to or rape. Scores on Burt's Rape Myth Acceptance Scale, which would be expected to reveal deficits in the rapists, did not discriminate these samples of child abusers and rapists.	High
Stefanska et al. (2015) [Stefanska 2015]	Aimed to examine pathways to sexual killing while also separating sexual killers based on whether they had a previous conviction for rape. Consideration was then given to whether the pathways to offending differ based on this distinction. Observational (cross-sectional).	150 sexual murderers	N/A	N/A	The Rape Myths Scale (Offending Behaviour Programmes Unit - now known as Operational Services & Intervention group [OS&IG], 1995)	Rape myths were not analyzed in isolation. However, upon reading the results tables, it is shown that of the men in the *high problem* group (offenders who were likely to report high levels of sexual entitlement beliefs, rape myths, have problems with being open to others and tend to believe that women are deceitful) 35% ($p < 0.001$) were found to endorse rape myths, whereas in the *low problem* group (those who did not report problems in the aforementioned areas), 13% ($p < 0.001$) endorsed rape myths.	Unclear
Webster et al. (2004) [Webster 2004]	Explored the impact of the Prison Service CORE Sex Offender Treatment Programme (SOTP) upon Black sexual offenders as compared with White sexual offenders. Observational (before-and-after).	52 Black sexual offenders	CORE SOTP	52 White sexual offenders Pre- and post-intervention	Burt's RMAS	Groups significantly improved on the RMAS post-treatment. There were no other significant within-subjects main effects or interactions. *It should be noted that when reviewing the statistics, it would appear that child molesters had significant change in their RMAS post-treatment as compared to the rapists that seem to exhibit very little/no change at all. Additionally, an interaction seems to be apparent here where the White rapists experienced a positive change in RMAS post treatment. As this was not the main focus of the study, the in-depth statistics for these were not reported.	Low

appropriateness of sexual partners, sexual acts, conditions or circumstances under which sex should occur" (Burt, 1980, p. 219). The final scale, Acceptance of Interpersonal Violence, contains 6 items which refer to the notion that force and coercion are legitimate ways to gain compliance, especially in sexual relationships (Cohen, 2012, p. 50).

Not much could be uncovered for this review on the properties of the Operational Services and Intervention Group's Rape Myths Scale as it appears to be integrated into Her Majesty's Prison Service's Sex Offender Treatment Programme psychometric battery which remains unpublished. Stefanska et al. (2015) offer a brief outline. The Rape Myths Scale is a 17-item measuring externalization around rape and a higher score indicates a greater acceptance of justifications for rape. A good internal consistency ($\alpha = 0.83$) and test re-test reliability ($r = 0.85$) were reported.

3.6. Risk of bias ratings of included studies

The risk of bias ratings across studies included in the review varied extensively. Two studies had what was deemed the least amount of bias, which can be judged to equate to a higher quality, with overall ratings of "low risk" of bias (Marshall & Hambley, 1996; Webster et al., 2004). On the other hand, five studies were rated as being at an overall "high risk" of bias (Beech et al., 2006; Cohen, 2012; Olver et al., 2014; Overholser & Beck, 1986; Pithers, 1994) for the following reasons: incomplete outcome data (Beech et al., 2006; Cohen, 2012; Olver et al., 2014; Overholser & Beck, 1986; Pithers, 1994); no blinding of the participants and/or personnel (Beech et al., 2006; Cohen, 2012; Olver et al., 2014; Overholser & Beck, 1986; Pithers, 1994); no blinding of the outcome assessment (Beech et al., 2006; Cohen, 2012; Olver et al., 2014; Pithers, 1994); lack of random sequence generation (Cohen, 2012; Overholser & Beck, 1986); lack of allocation concealment (Overholser & Beck, 1986); selective reporting (Pithers, 1994); and other biases (Cohen, 2012, Olver et al., 2014; Pithers, 1994).

3.7. Narrative data synthesis and key findings

The data extracted from the studies varied in aims, methodology, and participant group and consequently, the results of the studies are not homogenous. As such, it was deemed inappropriate to combine the results and evaluate them as part of a meta-analysis. Alternatively, a narrative data synthesis will be carried out for this review, highlighting key findings of the studies in relation to the aims outlined in the Introduction.

3.7.1. Can adult, male rapists be distinguished from adult, male child molesters, non-sexual offenders, or non-offenders on measures on RMA?

Two studies in this review explicitly compared rapists with non-rapists. Both studies employed Burt's Rape Myth Acceptance Scale as a measure of RMA with Cohen (2012) utilizing the 55-item scale. Overholser and Beck (1986) did not specify the number of items in the scale used, however, the results reported an effect on the "Sex Role Stereotyping Scale" (p. 686) and it may be reasonable to assume that the researchers used a similar scale to Cohen.

Cohen (2012) reported that for most measures of offence-related cognition non-offenders scored similarly to sexual offenders. However, on the measure of rape myth acceptance, sex offenders reported a significantly *lower* level of rape myth acceptance than non-offenders. They found that this response could not be attributed to social desirability, at least not in that particular study, any more than could be attributed to social desirability in non-offenders. Overall, Cohen (2012) found that cognitive distortions were present in both sex offenders and non-offenders. However, in sex offenders they found that these distortions interact with lack of sex-role satisfaction (a subscale of the Burt Rape Myth Acceptance Scale), high trait anger and trait anxiety, lending them an emotional salience not present in non-offenders. The study found that cognitive distortions manifest in high risk situations for offenders but not for non-offenders.

Overholser and Beck (1986) did not find any significant effect on Burt's Rape Myth Acceptance Scale, however they did find a significant main effect for group on the Sex Role Stereotyping subscale. Further analysis showed that child molesters displayed significantly higher levels of sex role stereotyping (i.e. more conservative) than did both the community control groups, however, the rapists did not appear to be distinguishable from the child molesters.

Although a comparison group was not included in the main study, Beech et al. (2006) compared their results to those of a similar study by Beazley Richards (2000) whose sample consisted of UK male non-offender employees of a civil engineering company and UK male non-sexual violent offenders. Beech and colleagues found that sexually motivated offenders (those who were primarily motivated to have sex and have used some form of force or violence against the victim to achieve this aim) scored higher on the Sex Role Stereotyping subscale than non-sexual violent offenders and community non-offending males.

Even though comparison between rapists and non-rapists was not the focus of the study, Pithers (1994) reported that Burt's Rape Myth Acceptance Scale did not discriminate between child abusers and rapists.

Overall, the results of these studies suggest that rapists may be distinguished from other non-sexual offenders and from community non-offending males on measures of rape myth acceptance, particularly on the sex-role subscales of the Burt Rape Myth Acceptance Scale. It could be that these subscales are more sensitive to differences that separate sexual offenders from non-sexual offenders and non-offenders. However, whether these results are reliable or consistent may be called into question. Furthermore, there still appears to be difficulty in discriminating between child molesters and rapists when relying on rape myth acceptance as the distinguishing factor. Also, an unexpected result was noted in the Cohen study in which sex offenders reported lower acceptance of rape myths than non-offenders.

3.7.2. Are there differences in levels of RMA between different sub-groups of rapists?

Two studies examined sub-groups of rapists. Beech et al. (2006) carried out typological comparisons by categorizing rapists into one of five main types as described by Knight and Prentky (1990): opportunistic, pervasively angry, vindictive, sexual non-sadistic and sexual sadistic rapists. The researchers found that groups differed significantly on their scores for the Adversarial Sexual Beliefs subscale of Burt's Rape Myth Acceptance Scale. Post-hoc analyses revealed that opportunistic rapists scored significantly higher than both the sexual sadistic and sexual non-sadistic types. Difference in scores between the sexual sadistic and non-sadistic types was not significant. Overall, the sexual non-sadistic rapists had the lowest scores on the scale. This finding may likely reflect that the sexual non-sadistic rapists hold less negative views about sexual relationships when compared to the other groups as Beech and colleagues found no correlation between measures of socially desirable responding and this subscale. Conversely, according to Burt (1980), the opportunistic rapists' scores revealed beliefs that sexual partners are manipulative, cunning, and not to be trusted. Vindictive rapists had the second largest mean scores on the Adversarial Sexual Beliefs subscale. Beech et al. (2006) do note, however, that despite these findings, the highest mean score overall ('opportunistic' sub-types: mean = 22.5) was only 0.4 of a standard deviation above the mean of 20.6 of a non-offender sample (from Beazley Richards, 2000). On the Acceptance of Interpersonal Violence subscale, the opportunistic rapists scored significantly higher than both the sexual sadistic and non-sadistic types. This time, the opportunistic rapists' mean score was over two-and-a-half standard deviations above the non-offending sample's mean score and 1.3 of a standard deviation above the mean of a sample of incarcerated non-sexual violent offenders. The researchers offer an explanation; this scale reflects the notion that coercion and force are

legitimate modes through which to gain compliance in intimate and sexual relationships, versus relationships in general.

Although not explicitly defined as a "subgroup" of rapists, sexual murderers will be included in this section of the review. In regards the sexual murderers, Beech et al. (2006) found no difference in RMA between rapists and sexual murderers.

Stefanska et al.'s (2015) study on sexual murderers' pathways to offending explored potential differences between rapists who were also sexual killers and sexual killers who had never solely committed rape (i.e. without killing the victim). It was difficult to gather data from this study as rape myths were not analyzed in isolation. However, upon reading the results tables, it is shown that of the men in the "high problem" group (offenders who were likely to report high levels of sexual entitlement beliefs, rape myths, have problems with being open to others and tend to believe that women are deceitful) 35% ($p < 0.001$) were found to endorse rape myths, whereas in the low problem group (those who did not report problems in the aforementioned areas), 13% ($p < 0.001$) endorsed rape myths. However, this data combines both groups of sexual murderers so it was impossible to extract and separate the data to allow for comparison between the two groups.

3.7.3. Can differences in levels of RMA discriminate between rapists who reoffend (recidivists) and those who do not?

Two studies examined recidivism outcomes (Beech et al., 2006; Olver et al., 2014), however, in the Beech study, recidivism was not explored in terms of RMA. In the Olver study, RMA was not found to be a significant predictor of sexual or violent recidivism nor did the study compare recidivists with non-recidivists.

3.7.4. Is RMA amenable to sex offender treatment programs?

Four studies examined the effect of treatment on rape myth acceptance. Beech et al. (2006) found no main effect of treatment in sexual murderers and no effect in rapists overall. However, when looking at the impact of treatment on typologies, Beech et al. (2006) grouped the Knight and Prentky (1990) typologies into three groups: sexually motivated (opportunistic and sexual non-sadistic rapists), anger motivated (vindictive and pervasively angry rapists) and sexual sadistic rapists remained a standalone "sadistic" type. Analysis revealed a significant change in scores overall on the Sex Role Stereotyping scale showing scoring actually increasing post-treatment. This is indicative of greater endorsement of stereotypical beliefs. There were no differences found between typologies. A result like this could indicate something inherent in the program that would change these scores for the worse, however it should be noted that Beech et al. (2006) mentioned than quite a few offenders were removed from this sample as they could not be grouped into the categories.

Olver et al. (2014) reported that the mean RMA score was approximately one full standard deviation below the normative mean for both offenders and non-offenders (as reported in Burt, 1980) at pre-treatment, and approximately two-thirds of a standard deviation lower at post-treatment. There was a significant decrease in rape myths endorsed within the sample from pre- to post-treatment which would suggest that the Clearwater Programme has the capabilities to effect positive change in cognitive distortions around rape. Similarly, Pithers (1994) reported a significant treatment effect and found that there was a reduction in acceptance of rape myths after completion of the program. This, perhaps, points to some effectiveness of victim empathy programs for use with rapists.

Although the major outcome for Webster et al.'s (2004) study was differences in psychometric data between Black versus White sexual offenders, the researchers did investigate changes in RMA and found that both groups significantly improved on rape myths. However, these data view look at child molesters and rapists combined. Also, it should be noted that when reviewing the statistics, it would appear that child molesters had significant, positive change in their rape myths post-treatment as compared to the rapists that seem to exhibit very little/no change at all. Additionally, an interaction seems to be apparent here where the White rapists experienced a positive change in rape myths post treatment. As this was not the main focus of the study, the detailed statistics for these were not reported.

4. Discussion and conclusions

4.1. Main findings of the review

This systematic review explored the relevance of rape myth acceptance as a treatment need for rapists. Overall, the literature on rape myth acceptance is quite vast, however, it was surprising how few studies were found that addressed this in rapists ($n = 8$) versus the general public. Only studies with identified specific measures of RMA were included in this review, which led to the exclusion of studies solely using the Bumby Rape Scale as there was disagreement amongst experts as to its use as a measure of RMA. Also included, were studies which had convicted rapists as participants. Studies focusing on rape proclivity with non-offending samples only were excluded. Being quite strict with measures of rape myth acceptance may have biased this review in a way as seven out of the eight studies included utilized the Burt Rape Myth Acceptance scale. Perhaps broadening the definition could allow for the inclusion of more studies. It may be worth mentioning that Payne, Lonsway and Fitzgerald's (1999) Illinois Rape Myth Acceptance Scale would have been suitable for this review, however, the author could not locate or gain access to studies which examined its use that met the inclusion and exclusion criteria. However, the search carried out was extensive. The search was conducted across three platforms, OvidSP, Web of Science, and Proquest and also included hand searching of the reference lists of included studies and a meta-analysis, and contacting many experts in the field for any published or unpublished literature. The response from experts was remarkable and four additional papers were garnered from this. Overall, there is confidence that most relevant research has been included in this review and that the consequent conclusions are from the synthesis of a solid evidence base.

The results indicate that while rapists can be distinguished from non-offenders and non-sexual offenders on measures of RMA, they cannot be significantly discriminated from child molesters by relying on these measures. Also, in regards to rapists and sexual murders, the two groups could not be distinguished using RMA scores. In analyzing differences that were found, Cohen (2012) had data in the opposite direction from what is to be expected and reported that sexual offenders scored lower on rape myth acceptance than non-offenders. In terms of differentiating between subgroups of rapists, Beech et al.'s (2006) findings were quite enlightening. The opportunistic rapists scored higher on the Adversarial Sexual Beliefs subscale with the sexual non-sadistic rapists scoring the lowest. Additionally, the opportunistic rapists as scored the higher on the Acceptance of Interpersonal Violence subscale. Knight and Prentky (1990) posited that the opportunistic rapist views violence as an instrument to be used if needed to succeed in a sexual attack. They are described as taking advantage of an opportune situation and are indifferent to any impact this may have on the victim. Beech et al.'s (2006) study is congruent with this assertion. Studies did not compare recidivists and nonrecidivists, nor was RMA found to be a significant predictor of sexual or violence recidivism (Stefanska et al., 2015). RMA did, however, appear to be affected by sex offender treatment programs in a positive manner. Studies reported significant positive change in RMA after sex offenders completed a treatment program.

4.2. Strengths and weaknesses of the review

The major weakness of this review has been mentioned in Section 4.1, namely the limitation on measures included in the review. Additionally, the assessment of quality guidelines used to judge the literature was quite strict and as such some of the studies reviewed may be deemed 'poor quality'; this is discussed further Section 4.3. However,

this review employed a comprehensive research strategy guided by the advice of experts within the field. Additionally, new information around the applicability of RMA in rapist typologies has been introduced and long held assumptions about RMA in sexual offenders versus in the general public have been challenged in this review. The impact for future directions is outlined in Section 4.3.

4.3. Implications for practice and future direction

The review adopts the Cochrane principles of systematic review; however, the randomized control trial is heralded as the belief is held that other types of trial evidence are likely to inflate the positive findings for the intervention (Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996). Due to the nature of the research being sought in this review, a completely "randomized controlled trial" would be impossible to attain. Even far more robust reviews struggle with adhering to Cochrane principles for judging sex offender treatment. The Hanson, Bourgon, Helmus, and Hodgson (2009) review of 129 sex offender treatment studies could rate none as 'strong' according to Collaborative Outcome Data Committee guidelines. For this reason, three quality assessors came to an informed agreement about the risk of bias present in studies. As this review is interested in relevance of an attitudinal variable rather than treatment efficacy or effectiveness as the primary outcome, attaining the "gold standard" as determined by Cochrane is out of reach. So, it is worth noting that for future reviews, studies examining this construct would best not be marked as high risk based on the fact that they are conceptually different from randomized control studies.

Out of the eight studies in this review, the most comprehensive was Beech et al.'s (2006) as it was a part of a large-scale project with the Prison Service. They found differences in scores on RMA subscales amongst rapists' typologies, which may be indicative of the differences in beliefs of each of the typologies. If this is the case, then it is important that these differences be identified to develop specific treatment programs to target these beliefs. Sex Role Stereotyping was a subscale on which sexually motivated offenders scored highly on. It could be possible that this stereotyping is linked to the feeling of one's entitlement to male dominance or power. It is evident that future work needs to target differentiating amongst sexual offenders and utilizing implicit measures to measure these associations may be a means of overcoming the transparency of using explicit measures alone.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Characteristics of included studies (ordered by study ID)

Beech 2006	
Methods	Design: observational (before-and-after) Measure (of interest): Burt's Rape Myth Acceptance Scale (RMAS; 1980)
Participants	Participants: male prisoners incarcerate for rape ($n = 112$) or sexual murder ($n = 58$) Sex: all male Age: rapists - mean 34.9 (SD 8.4) years; sexual murderers - mean 39.3 (SD 10.5) years Setting: 7 prisons Inclusion criteria: prisoner with conviction for sexual offence apart from convictions related to consensual sexual behaviour; prisoner falls into medium- or high-risk group as determined by Risk Matrix 2000 (RM2000; Thornton et al., 2003); prisoner with a homicide conviction with a clear sexual element to the homicide Exclusion criteria: Psychopathy Checklist - Revised (PCL-R; Hare, 1991) score > 30; IQ < 80; men suffering from current mental illness; men suffering from mental illness/brain damage at time of offence; men deemed 'not ready' for treatment (treatment not suitable for him at this time); total denial of the offence; refusal of treatment; does not speak English; physical disability incl. deafness or blindness; poor literacy; suicidal or self-harming
Interventions	CORE Sex Offender Treatment Programme (SOTP) - average treatment dose in study: 188 h (94 sessions); two to five sessions per week
Outcomes	Difference in RMA: • none found between rapists and sexual murderers • no within-group change in pre-treatment to post-treatment for both rapists and sexual murderers • rapists typologies differed on 'Adversarial Sexual Beliefs' subscales scores on RMAS • *sexually motivated* offenders scored higher of 'Sex Role Stereotyping' subscale of RMAS than non-sexual violent offenders and community non-offending males
Notes	N/A

Risk of bias

Bias	Author's judgment	Support for judgment
Random sequence generation	Low risk.	There is a clear risk of selection bias when the person recruiting participants knows in advance the clinical characteristics of a participant and which intervention they will receive. However, due to the study design, this is not very feasible to randomize. *Decision made to override 'high risk' rating.*
Allocation concealment	Low risk.	Lack of concealed randomized allocation increases the risk of selection bias. However, due to the study design, this is not possible to conceal. *Decision made to override 'high risk' rating.*
Blinding of participants and personnel	High risk.	Participants were not blinded. Personnel were not blinded. All interviews and the treatment program itself were conducted by the investigators.
Blinding of outcome assessment	High risk.	Outcome assessment was not blinded. All interviews and treatment program carried out by the investigators.
Incomplete outcome data addressed	High risk.	Rapists (pre-treatment): 86/112 completed pre-treatment assessments; of the remaining 16, 14 refused at interview to complete the questionnaires and 12 simply did not fill them out. Rapists (post-treatment): 65/86 that completed the questionnaires before treatment also completed them afterwards; the remaining 21 failed to complete and return the questionnaires. Sexual murderers (pre-treatment): 45/58 completed pre-treatment assessments; the remaining 13 either refused or failed to complete and/or return the questionnaires. Sexual murderers (post-treatment): 40/45 that completed the questionnaires before treatment, completed them after treatment; of the remaining 5, 2 refused to complete them and 3 failed to complete/return them.

Appendix A *(continued)*

Risk of bias		
Bias	Author's judgment	Support for judgment
Selective reporting	Low risk.	The published report included all expected outcomes, including those that were pre-specified
Other bias	Low risk.	The study appeared to be free of other sources of bias.

Cohen 2012	
Methods	Design: controlled observational (case control) Measure (of interest): Burt's Rape Myth Acceptance Scale (RMAS; 1980)
Participants	Participants: sex offenders (experimental group; $N = 46$ initially, $N = 44$ final); non-offender students (control; $N = 50$ initially, $N = 44$ final) Sex: all male Age: offenders - mean 36.3 (SD 14.78) years; students - mean 27.8 (SD 3.5) years Setting: treatment center in central-Israel and the Probation Department. Inclusion criteria: these were not explicitly outlined. However, the following information could be gathered from the article: "sex offender" in Israeli law (someone who has committed an offence according to Chapter 6 of the 1977 Criminal Code. These offences include rape, sodomy, sexual assault without penetration and 'flashing'. Also included are 'consensual' sexual conduct between an adult and a teenager under 16, between a therapist and a patient, or between an employer and an employee.); physical ability to complete questionnaires; appropriate reading ability; ability to read and comprehend Hebrew Exclusion criteria: these were not explicitly outlined. However, the following information could be gathered from the article: student's disclosure of any undetected sexual coercion; does not speak Hebrew; physical disability; poor literacy; colour-blindness
Interventions	N/A
Outcomes	Sex offenders and students had similar scores on the attribution subscale and the whole-form (both combined subscales) of the MCSDS. Sex offenders had significantly lower scores on the denial subscale of the MCSDS. For most inventories dealing with offence-related cognitions, the scores of non-offenders and sex offenders were similar, but sex offenders reported significantly lower levels of rape myth acceptance than did non-offenders. The sex offenders' self-reported levels of sex role satisfaction and sexual entitlement were correlated with measures of social desirability. Participants in both groups showed slower reaction times on the ES and the LD, but no difference in accuracy on the ES in response to "general threat" words compared to neutral words. Non-offenders showed slower reaction times on the ES and LD, but no difference in levels of accuracy on the ES in response to "general threat" words compared to neutral words and "uncontrollability", "entitlement", and "sexy children" words. Sex offenders reacted similarly to "general threat" and to "uncontrollability" and "entitlement", on the ES and LD. Reaction times to these words were slower on the ES and the LD, and but levels of accuracy on the ES did not differ significantly to these words than to neutral words. The Emotional Stroop showed significant differences in response for sex offender-specific words but not "general threat" words and the LD showed a difference for "uncontrollability" words only. Sex offenders who victimized children showed slower reaction times on the LD in response to "sexy children" words than offenders with older victims, but this was not the case for the Emotional Stroop.
Notes	N/A

Risk of bias		
Bias	Author's judgment	Support for judgment
Random sequence generation	High risk.	There is a clear risk of selection bias when the person recruiting participants knows in advance the clinical characteristics of a participant and which intervention they will receive. However, due to the study design, random sequence generation for the offenders was not very feasible. Although, there was some bias in the selection of students as they were only recruited from a particular department.
Allocation concealment	Low risk.	Lack of concealed randomized allocation increases the risk of selection bias. However, due to the study design, this is not possible to conceal.
Blinding of participants and personnel	High risk.	Participants were not blinded. Personnel were not blinded. All interviews and the treatment program itself were conducted by the investigators. Quote: "Almost all of the sex offenders who participated in the study were in under some form of judicial impetus to participate in this treatment." "Students were solicited in the usual way, through advertisements on bulletin boards in the Criminology departments. The students participated in the study in return for "signatures" that they partook in an experiment (a requirement for undergraduate students). As not enough respondents were recruited in this way, the author approached a colleague who taught a summer course in Criminology at Bar-Ilan University, and requested that she enlist her students to participate."
Blinding of outcome assessment	High risk.	Outcome assessment was not blinded. All interviews carried out by the investigators.
Incomplete outcome data addressed	High risk.	Sex offenders: 2/46 removed from study as they were unable to complete both the ES and LD tasks. 44 male sex offenders in the final experimental group. 1/44 refused to provide full demographic information on himself. Students: 6/50 rejected due to physical problem which affected their performance (3/6), recent immigration to Israel resulting in imperfect command of Hebrew (2/6), and one (1/6) admitted to having physically coerced a woman to have sex.
Selective reporting	Low risk.	The published report included all expected outcomes, including those that were pre-specified.
Other bias	Unclear risk.	Decision to use two implicit tests but in methodologically dissimilar ways. "The present study used a 'fast' ES task and an untimed LD task, and then compared the results. The decision to use the untimed LD rather than an untimed ES, was based on the necessity to minimize as much as possible the possibility of 'cheating' (e.g., conscious efforts to avoid focusing on word content). Despite the aforementioned difficulty of 'cheating' on the ES, it was decided that an untimed version of that task, which does not necessitate the actual reading of the target word, was more open to manipulation than an untimed LD, in which words must be read. Therefore, an untimed LD was used. While the decision to use the different tests is, I believe, the correct one, the fact that ES and LD do not work in exactly the same way cannot be ignored (for example, LD is more susceptible to variations in word frequency)."

Marshall 1996	
	Design: observational (cross-sectional) Measure (of interest): Burt's Rape Myth Acceptance Scale (1980)
Participants	Participants: male prisoners who had been convicted of sexually assaulting an adult female ($N = 27$) Sex: all male Age: range: 21–58 years. Mean age: 33.89 (SD 8.31)

Appendix A *(continued)*

Marshall 1996	
	Design: observational (cross-sectional) Measure (of interest): Burt's Rape Myth Acceptance Scale (1980)
	Setting: Canadian penitentiary unit or Canadian medium-security penitentiary. Inclusion criteria: male; incarcerated; convicted of sexually assaulting an adult female. Exclusion criteria: these were not explicitly outlined.
Interventions	N/A
Outcomes	Loneliness and intimacy scores were negatively correlated with and were found to share >60% variance in common. Similarly, hostility toward women and rape myth acceptance were significantly related, sharing 67% of their variance in common. Regarding the relationship between scores on the loneliness and intimacy measures on one hand and scores on the hostility and rape myth acceptance scales on the other; intimacy appeared to be more strongly related to hostility toward women ($r = 0.79$) than does loneliness ($r = 0.53$), and intimacy is also more strongly related to the acceptance of rape myths ($r = 0.68$) than is loneliness ($r = 0.39$).
Notes	N/A

Risk of bias

Bias	Author's judgment	Support for judgment
Random sequence generation	Low risk.	Study design did not allow for random sequence generation. *Decision made to override 'high risk' rating.*
Allocation concealment	Low risk.	Lack of concealed randomized allocation increases the risk of selection bias. However, due to the study design, this is not possible to conceal. *Decision made to override 'high risk' rating.*
Blinding of participants and personnel	High risk.	Neither participants nor personnel could be blinded.
Blinding of outcome assessment	High risk.	Outcome assessment was not blinded. All interviews carried out by the investigators.
Incomplete outcome data addressed	Low risk.	There were no missing outcome data.
Selective reporting	Low risk.	All outcomes reported as pre-specified.
Other bias	Low risk.	The study appeared to be free of other sources of bias.

Olver 2014	
Methods	Design: observational (case series) Measure (of interest): Burt's Rape Myth Acceptance Scale (1980)
Participants	Participants: male federal sex offenders who attended treatment services at a high intensity sex offender treatment program ($N = 267$) Sex: all male Age (at program admission): range: 18–66 years. Mean age: 32.22 (SD 8.99) Setting: maximum-security forensic psychiatric facility in Saskatoon, Saskatchewan, Canada: The Regional Psychiatric Centre (RPC) Inclusion criteria: male; one or more index convictions for contact sexual offences Exclusion criteria: these were not explicitly outlined.
Interventions	The Clearwater Sex Offender Programme: a cognitive-behaviourally based treatment program, approximately 6–8 months in duration, mandated to target moderate to high risk sex offenders.
Outcomes	The 257 offenders included in outcome analyses were followed up an average of 18.2 years (SD 4.7) post release. Employing a 20-year cap on follow-up time, 73 (27.3%) men were convicted for a new sexual offence and 135 (50.6%) were convicted for any new violent (including sexual) offence. Secondary Outcomes: the mean Rape Myth Acceptance (RMA) score was approximately one full standard deviation below the normative mean for both offenders and non-offenders (Burt, 1980) at pre-treatment, and approximately two-thirds of a standard deviation lower at posttreatment. There was a significant decrease in rape myths endorsed within the sample from pre- to post-treatment. Additionally, there was a significant convergent validity correlation pre-treatment between RMA and two of the three VRS-SO factors - criminality ($r = 0.16$, $p < 0.05$) and treatment responsivity ($r = 0.22$, $p < 0.01$). Post-treatment, RMA had no correlation with criminality ($r = 0.00$) and still, a significant positive correlation with treatment responsivity ($r = 0.21$, $p < 0.01$)
Notes	N/A

Risk of bias

Bias	Author's judgment	Support for judgment
Random sequence generation	Low risk.	Study design did not allow for random sequence generation. *Decision made to override 'high risk' rating.*
Allocation concealment	Low risk.	Lack of concealed randomized allocation increases the risk of selection bias. However, due to the study design, this is not possible to conceal. *Decision made to override 'high risk' rating.*
Blinding of participants and personnel	High risk.	Neither participants nor personnel could be blinded.
Blinding of outcome assessment	High risk.	Outcome assessment was not blinded.
Incomplete outcome data addressed	High risk.	Missing outcome data apparent in results table, however, no explanation offered.
Selective reporting	Low risk.	All outcomes reported as pre-specified.
Other bias	High risk.	Quote: "Sexual recidivism was defined as any conviction for a new sexual offence following first release to the community after program participation. Violent recidivism was defined as any conviction for a person involved offence (including sexual offences) with the potential for physical or psychological harm (e.g. non-sexual assault and robbery). Both violent and sexual recidivism were coded in a binary manner (i.e., 0"no recidivism; 1"recidivism)." Investigators chose to include sexual offences in violent recidivism results which may lead to skewed finding. For example, in the same offender it would be impractical compare sexual recidivism to their violent recidivism if there is cross-over of offences. Quote: "content of the risk need domains was constrained by the availability of measures used in the sex offender treatment program at the RPC. For instance, a measure of child molester cognitive distortions was not introduced until some years later into the program (Attitudes Toward Sex with Children) and could not be included owing to large amounts missing data. As a result, the attitudinal domain in the present study did not contain a measure of child molester cognitive distortions in contrast to other related studies (e.g. Allan, Grace, Rutherford, & Hudson, 2007; Craig, Thornton, Beech, & Browne, 2007; Thornton, 2002; Wakeling, Beech, & Freemantle, 2013)."

Overholser 1986	
Methods	Design: quasi-experimental Measure (of interest): Burt's Rape Myth Acceptance Scale (1980)
Participants	Participants: male inmates at a medium-secure prison Experimental: 12 rapists; 12 child molesters Controls: 12 non-sex-offender inmates; 12 community-based low SES men; 12 "minimal-dater" college students Sex: all male Age: rapists - $M = 34.5$ (SD 12.2); child molesters - $M = 38.8$ (SD 6.1); non-sex-offender prisoners - $M = 37.8$ (SD 9.5); low-SES volunteers - $M = 33.8$ (SD 8.5); college students - $M = 20.4$ (SD 1.3)

Appendix A *(continued)*

Overholser 1986

Methods	Design: quasi-experimental Measure (of interest): Burt's Rape Myth Acceptance Scale (1980)
	Setting: not specified. Inclusion criteria: Rapists: male; committed sexual offence involving nonconsensual sexual contact with a female nonrelative who was over the age of 17. Child molesters: male; committed sexual offence against a female, nonrelative who was under the age of 12; the offender was 18 years of age or older Non-sex-offender prisoners: male; no prior record of sexual offences; denied ever participating in coercive sexual activity Community-based men: low SES; (matched to prison participants) College-students: male; adult; single, no girlfriend; dated less than twice in the past month and less than four times in the past six months; reported feeling at least moderately anxious when in social settings with women Exclusion criteria: these were not explicitly outlined.
Interventions	N/A
Outcomes	No significant effect was found on the Rape Myth Acceptance Scale. A significant main effect for group was found for the Sex Role Stereotyping scale, $F(4, 55) = 4.00$, $p < 0.01$. A Newman-Keuls analysis indicated that child molesters displayed significantly higher levels (more conservative) of sex role stereotyping than did both the community-based low-SES men and the minimal-dater college students. Heterosocial skills deficits were observed in both child molesters and rapists, in comparison with the nonincarcerated control groups, while they participated in the naturalistic controlled interaction and in the role-play scenes. Rapists in the study displayed higher levels of physiological arousal in the assertive role-play scenes than did the other groups. College students who were minimal daters appeared more behaviourally anxious in the role-play scenes than did the other groups. Additionally, behavioural and physiological differences were found amongst the groups in interactions with the confederate, which suggests that the controlled interaction scene and the role-play scenes still appeared to provide assessments of all subjects' general ability to satisfactorily interact with a woman. Child molesters displayed significantly higher levels of fear of negative evaluations. Hostility and impulsivity as measured in this study did not differentiate child molesters and rapists from the control groups. In general, child molesters and rapists did not appear all that dissimilar on several diverse measures.
Notes	N/A

Risk of bias

Bias	Author's judgment	Support for judgment
Random sequence generation	High risk.	Person recruiting participants knew in advance the clinical characteristics of participants and which intervention they will receive.
Allocation concealment	High risk.	Lack of concealed randomized allocation increases the risk of selection bias.
Blinding of participants and personnel	High risk.	Participants blinded. Personnel only partially. Quote: "The female confederate was a 22-year-old undergraduate student who knew the purpose of the study but did not know the status of the prisoner subjects. Data from the college students and community men were collected at a university laboratory, and thus the confederate was aware of the status of the two nonprisoner groups."
Blinding of outcome assessment	Low risk.	Quote: "Three undergraduate students, who were naive as to the purpose and types of men in the study, were trained to observe and score the controlled interaction and role-play scenes on the molecular and global ratings of social skills and social anxiety. The observers were trained with practice tapes until their agreement was at least 80%."
Incomplete outcome data addressed	High risk.	There were no missing outcome data.
Selective reporting	Low risk.	All outcomes reported as pre-specified.
Other bias	Low risk.	The study appeared free of other sources of bias.

(continued)

Risk of bias

Bias	Author's judgment	Support for judgment

Pithers 1994

Methods	Design: observational (before-and-after) Measure (of interest): Burt's Rape Myth Acceptance Scale (1980)
Participants	Participants: convicted males ($N = 20$): paedophiles ($n = 10$); rapists ($n = 10$) Paedophiles that abused prepubescent males exclusively ($n = 4$) Paedophiles that abused prepubescent females exclusively ($n = 4$) Paedophiles that abused children of both genders ($n = 2$) All rapists had abused adult females ($n = 10$) Sex: all male Age: rapists - $M = 32.2$ (SD 7.53); paedophiles - $M = 36.3$ (SD 9.79) Setting: Northwest State Correctional Facility Inclusion criteria: these were not explicitly outlined. Exclusion criteria: these were not explicitly outlined.
Interventions	Survivor empathy group (as part of Vermont Treatment Programme for Sexual Aggressors; Pithers et al., 1989)
Outcomes	A univariate repeated-measures ANOVA did not identify a significant Group effect, $F(1, 18) < 1$, although the Treatment effect was significant, $F(1, 18) = 117.47$, $p < 0.001$. The Group × Treatment interaction was not significant, $F(1, 18) < 1$. Thus both groups displayed reduced acceptance of rape myths. Paedophiles and rapists did not differ in pre-treatment or post-treatment endorsement of cognitive distortions hypothetically related to or rape. Scores on Burt's Rape Myth Acceptance Scale, which would be expected to reveal deficits in the rapists, did not discriminate these samples of child abusers and rapists.
Notes	N/A

Risk of bias

Bias	Author's judgment	Support for judgment
Random sequence generation	Unclear risk.	Insufficient information about the sequence generation process to permit judgment of 'low risk' or 'high risk'
Allocation concealment	Low risk.	Lack of concealed randomized allocation increases the risk of selection bias. However, due to the study design, this is not possible to conceal. *Decision made to override 'high risk' rating.*
Blinding of participants and personnel	High risk.	Participants were not blinded. Personnel were not blinded
Blinding of outcome assessment	High risk.	Investigators were not blind to outcome assessment.
Incomplete outcome data addressed	High risk.	There were no missing outcome data.
Selective reporting	High risk.	Did not report relationship between RMAS and established measures of empathy change.
Other bias	High risk.	Used Rape Myth Acceptance Scale as an "indirect measure of empathy" but its use in this manner had not been validated (or there is no mention of its use in this way in the article). But states an alternative use: "or…assess[es] a construct that may be expected to vary as empathy varies".

Stefanska 2015

Methods	Design: observational (cross-sectional) Measure (of interest): the Rape Myths Scale (Offending Behaviour Programmes Unit, 1995)
Participants	Participants: sexual killers ($N = 150$); sexual killers with previous rape or attempted rape offence ($n = 44$) Sex: all male Age: range: 18–45 years. Mean age at the time of the offence: 25.87 (SD 7.23)

(continued on next page)

Appendix A *(continued)*

Stefanska 2015	
Methods	Design: observational (cross-sectional) Measure (of interest): the Rape Myths Scale (Offending Behaviour Programmes Unit, 1995)
	Setting: data retrieved from National Offender Management Service, OASys research database; Sex Offender Treatment Programme (SOTP) database; and Public Protection Unit Database (PPUD) Inclusion criteria: these were not explicitly outlined. However, the following information could be gathered from the article: male; sex offender (rapist or child molester); non-serial sexual killers (those convicted of killing one or two victims without an emotional cool-off period, e.g. two victims killed at the same time or within a period indicative of a single event) have been convicted and served or are serving a custodial sentence within HM Prison Service; victims are females aged 14, or older; a sexual killing includes murders and manslaughters where a sexual element and/or a sexual motivation was evidenced, suspected or admitted; completed the SOTP; appropriate reading ability; ability to read and comprehend English Exclusion criteria: these were not explicitly outlined. However, the following information could be gathered from the article: serial sexual killers; sexual murderers of men; sexual murderers of children; does not speak English; physical disability; poor literacy; negative attitude to treatment/low motivation
Interventions	N/A
Outcomes	Rape myths were not analyzed in isolation. However, upon reading the results tables, it is shown that of the men in the high problem group (offenders who were likely to report high levels of sexual entitlement beliefs, rape myths, have problems with being open to others and tend to believe that women are deceitful) 35% ($p < 0.001$) were found to endorse rape myths, whereas in the low problem group (those who did not report problems in the aforementioned areas), 13% ($p < 0.001$) endorsed rape myths.
Notes	N/A

Risk of bias

Bias	Author's judgment	Support for judgment
Random sequence generation	Unclear risk.	Insufficient information about the sequence generation process to permit judgment of 'low risk' or 'high risk'
Allocation concealment	Low risk.	Lack of concealed randomized allocation increases the risk of selection bias. However, due to the study design, this is not possible to conceal. *Decision made to override 'high risk' rating.*
Blinding of participants and personnel	High risk.	Neither participants nor personnel could be blinded.
Blinding of outcome assessment	High risk.	Outcome assessment was not blinded. All interviews carried out by the investigators.
Incomplete outcome data addressed	Low risk.	There were no missing outcome data.
Selective reporting	Unclear risk.	As the study was carried out for exploratory purposes, there were no predetermined hypotheses.
Other bias	Low risk.	The study appeared to be free of other sources of bias.

Webster 2004	
Methods	Design: observational (before-and-after) Measure (of interest): Burt's Rape Myth Acceptance Scale (RMAS; 1980)
Participants	Participants: Black sex offenders (experimental group; $N = 52$); White sex offenders (comparator; $N = 52$) Sex: all male Age: Black SOs - mean 31.22 (SD 12.14) years; White SOs - mean 35.38 (SD 10.54) years Setting: Her Majesty's Prison Service (data retrieved from the national database) Inclusion criteria: these were not explicitly outlined. However, the following information could be gathered from the article: male; sex offender (rapist or child molester); appropriate reading ability; ability to read and comprehend English

(continued)

Webster 2004	
Methods	Design: observational (before-and-after) Measure (of interest): Burt's Rape Myth Acceptance Scale (RMAS; 1980)
	Exclusion criteria: these were not explicitly outlined. However, the following information could be gathered from the article: does not speak English; physical disability; poor literacy; negative attitude to treatment/low motivation
Interventions	CORE SOTP
Outcomes	A main effect was not found for either ethnicity or type of victim. There was also no interaction effect between ethnic group and victim type. Within-subjects analysis showed that the groups significantly improved on rape myths, $F(1, 69) = 20.71$, $p < 0.001$. There were no other significant within-subjects main effects or interactions re: RMA.
Notes	N/A

Risk of bias

Bias	Author's judgment	Support for judgment
Random sequence generation	Unclear risk.	Insufficient information about the sequence generation process to permit judgment of 'low risk' or 'high risk'
Allocation concealment	Low risk.	Lack of concealed randomized allocation increases the risk of selection bias. However, due to the study design, this is not possible to conceal. *Decision made to override 'high risk' rating.*
Blinding of participants and personnel	Low risk.	Participants were not blinded. Personnel were not blinded. However, the outcome is not likely to be influenced by lack of blinding.
Blinding of outcome assessment	Low risk.	Investigators were blind to outcome assessment.
Incomplete outcome data addressed	Low risk.	There were no missing outcome data.
Selective reporting	Low risk.	The published report included all expected outcomes, including those that were pre-specified.
Other bias	Low risk.	The study appeared to be free of other sources of bias.

References

Allan, M., Grace, R. C., Rutherford, B., & Hudson, S. M. (2007). Psychometric assessment of dynamic risk factors for child molesters. *Sexual Abuse: A Journal of Research and Treatment, 19*, 347–367. http://dx.doi.org/10.1177/107906320701900402.

Beazley Richards, J. (2000). *A comparison of attitudes of convicted violent sexual offenders, convicted violent non-sexual offenders and non-offenders.* (Unpublished Masters dissertation) UK: University of Leicester.

Beech, A., Oliver, C., Fisher, D., & Beckett, R. C. (2006). *STEP 4: The sex offender treatment programme in prison: Addressing the needs of rapists and sexual murderers.* Birmingham: University of Birmingham Retrieved from www.hmprisonservice.gov.uk/assets/documents/100013DBStep_4_SOTP_report_2005.pdf

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.

Chapleau, K. M., & Oswald, D. L. (2010). Power, sex, and rape myth acceptance: Testing two models of rape proclivity. *Journal of Sex Research, 47*(1), 66–78. http://dx.doi.org/10.1080/00224490902954323.

Chiroro, P., Bohner, G., Viki, G. T., & Jarvis, C. I. (2004). Rape myth acceptance and rape proclivity: Expected dominance versus expected arousal as mediators in acquaintance-rape situations. *Journal of Interpersonal Violence, 19*(4), 427–442. http://dx.doi.org/10.1177/0886260503262081.

Cohen, D. A. (2012). *Attention bias to offense related stimuli in rapists and child molesters.* (Unpublished doctoral dissertation) Israel: Bar-Ilan University.

Cooke, A., Smith, D., & Booth, A. (2012). Beyond PICO the SPIDER tool for qualitative evidence synthesis. *Qualitative Health Research, 22*(10), 1435–1443.

Craig, L. A., Thornton, D., Beech, A., & Browne, K. D. (2007). The relationship of statistical and psychological risk markers to sexual reconviction in child molesters. *Criminal Justice and Behavior, 34*, 314–329. http://dx.doi.org/10.1177/ 0093854806291416.

Edwards, K. M., Turchik, J. A., Dardis, C. M., Reynolds, N., & Gidycz, C. A. (2011). Rape myths: History, individual and institutional-level presence, and implications for change. *Sex Roles, 65*, 761–773. http://dx.doi.org/10.1007/s11199-011-9943-2.

Feild, H. S. (1978). Attitudes toward rape: A comparative analysis of police, rapists, crisis counselors, and citizens. *Journal of Personality and Social Psychology, 36*(2), 156–179. http://dx.doi.org/10.1037/0022-3514.36.2.156.

Gray, J. M. (2006). Rape myth beliefs and prejudiced instructions: Effects on decisions of guilt in a case of date rape. *Legal and Criminological Psychology*, *11*(1), 75–80. http://dx.doi.org/10.1348/135532505X68250.

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.

Hanson, R. K., & Morton-Bourgon, K. (2004). Predictors of sexual recidivism: An updated meta-analysis. Retrieved from https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/2004-02-prdctrs-sxl-rcdvsm-pdtd/2004-02-prdctrs-sxl-rcdvsm-pdtd-eng.pdf

Hare, R. D. (1991). *The Hare psychopathy checklist-revised: Manual*. Multi-Health Systems, Incorporated.

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.

Higgins, J. P. T., & Green, S. (2011). *Cochrane handbook for systematic reviews of interventions version 5.1.0.* updated March 2011 The Cochrane Collaboration Available from www.cochrane-handbook.org

Knight, R. A., & Prentky, R. A. (1990). Classifying sexual offenders: The development and corroboration of taxonomic models. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.), *Handbook of sexual assault: Issues, theories, and treatment of the offender* (pp. 23–52). New York: Plenum.

Lonsway, K. A., & Fitzgerald, L. F. (1994). Rape myths: In review. *Psychology of Women Quarterly*, *18*(2), 133–164. http://dx.doi.org/10.1111/j.1471-6402.1994.tb00448.x.

Lonsway, K. A., & Fitzgerald, L. F. (1995). Attitudinal antecedents of rape myth acceptance: A theoretical and empirical reexamination. *Journal of Personality and Social Psychology*, *68*, 704–711.

Marshall, W. L., & Hambley, L. S. (1996). Intimacy and loneliness, and their relationship to rape myth acceptance and hostility toward women among rapists. *Journal of Interpersonal Violence*, *11*(4), 586–592.

Muehlenhard, C. L., & Linton, M. A. (1987). Date rape and sexual aggression in dating situations: Incidence and risk factors. *Journal of Counseling Psychology*, *34*, 186–196.

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

Overholser, J. C., & Beck, S. (1986). Multimethod assessment of rapists, child molesters, and three control groups on behavioral and psychological measures. *Journal of Consulting and Clinical Psychology*, *54*(5), 682.

Payne, D. L., Lonsway, K. A., & Fitzgerald, L. F. (1999). Rape myth acceptance: Exploration of its structure and its measurement using the Illinois Rape Myth Acceptance Scale. *Journal of Research in Personality*, *33*(1), 27–68http://dx.doi.org/10.1006/jrpe.1998.2238

Pithers, W. D. (1994). Process evaluation of a group therapy component designed to enhance sex offenders' empathy for sexual abuse survivors. *Behaviour Research and Therapy*, *32*, 565–570.

Pithers, W. D., Martin, G., & Cimming, G. F. (1989). The Vermont treatment program for sexual aggressors. In D. R. Laws (Ed.), *Relapse prevention with sex offenders* (pp. 292–310). New York: Guilford Press.

Reid, S., Wilson, N. J., & Boer, D. P. (2011). Risk, needs, and responsivity principles in action: Tailoring rapist's treatment to rapist typologies. In D. P. Boer, R. Eher, & M. H. Miner (Eds.), *International perspectives on the assessment and treatment of sexual offenders* (pp. 287–297). Chichester, UK: Wiley-Blackwell.

Sackett, D. L., Rosenberg, W., Gray, J. A., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't. *British Medical Journal*, *312*(7023), 71–72.

Stefanska, E. B., Carter, A. J., Higgs, T., Bishopp, D., & Beech, A. R. (2015). Offense pathways of non-serial sexual killers. *Journal of Criminal Justice*, *43*(2), 99–107.

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

Thornton, D., Mann, R., Webster, S., Blud, L., Travers, R., Friendship, C., & Erikson, M. (2003). Distinguishing and combining risks for sexual and violent recidivism. *Annals of the New York Academy of Sciences*, *989*(1), 225–235.

Offending Behaviour Programmes Unit (1995). *Rape myths scale*. Unpublished London, England: National Offender Management Service: Ministry of Justice.

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

Webster, S. D., Akhtar, S., Bowers, L. E., Mann, R. E., Rallings, M., & Marshall, W. L. (2004). The impact of the prison service sex offender treatment program on minority ethnic offenders: A preliminary study. *Psychology, Crime & Law*, *10*(2), 113–124.