A systematic review of the association between rape myth acceptance and male-on-female sexual violence

OPEN ACCESS

A Systematic Review of the Association Between Rape Myth Acceptance and

Male-on-Female Sexual Violence

Emma J. Yapp, M.A.[1*], Ethel Quayle, DClinPsy[2]

1. School of Philosophy, Psychology, & Language Science, University of

Edinburgh, UK

2. School of Health in Social Science, University of Edinburgh, Edinburgh, UK

* To whom correspondence should be addressed

Emma Yapp

Research Assistant

Section of Women's Mental Health

Instute of Psychiatry, Psychology and Neuroscience

King's College London

Room 3.06 David Goldberg Centre

De Crespigny Park

London SE5 8AF

Tel: +44(0)207 848 5093

Email: emma.yapp@kcl.ac.uk

Abstract

Rape myth acceptance is considered an established risk factor for male-on-female sexual violence, and is therefore the target of a number of primary prevention programmes. However, there is not a clear evidence base substantiating the role of rape myth acceptance in sexual violence, nor any reviews of recent literature. This review systematically searched relevant Psychology and Social Science databases in Autumn 2016, in order to collate cross-sectional and longitudinal research on the association between rape myth acceptance and self-reported sexual violence. The analysis established associations between these variables in all but one study (Warren, Swan, & Allen, 2015), and two longitudinal studies demonstrated that rape myth acceptance differentiates non-perpetrators from those who go on to exhibit sexual violence behaviours. These findings provide support for the targeting of rape myth acceptance in primary prevention strategies. However, a number of failings within this literature were also identified: instruments used to analyse rape myth acceptance were widely varied; the comprehensiveness of study reporting was universally flawed; measures were rarely taken to ensure participants were heterosexual men; and there remains a dearth of longitudinal evidence, as well as a lack of research outside of the United States. Future directions and other limitations are discussed.

Key words: sexual violence; rape myth acceptance; rape.

Abbreviations: Rape myth acceptance (RMA), Sexual Experiences Survey (SES).

1.1 **Introduction**

Male-on-female sexual violence is a public health problem that affects women worldwide (Abrahams et al., 2014), and is associated with a variety of negative mental health consequences, such as depression, anxiety, and alcohol use (World Health Organisation, 2013). In the United Kingdom, sexual violence victimisation rates have not changed significantly since 2005 (Office for National Statistics, 2018): around 20% of UK women experience sexual violence (Office for National Statistics, 2018), yet it is estimated that only 15% of these incidents are reported to the police (Ministry of Justice, 2013). As a result of this, most sexual violence perpetrators go undetected (Campbell, Patterson, & Bybee, 2012; Larcombe, 2011; Lisak & Miller, 2002), and often reoffend (e.g. Zinzow & Thompson, 2015). It is therefore essential that we obtain a better understanding of what drives this behaviour, in order that primary prevention strategies may be strengthened. Researchers (e.g. Walker & Bright, 2009; Bowes & McMurran, 2013) have emphasised the need to examine the cognitive distortions associated with sexual violence, as these distortions may have predictive validity for recidivism (Helmus et al., 2013), and targeting these variables is thought to enhance the effectiveness of treatment (see: Olver, Nicholaichuk, Kingston, & Wong, 2014).

1.1.1 Rape Myth Acceptance

The literature on male-on-female sexual violence typically focuses on cognitive distortions that manifest as "rape supportive attitudes"; attitudes that facilitate the justification of sexual violence, often serving either to blame the victim, to

exonerate the perpetrator, to minimise claims of rape, or to allude that only certain types of women are raped (Hust, Rodgers, Ebreo, & Stefani, 2017; Bohner, Eyssel, Pina, Siebler, & Viki, 2009; Burt, 1980; Lonsway & Fitzgerald, 1994, 1995). Within this literature is a concept referred to as "rape myths". Often inseparable from rape supportive attitudes, rape myths are defined as beliefs about rape that are generally false and widely held (Lonsway & Fitzgerald, 1994). For example, believing that a woman is totally or partly to blame for her sexual assault if she was out late at night, wearing a short skirt, and drunk, are all examples of rape myths, and a recent UK report indicated that 38% of all men, and 34% women endorse such beliefs (Fawcett Society, 2017).

It is generally thought that rape myth acceptance (RMA), the endorsement of beliefs such as these, is reflective of a cognitive distortion that can result in sexual violence behaviours, and there have been several attempts to capture this distortion to somewhat contested degrees of success. The first instrument to measure RMA was designed by Martha Burt in 1980 (Burt, 1980): the Rape Myth Acceptance Scale (RMAS). Burt's (1980) work built on the legacy of Hubert Feild, who developed the Attitudes Toward Rape Questionnaire in 1978, which similarly measured attitudes supportive of rape, but failed to discriminate rapists from police officers on a number of items (Feild, 1978). Burt's (1980) scale has motivated many related instruments since its conception, many of which are inspired by perceived short-comings of the tool (RAPE scale, Bumby, 1996) or frustrations at its lack of predictive validity (Rape Myth Scale, Lonsway & Fitzgerald, 1994; Illinois Rape Myth Acceptance Scale, Payne, Lonsway, & Fitzgerald, 1999; McMahon & Farmer, 2011).

Despite this heterogeneity in measurement tools, RMA has shown strong predictive validity in several studies (McDermott, Kilmartin, McKelvey, & Kridel, 2015; Süssenbach et al., 2013; Vega & Malamuth 2007; Abrams, Viki, Masser, & Bohner, 2003; Bohner et al., 2005). It is still considered a key risk factor for sexual violence perpetration in the prevailing model of this behaviour, the Confluence Model (Malamuth, Linz, Heavey, Barnes, & Acker, 1995); RMA is part of the model's wider construct of "hostile masculinity", which is hypothesised to interact with several other variables in order to culminate in sexual violence.

RMA has been a popular target for recent primary prevention programmes (e.g. Bolton-Holz, Fischer, & Daood, 2016; McMahon, Postmus, Warrener, & Koenick, 2014; Peterson et al., 2016), despite the fact that evidence published on this association since 2008 has not yet been reviewed (cf Tharp et al., 2013), and there are several other concerns over its validity as a predictor of sexual violence. RMA does not always significantly predict the onset sexual violence (e.g. Loh, Gidycz, Lobo, & Luthra, 2005), and could only arise post-perpetration, as a means of justifying past actions and alleviating guilt (Maruna & Mann, 2006) in which case primary prevention strategies may be ultimately misguided in targeting RMA.

1.1.2 Sexual violence

The bi-directionality concern over RMA arises out of two difficulties within the sexual violence literature: (a) it is difficult to measure an individual's future

sexual violence behaviours, and (b) most psychological research uses cross-sectional data. Some studies attempt to measure future sexual violence (e.g. Bohner et al., 2009); for example, Malamuth's (1981) Likelihood to Rape scale asks participants whether they would rape someone "if guaranteed they would not be caught or punished" (Malamuth, 1981, p.140). This is ultimately a proxy measure for sexual violence behaviours, as there is no guarantee that the behaviour will ensue. Self-report measures are therefore regarded as the best available instruments, as they provide actual measures of sexual violence, and thus provide higher external validity for the risk factors they identify. The Sexual Experiences Survey (SES, Koss & Oros, 1982) is the most popular of these instruments (McDermott et al., 2015; Porter & Critelli, 1992), yet when administered in cross-sectional studies, the SES necessarily measures retrospective perpetration of sexual violence, as items refer to past behaviours (e.g. Swartout, 2013), and therefore cannot rule out reverse causality. Given the popularity of this instrument, a significant literature exists on the association between RMA and sexual violence behaviours measured by the SES. Longitudinal evidence in particular may help establish whether RMA is predictive of the onset of sexual violence perpetration.

The present study will provide a systematic review of both cross-sectional and longitudinal literature published since 2008 on the association between RMA and male-on-female sexual violence in the general population, as measured by the SES, since other self-report measures are infrequently used (e.g. the Coercive Sexuality Scale, Rappaport & Buckhart, 1984; see: McDermott et al., 2015), and measures of rape proclivity do not provide assessments of actual sexual violence.

This work intends to clarify the current state of (a) the literature and (b) the relationship between RMA and sexual violence behaviour within the general population, in order to inform current and future primary prevention strategies.

1.2 Methods

1.2.1 Search strategy

After consulting with a professional librarian in order to establish best searching practice, a variety of potential search terms and databases were explored. Relevant databases containing work on both Psychology and the Social Sciences were identified on the basis of the librarian's advice, as well as by consulting similar reviews (e.g. Tharp et al., 2013). Search terms were honed via assessment of subject headings, in order to identify synonyms for the key terms (e.g. "sexual violence", "rape myth acceptance") within the literature.

The final search was conducted in Autumn 2016 on Embase, OVID, MEDLINE, PsychINFO, Applied Social Sciences Index and Abstracts (ASSIA), Criminal Justice Database, ERIC, International Bibliography of the Social Sciences, Social Sciences Database, Social Services Abstracts, Sociological Abstracts, Sociology Database, and ProQuest Dissertations & Theses Global, using the following search: "sexual" AND ("aggression" OR "coercion" OR "violence" OR "assault" OR "rape") AND "perpetrat*" AND ("rape myth" OR "rape-supportive attitudes" OR "rape supportive attitudes" OR "rape-supportive attitude" OR "rape supportive

attitude"). The inclusion of the database, "ProQuest Dissertations & Theses Global", was justified in an attempt to reduce publication bias.

1.2.2 Inclusion and Exclusion

Studies exclusively evaluating risk factors for sexual violence victimisation or intimate partner sexual violence were excluded, as intimate partner violence represents a particular subset of sexual violence and may not be representative of all sexual violence committed by the general population. Studies of prison rape and child sexual abuse were also excluded, as prison rape is exclusively committed against victims of the same gender, and child sexual abuse is characterised by a differing etiology to that of male-on-female sexual assault (see: Mann, Hanson, & Thornton, 2010; Casey & Lindhorst, 2009). Although some intervention and instrument development studies (e.g. Stephens & George, 2009) include measures of both RMA and sexual violence at baseline, the goals of these works are ultimately tangential to the aims of this review. As a result of this, studies of this nature were also excluded.

Table 1

Inclusion and exclusion criteria for the literature search

Inclusion Criteria	Exclusion Criteria
• Journal articles and theses	• Reviews and meta-analyses
• Available in English	• Intervention studies and instrument development works
• Published 2008-2016	• Studies that:

• Use of one or more measures that make explicit reference to RMA	a. Exclusively looked at victimisation
• Study sampled non-incarcerated men over the age of 14	b. Assessed prison rape
• Male-on-female sexual violence perpetration as measured by a version of the SES	c. Assessed intimate partner violence
• Analysis of the association between RMA and the Sexual Experiences Survey	d. Assessed childhood sexual abuse
• Cross-sectional or longitudinal study design	

Inclusion criteria were initially shaped by limits to the scope of this research: the search was restricted to journal articles and theses, on account of the limited time frame available, and the inclusion of non-English language papers was beyond the scope of this researcher, though this restriction should not incur bias (see: Moher, Pham, Lawson, & Klasson, 2003). Both cross-sectional and longitudinal works were included, for the bulk of the research within this literature is cross-sectional (see: Tharp et al., 2013), but longitudinal papers provide preliminary assessments of causality within this relationship. Papers were selected for review from 2008, as the aim of this review was to appraise the recent literature on RMA and sexual violence, and another systematic review of multiple risk factors has comprehensively assessed existing research on this association published up until 2008 (e.g. Tharp et al., 2013). Furthermore, studies since this time have been informed of certain potential pitfalls within

rape myth instruments: for example, the internal consistency of these measures

is rarely reported (Buhi, 2005), and the language used in traditional measures of

RMA may now be outdated for use in university age populations (McMahon &

Farmer, 2011).

1.2.3 Measure of RMA

In choosing which measures of RMA to include, the aim was to select

instruments that broadly assess an identical construct, so as to ensure

consistency in the variable being evaluated. Therefore, it was elected to

exclusively appraise studies that made use of established instruments that

explicitly claim to measure RMA (Burt, 1980; Bumby, 1996; Lonsway &

Fitzgerald, 1995; Payne et al., 1999; Gerger et al., 2007; McMahon & Farmer,

2011), as opposed to novel instruments (Kennair & Bendixen, 2011), or those

appraising wider rape-supportive attitudes (Feild, 1978). While this definition

excludes more seminal measures of RMA (e.g. Feild, 1978), such measures are

multi-dimensional, and cannot be conceptualised as providing a close

assessment of RMA when scored as a whole (Ward, 1988). Where researchers

exercised a subset of items from included measures, the use of this selection

must have been previously validated, in order to ensure that it reliably measures

the same construct.

1.2.4 Measure of Sexual Violence

This review is restricted to studies that operationalised a version of the Sexual

Experiences Survey (SES, Koss & Oros, 1982; Koss et al., 1987; Abbey et al., 2006;

Koss et al., 2007). This instrument has good internal consistency (Koss & Gidycz,

1985), and reliability (Kolivas & Gross, 2007) is strongly correlated with results

obtained in face-to-face interviews (Lisak & Roth, 1988; Koss et al., 1987; Koss &

Gidycz, 1985; Koss & Oros, 1982), and uses behaviourally-specific questions in

order to combat underreporting (see: Fisher, 2009; Crowell & Burgess, 1996;

Fisher & Cullen, 2000). It was elected to include studies using any version of the

SES, for while the original instrument has been criticised for adopting some

ambiguous items (Kolivas & Gross, 2007), this resulted in revisions that were

made only just before the beginning of the included timeframe (Koss et al.,

2007), and there was therefore concern that restricting inclusion to this version

might not retrieve sufficient works for discussion. In addition, by including all

versions of the same instrument, this should enable consistency in the

interpretation of results, and avoid the pitfalls of appraising a single instrument

that is flawed or faulty.

1.2.5   Final inclusion criteria

The selected instruments measuring both RMA and sexual violence subsequently

dictated further inclusion criteria. The SES inquires after sexual violence

committed since age 14, and thus the surveyed sample was necessarily required

to be men over this age threshold. The included measures of RMA are also

commonly aggregated with other variables to form a composite variable of

hostile masculinity, on account of the influence of the Confluence Model of sexual

aggression (Malamuth et al., 1995). Therefore, where studies had data on the relevant measures, but had aggregated this information into a composite variable, the authors were contacted in order to request data on RMA alone, and given one month within which to respond.

1.2.6 Quality assessment

As this review appraised both longitudinal and cross-sectional studies, it was important to choose a quality tool that would be comprehensive enough and applicable to both. Formal quality tools for the appraisal of cross-sectional studies are limited and minimal (Zeng et al., 2015), however, a new tool – referred to as "AXIS" – to assess cross-sectional studies was recently published in the BMJ (Downes, Brennan, Williams, & Dean, 2016), which proffers assessment that is more comprehensive than alternative options (see: Zeng et al., 2015), as it uses significantly more items, and addresses issues of both study reporting, and analysis. All articles that met inclusion criteria were appraised by using the AXIS tool, and where studies were longitudinal, these were additionally assessed for information reported on attrition by appraising loss to follow-up, as has been done in other clinical reviews (e.g. Gami et al., 2007), by extracting criteria from a flow-chart developed by Tooth and colleagues (2005, see Table 7 for item details). Two independent researchers conducted this process; this author first critically appraised all included studies, and then these were ordered alphabetically and numbered, such that a random number generator (Haahr, 1998) was used to select half of the papers to be assessed by an alternative

researcher. Next, assessments of quality were compared for discrepancies, and these were resolved by means of discussion.

1.3 Results

1.3.1 Search

A description of the search is represented in Figure 1. The initial search returned 1,112 online papers. After de-duplicating, 1,012 papers remained; these were subsequently screened on titles alone, in order to identify irrelevant works. The remaining 749 papers were screened on information in the abstract. The bulk of the papers removed at this stage either focused on the wrong population, such as women or homosexual individuals (e.g. Bryce, 2012), or exclusively examined the role of rape myths in victims of sexual violence, as opposed to perpetrators (e.g. Alberty, 2011).

This culminated in 67 papers to be assessed for inclusion by their full-text (see Figure 1 for full details). During the appraisal of full-texts, 17 authors were contacted either for further statistical information or to request access to the full manuscript. Two authors responded with statistical information (C. Anderson, personal communication, 30 January; P. Warren, personal communication, 4 February), and one (H. Zinzow, personal communication, 23 January, 2017)

identified 2 further papers for assessment[1]. This resulted in a full-text appraisal

of 69 articles in total.

---

[1] Though it might seem cause for concern that some potentially relevant articles were not captured by the search terms exercised herein, it should be noted that neither of these works met all inclusion criteria (see Table A1).
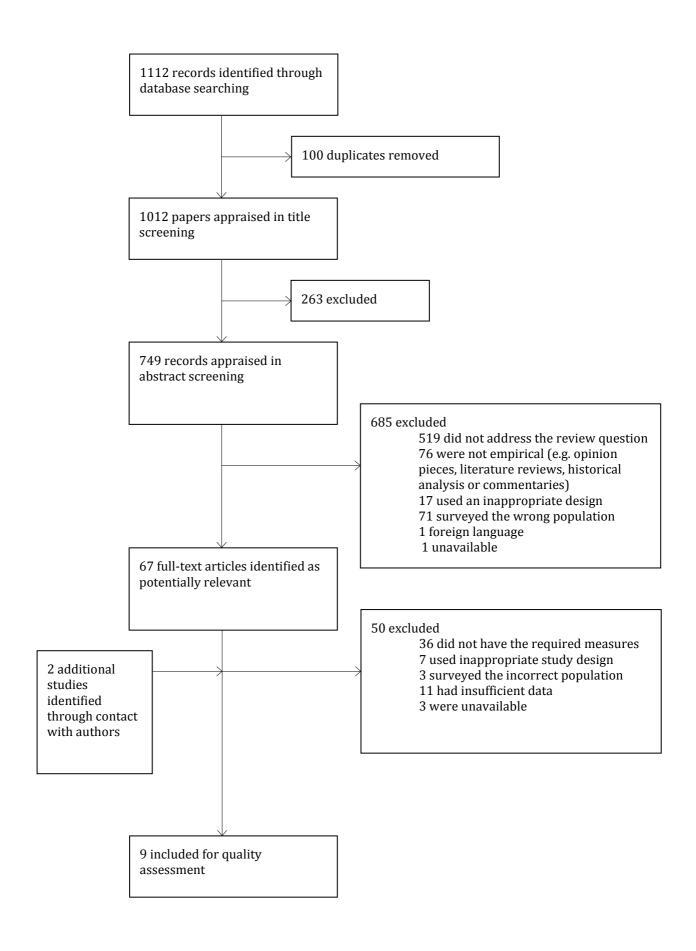
14

Figure 1. Flow of study search.

Out of these papers, 36 did not have appropriate measures of RMA or the SES, 3 surveyed the incorrect population, 7 used an inappropriate study design, 3 were unobtainable and failed to respond to email requests, and 11 reported insufficient statistical information for use within this review following an email requesting this information (see Table A1 for full details regarding full-text sift). Further, it should be noted that a number of the papers rejected for reporting insufficient information on the included variables (Thompson, Swartout, & Koss, 2013; Zinzow & Thompson, 2014; Zinzow & Thompson, 2015) analysed the same sample as an included paper (Thompson, Koss, Kingree, Goree, & Rice, 2011), which therefore raises concern over data mining.

Nine manuscripts met all inclusion criteria: seven published papers and two theses. Of these, it should be noted that two studies utilised only a selection of items from instruments measuring RMA (Anderson & Anderson, 2008; Abbey, Wegner, Pierce, & Jacques-Tiura, 2012), however, these selections were previously validated, as in one case the measure was piloted (Abbey et al., 2012), and in the other, the same selection of items has been justified and validated elsewhere (e.g. Romero-Sánchez & Megías, 2010; Truman, Tokar, & Fischer, 1996).

1.3.2 Study characteristics

Table 2 summarises the characteristics of the nine studies, as well as details regarding the RMA and SES instruments, and analysis of the association between

the two – either correlational, or when RMA was used to predict the sexual

violence outcomes. All studies were conducted in the United States.

Table 2

Key Information Extracted from Included Studies

Reference	Country of study, and Publication Status	Characteristics of final sample	Study design, directionality, and aims	Measurement of RMA	Version of SES and scoring method	Analysis (RMA to predict SES) and findings
Abbey, Wegner, Pierce, & Jacques-Tiura, 2012	United States, Published	423 Community men. Age at baseline: 18-35 (*M*=23, *SD*=4.95). Recruited using a commercial landline telephone list.	Longitudinal (prospective). To identify patterns of sexual aggression over time.	Combination of Bumby's (1996) RAPE scale, and Payne, Lonsway & Fitzgerald's (1999) IRMAS, pilot tested. α=0.85.	16 item modified SES (Abbey et al., 2006). α=0.84 at baseline, α=0.92 at follow-up. At baseline since age 14, at follow-up since baseline. Categorised participants into four mutually exclusive groups using data from T1 and T2: persisters, desisters, initiators, and non-perpetrators. For total frequency: α=0.84 at baseline and α=0.92 at follow-up.	Discriminant function analysis to discriminate between groups, ANOVA's to aid interpretation of group differences. Groups differed significantly on RMA (F(3,417)=13.27, p=0.0001). Specifically: persisters (M=2.97, SD=0.13) significantly (p<0.05) different from desisters (M=2.32, SD=0.11) and non-perpetrators (M=2.09, SD=0.08), and initiators (M=2.73, SD=0.19) significant different from non-perpetrators (M=2.09, SD=0.08).
Anderson & Anderson, 2008 (study 1 only)	United States, Published	194 Undergraduate men. Aged 18-36 (*M*=19.5). Midwestern University.	Cross-sectional (retrospective). To model sexual violence using the Confluence Model and the General	First eleven items from Burt's (1980) RMAS. α=0.742.	10 item SES (Koss et al., 1987). Frequency since age 14. Scoring method not defined. Internal consistency not reported.	Correlation between RMA and log(SES) obtained via contact with the author. *r*=0.336, *p*<0.0001.

			Aggression Model.			
Mouilso & Calhoun, 2008	United States, Published	308 Undergraduate men. Age (*M*=19.72, *SD*=1.55). Southeastern University.	Cross-sectional (retrospective). To evaluate the roles of RMA and psychopathy in sexual violence.	64 item IRMA (Payne et al., 1999). α=0.91.	10 item SES (Koss et al., 1987). Frequency since age 14. Classified into non-perpetrators, perpetrators of rape, and perpetrators of sexual assault. Internal consistency not reported.	Logistic regression found RMA significantly differentiates between all perpetrators and non-perpetrators ($\chi2(1, N = 286) = 7.83$, p = .005), explaining 4.5% variance (Nagelke's R^2), but regression no longer significant after accounting for variance explained by psychopathy (OR = 1.01, 95% CI = [.00, 1.02], p = .050).
Russell & King, 2016	United States, Published	489 Community men. Aged 18-66 (*M*=33.98). Recruited through MTurk.	Cross-sectional (retrospective). To evaluate the predictive power of RMA, hostility towards women, everyday sadism, and parental attachment in sexual violence. Makes reference to Confluence Model.	Rape myth scale (Lonsway & Fitzgerald, 1995). α=0.95.	Revised SES short form (Koss et al., 2007). Frequency since age 14. Categorised into perpetrators of sexual aggression and perpetrators of sexual coercion (not mutually exclusive). α=0.94 for both indicies.	RMA significantly contributed to a simultaneous multiple regression model using RMA, hostility towards women, sadism, and attachment style to predict both sexual aggression (*ß*=0.38, *r*=0.5) and sexual coercion (*ß*=0.33, *r*=0.45).
Russell, 2016	United States, unpublished	512 Community men and 100 undergraduate men. Community men aged 18-73 (*M*=34.8). MTurk. Undergraduate men	Cross-sectional (retrospective). To evaluate the relationship between psychopathy, sadism,	Rape myth scale (Lonsway & Fitzgerald, 1995), α=0.95.	Revised SES short form (Koss et al., 2007). Frequency since age 14. Categorised into perpetrators of sexual aggression and perpetrators of sexual	Aggressors (*F*(610)=66.09, *p*<0.001, *d*=0.77) and coercers (*F*(610)=69.44, *p*<0.001, *d*=0.80) differ significantly from non-perpetrators on RMA.

		aged 18-38 (*M*=20.3). Midwestern University.\n\nTotal *M*=32.88.	attachment, and the Confluence Model.		coercion (not mutually exclusive). α=0.96.	
Saenz, 2009	United States, unpublished	430 Undergraduate men. Age (*M*=22.10). Urban university.	Cross-sectional (retrospective). To integrate narcissism into the Confluence Model of sexual aggression.	IRMAS (Payne et al., 1999). α=0.89.	Modified SES (Abbey et al., 2006). α=0.87. Frequency since age 14. Scoring unclear, though the original Abbey paper summed across questions to get a total frequency, so can assume that went on here. α=0.87.	Data from the SES was skewed (4.82) and leptokurtotic (27.31) in nature, and was subsequently transformed using a base log 10 transformation. Significant correlation between RMA and log(SES) (*r*=0.17, *p*<0.001).
Swartout, 2012	United States, Published	341 University men. Age (*M*=18.9). Medium-sized public University.	Cross-sectional (retrospective). To integrate the role of peer networks into the Confluence Model of sexual aggression.	Rape Myth Acceptance Scale (Burt, 1980), α=0.86.	Combination of long and short form revised SES (Koss et al., 2007). Frequency since age 14. Constructed four indices of sexual violence: unwanted sexual contact (α=0.85), verbal coercion (α=0.87), attempted rape (α=0.97), and rape (α=0.98).	RMA significantly correlated with unwanted contact (*r*=0.24, p<0.01) and verbal coercion (*r*=0.16, p<0.01), but not rape (*r*=0.10, p>0.05) or attempted rape (*r*=0.11, p>0.05).
Thompson, Koss, Kingree, Goree, & Rice, 2011	United States, Published	652 Undergraduate men. Age at wave 1 (*M*=18.67), age at wave 2 (*M*=19.59). Large Southeastern university.	Longitudinal (prospective). To use the Theory of Planned Behaviour to examine prospective associations of	Rape Myth Scale (Lonsway & Fitzgerald, 1995), α=0.9.	Revised Sexual Experiences Survey (Koss et al., 2007). Frequency since age 14 at baseline, since T1 at follow-up. Scored from 0-15, where the order of severity goes from unwanted sexual contact, attempted	In the final path analysis model, RMA predicted perpetration status (standardised *ß*=0.23, *z*=2.36, *p*<0.05) in a model that also included perceived norms and perceived control

					coercion, completed coercion, attempted rape, completed rape, and participant frequency is recorded within each category (from one to three or more).	as predictors of wave 2 sexual violence.
Warren, Swan, & Allen, 2015	United States, Published	217 Undergraduate men. Large Southeastern University. Aged 18-46 (*M*=21.07, *SD*=3.3).	Cross-sectional (retrospective). To examine the relationship between comprehension of sexual consent and sexual violence in the context of a variety of other cognitive and social risk factors.	IRMAS (Payne et al., 1999), α=0.76.	The Sexual Experiences Survey short form (Koss et al., 2007). Frequency in past four months. Scored as a dichotomous perpetration variable.	In the final path analysis, the path from RMA to sexual violence was not significant (β =0.09, *p*=0.401) in a model where other predictors included conformity to masculine norms, comprehension of sexual consent, and peer support of abuse. P-value obtained via personal communication with the author.

1.3.3 Sample

Sample size was generally large across studies: the average sample size was 407, and the smallest was 194 (Anderson & Anderson, 2008). Undergraduate men comprised the samples for most studies, though three papers used samples of community men (Russell, 2016; Russell & King, 2016; Abbey et al., 2011), which is a greater proportion than has been established in similar systematic reviews (14%, Tharp et al., 2013). Participants were a minimum of 18 years old across studies, and age ranges were broad where specified (min=17, Abbey et al., 2012; max=55, Russell, 2016), though in four cases the age range was not reported (Mouilso & Calhoun, 2013; Saenz, 2009; Swartout, 2013; Thompson et al., 2011).

1.3.4 Aims

Many study aims were guided by existing theory, either seeking to expand upon the Confluence Model (Saenz, 2009; Anderson & Anderson, 2008; Swartout, 2013; Russell & King, 2016), theories about Delinquency (Abbey et al., 2012), and the Theory of Planned Behaviour (Thompson et al., 2011). Additionally, while the cross-sectional studies examined the association between RMA and sexual violence retrospectively (Anderson & Anderson, 2008; Mouilso & Calhoun, 2013; Russell, 2016; Russell & King, 2016; Saenz, 2009; Swartout, 2013; Warren et al., 2015), the two longitudinal studies examined the prospective role of RMA in both differing degrees of sexual violence (Thompson et al., 2011) and in different patterns of sexual violence perpetration (Abbey et al., 2012),

1.3.5 Measures of RMA and SES

A variety of instruments are still being used to measure RMA (see Table 3). The most common were Lonsway & Fitzgerald's (1995) Rape Myth Scale, and the Illinois RMA Scale (Payne et al., 1999), which was developed for use with University populations, and was therefore adopted appropriately in the these works (Mouilso & Calhoun, 2013; Saenz, 2009; Warren et al., 2015). Other papers administered Burt's (1980) RMA scale, either using all 19 items, in order to be consistent with previous investigations of the Confluence Model (Swartout, 2013), or by extracting the first 11 (Anderson & Anderson, 2008), as the remaining 8 items chiefly pertain to discrimination against particular types of women (e.g. "A person comes to you and claims they were raped. How likely would you be to believe that person if it was: an Indian woman?", Burt, 1980, p. 223). Finally, one paper combined items from Bumby's (1996) RAPE scale, and the Illinois RMA Scale (Payne et al., 1999), to create a 9-item measure of "Stereotypic attitudes about women that justify forced sex".

Table 3

Frequency of RMA Measures

Measure of RMA	Frequency
Burt (1980)	2
Lonsway & Fitzgerald (1995)	3
Payne et al., (1999)	3
Combination of Bumby (1996) and Payne et al., (1999)	1

Versions of the Sexual Experiences Survey also varied (see Table 4), though the most recent version of the Sexual Experiences Survey (Koss et al., 2007) has been adopted more often and more recently (Thompson et al., 2011; Swartout, 2013; Warren et al., 2015; Russell, 2016; Russell & King, 2016) than other versions (e.g. Abbey et al., 2006, Koss et al., 1987). Most studies used the SES to measure sexual violence perpetration since age 14 (Abbey et al., 2012; Anderson & Anderson, 2008; Mouilso & Calhoun, 2013; Russell, 2016; Russell & King, 2016; Saenz, 2009; Swartout, 2013; Thompson et al., 2011), and then again since baseline in the longitudinal cases (Abbey et al., 2012; Thompson et al., 2011). However, one paper solely examined sexual violence perpetrated within the past four months (Warren et al., 2015).

Table 4

Sexual Experiences Survey: version and scoring method

Measure of SES	Measure reference	Internal consistency	Period assessed	Scoring method	Year	Authors
Original SES	Koss et al., (1987)	-	Since age 14	Not reported.	2008	Anderson & Anderson
	Koss et al., (1987)	-	Since age 14	Classified into mutually exclusive categories: perpetrators of rape, perpetrators of sexual assault, and non-perpetrators.	2013	Mouilso & Calhoun
Modified SES	Abbey et al., (2006)	α=0.87	Since age 14	Not reported, though presumed that scores were summed to generate a total frequency of assault score, as in the original	2009	Saenz

			paper by Abbey et al., (2006).		
	Abbey et al., (2006)	α=0.84 at baseline, α=0.92 at follow-up	Since age 14 at baseline, then at follow-up asked to report on period since baseline.	Classified into three groups of longitudinal sexual violence patterns: desisters, persisters, initiators, and non-perpetrators.	2012	Abbey, Wegner, Pierce, & Jacques-Tiura
Revised SES	Koss et al., (2007) (unclear whether short or long form)	-	Since age 14 at baseline, then at follow-up asked to report on period since baseline.	Scored from 0-15, where the order of severity goes from unwanted sexual contact, attempted coercion, completed coercion, attempted rape, completed rape, and participant frequency is recorded within each category (from one to three or more).	2011	Thompson, Koss, Kingree, Goree, & Rice
	Koss et al., (2007) (combination of long and short form)	Unwanted sexual contact (α=0.85), verbal coercion (α=0.87), attempted rape (α=0.97), and rape (α=0.98).	Since age 14	Generated four indices of sexual violence: unwanted sexual contact, verbal coercion, attempted rape, and rape.	2012	Swartout
	Koss et al., (2007) (short form)	-	Past four months	Dichotomous perpetration variable: perpetrators and non-perpetrators.	2015	Warren, Swan, & Allen
	Koss et al., (2007) (short form)	α=0.94 for both indicies.	Since age 14	Categorised into perpetrators of sexual aggression, and perpetrators of sexual coercion (cf DeGue, DeLillo, & Scalora, 2010).	2016	Russell & King

Koss et al., (2007) (short form)	α=0.95 for all sexual assault.	Since age 14	Categorised into perpetrators of sexual aggression, and perpetrators of sexual coercion (cf DeGue et al., 2010).	2016	Russell

Methods to score the SES varied widely, which is partly on account of the flexibility in scoring the instrument; to use the most recent version of the SES (Koss et al., 2007) researchers are required to contact Dr Koss for permission, at which time she provides instructions for methods of scoring both the long and short forms of the instrument – some methods reflect the frequency or severity of sexual acts, while others distinguish between groups of perpetrators. Of the five studies that adopted the most recent version of the SES (Koss et al., 2007), four categorised participants into groups on the basis of their answers, whether comparing non-perpetrators with all perpetrators (Warren et al., 2015), or exercising more specific categories. For example, Mouilso & Calhoun (2008) contrasted perpetrators of rape with perpetrators of other sexual violence (Mouilso & Calhoun, 2013), while others examined RMA across sexual coercers and sexual aggressors (Russell & King, 2016; Russell, 2016), or across different categories of sexual violence perpetration in order of increasing severity (Swartout, 2013). The final study using this version of the SES (Koss et al., 2007) used a scoring method that took both severity and frequency into account: scores were constructed by generating categories of sexual violence in order of severity, and by counting the frequency of each act, in order to generate a final score (Thompson et al., 2011).

Two studies did not explicitly report how they scored the SES (Anderson &

Anderson, 2008; Saenz, 2009), though one made reference to the "Number of

sexual assault" (Saenz, 2009, p.36), and thus this paper presumably coded the

sexual violence variable as a continuous variable based on individual items. The

remaining longitudinal study used information on perpetration at both baseline

and follow-up to categorise participants according to their pattern of sexual

violence: into persisters, desisters, initiators, and non-perpetrators (Abbey et al.,

2012).

1.3.6 Study findings

Two studies analysed group differences in RMA across sexual violence

perpetration. Russell (2016) used ANOVAs to establish that both aggressors

($F(610)=66.09$, $p<0.001$, $d=0.77$) and coercers ($F(610)=69.44$, $p<0.001$, $d=0.80$)

differed significantly from non-perpetrators on their acceptance of rape-myths.

Similarly, the longitudinal study by Abbey et al., (2012), used Discriminant

function analysis to indicate significant group differences across groups on RMA;

ANOVAs revealed that persisters (M=2.97, SE=0.13) were significantly (p<0.05)

different from desisters (M=2.32, SE=0.11) and non-perpetrators (M=2.09,

SE=0.08); and initiators (M=2.73, SE=0.19) were significantly different from non-

perpetrators (M=2.09, SE=0.08).

Three works analysed the association using correlations. The strongest

correlation was obtained by Anderson & Anderson (2008) (r=0.336), though this

was also the paper with the smallest sample (n=194). One study found a

moderate correlation (r=0.17, Saenz, 2009), and the other assessed the relationship within perpetration groupings (Swartout, 2013). An interesting pattern emerged within this latter paper, as RMA had strongest correlations with the perpetration of acts that were less severe – being largest for unwanted sexual contact (r=0.24), slightly weaker for verbal coercion (r=0.16), weaker still for attempted rape (r=0.11), and the smallest association was found between RMA and the perpetration of rape itself (r=0.1).

The remaining papers analysed the association between RMA and sexual violence with more complex models. RMA was no longer a significant predictor of sexual violence in a model accounting for the variance explained by psychopathy (OR = 1.01, 95% CI = [.00, 1.02], p = .050, Mouilso & Calhoun, 2013), nor in one that controlled for: conformity to masculine norms, comprehension of sexual consent, and peer support of abuse ($ß$ =0.09, p=0.401, Warren et al., 2015). However, RMA significantly predicted sexual violence in a single regression model (Mouilso & Calhoun, 2013), explaining 4.5% of the variance in sexual violence behaviours. Further, RMA significantly contributed to two multiple regression-style models; in a prospective path analysis that controlled for perceived norms, perceived control, and sexual violence perpetration at baseline ($ß$ =0.23, z=2.36, p<0.05, Thompson et al., 2011), and RMA significantly predicted both sexual aggression ($ß$=0.38, r=0.5) and sexual coercion ($ß$=0.33, r=045) in a simultaneous multiple regression model (Russell & King, 2016) controlling for hostility towards women, sadism, and attachment style. All reported coefficients are standardised, and thus the predictive power of RMA is reasonably strong within these models.

1.3.7 Study Strengths, Limitations, and Recommendations

Mentioned strengths, limitations, and recommendations within the papers are shown in Table 5. Cited strengths of the research included the prospective design in the two longitudinal papers (Abbey et al., 2012; Thompson et al., 2011), the large sample sizes (Russell, 2016; Saenz, 2009), the representativeness of the target population (Saenz, 2009), and the advantages of the study selection processes, whether online surveys (Russell, 2016) or on account of using random digit dialing to obtain participants (Abbey et al., 2012).

Table 5

Cited strengths, limitations, and future directions within papers

Authors, Year	Cited strengths	Cited limitations	Future recommendations regarding RMA and sexual violence
Abbey, Wegner, Pierce, & Jacques-Tiura, 2012	Prospective design, 18-35 year old age range increases generalisability beyond the 18-22 year old age groups used in most research. Random digit dialling participant selection.	Age range encompasses several developmental stages that were unable to be investigated on account of the small sample. Completion of baseline questionnaire might have sensitised participants to the aims of the research.	The use of more cell phone and nationally representative samples. Development of more nuanced theories to explain different patterns of sexual aggression. Use of behaviourally specific questions during research. Further development of prevention programmes.
Anderson & Anderson, 2008 (study 1 only)	Results correspond with other longitudinal works.	Cross-sectional correlation study, so difficult to infer causality. Participants represent only a subset of sexual aggressors.	Require additional longitudinal study on male-on-female aggression looking to answer both specific and general questions.
Mouilso & Calhoun, 2013	Adds to the literature emphasising the importance of personality in sexual aggression.	Self-report data. Cross-sectional study, so unable to infer causation. Sample consists mostly in young Caucasian men, so findings may not be generalisable.	

Russell, 2016	Large sample size, online survey format (although also acknowledged as a weakness).	Small university sample. Online data collection. Self-report data. Cross-sectional sample.	Replication of these findings.
Russell & King, 2016		Self-report, cross-sectional data.	Would benefit from further longitudinal data.
Saenz, 2009	Ethnic diversity of sample. Study sample representative of college students. Large sample.	Findings not generalisable beyond university population. Did not investigate all potential predictor variables. Cross-sectional design cannot imply causality.	Replicate in community samples. Longitudinal studies. Investigation of risk factors within incarcerated individuals to supplement this.
Swartout, 2013		Measure of RMA was developed a long time ago. University sample. Some participants had no history of sexual activity.	More longitudinal research. Replicate findings in larger community samples. Integrate findings into prevention programmes.
Thompson, Koss, Kingree, Goree, & Rice, 2011	Prospective design. Emphasises importance of attitudes and norms in contributing to the risk of sexually violent behaviours.	Only two waves of data collection analysed. Included male students from only one university. No measure of rape proclivity.	Prevention programmes should incorporate strategies to alter attitudes and norms, whether for high-risk groups, or the general population.
Warren, Swan, & Allen, 2015		Self-report data. Did not assess prior history of abuse. Predictors might vary according to sexual violence severity. Convenience sample of university students. Cross-sectional design.	Findings in need of replication.

Cross-sectional studies cited the use of self-report measures (Mouilso & Calhoun, 2013; Russell, 2016; Russell & King, 2016; Warren et al., 2015) and cross-sectional data (Anderson & Anderson, 2008; Mouilso & Calhoun, 2013; Russell, 2016; Russell & King, 2016; Saenz, 2009; Warren et al., 2015) as key limitations.

Other mentioned limitations either pertained to flaws in the representativeness of the sample, or in the variables measured, and in the subsequent analysis. For example, some noted the lack of ethnic diversity in their samples (Mouilso & Calhoun, 2013), and the fact that a University sample was used (Swartout, 2013;

Russell, 2016; Warren et al., 2015), or in one case that the sample was only drawn from a single University (Thompson et al., 2011). Others noted that they may not have measured all key predictor variables (Saenz, 2009; Warren et al., 2015), or that some participants reported no history of sexual activity (Swartout, 2013), and were thus unlikely to report sexual violence perpetration. Further, one study discussed the possibility of differing sexual violence etiologies across developmental stages (Abbey et al., 2011), and another suggested that predictor variables might vary according to the severity of the sexual violence perpetrated (Warren et al., 2015).

Recommendations for further research included the replication of study findings (Swartout, 2013; Warren et al., 2015; Saenz, 2009; Russell, 2016), the need for more longitudinal research (Anderson & Anderson, 2008; Saenz, 2009; Swartout, 2013), and the integration of these findings into prevention efforts (Abbey et al., 2012; Swartout, 2013; Thompson et al., 2011). Other recommendations included the replication of findings within incarcerated populations (Saenz, 2009); more frequent use of random digit dialed sample selection, behaviourally specific questions, and the development of more comprehensive theories (Abbey et al., 2012); as well as the use of both specific and general research questions (Anderson & Anderson, 2008).

1.3.8 Quality appraisal

After arranging the papers in alphabetical order, and assigning each a number corresponding to this order, a random number generator (Haahr, 1998) was

used to select 5 papers for independent critical appraisal, as this covers more than half of the final set of works. The resultant numbers generated were: 9, 1, 8, 3, 9, and 7. As the number 9 was selected twice, only those papers corresponding to numbers 9, 1, 8, 3, and 7 were submitted to the external researcher for critical appraisal. Once both researchers had conducted critical appraisal, assessments were compared for discrepancies. Agreement was substantial across ratings ($κ$=0.62, see Table 6), and many of the disagreements were simply on account of the second researcher's lack of familiarity with the literature. For example, she felt unable to comment on the legitimacy of the instruments used, and therefore there were discrepancies on ten items for this reason alone (see Table A2).

Table 6

Inter-rata agreement on critical appraisal between researchers

First Researcher		Second Researcher		
		Yes	No	Unknown
	Yes	54	2	10
	No	5	6	1
	Unknown	4	0	26

Some of the remaining disagreements were swiftly resolved by pointing out information that was not immediately obvious, for example, the target population was mentioned in the title of some papers (Thompson et al., 2011; Warren et al., 2015). Others involved more in-depth discussion, for example, the second researcher deemed University samples appropriate without need of justification (Swartout, 2013; Thompson et al., 2011; Warren et al., 2015), on account of their widespread use within Psychology research. However, given the

high levels of sexual violence perpetration within samples of University males (e.g. Finley & Corty, 1993), we resolved that a University sample is appropriately representative only where University men were specified as the target population.

In general, discrepancies were resolved without change to the original quality appraisal, and this was therefore seen as a robust assessment of study quality.

1.3.9 Study Quality

There were several trends in the quality of the reporting of this research (see Table 7). Strengths of the literature included clear statement of the aims of the research, choosing an appropriate study design to achieve these aims, reporting basic data descriptions, as well as drawing conclusions that were sufficiently justified by these results, and reporting on the potential limitations of these conclusions. In addition, the measures of RMA and the SES used were appropriate for the aims of the studies, although this is partly owing to the stringent inclusion criteria adopted herein.

The instruments and consent materials were predominantly administered correctly – studies generally employed measures to ensure participants' privacy when answering sensitive questions and explicitly cited achieving consent, although instrument and consent administration was not detailed in one case (Anderson & Anderson, 2008). Further, in contrast to earlier findings (e.g. Buhi, 2005), internal consistency of RMA instruments was exemplary in most cases,

defined as >0.8 (Robinson, Shaver, & Wrightsman, 1999), and more than

adequate in the remaining two (>0.7, see: Ponterotto & Ruckdeschel, 2007). The

internal consistency of the SES was not reported in some cases (Anderson &

Anderson, 2008; Mouilso & Calhoun, 2013; Thompson et al., 2011; Warren et al.,

2015), but all papers that did report this information again demonstrated good

Table 7

Final Quality Appraisal across Studies

	Abbey et al., 2012	Anderson & Anderson, 2008	Mouilso & Calhoun, 2013	Russell & King, 2016	Russell, 2016	Saenz, 2009	Swartout, 2013	Thompson et al., 2011	Warren et al., 2015
Were the aims/objectives of the study clear?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was the study design appropriate for the stated aims?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was the sample size justified?	No – but cited as a strength in the discussion.	No	No - but acknowledged as a limitation	No	No – but cited as a strength in the discussion.	No	No – but described as sufficient for Structural Equation Modeling	No	No
Was the target/reference population clearly defined?	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Was the sample frame taken from an appropriate population base so that it closely represented the target/reference population under investigation?	Yes	No	Yes	Unknown – was not properly defined.	Yes	Yes	No	Yes	Yes
Was the selection process likely to select	Yes - ensured that	Unknown -	Unknown - demographic	Unknown - above.	Yes	Yes – checked for	Unknown - university	Unknown – difficult to	Unknown – convenience

subjects/participants that were representative of the target/reference population under investigation?	participants were single and had dated a woman recently.	university sample used to draw conclusions about men in general, but did check for heterosexuality.	s questions might have screened people according to their sexuality, but this is unclear.			heterosexuality	sample used to draw conclusions about men in general, but some screening checks exercised.	tell from reported information about selection.	sample, but University men seem to be the target population.
Were measures undertaken to address and categorise non-responders?	Unknown	Unknown	Unknown	n/a	n/a	n/a	Unknown	n/a	Unknown
Were the risk factor and outcome variables measured appropriate to the aims of the study?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Were the risk factor and outcome variables measured correctly using instruments/measurements that had been trialled, piloted, or published previously?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Is it clear what was used to determine statistical significance and/or precision estimates?	Yes	Yes	Yes	Yes	Yes	No - p-values not reported.	Yes	Yes	Yes
Were the methods sufficiently described to	Yes	Yes	Yes	Yes - but did not specify	Yes	No - unclear how SES was scored.	Yes	Yes	Unknown - mediation analysis is

enable them to be repeated?				that RMA was total score.					described in an ambiguous manner.
Were the basic data adequately described?	Yes	No	Yes - but little information reported on demographics questionnaire used.	Yes	Yes	Yes	Yes	No - did not include basic data on drinking, pornography use, and other descriptive data	Yes
Does the response rate raise concerns about non-response bias?	Unknown – not reported.	Unknown – not reported.	Unknown - not reported.	n/a	n/a	n/a	Unknown - not reported.	n/a	Unknown - not reported.
If appropriate, was information about non-responders described?	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown
Were the results internally consistent?	Yes	Yes	Unknown - not reported for SES.	Yes	Yes	Yes	Yes	Unknown - not reported for SES.	Unknown – not reported for SES
Were the results for the analyses described in the methods, presented?	Unknown - no analyses described in methods.	Unknown - no analyses described in methods	Unknown - no analyses described in methods.	Yes	Yes – but primarily described in introduction.	Unknown - no analyses described in methods.	Unknown - no analyses described in methods.	Unknown - no analyses described in methods.	Yes
Were the authors' discussions and conclusions justified by the results?	Yes	Yes	Yes	Unknown	Yes	Yes	Yes	Yes	Yes
Were the limitations of the study discussed?	Yes	Yes	Yes	Yes	Yes	Yes	Yes - and alternative	Yes – in brief.	Yes

								explanations discussed.	
Were there any funding sources or conflicts of interest that may affect the authors' interpretation of the results?	Unknown – not reported.	Unknown – not reported.	Unknown – not reported.	Unknown – not reported.	Unknown – not reported.	Unknown – not reported.	Unknown – not reported.	No – though this was the only study in the group to make this explicit, and some funding was disclosed, though unlikely to impact bias.	Unknown – not reported.
Was ethical approval or consent of participants attained?	Yes	Unknown – does not say, but participants were debriefed.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Participants excluded for ineligibility at baseline, and why	Unknown - the paper cites: Abbey, Jacques-Tiura, and leBreton (2011), but no further information is given in this work.	n/a	n/a	n/a	n/a	n/a	n/a	Unknown - no data provided, but this is partly to do with how data was collected.	n/a

Participants who did not consent, and why	"	n/a	n/a	n/a	n/a	n/a	n/a	No	n/a
Participants lost after consent, and why	"	n/a	n/a	n/a	n/a	n/a	n/a	No	n/a
Total number of participants at baseline	470	n/a	n/a	n/a	n/a	n/a	n/a	795	n/a
Loss to follow-up, and why	27 did not want to participate, 7 had another person refuse for them, 14 repeatedly missed appointments, 12 were inelligible either due to moving (10), hospitalisation (1), or incarceration (1).	n/a	n/a	n/a	n/a	n/a	n/a	82% responded at follow-up, but no reasons given	n/a
Total number of participants participating at wave 2 of data collection	425	n/a	n/a	n/a	n/a	n/a	n/a	652	n/a

internal consistency, with a minimum alpha coefficient of 0.84 (Abbey et al.,

2012).

Where critical appraisal items were not met across studies, this often either

pointed to failures in study reporting, or to limitations of the study sample. Only

one of the two longitudinal studies effectively reported on loss to follow-up

(Abbey et al., 2012); one paper generalised findings to extend from their

University sample to the male population in general (Anderson & Anderson,

2008); and despite the large average sample size across studies, justification of

sample size was never offered: one paper mentioned a "target" sample size (800,

Thompson et al., 2011), but offered no explanation for this figure. Further, many

of the studies failed to take measures during study selection to ensure that their

sample consisted in heterosexual men (Russell, 2016; Russell & King, 2016;

Swartout, 2013; Thompson et al., 2011; Mouilso & Calhoun, 2013), which would

be the population at risk of perpetrating male-on-female sexual violence.

In some cases where critical appraisal items were inconsistently met across

papers, this was because they were not easily applied to the collated works. For

example, it was difficult to comment on funding sources or conflicts of interest as

these were exclusively disclosed in one study (Thompson et al., 2011), and

assessing whether the intended analysis was executed was often impossible, as

no such intentions were described in the methods (Anderson & Anderson, 2008;

Mouilso & Calhoun, 2013; Saenz, 2009; Swartout, 2013; Thompson et al., 2011).

Similarly, none of the included papers provided information about non-responders, but this was not always plausible. In those studies that used flyers (Thompson et al., 2011) or advertisements (Saenz, 2009) to recruit participants, describing non-responders would be transparently difficult. Further, some studies (Russell & King, 2016; Russell, 2016) used the online subject pool MechanicalTurk, which does not provide information on individuals who previewed the study and declined. Most studies utilised University subject pools, and while these may encounter the same problems as above, no method of recruitment was described, and thus it was difficult to determine whether the omission of this information was legitimate or not (Anderson & Anderson, 2008; Mouilso & Calhoun, 2013; Russell, 2016; Swartout, 2013; Warren et al., 2015).

1.4   Discussion

This systematic review examined the current evidence on the association between RMA, as measured by instruments that explicitly reference RMA, and sexual violence behaviours, as measured by the SES. While there was a general paucity of recent evidence on this association, eight of the nine included works established a significant relationship between RMA and sexual violence. The only study not to find a significant association between RMA and sexual violence exclusively examined sexual violence behaviours perpetrated within the past four months (Warren et al., 2015), and two of the significant associations were found in longitudinal studies. The first of these longitudinal works found that RMA significantly differentiated non-perpetrators from "initiators", which

suggests that it significantly affects the onset of sexual violence behaviours; and the second controlled for sexual violence at baseline, so RMA maintained predictive power on top of previous behaviour, though from this analysis we are unable to establish whether RMA plays a role in maintaining or exacerbating these behaviours, rather than initiating them. RMA was no longer a significant predictor of sexual violence in two multivariate models – one controlling for psychopathy (Mouilso & Calhoun 2008), and one that factored in conformity to masculine norms, comprehension of sexual consent, and peer support of abuse (Warren et al., 2015). Yet, RMA remained significant in a model that controlled for "perceived norms", which likely overlaps with: "conformity to masculine norms" and "peer support of abuse". Therefore, the variables: psychopathy, and comprehension of sexual consent, might represent the best targets for future research on RMA and sexual violence, especially as there is preliminary evidence elsewhere to suggest an association between RMA and psychopathy (Debowska, Boduszek, Dhingra, Kola, & Meller-Prunska, 2015). This builds on the results of a previous systematic review of the association between RMA and sexual violence behaviours, in which 29 out of 31 cross-sectional studies and 2 out of 3 prospective works established significant associations (Tharp et al., 2013). Taken together, these works provide strong evidence for the association between RMA and sexual violence perpetration, and indicate that the wide variety of instruments currently used to measure RMA have predictive validity.

This review has also established several important issues in the quality of the current literature surrounding RMA and sexual violence perpetration. Many of these pertain to flaws in study reporting: longitudinal works failed to detail loss

to follow-up, sample size was never justified, and no studies provided details on non-responders – while this was not always appropriate, it should be acknowledged, particularly as volunteers for sexual research tend to be more sexually active (Strassberg & Lowe, 1995), and therefore might represent a special population. Furthermore, researchers frequently failed to detail their intended analysis, and couple this with the fact that many of those papers rejected during the full-text sift were analysing the same sample of men (Zinzow & Thompson, 2015; Zinzow & Thompson, 2014; Thompson et al., 2013), this raises significant concern over the potential for multiple testing and data mining.

Other issues were revealed in the instruments administered to subjects. Many studies failed to issue questions pertaining to sexuality, which would have helped to establish whether the sample constituted the portion of the male population at risk of perpetrating male-on-female sexual violence. In addition, although the internal consistency of RMA instruments was a significant improvement over previous work (cf Buhi, 2005), the instruments used were varied, with some studies using fractions or combinations of measures (Anderson & Anderson, 2008; Abbey et al., 2012).

There was considerably more consistency in the version of the SES used, though internal consistency was sometimes not offered. The newest version (Koss et al., 2007) has been favoured in recent research, which is encouraging, as this version eliminated items that had been criticised for ambiguity (Kolivas & Gross, 2007). However, within this instrument, there was little consensus over how to score the SES. Although testament to the instrument's versatility, this

heterogeneity is also somewhat problematic, as different methods of scoring

result in different rates of perpetration (Davis et al., 2014). In general, studies

coded the SES according to varying degrees of the severity of sexual violence, but

this scoring method might be considered particularly misleading, as coding by

the objective severity of acts correlates poorly with the subjective trauma

incurred by these actions (Testa VanZile-Tamsen, Livingston, & Koss, 2004) and

RMA may function differentially in accordance with sexual violence severity (a

concern expressed by: Warren et al., 2015), especially as one paper established

stronger associations between RMA and the perpetration of less severe acts

(Swartout, 2013). Similarly, coding according to frequency equates less severe

acts with worse ones (Koss et al., 2007), and in this review it was often unclear

how frequency was determined.

1.4.1 Recommendations

This review has therefore generated several recommendations for future

research. Included papers cited the need for more longitudinal evidence, random

sampling methods, and replication in community or incarcerated samples. This is

essential in order to establish generalisability of the association that is evidenced

herein, and to contribute a larger body of prospective evidence towards the

assertion that RMA facilitates sexual violence. Theorists should also establish

some consistency in how the SES is scored, and a scoring method that considers

both severity and frequency is recommended (cf Davis et al., 2014; e.g.

Thompson et al., 2011). Researchers should also make a concerted effort to

report on intended analysis, non-responders, the sexuality of participants, loss to

follow-up, and justification of the sample size. These measures would ease concern over the power of these findings, multiple testing, and response bias.

This review recommends further research into this association outside of the United States, as male-on-female sexual violence still remains a significant problem elsewhere (Office for National Statistics, 2018), and many organisations in the United Kingdom specifically campaign around rape supportive attitudes (see: Rape Crisis England and Wales, 2018). In light of the two multivariate models in which RMA was not a significant predictor of sexual violence, further research into the association between RMA and sexual violence with particular reference to psychopathy and comprehension of sexual consent, as well as the roles of conformity to masculine norms, and peer support of abuse, should also be investigated within this relationship. This might help to illuminate the cognitive profile of a perpetrator of male-on-female sexual violence, and would ease concern over whether RMA is a proxy predictor of sexual violence, and subsequently a redundant target of prevention programmes.

1.4.2 Limitations of this review

There are, however, several limitations of the conclusions to be drawn from this review. All of the included works were conducted in the United States, and while there is understandably high concern over sexual violence in the U.S., as it has the highest rape rate of any industrialised country (see: Black, Basile, Breiding, & Ryan, 2014), this still limits the scope of these conclusions to American perpetrators. Similarly, many of the studies utilised University samples, and

while this was a smaller proportion than has been established in other systematic reviews (e.g. Tharp et al., 2013) this too restricts the generalisability of these findings.

Furthermore, the inclusion criteria adopted herein were relatively stringent: in restricting the measure of RMA exclusively to instruments that explicitly refer to rape myths, some similar instruments were consequently excluded (e.g. Feild, 1978). In restricting the search to articles exercising the Sexual Experiences Survey, these conclusions are limited to a self-report instrument. During the full-text sift, 29 articles were rejected as they did not measure sexual violence with a version of the Sexual Experiences Survey (see Table A1), and therefore this review should be supplemented with more work describing these papers, and the association between RMA and sexual violence within them. Allowing unpublished works within the scope of the search meant the inclusion of one doctoral and one master's thesis, and as a result, these works may not have been conducted with appropriate rigor. However, the fact that both established significant associations between RMA and sexual violence is encouraging, as it indicates that this association, at least within this review, has not been compounded by publication bias. The second researcher who conducted the critical appraisal was from a biology background, and was relatively unfamiliar with the nuances within this literature. As a result of this, bias may have inadvertently been introduced into the quality appraisal.

1.4.3 Conclusion

In conclusion, the limited current literature suggests that RMA is associated with the perpetration of male-on-female sexual violence, and even temporally precedes sexual violence in two longitudinal works, which adds to other existing evidence of causality (cf the dose-response effect observed in DeGue et al., 2010). However, there are many pitfalls within this research, as it is all conducted in the United States; the SES is coded in a variety of different ways; and much of the important data on sample selection and justification is neglected in study reporting. Therefore, future research into this association could ease concern over the legitimacy of this relationship by providing more comprehensive justifications and details during reporting, as well as addressing the prospect of RMA being an indirect predictor of sexual violence.

References

Abbey, A., Parkhill, M. R., BeShears, R., Clinton-Sherrod, A. M., & Zawacki, T. (2006). Cross-sectional predictors of sexual assault perpetration in a community sample of single African American and Caucasian men. *Aggressive behavior*, *32*(1), 54-67.

Abbey, A., Wegner, R., Pierce, J., & Jacques-Tiura, A. J. (2012). Patterns of sexual aggression in a community sample of young men: Risk factors associated with persistence, desistance, and initiation over a 1-year interval. *Psychology of violence*, *2*(1), 1.

Abrahams, N., Devries, K., Watts, C., Pallitto, C., Petzold, M., Shamu, S., & GarcÍa-Moreno, C. (2014). Worldwide prevalence of non-partner sexual violence: a systematic review. *The Lancet*, *383*(9929), 1648-1654.

Abrams, D., Viki, G. T., Masser, B., & Bohner, G. (2003). Perceptions of stranger and acquaintance rape: the role of benevolent and hostile sexism in victim blame and rape proclivity. *Journal of personality and social psychology*, *84*(1), 111.

Alberty, J. (2011). Unwanted Sexual Experience: An Investigation of Emotion and Physiology. (Unpublished Masters thesis). Loma Linda University, California.

Anderson, C. A., & Anderson, K. B. (2008). Men who target women: Specificity of target, generality of aggressive behavior. *Aggressive Behavior*, *34*(6), 605-622.

Black, M. C., Basile, K. C., Breiding, M. J., & Ryan, G. W. (2014). Prevalence of

    sexual violence against women in 23 states and two US territories, BRFSS

    2005. *Violence against women*, *20*(5), 485-499.

Bohner, G., Jarvis, C. I., 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*(6), 819-

    828.

Bolton-Holz, K., Fischer, A., & Daood, C. (2016). The Role of Men's Beliefs in

    Shaping Their Response to a Sexual Violence Prevention Program.

    Psychology of Men & Masculinity, Psychology of Men & Masculinity, 2016.

Bowes, N., & McMurran, M. (2013). Cognitions supportive of violence and violent

    behavior. *Aggression and Violent Behavior*, *18*(6), 660-665.

Bryce, D. M. (2012). *Predicting men's relationship satisfaction with men through

    internalized homonegativity and restricted emotionality* (Unpublished

    Doctoral dissertation). University of La Verne, California.

Buhi, E. R. (2005). Reliability reporting practices in rape myth research. *Journal

    of School Health*, *75*(2), 63-66.

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.

Campbell, R., Patterson, D., & Bybee, D. (2012). Prosecution of Adult Sexual

    Assault Cases*. Violence Against Women, 18*(2), 223-244.

Casey, E. A., & Lindhorst, T. P. (2009). Toward a multi-level, ecological approach

  to the primary prevention of sexual assault prevention in peer and

  community contexts. *Trauma, Violence, & Abuse*, *10*(2), 91-114.

Crowell, N. A., & Burgess, A. W. (1996). Understanding Violence Against Women:

  Panel on Research on Violence Against Women. Washington, DC: National

  Research Council. *National Academy of Sciences*.

Davis, K. C., Gilmore, A. K., Stappenbeck, C. A., Balsan, M. J., George, W. H., &

  Norris, J. (2014). How to score the sexual experiences survey? A

  comparison of nine methods. *Psychology of violence*, *4*(4), 445.

Debowska, A., Boduszek, D., Dhingra, K., Kola, S., & Meller-Prunska, A. (2015).

  The role of psychopathy and exposure to violence in rape myth

  acceptance. *Journal of interpersonal violence*, *30*(15), 2751-2770.

Downes, M. J., Brennan, M. L., Williams, H. C., & Dean, R. S. (2016). Development

  of a critical appraisal tool to assess the quality of cross-sectional studies

  (AXIS). *BMJ open*, *6*(12), e011458.

Fawcett Society (2017). *Sounds Familiar*. Retrieved from:

  https://www.fawcettsociety.org.uk/News/fawcett-report-hostility-

  complacency-blame-culture-against-women

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.

Finley, C., & Corty, E. (1993). Rape on the campus: The prevalence of sexual

  assault while enrolled in college. *Journal of College Student

  Development*, *34*, 113-113.

Fisher, B. S. (2009). The effects of survey question wording on rape estimates: Evidence from a quasi-experimental design. *Violence Against Women*, *15*(2), 133-147.

Fisher, B. S., & Cullen, F. T. (2000). Measuring the sexual victimization of women: Evolution, current controversies, and future research. *Criminal justice*, *4*, 317-390.

Forbes, G. B., & Adams-Curtis, L. E. (2001). Experiences with sexual coercion in college males and females: Role of family conflict, sexist attitudes, acceptance of rape myths, self-esteem, and the Big-Five personality factors. *Journal of Interpersonal Violence*, *16*(9), 865-889.

Forbes, G. B., Adams-Curtis, L. E., & White, K. B. (2004). First-and second-generation measures of sexism, rape myths and related beliefs, and hostility toward women their interrelationships and association with college students' experiences with dating aggression and sexual coercion. *Violence against women*, *10*(3), 236-261.

Gami, A. S., Witt, B. J., Howard, D. E., Erwin, P. J., Gami, L. A., Somers, V. K., & Montori, V. M. (2007). Metabolic syndrome and risk of incident cardiovascular events and death: a systematic review and meta-analysis of longitudinal studies. *Journal of the American College of Cardiology*, *49*(4), 403-414.

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*(5), 422-440.

Haahr, M. (1998). RANDOM.ORG [Software]. Available from https://www.random.org/

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.

Hust, S., Rodgers, K., Ebreo, S., & Stefani, W. (2017). Rape Myth Acceptance,

Efficacy, and Heterosexual Scripts in Men's Magazines: Factors Associated

With Intentions to Sexually Coerce or Intervene. *Journal of Interpersonal*

Violence, doi:088626051665375.

Kennair, L. E. O., & Bendixen, M. (2012). Sociosexuality as predictor of sexual

harassment and coercion in female and male high school

students. *Evolution and Human Behavior*, *33*(5), 479-490.

Kolivas, E. D., & Gross, A. M. (2007). Assessing sexual aggression: Addressing the

gap between rape victimization and perpetration prevalence

rates. *Aggression and Violent Behavior*, *12*(3), 315-328.

Koss, M. P., Abbey, A., Campbell, R., Cook, S., Norris, J., Testa, M., ... & White, J.

(2007). Revising the SES: A collaborative process to improve assessment

of sexual aggression and victimization. *Psychology of Women*

Quarterly, *31*(4), 357-370.

Koss, M. P., & Gidycz, C. A. (1985). Sexual experiences survey: reliability and

validity. *Journal of consulting and clinical psychology*, *53*(3), 422.

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*(2), 162.

Koss, M. P., & Oros, C. J. (1982). Sexual Experiences Survey: a research instrument investigating sexual aggression and victimization. *Journal of consulting and clinical psychology*, *50*(3), 455.

Larcombe, W. (2011). Falling Rape Conviction Rates: (Some) Feminist Aims and Measures for Rape Law. *Feminist Legal Studies, 19*(1), 27-45.

Lisak, D., & Miller, P. M. (2002). Repeat rape and multiple offending among undetected rapists. *Violence and victims*, *17*(1), 73-84.

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*(10), 1325-1348.

Lonsway, K. A., & Fitzgerald, L. F. (1994). Rape myths in review. *Psychology of women quarterly*, *18*(2), 133-164.

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*(4), 704.

Malamuth, N. M. (1981). Rape proclivity among males. *Journal of social issues*, *37*(4), 138-157.

Malamuth, N. M., Linz, D., Heavey, C. L., Barnes, G., & Acker, M. (1995). Using the confluence model of sexual aggression to predict men's conflict with women: a 10-year follow-up study. *Journal of personality and social psychology*, *69*(2), 353.

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

Maruna, S., & Mann, R. E. (2006). A fundamental attribution error? Rethinking

cognitive distortions. *Legal and Criminological Psychology*, *11*(2), 155-

177.

McDermott, R. C., Kilmartin, C., McKelvey, D. K., & Kridel, M. M. (2015). College

male sexual assault of women and the psychology of men: Past, present,

and future directions for research. *Psychology of Men & Masculinity*, *16*(4),

355.

McMahon, S., & Farmer, G. L. (2011). An updated measure for assessing subtle

rape myths. *Social Work Research*, *35*(2), 71-81.

McMahon, S., Postmus, J. L., Warrener, C., & Koenick, R. A. (2014). Utilizing peer

education theater for the primary prevention of sexual violence on college

campuses. *Journal of college student development*, *55*(1), 78-85.

Ministry of Justice. (2013). *An Overview of Sexual Offending in England and Wales.*

Retrieved from: https://www.gov.uk/government/statistics/an-

overview-of-sexual-offending-in-england-and-wales

Moher, D., Pham, B., Lawson, M. L., & Klassen, T. P. (2003). The inclusion of

reports of randomised trials published in languages other than English in

systematic reviews. *Health Technol Assess*, *7*(41), 1-90.

Mouilso, E. R., & Calhoun, K. S. (2013). The role of rape myth acceptance and

psychopathy in sexual assault perpetration. *Journal of Aggression,

Maltreatment & Trauma*, *22*(2), 159-174.

Office for National Statistics. (2018). *Sexual offences in England Wales: year

ending March 2017.* Retrieved from:

https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjusti
ce/articles/sexualoffencesinenglandandwales/yearendingmarch2017

Olver, M. E., Nicholaichuk, T. P., Kingston, D. A., & Wong, S. C. (2014). A multisite
examination of sexual violence risk and therapeutic change. *Journal of
consulting and clinical psychology*, *82*(2), 312.

Payne, D. L., Lonsway, K. A., & Fitzgerald, L. F. (1999). Rape Myth Acceptance:
Exploration of Its Structure and Its Measurement Using theIllinois Rape
Myth Acceptance Scale. *Journal of Research in Personality*, *33*(1), 27-68.

Peterson, K., Sharps, P., Banyard, V., Powers, R. A., Kaukinen, C., Gross, D., ... &
Campbell, J. (2016). An evaluation of two dating violence prevention
programs on a college campus. *Journal of interpersonal violence*,
doi:0886260516636069.

Ponterotto, J. G., & Ruckdeschel, D. E. (2007). An overview of coefficient alpha
and a reliability matrix for estimating adequacy of internal consistency
coefficients with psychological research measures. *Perceptual and motor
skills*, *105*(3), 997-1014.

Porter, J. F., & Critelli, J. W. (1992). Measurement of sexual aggression in college
men: A methodological analysis. *Archives of sexual behavior*, *21*(6), 525-
542.

Rape Crisis England and Wales. (2018). *Myths vs realities.* Retrieved from:
https://rapecrisis.org.uk/mythsvsrealities.php

Rapaport, K., & Burkhart, B. R. (1984). Personality and attitudinal characteristics
of sexually coercive college males. *Journal of Abnormal Psychology*, *93*(2),
216-221.

Romero-Sánchez, M., & Megías, J. L. (2010). Alcohol use as a strategy for

    obtaining nonconsensual sexual relations: incidence in Spanish university

    students and relation to rape myths acceptance. *The Spanish journal of*

    psychology, *13*(02), 864-874.

Russell, T. D. (2016). *The PID-5, Everyday Sadism, and Parental Attachment*

    Predict Sexual Aggression (Unpublished Doctoral dissertation). The

    University of North Dakota, North Dakota.

Russell, T. D., & King, A. R. (2016). Anxious, hostile, and sadistic: Maternal

    attachment and everyday sadism predict hostile masculine beliefs and

    male sexual violence. *Personality and Individual Differences*, *99*, 340-345.

Saenz, C. (2009). *Integrating theories of sexual assault: Incorporating narcissistic*

    reactance theory into the confluence model of sexual

    aggression (Unpublished Doctoral dissertation). Wayne State University,

    Michigan.

Stephens, K. A., & George, W. H. (2009). Rape prevention with college men

    evaluating risk status. *Journal of Interpersonal Violence*, *24*(6), 996-1013.

Strassberg, D. S., & Lowe, K. (1995). Volunteer bias in sexuality research. *Archives*

    of sexual behavior, *24*(4), 369-382.

Süssenbach, P., Eyssel, F., & Bohner, G. (2013). Metacognitive aspects of rape

    myths: Subjective strength of rape myth acceptance moderates its effects

    on information processing and behavioral intentions. *Journal of*

    interpersonal violence, *28*(11), 2250-2272.

Swartout, K. M. (2013). The company they keep: How peer networks influence

    male sexual aggression. *Psychology of Violence*, *3*(2), 157.

Testa, M., VanZile-Tamsen, C., Livingston, J. A., & Koss, M. P. (2004). Assessing

women's experiences of sexual aggression using the sexual experiences

survey: Evidence for validity and implications for research. *Psychology of*

Women Quarterly, *28*(3), 256-265.

Tharp, A. T., DeGue, S., Valle, L. A., Brookmeyer, K. A., Massetti, G. M., & Matjasko,

J. L. (2013). A systematic qualitative review of risk and protective factors

for sexual violence perpetration. *Trauma, Violence, & Abuse*, *14*(2), 133-

167.

Thompson, M. P., Koss, M. P., Kingree, J. B., Goree, J., & Rice, J. (2011). A

prospective mediational model of sexual aggression among college

men. *Journal of Interpersonal Violence*, *26*(13), 2716-2734.

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.

Tooth, L., Ware, R., Bain, C., Purdie, D. M., & Dobson, A. (2005). Quality of

reporting of observational longitudinal research. *American Journal of*

Epidemiology, *161*(3), 280-288.

Truman, D. M., Tokar, D. M., & Fischer, A. R. (1996). Dimensions of masculinity:

Relations to date rape supportive attitudes and sexual aggression in

dating situations. *Journal of Counseling and Development: JCD*, *74*(6), 555.

Vega, V., & Malamuth, N. M. (2007). Predicting sexual aggression: The role of

pornography in the context of general and specific risk factors. *Aggressive*

Behavior, *33*(2), 104-117.

Walker, J. S., & Bright, J. A. (2009). Cognitive therapy for violence: reaching the parts that anger management doesn't reach. *The Journal of Forensic Psychiatry & Psychology*, *20*(2), 174-201.

Ward, C. (1988). The attitudes toward rape victims scale. *Psychology of women quarterly*, *12*(2), 127-146.

Warren, P., Swan, S., & Allen, C. T. (2015). Comprehension of sexual consent as a key factor in the perpetration of sexual aggression among college men. *Journal of Aggression, Maltreatment & Trauma*, *24*(8), 897-913.

World Health Organization. (2013). *Global and regional estimates of violence against women: prevalence and health effects of intimate partner violence and non-partner sexual violence*. World Health Organization.

Zeng, X., Zhang, Y., Kwong, J. S., Zhang, C., Li, S., Sun, F., ... & Du, L. (2015). The methodological quality assessment tools for preclinical and clinical studies, systematic review and meta-analysis, and clinical practice guideline: a systematic review. *Journal of Evidence-Based Medicine*, *8*(1), 2-10.

Zinzow, H. M., & Thompson, M. (2014). Factors associated with use of verbally coercive, incapacitated, and forcible sexual assault tactics in a longitudinal study of college men. *Aggressive behavior*, *41*(1), 34-43.

Zinzow, H. M., & Thompson, M. (2015). A longitudinal study of risk factors for repeated sexual coercion and assault in US college men. *Archives of sexual behavior*, *44*(1), 213-222.

Appendices

Table A1

Papers excluded after full-text sift

Authors	Year	Title	Reason for exclusion
Addison, William B.	2015	Embedding Sexual Assault Bystander Awareness Education in a Criminal Justice Course	Wrong measures: no measure of RMA nor SES.
Barnard, Sarah	2015	Police officers' attitudes about rape victims	Wrong measures: no measure of SES.
Berliant, Julia	2012	Sexual Assault Prevention: Changing Explicit and Implicit Cognitions of University Men	Wrong measures: no measure of SES.
Bethune, Kristen M.	2016	A perfect storm: A routine activity analysis of female undergraduate sexual assault	Manuscript unavailable
Bezouska, Saori	2014	Implicit Objectification and Sexual Aggression Myths in Japanese Culture	Wrong measures: no measure of SES.
Bliss, Beth A.	2013	Is it still a sexual offense if society doesn't find it offensive? Cultural constructions and rape proclivity	Wrong measures: no measure of SES.
Bradley, April R.; Yeater, Elizabeth A.; O'Donahue, William	2009	An Evaluation of a Mixed-Gender Sexual Assault Prevention Program	Wrong measures: no measure of SES.

Chapleau, Kristine	2009	Power, sex, and rape myth acceptance: Testing two models of rape proclivity	Insufficient data (nothing reported on relation between RMA and SV): no response to email
Cook-Craig, Patricia G.; Coker, Ann L; Clear, Emily R; Garcia, Lisandra S; Bush, Heather M; Brancato, Candace J; Williams, Corrine M; Fisher, Bonnie S.	2014	Challenge and opportunity in evaluating a diffusion-based active bystanding prevention program: Green Dot in high schools	Wrong measures: no measure of SES.
Dardis, Christina M.; Murphy, Megan J.; Bill, Alexander C.; Gidycz, Christine A.	2016	An investigation of the tenets of social norms theory as they relate to sexually aggressive attitudes and sexual assault perpetration: A comparison of men and their friends	Wrong population: only men who were already considered to be moderately coercive, so not representative of full male population.
Darlington, Erin Marie	2014	Decreasing misperceptions of sexual violence to increase bystander intervention: A social norms intervention	Wrong design: Intervention study
Davis, Kelly C.; Gilmore, Amanda K.; Stappenbeck, Cynthia A.; Balsan, Michael J.; George, William H.; Norris, Jeanette	2014	How to score the sexual experiences survey? A comparison of nine methods	Wrong design: Instrument development
Davis, Kelly C.; Logan-Greene, Patricia	2015	Background Predictors and Event-Specific Characteristics of Sexual	Wrong measures: no measure of SES.

		Aggression Incidents: The Roles of Alcohol and Other Factors	
Davis, Kelly Cue; Danube, Cinnamon L; Stappenbeck, Cynthia; Norris, Jeanette; George, William H.	2012	Young Men's Aggressive Tactics to Avoid Condom Use: A Test of a Theoretical Model	Wrong measures: no measure of SES.
DeGannes, Asha Brown	2009	Evaluation of a sexual assault and dating violence prevention program for middle school students	Wrong measures: no measure of SES.
DePuy, Jacqueline; Hamby, Sherry; Lindemuth, Caroline.	2014	Teen Dating Violence in French-speaking Switzerland: Attitudes and Experiences	Wrong measures: no measure of RMA nor SES.
Dyshniku, Fiona	2014	Effect of Deviant Sexual Fantasies on Aberrant Sexual Behaviours	Wrong measures: no measure of SES.
Elias-Lambert, Nada	2013	Bystander sexual violence prevention program: Implementation and evaluation with high-risk university males	Wrong design: Intervention study
Fahs, Breanne; Swank, Eric; Clevenger, Lindsay.	2015	Troubling Anal Sex: Gender, Power, and Sexual Compliance in Heterosexual Experiences of Anal Intercourse	Wrong measures: no measure of SES.
Giovannelli, Thorayya Said	2012	Beliefs of Safety: Sexual Violence Perceptions among Christian College Students	Wrong measures: no measure of SES.
Gonsalves, Valerie M.	2010	Exploring online sexually explicit material: What is the relationship to sexual coercion?	Insufficient data (nothing reported on relation between RMA and SV): no response to email

Hackman, Christine Louise	2015	Investigating multiple layers of influence on sexual assault in a university setting	Insufficient data (stats reported on perpetrators could include women and transgender individuals, not necessarily just males): no response to email.
James, Belinda-Rose	2012	Beliefs of and Attitudes towards Sexual Violence by a Diverse Group of Self-identified Male Collegiate Athletes	Wrong measures: no measure of SES.
Johnson, Rachel	2009	An Investigation of Sexual Narcissism As a Predictor of Sexual Assault	Wrong population: participants in a sex-ofender treatment programme so high risk.
Kilimnik, Chelsea Dawn	2015	Sexual consent: The role of nonconsensual sexual experiences, identification, and affective sexuality	Wrong measures: no measure of SES.
Kingree, J. B.; Thompson, Martie	2015	A Comparison of Risk Factors for Alcohol-Involved and Alcohol-Uninvolved Sexual Aggression Perpetration	Insufficient data: Composite measure of RMA, AND used same sample as included paper by Thompson, Koss, Kingree, Goree, & Rice, 2011
Klein, Carolin; Kennedy, M Alexis; Gorzalka, Boris B	2009	Rape Myth Acceptance in Men Who Completed the Prostitution Offender Program of British Columbia	Wrong measures: no measure of SES.
Malamuth, Neil M.; Hald, Gert Martin; Koss, Mary	2012	Pornography, Individual Differences in Risk and Men's Acceptance of Violence Against Women in a Representative Sample	Wrong measures: no measure of SES.
McCauley, H. L.; Tancredi D.; Silverman	2013	Sport, gender-equitable attitudes and abuse perpetration among a	Insufficient information: no data reported on RMA specifically.

J.; Decker M.; Virata M.C.; O'Connor B.; Miller E.		sample of high school studentathletes	
McGeeney, Ryan	2009	Internet pornography and its effects on the sexual attitudes and behaviors of college students	Wrong measures: invented for study.
Mercer, Mary Catherine	2014	Impact of social influences on men and women's risk recognition of sexual assault	Wrong design: measures administered after an intervention so no pre-test data.
Moor, Avigail PhD	2011	The Efficacy of a High School Rape Prevention Program in Israel	Wrong measures: invented for study.
Moynihan, Mary M.; Banyard, Victoria L.	2008	Community responsibility for preventing sexual violence: A pilot study with campus Greeks and intercollegiate athletes	Wrong measures: no measure of SES.
Oglesby-Taylor, Suzanne F.	2015	Measuring perceptions and attitudes towards rape victims of military members who had sexual assault training	Wrong measures: no measure of SES.
Ojjeh, Falak	2015	Sexual Assault on the College Campus: A Partial Test of Male Peer Support Theory	Wrong measures: no measure of SES.
Ottesen Kennair, Leif Edward; Bendixen, Mons.	2012	Sociosexuality as predictor of sexual harassment and coercion in female and male high school students	Wrong measures: invented for study.
Pape, Erin E.	2009	An analysis of sexually aggressive behavior among college age athletes and members of social fraternities and sororities	Wrong measures: invented for study.

		DOES ALCOHOL CONTRIBUTE TO	
Parkhill, Michele R.;		THE CONFLUENCE MODEL OF	Wrong measures: no measure of
Abbey, Antonia	2008	SEXUAL ASSAULT PERPETRATION?	RMA.
		Incorporating social norms into	Wrong population: men recruited
		sexual assault interventions: Effects	for study had to endorse rape
		on belief and behavior change among	myths to a moderate extent, so not
Paul, Lisa A.	2010	college men	representative.
		Predictors of Harmful Sexual	Wrong measures: no measure of
Poinsett, Matthew A.	2015	Behaviors in a Normative Population	SES.
Reyes, H. Luz		Sexual dating aggression across	
McNaughton; Foshee,		grades 8 through 12: Timing and	Wrong measures: no measure of
Vangie A.	2013	predictors of onset	SES.
		Alcohol Use as a Strategy for	
		Obtaining Nonconsensual Sexual	
		Relations: Incidence in Spanish	Insufficient data (nothing reported
Romero-Sánchez,		University Students and Relation to	on relation between RMA and SV):
Mónica; Megías, Jesús L	2010	Rape Myths Acceptance	no response to email
		Labeling acts of sexual violence:	
		What roles do assault characteristics,	
Sasson, Sapi; Paul, Lisa		attitudes, and life experiences play?	Wrong measures: no measure of
A.r	2014	[References]	SES.
		Blaming the victim: The role of	
		assault characteristics and victim	Wrong measures: no measure of
Sasson, Sapir	2014	attractiveness	SES.
		Enhancement of sexual boundaries:	Wrong measures: no measure of
Sisco, Melissa M.	2010	An online awareness project	SES.
		Explicating the Role of Sexual	
Starfelt, Louise C.;		Coercion and Vulnerability Alcohol	Wrong measures: no measure of
Young, Ross McD;	2015	Expectancies in Rape Attributions	SES.

White, Katherine M; Palk, Gavan RM			
Steel, Jennifer H.	2016	Attitudes and characteristics of military serial rapists: A comparison with their civilian counterparts	Wrong measures: no measure of RMA.
Stephens, K. A.; George, R.H.	2009	Rape prevention with college men: Evaluating risk status	Wrong design: Intervention study
Stephens, Kari A	2009	Rape prevention with Asian/Pacific Islander and Caucasian college men: The roles of culture and risk status	Unavailable.
Thompson, Martie P.; Swartout, Kevin M.; Koss, Mary P.	2012	Trajectories and predictors of sexually aggressive behaviours during emerging adulthood	Insufficient data: Composite measure of RMA, AND used same sample as included paper by Thompson, Koss, Kingree, Goree, & Rice, 2011
Vogt, Taylor Victoria	2015	Exploring Negative Sexual Experiences, Attitudes, and Behaviors by Auditory Status	Unavailable.
Ward, Rose Marie; Matthews, Molly R; Weiner, Judith; Hogan, Kathryn M; Popson, Halle C.	2012	Alcohol and sexual consent scale: Development and validation	Wrong design: Instrument development
Widman, Laura; Olson, Michael	2013	On the Relationship Between Automatic Attitudes and Self-Reported Sexual Assault in Men	Insufficient data: Composite measure of RMA.
Widman, Laura; Olson, Michael A; Bolen, Rebecca M.	2013	Self-reported sexual assault in convicted sex offenders and community men	Wrong design: Instrument development AND insufficient data

			(stats reported on community men and sex offenders combined)
Wiscombe, Karla	2012	An Exploratory Analysis of Sexual Violence and Rape Myth Acceptance at a Small Liberal Arts University	Insufficient data (stats reported on perpetrators could include women).
Yanagida-Ishii, Dailyn Yukimi	2009	Program content in a men-only sexual assault prevention program: The relationship between factual knowledge, familiarity with a victim, and self-reported behavior	Wrong measures: no measure of SES.
Young, B. R.; Desmarais SL; Baldwin JA; Chandler R.	2016	Sexual Coercion Practices Among Undergraduate Male Recreational Athletes, Intercollegiate Athletes, and Non-Athletes	Wrong measures: no measure of SES.
Zinzow, H. M.; Thompson, M.	2014	Factors associated with use of verbally coercive, incapacitated, and forcible sexual assault tactics in a longitudinal study of college men	Insufficient data: Composite measure of RMA, AND used same sample as included paper by Thompson, Koss, Kingree, Goree, & Rice, 2011
Zinzow, Heidi M.; Thompson, M.	2015	A Longitudinal Study of Risk Factors for Repeated Sexual Coercion and Assault in U.S. College Men	Insufficient data: Composite measure of RMA, AND used same sample as included paper by Thompson, Koss, Kingree, Goree, & Rice, 2011

Table A2

Discrepancies during critical appraisal, and resolution

Author	Item	First researcher	Second researcher	Resolution
Abbey et al., 2012	Was the sample size justified?	Yes	No	Yes – justified in Discussion.
Warren et al., 2015	Was the sample size justified?	No	Yes	No – it was not justified, but the second researcher acknowledged that it was admitted as a limitation, using a "convenience sample", and this was highlighted within the text as a result.
Thompson et al., 2011	Was the target/reference population clearly defined?	Unknown	Yes	Yes – was clearer than originally thought.

Swartout, 2012	Was the sample frame taken from an appropriate population base so that it closely represented the target/reference population under investigation?	No	Yes	No – University men represent a specific population within this literature due to their high perpetration rates (e.g. Finley & Corty, 1993)
Thompson et al., 2011	Was the sample frame taken from an appropriate population base so that it closely represented the target/reference population under investigation?	Unknown	Yes	Yes – evident that college men were the target population.
Warren et al., 2015	Was the sample frame taken from an appropriate population base so that it closely represented the target/reference population under investigation?	No	Yes	No – University men represent a specific population within this literature due to their high perpetration rates (e.g. Finley & Corty, 1993)

Thompson et al., 2011	Was the selection process likely to select subjects/participants that were representative of the target/reference population under investigation?	Unknown	Yes	Unknown – no details provided about data collection, though second researcher has alerted us to the fact that college men were the target population.
Warren et al., 2015	Was the selection process likely to select subjects/participants that were representative of the target/reference population under investigation?	No	Yes	Yes – second researcher again pointed out target population in title as "college men", so a University sample is appropriate.
Abbey et al., 2012	Were the risk factor and outcome variables measured appropriate to the aims of the study?	Yes	Unknown	Yes – second researcher unfamiliar with appropriate instruments.
Mouilso & Calhoun, 2008	Were the risk factor and outcome variables	Yes	Unknown	Yes – second researcher unfamiliar with

	measured appropriate to the aims of the study?			appropriate instruments.
Swartout, 2012	Were the risk factor and outcome variables measured appropriate to the aims of the study?	Yes	Unknown	Yes – second researcher unfamiliar with appropriate instruments.
Thompson et al., 2011	Were the risk factor and outcome variables measured appropriate to the aims of the study?	Yes	Unknown	Yes – second researcher unfamiliar with appropriate instruments.
Warren et al., 2015	Were the risk factor and outcome variables measured appropriate to the aims of the study?	Yes	Unknown	Yes – second researcher unfamiliar with appropriate instruments.
Abbey et al., 2012	Were the risk factor and outcome variables measured correctly using instruments/measurements that had been trialled, piloted, or published previously?	Yes	Unknown	Yes – second researcher unfamiliar with appropriate instruments.
Mouilso & Calhoun, 2008	Were the risk factor and outcome variables measured correctly using instruments/measurements that had been trialled,	Yes	Unknown	Yes – second researcher unfamiliar with appropriate instruments.

	piloted, or published previously?			
Swartout, 2012	Were the risk factor and outcome variables measured correctly using instruments/measurements that had been trialled, piloted, or published previously?	Yes	Unknown	Yes – second researcher unfamiliar with appropriate instruments.
Thompson et al., 2011	Were the risk factor and outcome variables measured correctly using instruments/measurements that had been trialled, piloted, or published previously?	Yes	Unknown	Yes – second researcher unfamiliar with appropriate instruments.
Warren et al., 2015	Were the risk factor and outcome variables measured correctly using instruments/measurements that had been trialled, piloted, or published previously?	Yes	Unknown	Yes – second researcher unfamiliar with appropriate instruments.
Warren et al., 2015	Were the methods sufficiently described to enable them to be repeated?	Unknown	Yes	Unknown – as this answer reflects lack of description of intentions in Methods section.

Thompson et al., 2011	Were the basic data adequately described?	No	Yes	No – second researcher had not noticed that a demographics questionnaire was administered, as it is only briefly mentioned.
Thompson et al., 2011	Were there any funding sources or conflicts of interest that may affect the authors' interpretation of the results?	No	Unknown	No – second researcher had missed disclosing of funding conflicts, as this is right at the end of the paper.
Thompson et al., 2011	Participants lost after consent, and why	Yes	No	Yes – second researcher missed this information on first read.