Rape myth acceptance: Investigating the dimensionality of the Illinois Rape Myth Acceptance Scale and the Male Rape Myth Scale

by
Demi Louise Geldenhuis

Thesis presented in fulfilment of the requirements for the degree of Master of Arts (Psychology) in the Faculty of Arts and Social Sciences at Stellenbosch University

Supervisor: Dr. Zuhayr Kafaar

December 2023

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that the reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

December 2023

Acknowledgements

Out of my entire thesis, this was probably the most difficult section to write. I lost count of how many times I started sobbing with overwhelming gratitude for the love, support, and encouragement I have received over the past four years.

There was a large part of my learning inspired by grief. Since starting on this journey I lost my beautiful Granny Helen and Oupa George and moved halfway across the world. Faced with the grief of loss and change, there were months at a time when I found myself unable to carry on researching rape myths. To keep momentum, I turned to technical skills such as coding and psychometry.

To my family, thank you for all the hugs, coffees, words of encouragement, and unwavering faith in me. There were many times when this thesis seemed too great, and felt I had bitten off more than I could chew.

To my supervisor, Dr. Kafaar, thank you for believing in me when I did not believe in myself.

Abstract

The measurement of rape myth acceptance has thus far been grounded in classical test theory, and rape myth acceptance scales have historically focused exclusively on either male or female rape myths. My research is an exploratory investigation into the dimensionality of the Illinois Rape Myth Acceptance Scale, the Male Rape Myth Scale, and the combined item pool of both measures. Using convenience sampling, I recruited 2,536 students from Stellenbosch University in South Africa. I conducted a series of dimensionality assessments under a bifactor analytic framework, to determine whether (a) one or both scales could be considered essentially unidimensional, and (b) whether both sets of items tap into the same underlying construct. The results of my dimensionality investigation indicate that both the Illinois Rape Myth Acceptance Scale and the Male Rape Myth scale can be treated as unidimensional under an item response theory framework, but not under a factor analytic framework. Furthermore, the male rape myths and female rape myths included in this study appear to tap into the same general rape myth acceptance dimension.

Key words: rape myth acceptance, bifactor modelling, dimensionality assessment, item response theory

Opsomming

Die meting van verkragtingsmite-aanvaarding is tot dusver gegrond binne klassieke toetsteorie, en verkragtingsmite-skale het histories uitsluitlik óf op manlike óf vroulike verkragtingsmites gefokus. My navorsing is ‘n verkennende ondersoek rondom die dimensionaliteit van die Illinois Rape Myth Acceptance Scale, die Male Rape Myth Scale, en die gekombineerde poel van items van beide meetinstrumente. By wyse van gerieflikheids-monsterneming het ek 2,536 studente van die Universiteit Stellenbosch in Suid-Afrika gewerf. Ek het ‘n reeks dimensionaliteitsassesserings onderneem binne ‘n bifaktor analitiese raamwerk om vas te stel of (a) een of albei skale basies as eendimensioneel beskou kan word, en (b) of beide stelle items op dieselfde onderliggende konstruk berus. Die resultate van my dimensionaliteitsondersoek dui aan dat beide die Illinois Rape Myth Acceptance Scale en die Male Rape Myth Scale binne ‘n itemresponsteorie-raamwerk as eendimensioneel behandel kan word, maar nie binne ‘n faktor-analitiese raamwerk nie. Verder berus die manlike verkragtingsmites en vroulike verkragtingsmites wat binne hierdie studie ingesluit is klaarblyklik op dieselfde algemene verkragtingsmite-aanvaardingsdimensie.

Sleutelwoorde: verkragtingsmite-aanvaarding, bifaktor-modellering, dimensionaliteitsassessering, itemrespons-teorie

Contents

List of Figures

List of Tables

Nomenclature

Acronyms and Abbreviations

CI	confidence interval
CTT	classical test theory
EFA	exploratory factor analysis
FA	factor analytic (framework)
GIRMA	Gender Inclusive Rape Myth Acceptance Scale
Hull-CAF	Hull method based on common parts accounted for
ILFA	item level factor analysis
IRMAS	Illinois Rape Myth Acceptance Scale
IRT	item response theory
MRMS	Male Rape Myth Scale
PA	parallel analysis
PA-MRFA	parallel analysis based on minimum rank factor analysis
PC	principal components
PCA	principal components analysis
PEBI	pure exploratory bifactor, bifactor modelling procedure
RMA	rape myth acceptance
RULS	robust unweighted least squares, factor estimation procedure
SL	Schid-Leiman, bifactor modelling procedure
SU	Stellenbosch University
ULS	unweighted least squares, factor estimation procedure

Metrics and indices

α	Cronbach's alpha
ω_h	omega hierarchical (OmegaH)
ω	McDonald's ordinal omega
X^2	chi-squared
a	slope/discrimination parameter in item response theory
λ	factor loading
AGFI	adjusted goodness of fit index.
ECV	explained common variance
GFI	goodness of fit index
GLB	greatest lower bound, reliability index
h^2	item communality
I-ECV	item-level explained common variance
I-REAL	item residual absolute loadings
I-UniCo	item-level unidimensional congruence
item-MSA	item-level measure of sampling adequacy.
KMO	Kaiser-Meyer-Olkin measure of sampling adequacy.
MIREAL	mean of item residual absolute loadings
RMSEA	root mean square error of approximation
SRMR	standardized root mean square residual
TLI	Tucker Lewis Index
u^2	item uniqueness
UniCo	unidimensional congruence
WRMR	weighted root mean square residual

CHAPTER 1: INTRODUCTION

South Africa is known to have a significantly high prevalence of rape. In the space of just one year, ranging from October 2021 to September 2022, there were 42,239 reported cases of rape (South Africa Police Service, 2021a, 2021b, 2022a, 2022b). Alarmingly, these figures likely only represent a fraction of the true number of rape cases, as rape is severely under-reported (Jewkes et al., 2010; Keehn et al., 2014). Sexual violence only reaches prolific proportions when it has the necessary support base to sustain it. In a society where sexual violence is considered a norm, perpetrators are not the only role players; the attitudes, beliefs, and actions of individuals contribute to a collective societal view that sees victims of sexual assault faced with more prejudice than their assailants and rape perpetrators not held accountable for their crimes (Beres et al., 2019).

My research focuses on the measurement of a particular contributor to the normalisation of sexual assault at the individual level: rape myth acceptance. *Rape myth acceptance* (RMA) can be understood as a general cognitive schema that actively guides information processing and organises how people interpret cases of rape (Eyssel & Bohner, 2011; Nyúl et al., 2018; Süssenbach et al., 2013). RMA essentially skews people's perceptions of who can be a victim, who can be a perpetrator, and under what circumstances rape can occur (Burt, 1980; Cherniawsky & Morrison, 2022; Gerger et al., 2007). The cognitive schema manifests as cognitive distortions called *rape myths* (Yapp & Quayle, 2018).

It is through rape myths that rape myth acceptance is measured. Established quantitative RMA measures consist of items that capture various rape myths, presented as statements, that respondents are asked to agree or disagree with to some extent. The exact rape myths used vary from measure to measure, but typical indicators of RMA include statements that (a) attribute a causal role or blame to victims, (b) excuse or justify the actions or motivations of rapists, (c) trivialise or minimise the scope or impact of rape, and (d) deny the possibility of rape under specific contexts or circumstances (Hine et al., 2021; Parratt & Pina, 2017; Payne et al., 1999; Turchik &

Edwards, 2012).

I limited the scope of my research to two RMA measures that have a psychometric foundation and are well-established in international literature. First, the Illinois Rape Myth Acceptance Scale (IRMAS; Payne et al., 1999) presents rape myth acceptance as multidimensional with seven domains. The IRMAS items centre exclusively on rape myths that pertain to scenarios where the victim is female, and the perpetrator is male. Second, the Male Rape Myth Scale (MRMS; Kerr Melanson, 1998), which was intended to be a unidimensional measure of rape myth acceptance. In contrast to the IRMAS, the MRMS item content centres around male rape victims.

For my research, I used the IRMAS and MRMS to explore two distinct avenues, each with the potential to expand on how rape myth acceptance is measured, particularly in the South African context. The first avenue concerned expansion in terms of measurement theory. Over the course of two dimensionality investigations, I explored whether the MRMS and IRMAS can be considered unidimensional for item response theory applications (IRT).

The second avenue concerned conceptual expansion. In terms of both theoretical discourse and practical data handling, researchers have historically treated rape myths anchored (explicitly or tacitly) to female victims as distinct from rape myths anchored to male victims (Hogge & Wang, 2022; Turchik & Edwards, 2012; Walfield, 2018). While the distinction makes sense at the item level, it has resulted in stratification at the global RMA measurement level, which I do not believe is adequately motivated by the existing literature. Therefore, I conducted a third dimensionality investigation, exploring the combined item pool of the IRMAS and MRMS to see whether their items were indicators of the same global rape myth acceptance construct.

The remainder of this chapter will consist of the following. I first elaborate on rape culture in South Africa, as it is integral to why I am passionate about my research. I also discuss why dimensionality investigations of rape myth acceptance measures are relevant for potential IRT applications. I then delve into the gendered nature of rape myth research and the stratification of global RMA measurement. I conclude this chapter with my research questions, hypotheses, and an overview of the coming chapters.

Rape Myth Acceptance and Rape Culture

Wide-scale acceptance of rape myths is one of the formative conditions for a rape culture (Beres et al., 2019; Lankster, 2019). *Rape culture* can be understood as the systematic dismissal of sexual violence, where victims receive more blame for their assault than perpetrators, and there is a general failure to take appropriate action against rapists (Beres et al., 2019).

When rape myths are deeply embedded in the attitudes of many everyday citizens, there can be layered implications contributing to rape culture. Some of the things that RMA can affect include; whether rape is recognized by both bystanders and victims (Beres et al., 2019; Wilson & Newins, 2019); whether rape is reported (Egan & Wilson, 2012); whether rape cases are handled appropriately by law enforcement officials (Du Plessis et al., 2009; S. E. Mgolozeli & Duma, 2020a); and the quality of support services offered to rape victims (Jina et al., 2013; Kassing & Prieto, 2003).

South African feminist legal studies have also highlighted how rape myths have been used as grounds for justification in court judgements to reduce the sentencing of perpetrators or to deny that a crime took place at all (Karimakwenda, 2021; Modri, 2014). I will elaborate on all these points further in my literature review.

In South Africa, there is an ever-growing frustration with the lack of progress made by the government, police and educational institutions to curb rampant sexual violence in the country (Fernando, 2019; Francke, 2019; Hartmann, 2019). In 2016, the #RUReferenceList and #EndRapeCulture campaigns were landmark student protests against the normalisation of sexual assault on university campuses that drew attention to rape culture in South Africa (Gouws, 2018; Macleod et al., 2018). The RU Reference List was a short list with only 11 names. The names listed were of students at Rhodes University who had allegedly committed acts of sexual violence with no recourse (Bashonga & Khuzwayo, 2017; Macleod et al., 2018). The list of names was called a reference list to highlight that the university appeared to have a more punitive stance towards plagiarism than sexual violence (Gouws, 2018; Macleod et al., 2018). The EndRapeCulture campaign followed shortly after, which led to a marked increase in discourses on rape, rape myths and rape culture on social media platforms such as Twitter and Facebook (Bashonga & Khuzwayo, 2017; Gouws, 2018; Orth et al., 2020).

I am interested in contributing to the improved measurement of rape myths, as I concur with Beres et al. (2019) that it is not enough to address sexual violence in and of itself. As has been so well vocalised and highlighted during the protests mentioned above, it is critically important to address rape culture and the collective attitudes and beliefs that enable the continued proliferation and normalisation of sexual assault.

Rape Myth Acceptance and Item Response Theory

Published studies investigating RMA as a predictor variable in the South African context have tended to rely on RMA measures created on an ad hoc basis. This ad hoc usage has manifested in one of two ways. In the first instance, researchers relied on a well-known RMA measure, but instead of using the entire measure, they only used a handful of the items (Jewkes et al., 2011; Kalichman et al., 2005). In the second instance, some researchers thought selecting items from a range of existing scales available to access a broader range of item content (Finchilescu & Dugard, 2021) would be more appropriate. Unfortunately, both practices result in unstandardised and widely varying measures, which makes it more difficult to draw meaningful comparisons between quantitative results (Fakunmoju et al., 2019).

To my knowledge, rape myth acceptance research (both international and local) has almost exclusively relied on summated scores and been grounded in classical test theory. However, considering how ad hoc rape myth items have been included in SA research studies, there could be a substantial long-term benefit in directing research focus to the item-level attributes of rape myth acceptance measures. Data analytic techniques available under an item response theory framework would be particularly well suited to understanding the qualities and attributes of individual items within a particular RMA measure.

In light of the numerous practical applications that unidimensional IRT modelling offers (which I elaborate on in my literature review), I believe it is worth investigating whether the IRMAS and the MRMS can be treated as unidimensional. Knowing more about the dimensionality of the IRMAS and the MRMS when administered in South Africa could potentially give South African RMA researchers the confidence or assurance needed to utilise IRT as an alternative measurement theory.

Male Rape Myths and Female Rape Myths

A well-vocalised criticism of RMA research is the mainstream tendency to focus exclusively on female victims and male perpetrators (Maxwell & Scott, 2014; Turchik & Edwards, 2012). One of the earliest definitions of rape myths describes them as “attitudes and beliefs that are generally false but widely and persistently help, and that serve to deny and justify male sexual aggression against women” (Lonsway & Fitzgerald, 1994, p. 134). However, contrary to what was originally believed, the impacts of rape myths are not experienced exclusively by women (Maxwell & Scott, 2014; Turchik & Edwards, 2012; Urban & Porras Pyland, 2022).

Walfield (2018, p. 5) has highlighted that research on "rape myths for male victims" was conducted separately and parallel to the mainstream RMA literature that focused on rape myths anchored to female rape victims. The extent of the stratification is made clear in the recent work of Hogge and Wang (2022, p. 422), who define male rape myths as "stereotypes and false beliefs about rape in instances when the victim is male.".

Highly gendered understandings of rape myth acceptance and rape myths served as a point of departure which has had far-reaching consequences. Initially, scales were developed based on the premise that the only indicators of RMA are rape myths that posit women as victims and men as perpetrators, referred to as *female rape myths* (Walfield, 2018). The most commonly cited measures in RMA literature are the Rape Myth Acceptance Scale (Burt, 1980), the Illinois Rape Myth Acceptance Scale (Payne et al., 1999), and the Updated Rape Myth Scale (McMahon & Farmer, 2011), which is a revised version of the Illinois Rape Myth Acceptance Scale. All three assume references to male rape victims are unnecessary to measure rape myth acceptance well.

There are at least three scales that extended the measurement of RMA to include male victims by exclusively relying on *male rape myths*. The assumption made by researchers was that RMA that prejudices male victims could be measured by ensuring the rape myths (the indicators) were male-centric. Currently, there appear to be three such measures with a psychometric basis: the Male Rape Myth Scale (Kerr Melanson, 1998), the Male Rape Myth Acceptance Scale (Hine et al., 2021), and the Male Rape Myths Scale Revised (Hogge & Wang, 2022). In all three cases, the developers assumed that accepting male rape myths differed from the mainstream RMA measurement that focused exclusively on female victims.

Consequently, whether female rape myths and male rape myths measure the same underlying RMA construct has historically been left up to interpretation. The ambiguity is a red flag for potential construct proliferation and biased parameter estimates (Hughes, 2018). Therefore, I intend to contribute to RMA literature by following a suggestion made by both Davies et al. (2012) and Walfield (2018) to investigate whether female rape myths and male rape myths are indeed tapping into the same construct.

Before moving on, I would like to acknowledge that in the past two years there have been several attempts to approach RMA measurement in a more gender-inclusive way (Canan et al., 2023; Johnson et al., 2023; Urban & Porras Pyland, 2022). These developments emerged well after I had planned my research and collected my data, but are nonetheless relevant to my research question, and are discussed in more detail in my literature review.

Research Rationale and Research Questions

Rape-supportive attitudes, including RMA, should be practically investigated as a matter of urgency. Despite its conceptual relevance, South Africa has a shortage of quantitative RMA research. Fakunmoju et al. (2019) suggest that a way to stimulate research is to start with investigating existing measures.

Aims

My research was exploratory, and the main aim was to investigate the dimensionality of rape myth acceptance measures in the South African context. For this exploratory investigation, I focused on two RMA measures that have been well-established in international literature, namely the Illinois Rape Myth Acceptance Scale (IRMAS; Payne et al., 1999) and the Male Rape Myth Scale (MRMS; Kerr Melanson, 1998). Although both scales claim to measure rape myth acceptance, the IRMAS and MRMS represent two distinct research streams and are sometimes treated as measuring separate constructs. My aims were as follows:

- **Aim 1:** Determine whether the MRMS can be considered unidimensional under an item response theory framework.
- **Aim 2:** Determine whether the IRMAS can be considered unidimensional under an item response theory framework.
- **Aim 3:** Determine whether the IRMAS and MRMS items may be tapping into the same underlying construct.

Rape myth acceptance has primarily been investigated using traditional psychometric methods and classical test theory. However, the limited South African quantitative rape myth acceptance literature highlights the need for flexible and adaptive measurement. Item response theory is a powerful form of psychometric analysis that, to my knowledge, has not yet been applied to the measurement of rape myth acceptance. Therefore, I am attempting to open up the field by investigating whether it is feasible to go beyond modelling summated scale scores and instead apply IRT in rape myth acceptance measurement.

Exploring whether male rape myths and female rape myths may be indicators of the same construct has practical relevance. Researchers interested in developing a new RMA measure in the South African context would benefit from knowing whether to include items spanning both male and female rape myth literature in the initial item pool. Additionally, if male rape myths and female rape myths tap into the same construct, comparing male rape myth and female rape myth scores could be possible.

Research Questions

By thoroughly investigating the dimensionality and general factor strength of the IRMAS and the MRMS individually, as well as examining their combined item pool, it was possible to test the following three research questions:

- **Research Question 1:** Can the MRMS be treated as essentially unidimensional when administered to a student sample in South Africa?
- **Research Question 2:** Can the IRMAS be treated as essentially unidimensional when administered to a student sample in South Africa?
- **Research Question 3:** Is it plausible that the indicators of the IRMAS and MRMS tap into the same underlying construct?

It is important to ground an investigation in practical considerations. As noted by Walfield (2018), student sampling has characterised RMA research. I posit that, given the psychometric focus of my study and the urgent need for more research in this field in the South African context, it is important to know how the IRMAS and the MRMS operate within the sample type where they are most likely to be used in the near future, viz., student populations.

Hypotheses

The theoretical framework and necessary psychometric basis for investigating and answering these two research questions will be explained in more detail in my literature review in Chapter 2. I present my hypotheses here as follows:

- **$H1_1$:** The MRMS can be treated as essentially unidimensional.
- **$H1_0$:** The MRMS cannot be treated as essentially unidimensional.
- **$H2_1$:** The IRMAS can be treated as essentially unidimensional.
- **$H2_0$:** The IRMAS cannot be treated as essentially unidimensional.
- **$H3_1$:** The items of the IRMAS and MRMS tap into the same RMA construct.
- **$H3_0$:** The items of the IRMAS and MRMS do not tap into the same RMA construct.

Rape myth acceptance is a construct with interdisciplinary relevance, and there is much to be learned about rape myths in the South African context (Abrahams et al., 2013).

By investigating the feasibility of an IRT approach to RMA, as well as whether male rape myths and female rape myths are indicators of the same construct, I hope that my research can offer a more informed point of departure for quantitative RMA research in the South African context; particularly with regards to factor analytic and item response theory applications. I hope to enable future researchers to circumvent some limitations and circular findings that have plagued the international scientific literature thus far. May this bring us closer to understanding rape myth acceptance and how it can be optimally measured in the general South African population.

Chapter Overview

Following this introductory chapter, the remainder of my thesis consists of the following:

- **Chapter 2:** Literature Review. The literature review is presented in two sections. The first section covers rape myth acceptance literature, and the second covers psychometric literature relevant to my research questions.
- **Chapter 3:** Methodology. In my methodology chapter, I cover my sampling strategy, research procedure, and ethical considerations and provide a detailed breakdown of the data analysis phase of my research.
- **Chapter 4:** Results and Data Analysis. In this chapter, I report the results of three separate dimensionality investigations. The dimensionality investigations cover the (1) MRMS, (2) the IRMAS and (3) the combined item pool of the IRMAS and MRMS.
- **Chapter 5:** Discussion. In this chapter, I discuss whether my hypotheses were supported or disproved by the results of my investigation. This chapter also includes recommendations for future researchers and the study limitations.

CHAPTER 2: LITERATURE REVIEW

My research focuses on the measurement of rape myth acceptance. There were two significant components I needed to research to make my investigation possible. The first component was rape myth acceptance literature. The second component was psychometric literature pertaining to dimensionality investigations and bifactor modelling. Therefore, the following literature review is presented in two sections. In Section 1, my focus is rape myth acceptance (RMA) literature. I have limited my focus to literature that captures the overlap between male rape myths and female rape myths and links the impacts and effects of rape myths and rape myth acceptance to the South African context. In Section 2, my focus is on psychometric literature relevant to conducting my dimensionality investigations.

Section 1: Rape Myth Acceptance Literature

For section 1, my interest is in presenting rape myth acceptance literature in a way that clarifies how this research is relevant in South Africa. I have focused on themes that allow me to cover a broad range of rape myths and situate the impacts of rape myth acceptance within the South African context. The two main themes are (1) types of rape myths and (2) the impacts and effects of rape myth acceptance.

Throughout this section, I will highlight various ways rape myths can manifest. As noted by Vetten (2017), the mere fact that rape myths exist does not give them power. Instead, rape myths draw on various processes, strategies and techniques that transform them from unfounded claims into persuasive and believable statements that appear objective and impartial (Karimakwenda, 2021; Vetten, 2017). Rape myths can therefore serve multiple functions and have various implications. Although research has shown that rape myths vary to the extent that they reflect the cultural norms and values of individual societies (Karimakwenda, 2021; Lee et al., 2010), there is a consensus in research that rape myths are relied on to blame rape victims for what happened to them, justify the actions of perpetrators, trivialise the violence of rape itself, or deny rape occurred at all (Chapleau et al., 2007; Turchik & Edwards, 2012).

I have grouped the literature into three main themes: (a) rape myths that narrow the scope of rape and consent, (b) rape myths that attempt to shift or reallocate blame away from perpetrators and towards victims, and (c) rape myths that minimise or trivialise rape or its impacts in some way. Interwoven into these themes are key concepts, specific rape myth examples, and the impacts of rape myths and rape myth acceptance highlighted by South African researchers across various disciplines. Many of these themes are interlinked, and some rape myths may feature more than once under different headings. By the end of this section, a broad range of rape myths and their effects will be clear.

Rape Myth Acceptance and Narrow Definitions of Rape and Consent

When people have high rape myth acceptance, they are more likely to have a narrow understanding of rape and consent. Narrow understandings of rape and consent can have far-reaching impacts, such as affecting whether victims and members of the criminal justice system recognise rape. In the following section, I will elaborate on how and why rape myths can contribute to narrow subjective understandings of rape and consent.

The Legal Definition of Rape in South Africa. In South Africa, the Criminal Law (Sexual Offences and Related Matters) Amendment Act 32 of 2007 defines rape as an offence that occurs when any person unlawfully and intentionally commits an act of sexual penetration with another without the latter's consent (Ngubane-Mokiwa & Chisale, 2019). The Act repealed the common law offences of rape and indecent assault. It replaced them with a broader range of statutory offences, defined victims and perpetrators in gender-neutral terms and created a hierarchical structure of sexual offences (Naylor, 2008). The term 'sexual penetration' replaced the term 'vaginal penetration', with the effect that forced anal penetration now also constitutes rape, and rape is no longer a crime that can be committed only against women (Naylor, 2008).

The amendment also redefined/broadened the scope of the word *penetration*, which was previously limited to penile penetration but now extends to penetration by an object. The implication is that potential perpetrators are no longer limited to men (Naylor, 2008; Ngubane-Mokiwa & Chisale, 2019). However, it is problematic that rape is still defined in terms of penetration because it limits the perpetrator to one who penetrates and the victim to one who is penetrated (Lowe & Rogers, 2017; Pearson & Barker, 2018). In other words, if someone with a penis was forced to penetrate another, the former would not be recognised as a rape victim.

Consent. A major criticism of the legal definition of rape in South Africa is that it requires the prosecution to prove a lack of consent, which places the focus in legal cases on the behaviour and reaction of the victim instead of the perpetrator (Adoch, 2022; Modri, 2014; Naylor, 2008). Rephrased, legal actors place an undue amount of focus on victim behaviour to determine whether rape instead of non-criminal sex took place (Adoch, 2022; Modri, 2014).

Various rape myths capture biased ideas of what constitutes consent, which contributes to narrowing the scope of rape. On the one hand are myths that wrongly infer consent, such as rape myths that conflate assumed pleasure with consent. For example, an item from the Male Rape Myth Acceptance scale is "Even if force is used to initiate sex, the victim's erection can be interpreted as pleasure" (Hine et al., 2021, p. 12). As a contrasting example, an item from the Illinois Rape Myth Acceptance Scale is "Many women actually enjoy sex after the guy uses a little force." (Payne et al., 1999, p. 50). Such rape myths are based on the problematic assumptions that (a) men always want and are ready for sex and (b) women secretly desire to be forced into sex.

However, assumed pleasure and involuntary physiological responses should not be conflated with active and willing consent. According to McLean (2013), erection and ejaculation can occur even under extreme duress and do not prove that someone consented to or wanted a given sexual interaction. Similarly, Levin and Van Berlo (2004) have highlighted that sexual organs respond to physical stimulation and genital arousal cannot and should not be conflated with consent. Such a physiological response can be very distressing for victims, as it can fuel the misbelief that they must have been willing participants in their assault (Levin & Van Berlo, 2004; Stern et al., 2015).

On the other hand are rape myths that invalidate expressions of non-consent or prescribe that a lack of consent should be expressed in a particular way. For example, item 5 from the Updated Rape Myth Acceptance scale is “When girls are raped, it’s often because the way they said ’no’ was unclear,” (McMahon & Farmer, 2011, p. 77). As noted by Naylor (2008), it is not uncommon for women who verbalise that they do not consent to be seen as having consented. A particular contributor to this idea that ’no’ can mean ’yes’ is the idea of *token resistance*, the belief that women reject sexual advances while intending to engage in them (Setia et al., 2020; Shafer et al., 2018). Modri (2014) highlights that South African judges have referred to victims’ lack of physical resistance as implied consent, drawing on rape myths in the delivered judgments. A similar example from male rape myth literature is an item from the Male Rape Myth Acceptance Scale, “A man who has been raped did not set sexual limits understood by the perpetrator” (Hine et al., 2021, p. 12).

Rape myths that blur the lines of what constitutes consent (1) introduce the idea that rape can be the unfortunate, unintended result of a miscommunication, and (2) are used to shift blame to victims for not setting clear enough boundaries. As consent is such a crucial part of legal proceedings, when judges rely on rape myths to justify inferring consent or invalidating victims’ experiences, it can mean the difference between whether alleged rapists are convicted for their crimes and whether convicted rapists serve a full or mitigated sentence (Karimakwenda, 2021; Modri, 2014; Vetten, 2017).

Rape Scripts. While legal definitions of rape can be broad, the subjective definitions of rape held by everyday people can be much narrower by comparison. For example, in a study by Adams-Clark and Chrisler (2018), participants were less likely to recognize a vignette was referring to a rape scenario when it detailed forced oral sex or non-penile penetration (e.g., with fingers or an object). Adams-Clark and Chrisler suggest this may have been because only penile-vaginal penetration fits traditional notions of rape and sex. Thus, participants who overlooked oral rape and non-penile penetration likely had a narrower subjective definition of rape.

One explanation for why RMA results in narrow definitions of rape is because rape myths reinforce and are reinforced by *rape scripts* (Davies et al., 2013; Peterson & Muehlenhard, 2004). According to Peterson and Muehlenhard (2004), people hold stereotypic rape scripts based on their impressions of what typically occurs during rape. For example, a rape script might involve a young woman walking alone late at night on a dark-lit side street, ambushed by a stranger and raped at knifepoint while putting up an intense struggle. Such a rape script would reinforce and be reinforced by rape myths that prescribe or limit rape to only occurring within arbitrary contexts, under certain conditions or between specific parties. Such a rape script would arguably reinforce and be reinforced by rape myths such as "If the rapist does not have a weapon, you really can't call it rape" and "A rape probably didn't happen if the woman has no bruises or marks" (Payne et al., 1999, p. 49).

Norton and Grant (2008) support the association between stereotypic rape scripts and rape myths with the finding that, when accounts of rape are inconsistent with rape stereotypes, accounts may seem less credible in the eyes of investigators and judges with high RMA. Karimakwenda (2021) noted that South African courts still look to evidence of physical injury to determine the severity of rape instead of consistently recognising rape as inherently violent. This is confirmed by Vetten (2017), who notes that lack of physical injury to victims continues to be used by South African courts as a reason to reduce the sentences of perpetrators.

Rape Acknowledgment. Rape acknowledgement is a distinct but related field of study. In a meta-analysis by Wilson and Miller (2016), 60% of a sample of 5,917 women drawn from 28 studies met the criteria for being an unacknowledged rape victim. An *unacknowledged rape victim* refers to someone who does not see or conceptualise themself as a rape victim, despite having experienced what would legally qualify as rape (Koss, 1985). An example provided by Wilson and Miller (2016, p. 149) is that people may have an experience that fits the definition of rape, but instead label it as a "miscommunication" or "bad sex".

RMA's inter-relatedness with rape acknowledgement is evidenced by Wilson et al. (2018), who argue that survivor outcomes are more interpretable when differences in RMA are considered. Furthermore, in a study by Reed et al. (2020), students who fit the criteria for having experienced rape had higher odds of being an unacknowledged victim if they were male, which the authors attributed to greater rape myth acceptance among men.

I now provide an example of how rape myths that narrow the definition of rape can impact rape acknowledgement among South Africans. To the hegemonic South African masculinity, male rape contradicts what is expected of men, with a victim status seen as emasculating, disempowering and weak (Ngubane-Mokiwa & Chisale, 2019). Rape myths that play directly into this masculine ideology are the myths that (a) men cannot be raped and, more commonly, that (b) women cannot sexually assault men (Pearson & Barker, 2018; Turchik & Edwards, 2012; Walfield, 2018). The two rape myths work together to render male rape invisible and unrecognisable by denying its very possibility (Hine et al., 2021; Javaid, 2019).

Rape myths that deny the possibility that women can sexually assault men can and do result in male victims remaining unaware that they have been raped. For example, Ngubane-Mokiwa and Chisale (2019) looked at how a group of disabled Zulu-speaking men perceived sexual interactions between them and their female caregivers (whom their communities had identified as predatory). Many of the participants held the belief that women cannot rape men, and even though participants described forceful sexual encounters, the experiences were perceived as being sexual favours (Ngubane-Mokiwa & Chisale, 2019). When rape is not recognised, people are vulnerable to repeated victimisation, and perpetrators are not held accountable for their crimes.

Rape Myth Acceptance and Blame Attribution

One of the critical components of rape myth acceptance is that it involves a degree of blame redistribution. Many rape myths allocate blame to the victim or remove responsibility from the perpetrator by justifying the latter's actions. In the following section, I will elaborate on how victim blaming and rape defending can manifest in rape myths.

Victim Blaming. A strong relationship between victim blaming and rape myth acceptance has been highlighted consistently by researchers (Adolfsson et al., 2020; Davies & Rogers, 2006; Grubb & Turner, 2012). *Victim blaming* is a seemingly counter-intuitive response, which entails blaming victims for 'allowing' the rape or even holding them directly responsible for the assault (Grubb & Turner, 2012). Victim blaming is a form of secondary victimisation that causes victims to internalise feelings of guilt and shame (Du Plessis et al., 2009; Grubb & Turner, 2012; van der Bruggen & Grubb, 2014). Charly and Reddy (2019) highlight that rape myth acceptance is one of the processes underlying victim blaming by legal professionals, indirectly influencing their decision-making processes.

Rape myths that lend themselves to victim blaming tend to be focused on how a victim behaved or dressed before their assault. For example, in female rape myth literature, victim blaming has been highlighted in rape myths that suggest women can attract or provoke rapists by behaving promiscuously or dressing provocatively (Ngubane et al., 2022; Selepe et al., 2021; Swemmer, 2019). According to Cherniawsky and Morrison (2022), the idea that women should dress or behave a certain way to avoid being targeted may have been unintentionally reinforced by well-intentioned prevention tips and advice spread through word-of-mouth, social media and dedicated rape resources and interventions. For example, common tips involve taking certain preventive measures, such as only walking around in daylight and carrying mace spray. As argued by Cherniawsky and Morrison, the proliferation of safety tips and advice that only posits women as potential victims may have had the unintended effect of reinforcing the ideas that (a) rape is inevitable and yet simultaneously that (b) the onus is on individuals to prevent themselves from being raped.

Victim blaming is also particularly salient in rape myths that focus on how victims react when faced with the threat of sexual assault. For example, an item from the Male Rape Myth Scale is "Any healthy man can successfully resist a rapist if he really wants to" (Kerr Melanson, 1998, p. 61). When people believe that masculinity is defined by physical strength, it creates an expectation that men should always be able to fight back against and ward off potential rapists (Javaid, 2018; Mkhize & Sibanyoni, 2019; Spruin, 2018). This expectation aligns with the hegemonic masculinity in South

Africa, according to which men are always expected to be strong and physically dominant (Morrell et al., 2012; Ngubane-Mokiwa & Chisale, 2019).

In a recent study, Jina et al. (2020) analysed a nationally representative sample of male rape cases in South Africa to understand more about the causes, prevalence, and general reporting of male rape. Based on 209 case dockets sampled from the South African Police Service, Jina et al. reported that 84% of the victims attempted to fight off or escape their attacker. However, it is likely a larger percentage of men do not resist their attackers as, according to Javaid (2018), men who do not physically resist their attack are less likely to file a police report or seek health care out of fears that their masculinity will be questioned. Indeed, Fisher and Pina (2013) have suggested that the under-reporting of male rape may be even higher than that of female rape.

In the previous section, I highlighted that a lack of resistance has been used in courts to deny rape occurred. In this section, rape myths that focus on the degree of resistance have also been presented as shifting blame to victims for failing to prevent their assault. I want to clarify that someone's resistance level should not be used as an excuse to shift the blame to victims or as a criterion for whether a rape occurred. A common reaction to sexual assault is for victims of any gender to go into a state of tonic immobility, a form of trauma-induced paralysis where the body freezes involuntarily, and the mind shuts down (Kalaf et al., 2015; S. E. Mgolozeli & Duma, 2020b). Even if someone does fight their attacker, it is possible to be overpowered if the perpetrator is stronger, if confronted with multiple attackers at once, or if restrained (S. E. Mgolozeli & Duma, 2020b). Furthermore, if someone of any gender is drugged or inebriated, this can leave them incapacitated and unable to resist rapists regardless of their physical strength or whether the rapist has a weapon (Fisher & Pina, 2013; Stern et al., 2015).

When rape myths characterised by victim blaming are internalised, it can lead to victims feeling personally responsible for their attack (Lowe & Rogers, 2017; S. E. Mgolozeli & Duma, 2020b). According to Jina et al. (2020), the first three days are crucial for reporting. However, victim blaming has a genuine effect on the willingness of victims to report rape to the police. In a study by Egan and Wilson

(2012), rape victims who had not reported their rape to the police had significantly higher levels of RMA than those who had reported it. This is in line with previous research that has suggested survivors of rape who endorse and believe rape myths are far less likely to report the crime, either due to blaming themselves or to avoid the shame, stigma, and blame that they may have to endure from others (Dartnall & Jewkes, 2013; Lowe & Rogers, 2017). Victims can also experience pressure from perpetrators and community members to keep quiet, which further contributes to under-reporting (Fleming & Kruger, 2013; Makongoza & Nduna, 2017).

According to Hine and Murphy (2019), police officers with high rape myth acceptance are more likely to blame victims and less likely to see cases as worth investigating further unless the case details conform to their existing biases. In a study by Mkhize and Sibanyoni (2019), a sample of gay and lesbian university students were asked about their experiences reporting rape to the South African Police Service. Most participants did not even attempt to report to the police due to a lack of faith in the Criminal Justice System. Those who tried to file a report recounted being laughed at by the police, being asked derogatory questions, and feeling neither protected nor supported. South African researchers continue to highlight inappropriate case handling by South African police as a significant barrier to reporting rape and sexual assault (Du Plessis et al., 2009; Keehn et al., 2014; S. E. Mgolozeli & Duma, 2020a).

Rape Defending. Rape myth acceptance research has primarily focused on rape myths anchored to victims. However, most RMA measures also include items with a perpetrator focus (Hine et al., 2021; McMahon & Farmer, 2011; Urban & Porras Pyland, 2022). Amongst items with a perpetrator focus are various rape myths that read as defending, justifying, or excusing the actions or intentions of the perpetrator. For example, the myth that men can become so overwhelmed by uncontrollable sexual impulses that they can rape someone unintentionally (Javaid, 2019; Selepe et al., 2021).

The myth that rape can result from a man's sex drive spiraling out of control goes hand in hand with victim-blaming rape myths that suggest women can provoke rapists by dressing or acting promiscuously (Ngubane et al., 2022). Selepe et al. (2021) interviewed incarcerated rapists and asked them to account for/explain their actions. Many men relied on rape myths, saying they could not help themselves and that they were provoked by women who dressed/acted in a certain way. Both kinds of rape myths attribute the cause of rape outside the scope of a rapist's intentions, instead suggesting that (a) rape is provoked and (b) rape is an inevitable reality of the supposed inherently hyper-sexual nature of men (Ngubane et al., 2022; Selepe et al., 2021). However, in a study by Jewkes et al. (2010), rapists admitted to targeting women who dressed or acted promiscuously not because they were aroused but because they felt angry at the women for defying the gender roles and sexist expectations the rapists held.

The reality is that perpetrator motivations are complex and widely varying. Some rape motivations uncovered in South African research studies include boredom (Jewkes et al., 2010), and a desire to assert masculinity or power through violence (Selepe et al., 2021), which includes retributive rape by women (S. Mgolozeli & Duma, 2019). A highly problematic belief noted among rapists is a sense of sexual entitlement, according to which rapists feel they are owed sex by women, particularly women they have relationships with or have bought things for (Jewkes et al., 2011; Ngubane et al., 2022; Selepe et al., 2021). A study by Lankster (2019) provides evidence that male sexual entitlement is instilled from a young age, and the normalisation of rape is evident even in the discourses of adolescents. Lankster conducted group interviews with roughly 260 school-going South African males, where participants discussed two rape vignettes in depth. Lankster found that across both vignettes, rape was seen as a "viable option when males are presented with sexual opposition" (p. 139). Furthermore, Lankster noted that victim-blaming rape myths were particularly prevalent in the group discussions.

A final motivation I would like to highlight is the desire to punish others for behaving or even simply existing in a way that the interviewed rapists found disagreeable or offensive (Jewkes et al., 2015; S. Mgolozeli & Duma, 2019; Selepe et al., 2021). Examples of this include (1) *corrective rape*, where rapists target people based on their sexual orientation or gender expression out of hatred and intolerance, particularly lesbian women (Mayeza, 2022; S. Mgolozeli & Duma, 2019), and (2) *streamlining*, a prominent motivation in instances of gang rape where peers and strangers come together to rape girlfriends and partners seen as unfaithful or non-compliant (Jewkes et al., 2015; Jewkes et al., 2010).

By now it should be evident that blame attribution forms a key component of RMA, as many rape myths involve shifting away from perpetrators and towards victims. Two fundamental theories that attempt to explain how and why people assign blame in cases of rape are the *defensive attribution hypothesis* and the *just world theory* (Cherniawsky & Morrison, 2022; Russell & Hand, 2017).

According to the defensive attribution hypothesis, when making sense of a negative experience or situation, people are less likely to attribute blame to someone they perceive as being similar to themselves and more likely to allocate blame to someone they perceive to be dissimilar (Landström et al., 2016; van der Bruggen & Grubb, 2014). The implication of this theory in cases of rape is that people who identify more with a perpetrator or the perpetrator's position would attribute less blame to the victim to defend against being helpless lest they find themselves accused of a similar crime (van der Bruggen & Grubb, 2014). In a study investigating the relationship between rape myths and the defensive attribution hypothesis, Kahn et al. (2011) found that the more participants identified with the assailants in a set of rape scenarios, the less blame they attributed to the assailants, and the more blame they placed on the victims.

By contrast, just world theory proposes that people want to believe the world is a fair and just place and, therefore, may search for ways to blame victims of sexual assault to preserve their belief that people only get what they deserve (Adolfsson et al., 2020; Landström et al., 2016; van der Bruggen & Grubb, 2014). According to this perspective, attributing blame towards a victim enables observers to maintain the erroneous belief that they can prevent being raped purely through their own behaviour.

Despite their clear theoretical relevance, neither defensive attribution nor belief in a just world has been consistently found to be a predictor of rape myth acceptance as a whole (Cherniawsky & Morrison, 2022; Egan & Wilson, 2012; Vonderhaar & Carmody, 2015). However, this may be because while blame attribution is a key component of rape myth acceptance, not all rape myths entail assigning blame to a victim or exonerating a perpetrator. As I will now discuss, many rape myths pertain to the context or consequences of sexual assault and do not necessarily entail blame attribution.

Rape Myth Acceptance and Minimisation of Rape and its Impacts

In this next section, I will cover various ways in which rape myths and rape myth acceptance attempt to minimise rape and its impacts. Based on the literature I have reviewed, this minimisation can be captured under two separate but related themes: trivialisation, and othering.

Trivialisation. In a study by S. E. Mgolozeli and Duma (2020b), 11 South African men narrated how they had been sexually violated and described the experience as "forceful, unwanted, painful and disgusting" (p. 4). Furthermore, the participants described rape as a form of torture. Yet, RMA literature has shown that people can believe rape is not as severe as other crimes or that victims exaggerate how much they are affected by rape to get attention (Leverick, 2020).

Rape is a highly traumatic experience for rape victims, where their bodies and sense of self are violated (Pretorius, 2009). Physically, victims can have internal and external injuries (Jina et al., 2020) and suffer from sexual dysfunction and impotence in the long term (Fisher & Pina, 2013). In addition, South African rape victims have the added fear of contracting HIV/AIDS (Kalichman et al., 2007; S. E. Mgolozeli & Duma, 2020b; Stern et al., 2015). Psychological effects include depression, anxiety, post-traumatic stress, antisocial behaviour, substance abuse, self-harm and suicidality (Emezue & Udmuangpia, 2022; Young et al., 2016).

Writing on rape myths in legal discourse, Karimakwenda (2021) argues that rape myths function as neutralisation techniques that dehumanise victims and deny the inherent violence of rape. Specific rape myth examples from the literature that highlight this trivialisation include “Male victims of rape have very little emotional trauma to cope with” (Hine et al., 2021, p. 12) and “Women tend to exaggerate how much rape affects them.” (Payne et al., 1999, p. 49).

While these are examples of myths where the trivialisation of rape is explicit, rape myth acceptance comes with a degree of denial surrounding the severity of rape and its consequences. Many rape myth acceptance scales include statements that suggest rape is often fabricated, lied about or weaponised, especially by women (Grubb & Turner, 2012; Leverick, 2020). Such rape myths pertain to the validity and weight of the average rape allegation. Unfortunately, many legal actors, police officers, and everyday citizens believe that false rape allegations are far more common than they are (Dewald & Lorenz, 2022; Leverick, 2020; Stabile et al., 2019). In a study by Walfield (2018), 1,220 participants were asked to indicate how prevalent they believed false reports of rape to be. A tenth of the participants indicated they believed over 20% of reported male rape are false, and a fifth indicated that over 20% of reported female rape cases are false. However, researchers estimate that only 2%-10% of rape allegations are false, with the upper limit being a conservative overestimation (Dewald & Lorenz, 2022; Lisak et al., 2010; Norton & Grant, 2008).

As highlighted by Javaid (2019), voluntary agencies are vital to supporting rape victims and assisting with referrals to appropriate services, especially when victims do not want to go to the police. However, rape myth acceptance has a very real impact on both help-seeking behaviour and the quality of support services available to rape victims (Javaid, 2019; Jina et al., 2013).

Victims who accept rape myths are less likely to seek help and support, which can have devastating effects on mental health as victims try to navigate their trauma alone (Abrahams et al., 2013; Fleming & Kruger, 2013). Furthermore, researchers have noted that especially when training is lacking, counsellors and related health professionals can implicitly reinforce rape myths, such as those which posit men as being less traumatised by rape than women (Emezue & Udmuangpia, 2022; Javaid, 2017; Kassing & Prieto, 2003).

The impact of rape myth acceptance on the quality of support services is clearly illustrated in a study by Jina et al. (2013), who discovered that many South African healthcare providers were highly confident about their ability to deliver post-rape care, despite their actual knowledge being severely lacking. According to Jina et al., participants with the greatest post-rape care knowledge had the lowest rape myth acceptance scores.

Othering. In a study by Dosekun (2013), 15 women who had never experienced rape were asked to comment on the rape crisis in South Africa. Dosekun highlights that the participants effectively tried to 'other' rape and relied on rape myths to construct ideas of 'typical' victims and perpetrators as less educated and belonging to a lower socio-economic class or a different racial or cultural group. Such thinking essentially minimises the scope of rape by relegating it to a particular set of circumstances, contexts and people. As argued by Karimakwenda (2021, p. 380), "[...]rape myth discourse stems from retrogressive perceptions of race, rural living, and culture in South Africa."

A significant body of international literature suggests that high RMA is strongly associated with other oppressive belief systems, such as religious intolerance, ageism, classism and racism (Aosved & Long, 2006; Suarez & Gadalla, 2010). There is also a strong link between rape myth acceptance and hostile and benevolent sexism (Chapleau et al., 2007, 2008; Obierefu & Ojedokun, 2017). Furthermore, correlation studies have consistently highlighted a strong association between RMA and various attitudes relating to sexual relations, such adversarial sexual beliefs, and sex-role stereotyping (Barnett et al., 2017; Lee et al., 2010; Shafer et al., 2018). Gender-role conflict and homophobia have also been highlighted as significant RMA correlates in male rape myth literature specifically (Hine et al., 2021; Kassing et al., 2005; Walfield, 2018). On a side note, according to a systematic review by Yapp and Quayle (2018), RMA has been shown to predict rape proclivity and sexual violence; however, this predictive power has only been established post-perpetration.

Specific rape myths that arguably serve to either socially or spatially distance rape have featured in both male rape myth literature and female rape myth literature.

Many of these rape myths arguably tie into rape scripts, which I covered in an earlier section. For example, the IRMAS contains an entire subscale of rape myths items that suggest rape is a deviant event that only happens on the fringes of society (Payne et al., 1999). The subscale Rape is a Deviant Event includes statements such as “Rape almost never happens in the woman’s own home” and “Rape mainly occurs on the ‘bad’ side of town” (Payne et al., 1999, pp. 49 50). It is, however, important to note that many rape myths are inherently geared towards the othering of rape victims, whether the othering is explicit from the myth content or in the way that rape myths are wielded in a particular setting. For example, as argued by Swemmer (2019), when the South African judiciary relies on victim-blaming rape myths in sentencing, it reinforces rape myths in a way that extends well beyond the courtroom, which can and does result in victims becoming isolated from and abandoned by their communities.

Another prime example of an othering rape myth is the myth that male rape only happens in prisons (Javaid, 2019; Turchik & Edwards, 2012). The research on student male rape victims alone is enough to refute this myth, confirming that rapists most certainly operate outside of prisons (Mkhize & Sibanyoni, 2019; Reed et al., 2020). Interestingly, in South Africa, most reported male rapes are rapes that occurred in penal institutions (Jina et al., 2020). This may be because, as uncovered by S. E. Mgolozeli and Duma (2020a), men are more likely to report rape in an institutional setting due to the fear of becoming a target to additional inmates or of being re-victimised.

In male rape myth literature, various rape myths centre around the sexual orientation of the victim or rapist and are largely homophobic. These rape myths relegate rape to something that only involves homosexual men (Kassing et al., 2005; Pearson & Barker, 2018). Rape myths centering around homophobic content arguably serve to other male rape victims and distance male rape to taking place within a social minority group. However, researchers who have focused on rapist motivations have shown that rape, regardless of who it is committed against, is primarily motivated by a desire to display power and dominance over another human being (Javaid, 2018; Jewkes et al., 2010), and rapists and rape victims are not limited to a particular gender or sexual orientation (Fisher & Pina, 2013).

Section 1 Conclusion

In this section, I have discussed various ways in which rape myth acceptance can manifest. I also provided an overview of the kinds of rape myth content seen in RMA measures. The main point I would like to emphasise is that rape myths can serve many functions, and the content of these myths is widely varying. However, there can be little doubt that rape myths facilitate victim blaming, rape defending, and the trivialising of rape and its impacts. Now that the consequences and various manifestations of rape myth acceptance have been covered, I will now move on to my review of psychometric literature, in the hope of elucidating avenues for improving the measurement of rape myth acceptance.

Section 2: Psychometric Literature

In this section, I cover literature with a specific psychometric focus. The literature will be presented in four parts: (a) key definitions, theory and concepts, (b) the recent shift towards gender-inclusivity in RMA measurement, (c) a review of some dimensionality assessment tools, and (d) practically focused literature on how to conduct exploratory bifactor modelling.

Key Concepts

Factor analytic procedures and item response theory assume that the items are *effect indicators* of the latent variable of interest (Bollen, 2002; Bollen & Bauldry, 2011). Therefore, my research relies on the following assumptions: (1) rape myth acceptance is a causal construct that can be conceptualised as a latent variable, and (2) rape myths are effect indicators of rape myth acceptance.

When working with effect indicators, the assumption is that the latent variable of interest manifests and results in changes in the indicators, but changes to the indicators themselves do not cause changes in the latent variable (Bollen, 2002; Bollen & Bauldry, 2011). This direction of change makes effect indicators markedly different from *causal indicators*, which have a direct or structural effect on a variable of interest (Bollen & Bauldry, 2011; Bollen & Ting, 2000).

For an item to constitute an effect indicator of a latent variable, the item should capture a particular manifestation of that specific construct (Bollen & Bauldry, 2011). For example, rape myth acceptance manifests in how people attribute blame in rape scenarios, such as attributing rape to victims instead of perpetrators. Therefore, belief in a rape myth that shifts the blame to a victim can be treated as an effect indicator because it demonstrates the respondent's underlying rape myth acceptance.

Unidimensionality is defined here as "the existence of one latent trait underlying the data" (Reise et al., 2015, p. 14). For data to be unidimensional, the variation in item responses for a particular measure should be able to be explained by a single common factor. If the common factor is extracted, the remaining response matrix should be locally independent, or item residuals should have zero correlation (Reise et al., 2015). A scale is considered multidimensional when the items violate the assumption of local independence and have commonality above and beyond a single common factor (Reise et al., 2015).

A *bifactor model* is one of many ways that multidimensionality can be modelled. Under a bifactor framework, common variance at the item level is partitioned into competing sources (DeMars, 2013; Rodriguez et al., 2016). A bifactor model captures common variance shared by all (or most) of the items in what is known as a *general factor* (DeMars, 2013; Reise, 2012). In addition, a bifactor model will include *group factors*, which capture residual variance that is common to specific clusters of items over and above the common variance already accounted for by the general factor (DeMars, 2013; Reise, 2012).

Multidimensional data can be treated as *essentially unidimensional* if (a) there is a strong common factor underlying the items, and (b) the multidimensionality does not distort the parameters of a unidimensional model of the data to an unacceptable degree (Reise, Bonifay, & Haviland, 2013; Reise, Scheines, et al., 2013; Reise et al., 2015; Rodriguez et al., 2016). There are many reasons why it is desirable to work with unidimensional models. According to Dowling et al. (2020), unidimensionality is a key component of test equating methodology. Furthermore, while multidimensional item response theory models are powerful psychometric tools, unidimensional IRT models are much easier to interpret and can be used in a wider range of applications (Ip & Chen, 2015).

One of the potential applications for which unidimensional IRT models are particularly well suited is computer adaptive testing. The benefits of computer adaptive testing include being able to draw from an item pool of hundreds and potentially even thousands of indicators once the psychometric properties for each item have been established (Furr & Bacharach, 2014). The process involves automating the selection of the best items for a particular respondent so that the best items are selected in real-time, which can drastically reduce the number of items required to assess the construct of interest, which decreases respondent burden and is a very efficient method of test administration (Dima, 2018). However, computer adaptive testing is better suited to unidimensional IRT models, with multidimensionality seen as nuisance variation that detracts from measuring the construct of interest (Reise et al., 2015).

Furthermore, the ability to calculate item characteristics curves under an item responses theory framework makes IRT particularly powerful for investigating differential item functioning (Foxcroft & Roodt, 2013). An item has differential functioning if it has different item properties for different groups of people, despite them having the same trait level (Foxcroft & Roodt, 2013; Gamerman et al., 2018). However, detecting meaningful differential item functioning becomes substantially more difficult in the presence of multidimensionality (Gamerman et al., 2018).

Gender Inclusivity in RMA measurement

As I noted in my introduction, RMA measurement has historically been gendered and stratified into literature focusing on male rape myths and female rape myths. However, in the past two years there has been a shift towards measuring rape myth acceptance more gender-inclusively, with two notably different approaches.

One approach has been to reword well-established RMA measures to have *gender-neutral* item content (Canan et al., 2023; Johnson et al., 2023). This approach entails removing all gendered terms from the scale items, and replacing them with gender-neutral terms such as ‘the individual’ and ‘they’ (Canan et al., 2023; Johnson et al., 2023). For example, “If an individual is drunk, they might rape someone unintentionally” (Johnson et al., 2023, p. 212).

However, using gender-neutral terms does not necessarily translate into gender-neutral interpretation. As noted by Rosenstein and Carroll (2015), using gender-neutral items may see respondents revert to thinking in terms of dominant discourses. In other words, respondents may still interpret gender-neutral items in gendered ways or apply schematic thinking that posits women as victims and men as perpetrators. A counterpoint raised by Canan et al. (2023), however, is that gender-neutral items broaden the potential to place a greater focus on variables relevant to rape myths other than gender, such as the role of drugs and alcohol in rape myths, and rape myths that link into other oppressive belief systems such as classism and racism.

Another move towards gender inclusivity in rape myth acceptance measurement was the development of the Gender Inclusive Rape Myth Acceptance Scale (GIRMA) by Urban and Porras Pyland (2022). As opposed to complete gender neutrality, Urban and Pyland instead included in the scale a variety of rape myths anchored to female, male and transgender individuals. Some examples of item content include "When transgender people are raped, it’s not that serious" and "Adult men do not experience rape" (Urban & Porras Pyland, 2022, NP20641). Urban and Pyland stated that the variously gendered myths included in the scale measure the same construct because (a) the reported unidimensional model was a good fit to the data and (b) because the GIRMA highly correlated with the Illinois Rape Myth Acceptance Scale (Payne et al., 1999) in a subsequent validation study. However, I would like to note that Urban and Pyland did not provide a comparison or alternative model to the unidimensional solution, and good model fit does not necessarily prove unidimensionality (Ferrando & Lorenzo-Seva, 2018; Reise, Scheines, et al., 2013).

Davies et al. (2012) and Walfield (2018) have also suggested that rape myths and female rape myths may be indicators of the same construct after finding female rape myth and male rape myth measures to be highly correlated. A high correlation between scales indicates convergent validity, one of the key facets of construct validity (Fine, 1992). That is why, according to Fine (1992), a high correlation (e.g., $r = .8$) between two scales can serve as an indication that they may measure either the same construct or a very similar construct. However, I argue that to assume the items of two scales measure the same construct purely because of a high correlation is to assume causation, as the assumption is that the correlation is due to the same causal latent variable acting on both sets of items.

With that said, according to Bollen and Bauldry (2011), when two or more items are effect indicators of the same latent variable, they will generally be associated with one another due to how that variable influences them. Bollen and Bauldry further say that when two or more items are effect indicators of the same latent variable, the association between them should see them load on the same dimension of the latent construct. Therefore, if female rape myths and male rape myths are indicators of the same latent variable, then gendered rape myths should be associated and correspond to the same general rape myth acceptance dimension. Based on the research I have done, one of the best ways to investigate whether and to what extent items are associated with a common dimension is through bifactor modelling (Bianchi, 2020; Chen et al., 2013; Reise, Ventura, et al., 2011).

Chen et al. (2013) used a bifactor model to portion out the common variance between the items of two well-being scales to determine whether they were measuring two distinct constructs and discovered through the bifactor model that a general dimension loaded well on all the items. Bianchi (2020) also used a bifactor modelling approach when investigating whether burnout and depressive symptoms form part of the same syndrome. Bianchi looked for whether the items had substantial loadings on the general factor, how much of the common variance was explained by the general factor, and whether the general factor had an omega hierarchical value greater than .8.

Howe et al. (2019) talk about testing for common cause under an IRT framework for conducting a meta-analysis. In this study, Howe et al. fitted over 100 items from 7 different measures of depressive symptoms to a single unidimensional graded response model and looked at which loadings were significant and in the same direction. Items that met these criteria were taken by Howe et al. as direct effect indicators of the same underlying construct.

However, I prefer the approach taken by Reise, Ventura, et al. (2011) who looked into whether two separate measures of schizophrenia could be modelled in the same unidimensional IRT model, which required the scales to be measuring the same construct. Reise et al. first performed an exploratory bifactor analysis of each scale and the combined item pool to determine the strength of the underlying general factor. This formed an important part of the motivation for analysing the items from both scales under the same unidimensional IRT model. I want to acknowledge that the investigation by Reise, Ventura, et al. (2011) was what inspired the methodological direction of my research, as it was upon reading that paper that I came up with the idea for investigating the IRMA and MRMS using bifactor modelling to see whether they tap into the same general construct.

Bifactor Approach to Dimensionality Assessment

Bifactor models are a powerful psychometric tool for dimensionality assessment (Reise, 2012). Bifactor models partial the common variance among items into general and specific factors, making it possible to look at those components' relative strength and reliability individually (Rodriguez et al., 2016). Indices that speak to the strength of the general factor of a bifactor model are the explained common variance (ECV) and Omega Hierarchical (OmegaH).

Explained common variance indicates how sizeable the first factor loadings are when compared to the entire factor solution (Stucky et al., 2013). In a bifactor context, the first factor is assumed to be the general factor, and therefore, the ECV indexes what percentage of common item variance is captured by the general factor (Stucky et al., 2013). Ideally, a unidimensional solution should have an ECV value greater than .85; however, the cut-off value can be as low as .70 (Ferrando & Lorenzo-Seva, 2018). Furthermore, according to (Stucky et al., 2013, p. 51), when ECV is calculated at the item level (I-ECV), the I-ECV indicates "the extent to which an item's responses are accounted for by variation on the latent general dimension alone, and thus acts as an assessment of unidimensionality at the individual item level."

ECV can be sensitive to the size of the group factors in a bifactor solution (Reise, Scheines, et al., 2013; Rodriguez et al., 2016). Therefore, a recommended complementary index of general factor strength is omega hierarchical, which is technically a reliability index (Flora, 2020; Garcia-Garzon et al., 2021; Reise et al., 2018). Rodriguez et al. (2016, p. 145) describe omegaH as "the percent of total score variance attributable to a single general factor". According to Reise, Scheines, et al. (2013), when working with a bifactor model omegaH is a more direct index of the strength of the general factor than ECV, because omegaH is not as sensitive to the size of group factors. Rodriguez et al. (2016) recommend that an omegaH greater than .8 can be taken as evidence of essential unidimensionality.

Additional indicators of dimensionality worth considering are the eigenvalues, unidimensional congruence (UniCo; Ferrando & Lorenzo-Seva, 2018), and the mean of item residual absolute loadings (MIREAL; Ferrando & Lorenzo-Seva, 2018). If the first-to-second eigenvalue ratio is equal to or greater than 3:1, there is likely a general factor worth investigating further (Reise, Ventura, et al., 2011). However, it is important to stress that eigenvalues are rooted in principal components theory. UniCo is rooted in factor analytic theory and compares the final loading matrix of a factor model with the loading matrix that would be expected in a perfectly unidimensional solution (Lorenzo-Seva & Ferrando, 2021). The degree of congruence is captured as a number ranging from 0-1, and a value above .95 indicates that the data is likely unidimensional (Lorenzo-Seva & Ferrando, 2021). However, the UniCo is model-dependent and therefore relies critically on the specified model. By contrast, the MIREAL is a model-independent indicator rooted in factor analytic theory, that provides a scale-level indication of departure from unidimensionality (Ferrando & Lorenzo-Seva, 2018). The MIREAL is the average of the first-to-second factor loading ratio for each item, and ranges between 0 and 1, with a value greater than .3 indicating a substantial departure from unidimensionality (Ferrando & Lorenzo-Seva, 2018).

When interpreted within the appropriate context, dimensionality indices can provide valuable information in a dimensionality assessment. However, many dimensionality indices are model-dependent, so if a model is misspecified its subsequent dimensionality indicators can be very misleading. Before assuming a unidimensional model is appropriate, it is important to check whether the parameter estimates have been biased by inherent multidimensionality (Reise, Moore, & Maydeu-Olivares, 2011; Reise, Scheines, et al., 2013; Reise et al., 2015). According to the comparison modelling approach outlined by Reise et al. (2015), the best way to check whether a unidimensional model is biased is to compare it with a less restricted, alternative model that better captures the multidimensionality. Reise et al. (2015) recommend comparing the item-level loadings of a one-factor model to the general factor loadings of an appropriate bifactor model; if the difference between the loadings is relatively small it serves as a strong justification for treating the data as essentially unidimensional. To quantify a substantial difference in loadings across FA models, Rodriguez et al. (2016, p. 145) calculated the relative parameter bias as "the difference between an item's loading in the one-factor solution and the general factor loading in the bifactor, divided by the general factor loading in the bifactor". If the parameter bias is less than 15% (at the most), the unidimensional model can be considered relatively unbiased (Rodriguez et al., 2016).

The comparison modelling approach outlined by Reise, Moore, and Maydeu-Olivares (2011) and Reise et al. (2015) can, and was intended to be, extended to interpretation under an IRT framework. Without getting too technical, item level-factor analytic (ILFA) models are equivalent to two-parameter normal-ogive models (DeMars, 2013; Reise et al., 2015). Therefore, the loadings of models computed under an ILFA framework can be converted to IRT slopes of a normal-ogive model, making it possible to compare the models under IRT parameters (Reise et al., 2015). However, Reise et al. (2015) refrained from indicating what constitutes a substantial difference in item slopes, arguing that what is considered acceptable would depend on the intended IRT application.

As an alternative, I would like to highlight a series of slope descriptors by Baker and Kim (2017, p. 16), which were intended to convey the interpretation of the item discrimination to a non-technical audience. According to Baker and Kim, item discrimination values under a normal ogive model can be interpreted as follows: very low discrimination ($a < .21$), low discrimination (a: .21 - .37), moderate discrimination (a: .38 - .79), high discrimination (a: .79 - .99), and very high discrimination ($a \geq 1$).

I argue that the slope descriptors by Baker and Kim (2017) could serve as an indication of what increments might constitute a substantial difference in the discriminating power of an item. For example, I posit that a slope difference less than .21 could be considered small, as an item with a slope less than .21 is considered to have very little discriminatory value. By contrast, an item slope of .38 places an item in the moderately discriminating range, and so I argue that a difference of that magnitude would be more likely to lead to an item holding a different psychometric relevance under a unidimensional model. Therefore, I posit that a slope difference of .38 or higher should be considered substantial. However, as noted by Reise et al. (2015), even a slope difference of .5 can be acceptable depending on the intended IRT application.

Exploratory Bifactor Modelling Steps

Exploratory bifactor models formed a critical component for all three of my research aims. There are five major components to exploratory bifactor modelling: (1) prepping and testing the underlying matrix, (2) selecting an appropriate factor estimation method, (3) selecting how many factors to retain, (4) choosing an appropriate factor rotation method, and (5) choosing an appropriate bifactor estimation procedure (DeMars, 2013).

The Underlying Matrix. Factor analysis is based on a matrix of either covariances or correlations between scale items. A correlation matrix is essentially a table of standardised covariances. This matrix can be seen as the foundation layer for building a common factor model. When working with ordinal data, the foundation matrix for a factor analysis should be a tetrachoric correlation matrix if working with binary data, and a polychoric correlation matrix if working with three or more Likert-style response categories (Timmerman & Lorenzo-Seva, 2011). Polychoric correlations make it possible to treat the underlying variable as continuous, even in the presence of non-normally distributed data observations (Timmerman et al., 2018). Pearson correlations should only be considered as a last resort if the polychoric correlation matrix fails to converge (Timmerman & Lorenzo-Seva, 2011; Timmerman et al., 2018).

Once a foundation correlation matrix is estimated, it is important to test whether there is enough variance to justify a factor analysis (Taherdoost et al., 2014). According to Taherdoost et al. (2014), there are two relevant statistics for this phase of analysis: the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO; Kaiser, 1974) and Bartlett's test of sphericity (Bartlett, 1951).

The KMO is a measure of sampling adequacy, and tests whether there are a significant number of factors in the dataset (Kaiser, 1974). The KMO returns a value between 0 and 1. Although there are some varying opinions on what a desired value is, according to Taherdoost et al. (2014) the absolute minimum requirement is a value greater than .5. It should be noted however that, according to the criteria suggested by Kaiser (1974), less than 0.5 is unacceptable, 0.5 - 0.6 is miserable, 0.6 - 0.7 is mediocre, 0.7 - 0.8 is middling, 0.8 - 0.9 is meritorious, and 0.9 - 1.0 is marvellous.

Bartlett's sphericity test (Bartlett, 1951) assesses whether the matrix is an identity matrix. In an identity matrix, variables are unrelated and therefore cannot be submitted to a date reduction technique such as factor analysis. Therefore, the null hypothesis is that the variables are uncorrelated. To pass this test, the null hypothesis should be rejected and therefore requires a significant p-value ($p < .05$). However, it is important to note that this statistic requires multivariate normality and relies on chi-square distribution (Bartlett, 1951).

Factor Estimation Methods. Rape myth acceptance measures such as the Rape Myth Acceptance Scale (Burt, 1980), the Illinois Rape Myth Acceptance Scale (Payne et al., 1999) and the Male Rape Myth Scale (Kerr Melanson, 1998) have all been noted for producing skewed data distributions when administered to student samples (Gerger et al., 2007; Walfield, 2018). Therefore, I focused on finding estimation methods that could accommodate positive skews and non-normal distributions.

Maximum likelihood estimation is the most widely used factor estimation method, as it is hailed as particularly consistent and efficient (Cai & Moustaki, 2018). However, according to Cai and Moustaki (2018), many of these benefits only hold for large sample sizes. More concerning for my study is that the maximum likelihood method requires strict distributional assumptions about the data.

By contrast, a strength of unweighted least squares (ULS) estimation is that it does not require any distributional assumptions (Joreskog, 2003). According to Joreskog (2003, p. 1), ULS is particularly suited for exploratory factor analysis “where only parameter estimates (and not standard error estimates or chi-square values) are of interest”. This makes it a perfect fit for studies like mine, where the focus is on parameter estimates, and chi-square values may be overinflated due to a large sample size.

Factor Retention Indices. The number of factors to extract is user-specified. It is a very important step, as specifying too few as well as too many factors leads to poor factor loading interpretation and potentially suggesting a latent variable with little to no substantive meaning (Hayton et al., 2004; Lorenzo-Seva et al., 2011). The number of factors to retain is a question I have often seen glossed over in factor analytic studies, yet there is an entire chapter dedicated to this question alone in the Wiley Handbook of Psychometric Testing (Timmerman et al., 2018).

Commonly reported indices in factor analytic papers are the Kaiser criterion, the Skree test, and parallel analysis (Lorenzo-Seva et al., 2011; Taherdoost et al., 2014). There is a fundamental difference between principal components, which form the basis of principal components analysis (PCA), and common factors, which form the basis of exploratory factor analysis (EFA). As noted by Lorenzo-Seva et al. (2011), procedures that are based on PCA should not be used in an EFA context. The Kaiser criterion, Skree test and parallel analysis are all based on eigenvalues (Braeken & Van Assen,

2017) and are therefore grounded in PCA and not EFA theory. Eigenvalues can provide insight for the interpretation stage of an empirical analysis, as PCA and EFA can often have similar results (Lorenzo-Seva et al., 2011). However, in the initial modelling stages, it is better to use factor retention indicators that are based on an understanding of common factors (Timmerman et al., 2018).

Timmerman and Lorenzo-Seva (2011) developed and tested a variant of parallel analysis based on minimum rank factor analysis (PA-MRFA) that is suitable when factor analysing ordinal data. Whereas traditional parallel analysis is based on eigenvalues, the PA-MRFA is based on variance. In both forms of parallel analysis, the number of factors recommended is based on comparing the actual polychoric correlation matrix with randomly generated matrices (usually 500+) to figure out which factors are substantively meaningful as opposed to pseudo-factors that are merely made up of noise. In the PA-MRFA assessment, the explained common variance of successive common factors in the observed data is compared with the explained common variance from the generated matrices, and only factors where the percentage of explained common variance in the real data is greater than the 95th percentile of random variance are flagged as genuine (Timmerman & Lorenzo-Seva, 2011).

However, Timmerman et al. (2018) have pointed out that particularly when dealing with large samples, the PA-MRFA can overestimate the number of factors to extract, and therefore recommend running a variant of the Hull test as well to make a more informed decision. A variant of the Hull test specifically intended for use in an FA context is based on an expression of the common part accounted for by a common factor model, and so is known as the Hull common part accounted for (Hull-CAF; Lorenzo-Seva et al., 2011). The Hull-CAF is intended to determine only major factors, and ignores minor factors, minimising the risk of over-extraction (Lorenzo-Seva et al., 2011).

A recent alternative suggested by Finch (2020) is to look to changes in model fit statistics such as the root mean square error of approximation (RMSEA) to determine the appropriate number of factors to retain. Model fit statistics are intended for use under a CFA context, but are often cited in EFA literature to provide additional support for a model (Finch, 2020). However, comparisons of model fit are arguably not appropriate in a bifactor modelling context. One of the criticisms of bifactor models is that they can appear to be better fitting than more appropriate models due to their flexibility and ability to accommodate implausible responses (Bornovalova et al., 2020; Reise et al., 2016). Nonetheless, the following can be interpreted as indicating good model fit: a weighted root mean square residual (WRMR) value less than 1 (DiStefano et al., 2018) and a root mean square error of approximation (RMSEA) value of .06 or lower (DiStefano et al., 2019), as well as a comparative fit index (CFI) value of .9 or higher, a standardized root mean square residual (SRMR) value greater than 0.1, and a Tucker Lewis index (TLI) value above .9 (Finch, 2020).

Factor Rotation Methods. Once a factor solution has been calculated by a computer, it needs to be interpreted by a human being. Rotation methods aid in this substantive interpretation (Jennrich, 2018). This is a standard consideration in FA, and even bifactor modelling methods (which will be elaborated on further in the next section) require a decision about the kind of rotation used in the process.

Rotation does not change the substantive nature of the solution; it just aids in interpretation. There are two main types of rotation: oblique, and orthogonal. Orthogonal rotations assume that the group factors are uncorrelated. Oblique rotations allow for correlation between factors (Jennrich, 2018). In a bifactor model, the group factors are required to be orthogonal to the general factor, but the group factors themselves can be correlated.

A wide range of rotation options are available. According to Baglin (2014), orthogonal rotation should only be used when there is a strong motivation for assuming factors are uncorrelated. Given that I did not have a strong motivation for assuming any factors implicit in the data were uncorrelated, I narrowed my focus to finding an applicable oblique rotation method. Robust Promin (Lorenzo-Seva & Ferrando, 2019b) in particular is a stand-out oblique rotation method. Robust Promin is a method for diagonally weighted factor rotation, and the rotated loading matrix obtained is expected to be simple and stable across samples as the procedure gives more weight to the most stable correlations (Lorenzo-Seva & Ferrando, 2019b). This is exceptional, as most rotation methods focus exclusively on maximising factor simplicity. According to Lorenzo-Seva and Ferrando (2019a), researchers tend to make use of rotational methods that have been used in previous studies, to allow for continuity and vertical growth of literature. Robust Promin is a relatively new rotation method but appears to be the rotation method most likely to produce a factor solution that is replicable.

Bifactor Estimation Procedures. With the increased interest in bifactor models, newer, faster, and more powerful bifactor modelling techniques have become available. As a result, there are more choices to be made than ever before particularly with regard to the kinds of bifactor modelling transformations available.

Common FA rotations aim to identify simple structures, where each item loads on only one factor (Jennrich, 2018). This is a problem when wanting to explore a bifactor structure because each item is allowed to load on a general factor as well as one or more specific group factors (DeMars, 2013; Reise, 2012). For this reason, there are specific methods used to produce bifactor models, which I have grouped into three main approaches.

In the first approach, traditional rotation methods are modified to rotate to a bifactor criterion. Two noteworthy bifactor rotations are the bi-quartimin rotation (Jennrich & Bentler, 2011), and the bi-geomin rotation (Jennrich & Bentler, 2012). These are known as direct analytic bifactor rotations and are used for handling orthogonal and oblique cases respectively. However, the results of a bifactor method comparison by Abad et al. (2017) indicate that the rotations perform poorly when items have significant cross-loadings, i.e., when items have substantial loadings on two or more group factors. Being unable to accommodate cross-loadings dependably is a significant criticism, and therefore the analytic bifactor rotations will not be considered further here.

In the second approach, instead of rotating to a bifactor model directly, the model is calculated in multiple stages. A second-order FA model is first fitted to the data and only then transformed into a bifactor model using the Schmid-Leiman (SL) orthogonalization procedure (Reise, Moore, & Maydeu-Olivares, 2011). A major limitation of the SL procedure is that it is subject to a proportionality constraint, which is a result of the multiple stages required in the factor estimation process. A detailed explanation of the proportionality constraint is available in a paper by Mansolf and Reise (2016). In a nutshell, the SL procedure struggles to accurately recover parameter estimates when group factors disproportional to one another, or when there are cross-loadings present. That is, it only works well when group factors consist of roughly the same number of items, and the items within a group have similar loadings on the group and general factor (Mansolf & Reise, 2016).

There have been several notable attempts to work around the inherent proportionality constraint of the SL procedure. This leads into the third and final bifactor modelling approach: methods that use a semi-specified target matrix in the modelling procedure. A target matrix is a semi-specified pattern matrix that provides a guiding outline, or general structure, for the rotation procedure. In the case of bifactor modelling, an item will only ever have two unspecified elements in a target matrix: one unspecified group factor, and the general factor (Lorenzo-Seva & Ferrando, 2019a). The remaining elements are specified as zero (Lorenzo-Seva & Ferrando, 2019a). Figure 1 is an example of what a target matrix looks like.

Table 1

Example of a semi-specified pattern matrix

Item	F1	F2	F3	GF
1	0			
2	0			
3		0		
4		0		
5			0	
6			0	

Note. Note. Zero = specified element. Non-specified elements indicated with a "—". Table 1 is a made-up matrix for illustrative purposes only.

Reise, Moore, and Maydeu-Olivares (2011) recommend using the SL procedure to define the target matrix for a bifactor model and then make use of the target Procrustes rotation to generate the final solution, as the combination removes the proportionality constraint inherent to the SL procedure, and significantly improves the accuracy of parameter estimation. Abad et al. (2017) realised that specifying multiple target patterns in an iterative process would improve upon the process even further, and introduced the Schmid-Leiman with iterative target rotation (SLi).

Waller (2018), on the other hand, went in a different direction and motivated bypassing the multi-stage FA process that results in the proportionality constraint. He therefore introduced a new method where all that is required to generate a bifactor SL model is a single factor analysis, followed by a Procrustes rotation. Waller's method is the latest contender in the ever-growing line of SL-based procedures, and is called the Direct Schmid-Leiman transformation (Waller, 2018).

The final bifactor modelling method I wish to highlight is the Pure Exploratory Bifactor (PEBI; Lorenzo-Seva and Ferrando, 2019a). According to a simulation study by Lorenzo-Seva and Ferrando (2019a) that compared several bifactor modelling methods, the PEBI is the most accurate approach when (a) a general factor is not present in the population and (b) when group factors are correlated. The first step of the PEBI is to define a partially specified matrix. A strength of the procedure is that in addition to the default method (Promin) for building the target matrix, there are several other options available. A researcher can propose a target based on previous research or use an SL-based target matrix both the target matrix proposed by Reise, Moore, and Maydeu-Olivares (2011) and the SLi target by Abad et al. (2017) are possibilities. The remaining three steps are identifying the general factor loadings, rotating the group loadings, and arriving at a final bifactor solution (Lorenzo-Seva & Ferrando, 2019a).

Something that stood out to me about the PEBI method is that it is the most inherently exploratory method of all the ones I reviewed. In addition to the strength of starting with a semi-specified target matrix, it was the only method in which the group loading rotation is also semi-specified so that the nonzero loadings in the target matrix are freely estimated before arriving at the final solution. By contrast, the other SL-based methods make use of Procrustes rotation, in which all nonzero loadings are given the same weight (Lorenzo-Seva & Ferrando, 2019a).

As a final note for this section, it is best to consider the EFA bifactor modelling approach early on, as the method will more than likely determine which software program is used to run the bifactor analysis (or at least exclude software packages that cannot run it). For example, to my knowledge, the pure exploratory bifactor modelling procedure is currently only available in FACTOR (Lorenzo-Seva & Ferrando, 2021). Furthermore, it is worth noting that many of the functions relevant to calculating indices in R are based on SL transformations. For example, the omega() function in the *psych* package by Revelle (2021) calculates omega based on an SL orthogonalization of the data.

Section 2 Conclusion

In this second major section of my literature review, I covered some of the many ways to investigate dimensionality. I covered key concepts and assumptions for my investigation to follow, with a major focus on exploratory bifactor modelling applications and the practical steps involved. By now, the reader should have an understanding of how and why I will be investigating whether data can be considered unidimensional, as well as how bifactor models can be used to assess whether indicators may be tapping into the same general construct.

When it comes to bifactor modelling and dimensionality investigations, there is a wide range of procedures and modelling options available, and understanding how the different elements fit together makes it easier to investigate measures with intention. There is no one way to investigate the dimensionality of a scale, and no index that will perfectly explain data structure. However, by relying on several dimensionality indices, and by comparing how the data loads under different models, the dimensionality of a measure can be better understood. With such an understanding comes the potential to explore data and measures in new ways, and that may well be what makes all the difference when investigating rape myth acceptance in the South African context.

CHAPTER 3: METHODOLOGY

From the very beginning of my research planning four years ago, I made the conscious decision to track my methodology in as much detail as possible. I remember reading countless papers I felt were too vague to learn from or potentially hope to replicate. With an ever-growing awareness of the replication crisis in psychological research, I have recorded in this chapter every step of my methodology with as much information and specificity as possible. Contained in this chapter are the details of my participant recruitment and sampling strategy, the measures and incentives used to gather my data, and the various institutional, ethical and data-handling procedures involved in turning my research from an idea to a reality.

Participants

The study population was students attending Stellenbosch University (SU). I have been an SU student for over six years and have witnessed first-hand the need for and interest in 'ending rape culture' by my fellow students. The recruitment email for both the main study (see Appendix A) and the pilot study (see Appendix B) included a link to the online survey. After being presented with a study brief, respondents needed to give informed consent to participate (see Appendix C).

Inclusion and Exclusion Criteria

To be considered for inclusion in the study and receive a recruitment email, a potential participant had to (a) have an active email address with Stellenbosch University and (b) be registered for an undergraduate degree programme at Stellenbosch University. Recipients who expressed an interest in the study but met any of the following criteria were excluded from the study: (1) anyone under 18 years old, (2) anyone studying an undergraduate degree at a postgraduate level and (3) recipients who were not studying an undergraduate degree.

The inclusion criteria were exercised at the mailout stage. For my recruitment, I used mailing lists compiled by Stellenbosch University to send out general communications to all registered undergraduate students. The exclusion criteria were exercised at the questionnaire stage. Potential participants needed to give their informed consent to proceed with the questionnaire. One of the declarations required was "I am over 18 years of age". If the declaration was not ticked, the questionnaire did not launch, and the viewer was automatically excluded from the study.

In addition, the demographic section of the questionnaire contained a screening question. The screening question was multiple choice and asked whether the reader was currently studying towards (a) an undergraduate degree, (b) a postgraduate degree or (c) other. Only participants who indicated they were studying towards an undergraduate degree could proceed to the main body of the questionnaire. Participants who did not indicate they were studying towards an undergraduate degree were redirected to the end of the survey, thereby excluded from participating.

Sampling

I used a comparison modelling procedure outlined by Reise et al. (2015) to answer my research questions. The comparison modelling procedure entails fitting a bifactor model under a factor analytic (FA) framework and interpretation of the model under FA as well as item response theory (IRT) parameters. Reise, Moore, and Maydeu-Olivares (2011) indicate that even when data are well structured, with strong item loadings on the general factor, a sample size of no less than 500 participants is required for optimal comparison modelling. Furthermore, Embretson and Reise (2000) and Woods (2015) recommend sample sizes greater than 500 for accurate IRT parameter estimation.

According to Sinclair et al. (2012), internet surveys have a much lower response rate than postal surveys, reporting an internet response rate between 2% and 5% depending on the degree of personalisation involved in the recruitment approach. Therefore, to ensure I would recruit the number of participants required for my analyses, I used convenience sampling. Convenience sampling was utilised to yield a sample size greater than 500 participants. The final sample size retained for analysis was 2,536 participants. Please note that the final sample did not include the 41 responses received for the pilot study.

The large sample size required consideration in the data analysis stage. There are many statistics, such as the chi-square (X^2), that are sensitive to sample size (Mulaik, 2015). Tiny deviations are multiplied by the number of people in the sample; therefore, the resulting statistical power is often too large in big samples. Given that the sample size was so big, I have avoided statistical procedures based on chi-square (X^2) where possible. This is relevant because chi-square is commonly reported in factor analytic papers.

Participation Incentives

To encourage participation, participants who completed the questionnaire had the option to enter a series of lucky draws to win (a) a bicycle (valued at R2500), (b) one of 6 R500 cash prizes, or (c) one of 10 moonstone necklaces.

Pilot Study

To determine the reliability of the RMA measures in a South African context, I piloted the questionnaire with 40 students using the same methodology as the main study, with two exceptions: (a) I made use of snowball sampling using a tailored recruitment email and (b) there was one lucky draw for a cash prize of R200.

The snowball sampling started with first-year psychology students and was forwarded to other undergraduates. At the time of recruitment, I tutored for the Psychology Department and tutored three groups of roughly 20 students enrolled for Psychology 144, an introductory psychology course. Mailing lists were linked to each of my three tutorial groups, which served as the point of departure for the snowball recruitment process. To protect the privacy of the recipients, I specified the three mailing lists in the blind carbon copy (Bcc) field. The recruitment email asked recipients to either forward the survey to an undergraduate student they knew and/or complete the survey.

Measures

A single electronic questionnaire was the sole data collection instrument (see Appendix D). The questionnaire had three sections consisting of (1) demographic items, (2) the Illinois Rape Myth Acceptance Scale (Payne et al., 1999) and (3) the Male Rape Myth Scale (Kerr Melanson, 1998).

Demographic Variables

There were four items in the demographic section. Only three demographic variables were included: age, gender and home language. These items were included to report on sample characteristics and to examine possible confounding covariates. The fourth item included in the demographic section of the questionnaire asked whether participants were studying towards an undergraduate degree at the time and was utilised as a screening question.

The Illinois Rape Myth Acceptance Scale

The Illinois Rape My Acceptance Scale (IRMAS) is a 45-item self-report measure of rape myth acceptance and was developed by Payne et al. (1999) over the course of 6 empirical studies. The IRMAS has seven subscales: (1) She Asked For It, (2) It Wasn't Really Rape, (3) He Didn't Mean To, (4) She Wanted It, (5) She Lied, (6) Rape is a Trivial Event, and (7) Rape is a Deviant Event (Payne et al., 1999). In the development paper, Payne et al. (1999) sampled undergraduate students ($n = 604$) from a university in the United States of America. Reported Cronbach's α was .93 for the overall scale and ranged from .74 to .84 for the subscales.

The scale items are administered in a specific order and not grouped according to the subscales. Only 40 of the items are rape myths, such as "Many women secretly desire to be raped" (Item 8), and "In reality, women are almost never raped by their boyfriends" (item 28). The remaining five items are filler items, e.g., "It is preferable that a female police officer conduct the questioning when a woman reports a rape" (item 32). Participants respond to each item on a 7-point Likert scale, ranging from 1, "not at all agree", to 7, "very much agree". The item scores are summed to provide a total score, with higher scores indicating greater acceptance of rape myths.

The Male Rape Myth Scale

The Male Rape Myth Scale (MRMS) is a 22-item scale that measures false or stereotypical beliefs about male rape (Kerr Melanson, 1998). The six-point Likert scale ranges from 1, "strongly disagree" to 6, "strongly agree", with no neutral category. Scores are summed, with a higher value indicating greater acceptance of male rape myths. Item 1, item 6 and item 19 are reverse scored.

In the development study, the sample (n = 303) consisted of undergraduate students drawn from a university in Canada. The scale had high internal consistency (α = .90) and 4-week test-retest reliability (r = .89).

The MRMS was developed 20 years ago, and some items' wording is dated. For example, item 9 is "If a man engages in necking and petting and he lets things get out of hand, it is his own fault if his partner forces sex on him" (Kerr Melanson, 1998). I, therefore, replaced the wording for some of the items in the scale with more current terminology. For example, item 9 was modified to "If a man engages in kissing and foreplay and allows things to get out of hand, it is his own fault if his partner forces sex on him."

Procedure

This study underwent two rounds of ethical review, first by the Psychology Departmental Ethics Screening Committee and then by the Research Ethics Committee for Stellenbosch University (SU). I also needed to apply to the Institutional Governance of SU to collect data from SU students. I requested permission to indirectly access the mailing lists used by the university to send out communications to undergraduate SU students. I received permission from Institutional Governance for indirect access to the email addresses in the form of mail-out lists for my study recruitment, permission to collect, store and analyse data from undergraduate SU students, as well as permission to utilise the SUNSurveys platform.

SUNSurveys is an online survey service available to students and staff of Stellenbosch University for academic research. SUNSurveys utilises the specialised survey software Checkbox, and collected data is stored on Stellenbosch University's secure institutional servers.

Upon approval from the relevant bodies, I successfully registered for the SUNSurveys service. In October 2020, I set up my questionnaire for electronic administration. The mailing lists needed to send out my recruitment email to all SU undergraduate students were uploaded to the Checkbox platform by one of the

SUNSurveys administrators. Therefore I did not handle the mailing lists directly. There were ten mailing lists, with roughly 2000 email addresses per list. The recruitment invitation was sent out to a total of 19,968 student email addresses.

There were two rounds of data collection: one for the pilot study and one for the main study. Potential participants received a recruitment email with a link to the relevant study. Participants were provided with a briefing sheet and required to give informed consent before proceeding to the questionnaire. Informed consent was given by checking three boxes to indicate that the respondents: (a) were over 18 years old, (b) understood the various aspects and the nature of the research, and (c) wished to partake in the study.

I piloted the survey using snowball sampling from 19 October 2020 to 30 October 2020, with 41 responses received. Preliminary analyses indicated sufficient scale reliability to proceed with data collection. The recruitment email for the main study was sent out to all undergraduate students registered with Stellenbosch University on 22 March 2021. One follow-up email was sent two weeks later, on 6 April 2021, and the survey was closed on 21 April 2021. The data collection period ran for one month. Data collection comprised a single online questionnaire with three sections: demographic variables, the IRMAS and the MRMS.

Under the demographic variables was a screening question. By using conditional logic, if a participant indicated they were studying towards anything other than an undergraduate degree, they were immediately rerouted to the end of the survey. I used further conditional logic to make sure that only people who had completed the questionnaire in its entirety were able to see and interact with the dialogue box linking to the lucky draw entry.

There were several checks included to prevent missing data. Moving on to the next page of the questionnaire was only possible once every question had a response. This ensured no individual items could be skipped. Logic conditions were also in place to prevent accidental skipping of an entire page of questions. Participants were allowed to withdraw from the study at any point. Incomplete responses due to discontinuation were taken as a sign of withdrawal from the study and deleted.

The survey took approximately 10 - 15 minutes to complete. Once participants finished and submitted the survey, a new browser window opened. There was a field where participants could leave their contact information. If they left their contact information, they needed to indicate whether they (a) wanted to enter the lucky draw and (b) if they wanted to find out the study's results once completed. The new browser separated the contact info provided from the answers submitted, maintaining anonymity.

I exported the data directly from SUNSurveys onto my computer as comma-separated values (.csv) for analysis. I used Excel to handle the exported data. I exported three Excel spreadsheets, one for my raw data for the main study, one for my pilot data, and another for the lucky draw entries.

Once data collection closed, the lucky draws were done immediately. I randomised the contact details in a password-protected Excel spreadsheet. I used a random number generator to determine the winning numbers, with the winner in the corresponding Excel row. This method did not require personal information to be entered into third-party sites, ensuring it remained protected and confidential. The contact information collected from lucky draw entrants was not used in the analyses and was destroyed as soon as all the winners had accepted their prizes.

Factors that Contributed to the Large Sample Size

The recruitment email was sent out to 19,968 SU students. The response rate for this study was 12.82%. Note that the response rate does not include incomplete responses submitted. The lucky draws appear to have encouraged participation in the study, with 96% of the respondents entering the lucky draws. Of the 2,536 completed questionnaires received, 2,435 participants went on to enter the lucky draws.

I took into consideration the timing of my recruitment. My supervisor advised me that the best times of year to send out a survey are at the beginning of each semester. This may be because students check their emails regularly for new course communications and have more time to participate in surveys.

I also considered the time of day that my email was sent out. I began emailing at 8 pm South African Standard Time (SAST). In both mailout rounds, potential recipients received the emails either at the end of the day or first thing in the morning. I believe this may have led to more people seeing the email at a time of day when they were more likely to have the time to complete it immediately.

The standard advice given to SUNSurvey users is to break up the mailout into chunks. Sending all the invites at once can cause the server to be put under too much strain. I scheduled my invite groups 10 to 20 minutes apart for the first recruitment email round. The interval was not long enough. The servers became backlogged, and the mailout took roughly 12 hours. This meant some students received the recruitment email at 7 pm on Tuesday, and others received it as late as 6 am on Wednesday.

Since I had scheduled my mailout at night, I did not backlog the servers during the day. For the reminder email, I spaced out my email batches hourly to reduce the strain placed on the server.

Ethics

This study fits the Research Ethics Committee's criteria for medium risk. Rape is highly prevalent in South Africa; the likelihood that this questionnaire would be completed by someone who had either been or knew a rape victim was high. My study did not require participants to recount or elaborate on rape-related experiences. However, it did require students to report on beliefs and attitudes related to the causes, context, and consequences of rape. Thus, there was a substantial risk that the questionnaire would expose a student to emotional distress or trigger a recall of previous trauma.

In both the recruitment email and informed consent form, I (a) included a forewarning that the questionnaire dealt with sensitive content that some students might find distressing or triggering, (b) clearly stated that the questionnaire had items about rape and sexual assault, and (c) listed the contact details of free and immediate support services.

Listed supported services were the telephonic crisis helplines from the Centre for

Student Counselling and Development, the Rape Crisis Centre, as well the 24/7 interactive SMS service provided by the TEARS Foundation. I also enlisted the counselling services of both the Welgevallen Community Psychology Clinic and the Centre for Student Counselling and Development.

The survey was hosted on SUNSurveys, and stored on the institutional server, which is secure. SUNSurveys makes use of the online survey tool Checkbox. The data collected was exported from SUNSurveys directly as a .csv file and stored on my password-protected laptop, to which I alone have had access. A OneDrive storage account contains the backup of the anonymous survey data. The account is set up to a unique email address to prevent syncing across my devices. My supervisor will store the data for five years and then it will be destroyed.

Data Analysis

Software

For this research I used PSPP (Free Software Foundation, 2018), FACTOR (Lorenzo-Seva & Ferrando, 2021), and R (R Core Team, 2020). All three software are available for free via the links provided.

PSPP is a freeware alternative to the Statistical Products and Service Solutions software, more commonly known as SPSS. SPSS used to stand for *Statistical Package for the Social Sciences* but was changed in light of a target market extension beyond social sciences (George & Mallery, 2016). PSPP does not have an official acronym expansion. The graphical user interface of PSPP follows that of SPSS and is easy to use. The descriptive statistics feature is well suited for quick and efficient sample analysis and produces an output that is easy to interpret and report. PSPP is free to download, use and distribute (Free Software Foundation, 2018).

FACTOR is a software specialised for factor analysis. FACTOR has the following strengths: (a) FACTOR is free, (b) it runs robust factor estimation and generates millions of comparison matrices to calculate confidence intervals, (c) FACTOR is executed as a file and does not require installation, and (d) once the file has been downloaded, no further internet connection is required for analyses to be carried out.

R is an open-source software environment and programming language that can be used for various statistical and data analytic applications (R Core Team, 2020). R was vital for prepping my data, obtaining detailed item-level statistics, generating the polychoric correlation matrix for the combined item pool of IRMAS and MRMS items, and computing omega hierarchical. In other words, I used R as a supplementary tool to run specific, directed commands.

Phases of Data Handling

The data handling for my investigation was broken down into three key phases; (1) descriptive statistics, (2) exploratory factor analyses, and (3) model comparisons. In **phase 1**, my focus was on prepping the raw data for analysis and computing traditionally reported psychometrics and statistics, such as the sample composition and item level distributions. For **phase 2**, the focus was on exploring the underlying data structure of the measures using exploratory factor analysis (EFA). In particular, the aim was to compute two viable EFA models per measure: a one-factor and a bifactor model. In **phase 3**, the focus was on calculating item-level differences between the EFA models generated in phase two to determine the extent of parameter distortion that occurred when multidimensionality was not appropriately modelled.

Phase 1: Descriptive Statistics. Before any analyses could be conducted, the data needed to be prepped. When first exported from the SUNSurveys platform, my data was available as a comma-separated value file (.csv). The raw data file was loaded directly into R. In R, I prepped the data by removing the filler items from the IRMAS data and reverse scoring items 1, 6 and 9 in the MRMS data. Using the *psych* package (Revelle, 2021), I generated an output that included global and item-level reliability indices and item-level descriptive statistics.

The data file prepped in R was re-exported as a .csv file and loaded into PSPP. In PSPP, I ran frequency and descriptive statistics for the sample, which focused on the demographic variables from my questionnaire. I also used PSPP to compute scale-level distribution statistics for the MRMS and the IRMAS data.

Phase 2: Exploratory Factor Analyses. I began phase 2 with a preliminary investigation. Using the MRMS data, I generated a polychoric correlation matrix in R. I tested whether the data were suitable for factor analysis using the Kaiser-Meyer-Olkin measure of sampling adequacy and Bartlett's test of sphericity.

Once I had confirmation that the polychoric correlation matrix was suitable for factor analysis, I loaded the prepped MRMS data into FACTOR, along with a text file containing the variable labels. Using FACTOR, I generated two sets of outputs that included preliminary indicators of (a) whether there was a substantial common factor underlying the data and (b) how many factors to specify for the subsequent EFA models. Specifically, I focused on (a) eigenvalues, (b) the Hull Common Parts Accounted For (Hull-CAF; Lorenzo-Seva et al., 2011) analysis, and the Parallel Analysis - Minimum Rank Factor Analysis (PA-MRFA; Timmerman and Lorenzo-Seva, 2011).

I computed two EFA models based on the MRMS data. The first was a one-factor model, where the data was restricted to loading on a single factor. The second was a bifactor model, where the data was able, but not forced, to load on a general factor and specific group factors. Once I had determined that both models fit the data well, I focused more intently on the bifactor model. All reported model-fit indices were included in the EFA outputs produced by FACTOR.

When analysing the bifactor model, I was interested in two indicators of general factor strength: the explained common variance (ECV), and omega hierarchical(OmegaH). The ECV was included in the bifactor model output produced by FACTOR. I calculated OmegaH in the R environment, and elaborate more on how I calculated it towards the end of this chapter.

Although my primary focus was on the general factor, I was also interested in whether the specific group factors lent themselves to substantive interpretation. Substantive interpretation is part of why a particular model should be considered plausible. Under a bifactor framework, specific group factors capture variance over and above that which is common among all items and therefore, elucidates avenues for reducing nuisance variation.

The procedures outlined in phase 2 were then repeated for (a) the IRMAS data and (b) the combined item pool of the IRMAS and the MRMS. However, for the combined item pool, there was one key difference: instead of loading the raw data to FACTOR, I loaded a pre-generated polychoric correlation matrix that I computed in R. When FACTOR looks at raw data, it assumes that all items are rated on the same Likert scale, and generates the polychoric correlations based on that assumption. To overcome this limitation, I made use of the polycor() function available in the *psych* package (Revelle, 2021), which is used to calculate polychoric correlation matrices in the R environment. The function has a specific argument that allows users to indicate whether items have been rated on different rating scales.

Phase 3: Model Comparisons. I began phase 3 by creating an Excel spreadsheet containing the FA and IRT parameters for the one-factor and bifactor models I had generated based on the MRMS data. I focused on the FA model parameters first, particularly (a) the one-factor model loadings and (b) the general factor loadings of the bifactor model. I quantified the extent to which the parameters had distorted under the one-factor model based on a written description by Rodriguez et al. (2016), which I have represented as Equation 1;

$$B_i = \frac{\lambda_{1F} - \lambda_{GEN}}{\lambda_{GEN}} \tag{1}$$

where the item-level bias (B_i) was calculated by subtracting the item loading on the general factor of the bifactor model (λ_{GEN}) from the same item's loading on the one-factor model (λ_{1F}), and divided by the item loading on the general factor (Rodriguez et al., 2016). My interpretation was structured as follows: $B_i > 15\% =$ unacceptable.

Next, I focused on the IRT transforms of the EFA models to determine the extent to which interpretation under a unidimensional model would rely on distorted item parameters. Item-level changes in slope (Δ_α) were calculated through simple subtraction, captured in Equation 2:

$$\Delta_\alpha = \alpha_{Uni} - \alpha_{Gen} \tag{2}$$

where the item slope for the general dimension of the bifactor IRT model (α_{Gen}) was subtracted from the item slope in the unidimensional IRT model (α_{Uni}). I based my interpretation of what constituted a small versus a big difference in item slope based on a series of slope descriptors by Baker and Kim (2017, p. 26), which I covered in more detail in my literature review. My interpretation was structured as follows: small/negligible ($\Delta_\alpha \leq .21$); moderate/noteworthy ($.22 \leq \Delta_\alpha \leq .37$); large/substantial ($\Delta_\alpha \geq .38$).

Once I had calculated the item slope differences, I had the necessary data to analyse to determine whether the MRMS and IRMAS could be considered essentially unidimensional (Reise et al., 2015). According to Reise et al. (2015), data that presents as multidimensional can be modelled and treated as *essentially unidimensional* if (a) there is a strong common factor underlying the items, and (b) the multidimensionality does not distort the model parameters to an unacceptable degree. If there is a strong general factor underlying the data, and model parameters under a one-factor model (FA) or unidimensional model (IRT) are not significantly distorted, it provides a strong motivation for treating the data as essentially unidimensional, despite the presence of multidimensionality (Reise et al., 2015).

The IRT parameters were included in their respective FACTOR outputs for the models generated for the IRMAS and MRMS individually. However, I had to transform the FA item loadings to IRT item slopes for the combined item pool myself. FACTOR did not produce the transform in the output in this case, likely because I used a polychoric correlation matrix generated in R instead of the raw data. I transformed the factor loadings to IRT slopes of a normal-ogive model based on the following equation;

$$\alpha = \frac{\lambda}{\sqrt{1 - h^2}} \tag{3}$$

where the slope (α) for each item was calculated by taking the item's loading (λ) in a particular model and dividing it by the square root of the item's uniqueness (u^2). An item's uniqueness is calculated by subtracting an item's communality (h^2) from 1. Item communalities are model dependent and included in FACTOR outputs. Equation 3 is based on a more technical version outlined by Reise (2012), which I have simplified to make it more accessible to lay audiences.

Once I had transformed the FA loadings to IRT slopes, I calculated the item slope difference in the same way I did for the IRMAS and MRMS models. By the end of phase 3, I had the necessary data to analyse to determine whether the items of the MRMS and the items of the IRMAS were tapping into the same latent construct.

Phases of Data Analysis

My phases of data analysis were broken up according to the datasets I worked with. I first looked at the MRMS data, then the IRMAS data, and finally the complete data set of the combined item pool.

MRMS Dimensionality Assessment. The aim was to determine whether the MRMS could be considered essentially unidimensional for IRT applications. I was looking for the following pieces of evidence: (a) a high first-to-second eigenvalue ratio, (b) a well-fitting one-factor model, (c) a well-fitting bifactor model with a strong general factor, and (d) an acceptable difference in item-level parameter estimation between the one-factor model and the bifactor model.

IRMAS Dimensionality assessment. The aim was to determine whether the IRMAS could be considered essentially unidimensional for IRT applications. I was looking for the following pieces of evidence: (a) a high first-to-second eigenvalue ratio, (b) a well-fitting one-factor model, (c) a well-fitting bifactor model with a strong general factor, and (d) an acceptable difference in item-level parameter estimation between the one-factor model and the bifactor model.

Combined Item Pool. The aim was to determine whether the items of the IRMAS and MRMS tap into the same latent construct. I was looking for three pieces of evidence: (a) a high correlation between the IRMAS and MRMS, (b) when fitted to a bifactor model, items from both measures should load well on the same general factor, and the general factor should be strong. In addition, an even stronger indication that the items are tapping into the same construct would be if the data from the combined item pool could be characterised as essentially unidimensional under an FA and an IRT framework.

Calculating Omega Hierarchical

Omega hierarchical (omegaH) is a key indicator of general factor strength when working with a bifactor model (Reise, Scheines, et al., 2013; Rodriguez et al., 2016). However, FACTOR does not include omegaH in bifactor modelling outputs. I therefore calculated the omegaH values for each of my bifactor models using R.

My R code for the entire project is available in Appendix E. Below is a snippet from the project code, which highlights the portion of code used to calculate omegaH for the MRMS. The same code was used to calculate omegaH for the IRMAS bifactor model, and the bifactor model of the combined item pool of the IRMAS and MRMS, apart from the file names and object names unique to each data set.

Listing 1: Omega Hierarchical Code: MRMS

```
mrms_comm<-read.csv("mrms_comm.csv", header = TRUE, sep = ",", row.names = 1)
mrms_PM<-read.csv("mrms_PM.csv", header = TRUE, sep = ",", row.names = 1)
genload <- mrms_PM[, 1]
grpload <- mrms_PM[, 2:4]
sum_genload_sq <- (sum(genload))^2
sum_grpload_sq <- (sum(grpload))^2
uniq <- sum(mrms_comm[, "u2"])
Tot_Var <- (sum_genload_sq + sum_grpload_sq + uniq)
omegah_mrms = sum_genload_sq/Tot_Var
omegah_mrms
```

I have also included the output for the code snippet included here in Appendix E, to provide a clear example of what the data components looked like that I imported into R for the omegaH calculation.

The steps were as follows. First, I saved the pattern matrix from my FACTOR output as a .csv file. Second, I saved the item communalities (h^2) in a separate .csv file, added a second column and calculated the uniqueness (u^2) for each item. The calculation was based on one provided by Reise et al. (2018), $u^2 = 1 - h^2$. I imported both .csv files into R, and then created multiple objects to match the components outlined in Equation 4.

I want to acknowledge two key sources that enabled me to understand the components of omegaH well enough to compute it in R. First, Reise et al. (2018, p. 691) provide a simplified equation for calculating omega hierarchical in the *Wiley Handbook of Psychometric Testing*, as follows:

$$\omega_h = \frac{\left(\sum \lambda_{Gen}\right)^2}{\left(\sum \lambda_{Gen}\right)^2 + \left(\sum \lambda_{Group}\right)^2 + \sum \left(1 - h^2\right)} \tag{4}$$

where λ_{Gen} refers to all the general factor loadings in the bifactor solution, and λ_{Group} refers to all the group factor loadings in the bifactor solution, and h^2 refers to all the item communalities in the bifactor solution.

Second, to understand what such a calculation should look like in R, I located the full code behind the function omegah() function available in the *psych* package by Revelle (2021), and printed out four copies. I broke down the function code into segments using pens, pencils and highlighters. Most of the code in the omegah() function from the *psych* package was dedicated to parsing the data, calculating a bifactor model based on a Shmid-Leiman transformation, calculating related indices, and feeding in information from other functions. Based on my learnings from code, I wrote my own basic R code using object labels that made intuitive sense to me.

I originally considered using the omegah() function available in the *psych* package by Revelle (2021). I, however, opted to write my own code because the function by Revelle calculates omegaH based on an independently calculated bifactor model, and I wanted the omegaH values I reported to be calculated based on the exploratory bifactor models I had fitted to the data.

Conclusion for Methodology

My goal with my methodology was to be as clear and detailed as possible, to facilitate replication should a future researcher desire to do so. In the following chapter, I report my results according to the structure outlined in the phases of data analysis section of this chapter.

CHAPTER 4: RESULTS

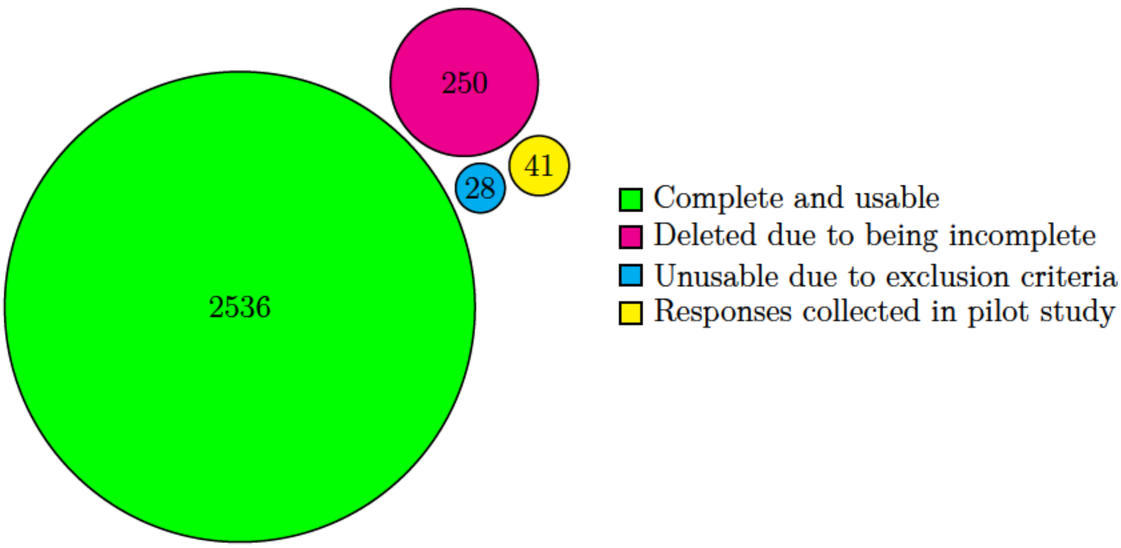
The results chapter is presented in four main parts. First, I report the sample statistics. Second, I present the results of the dimensionality assessment of the MRMS. Third, I present the results of the dimensionality assessment of the IRMA. Finally, I present the results of the dimensionality assessment conducted on the combined item pool of the IRMA and MRMS.

Sample Statistics

The study sample consisted of 2,536 participants. However, a total of 2,814 responses were received between 22 March 2021 and 21 April 2021. Figure 1 below provides a breakdown of the received responses. Of the 2,814 received responses, 28 participants met the exclusion criteria and were unable to progress to the RMA questionnaire. A further 250 responses were incomplete, which was about 10% of the total number of responses received. An incomplete response was one where a participant met the inclusion criteria for the survey and started to respond but did not finish the questionnaire in its entirety. All incomplete response sets were taken as a sign of withdrawal from the study and deleted. Despite the excluded and unusable responses, the response rate for this study was 12.76%.

Figure 1

Graphical Representation of the Number and Usability of Received Responses

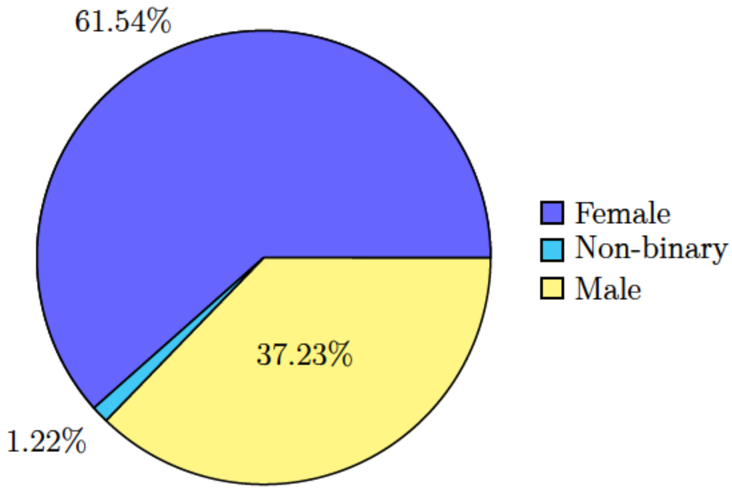


Note. Pilot study responses have been included in this representation to give a sense of scale.

The mean age of the sample was 20.19 years (SD = 2.37), and participants ranged from 18 to 51 years old. Of the 2,536 participants, 61.54% indicated they identified as female, 37.23% indicated they identified as male, and 1.22% indicated they identified as non-binary. Consistent with RMA studies that have sampled from student populations, a two-thirds majority of responses were received from women (see Figure 2 below).

Figure 2

Pie Chart: Gender Composition of the Study Sample



Note. Please be mindful that "female" is not limited to cisgender women, and "male" is not limited to cisgender men.

In total, 26 different home languages were recorded by the sample. Only half the participants (n = 1,276) indicated English as their home language. Afrikaans was the second most commonly reported home language (34.4%; n = 874), and the third was isiXhosa (5.4%; n = 137). All 11 of South Africa's official languages were recorded by the sample; the exact frequencies are reported in Table 2 on the following page. A further 14 home languages were reported by the sample, comprising 1.9% of the responses received, as follows: Shona (n = 21), German (n = 12), Mandarin (n = 4), and Lingala (n = 2), as well as Swahili, French, Spanish, Arabic, Dutch, Malagasy, Hindi, Oshindongo, Oshiwambo, and Oshikwanyama, with one response each.

Table 2

Frequency Table: Home Languages Recorded by the Sample.

Language	Frequency	%
English	1276	50.32
Afrikaans	874	34.46
isiXhosa	137	5.4
isiZulu	63	2.48
Sepedi[a]	32	1.26
Setswana	29	1.14
Sesotho	28	1.1
Tshivenda	19	0.75
Xitsonga	14	0.55
isiNdebele	10	0.39
siSwati	4	0.16
Other	49	1.9

Note. The specific languages reported here are all official languages of South Africa. The "Other" category comprises 14 home languages, for which further details are available in the main body of the text.

[a]Sepedi is also known as Sesotho sa Leboa/Northern Sotho.

Male Rape Myth Scale Dimensionality Investigation Results

The following section contains the results of my dimensionality investigation into the MRMS. In this section, I report the following: (a) traditional psychometrics for the MRMS; (b) exploratory factor analyses of the MRMS data, which include a one-factor model and a bifactor model fitted to the data; and (c) I present the results of a model comparison between the one-factor and bifactor model as well as IRT transforms of each, to elucidate potential item-level parameter distortion.

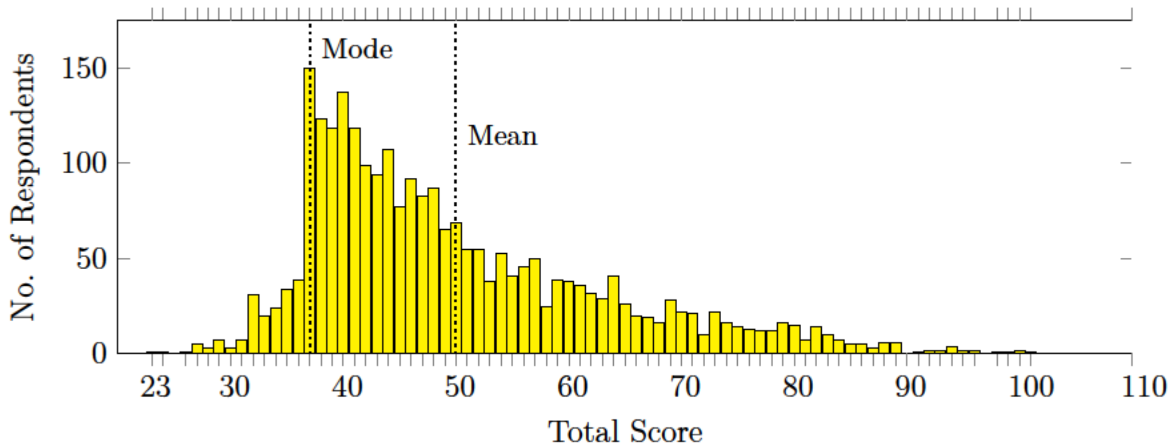
MRMS: Traditional Psychometrics

Cronbach's coefficient alpha was excellent for the MRMS in the piloting stage (n = 41, $\alpha = .9$), as well as for the main study (n = 2,536, $\alpha = .9$). Overall, the scale demonstrated excellent reliability (GLB = .97, $\omega = .95$). The MRMS data was non-normally distributed (n = 2,536, $M = 49.86$, $SD = 13.13$), with a skewness of 1.15 (SE = 1.15) and a kurtosis of 1.30 (SE = .10). The non-normal distribution of the data is visible in Figure 3 below, characterised by a positive skew and a single peak.

The distribution trends at the scale level were also observed at the item level, with no notable exceptions. I observed low item means across all 22 MRMS items (range: 1.35 to 2.77). Item-level skewness ranged from 0.45 to 3.04, and kurtosis (zero-centred) ranged from -1.02 to 9.52. Detailed item-level distribution statistics are available for all 22 MRMS items in Appendix F.

Figure 3

Histogram: Distribution of MRMS Total Scores

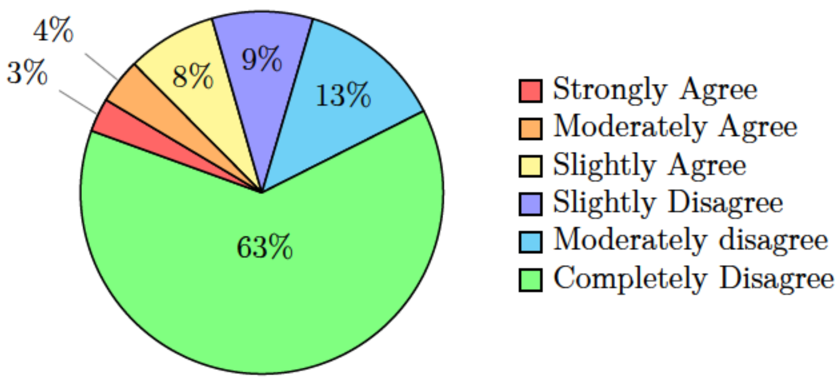


Note. Mean = 49.86, Mode = 37

Figure 4 below shows the percentage of the total responses that fell into each response category. Most responses fell within the first three response categories, which captured disagreement with the rape myth statements presented. Overall, 15% of the responses indicated a degree of agreement with the rape myth statements.

Figure 4

Pie Chart: MRMS Response Frequency by Category



Note. This pie chart was generated by summing the responses in each category across all MRMS items. Items 1, 6 and 19 were reverse-scored prior to summation.

MRMS: Exploratory Factor Analyses

Preliminary Statistics. The exploratory factor analysis for the MRMS was based on a polychoric correlation matrix of the response data. The correlation matrix demonstrated suitability for scale and item-level factor analysis. Bartlett's test of sphericity was significant ($X^2(231) = 29091.2$, $p < .00001$), and the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy result was marvellous (KMO = .96, CI = .95-.96). All items recording an Item-MSA value of .92 or higher, indicating excellent sampling adequacy at the item level.

Table 3

MRMS Eigenvalues

Variable	Eigenvalue	Proportion of variance	Cumulative variance
1	10.40	.47	.47
2	1.38	.06	.54
3	1.07	.05	.58

Note. Only eigenvalues > 1 are included here.

Table 3 on the previous page captures the first five eigenvalues. The first to second eigenvalue ratio was 7.5:1, over double the recommended criterion of 3:1 (Reise, Ventura, et al., 2011). In addition, the PA-MRFA and Hull-CAF procedures both advised one dimension underlying the MRMS data.

MRMS One-factor Model. The one-factor model was an excellent fit to the data (RMSEA = .047, CI = .045 - .048; GFI .989, CI = .989 - .991; AGFI = .988, CI = .987 - .990). Few residuals were unaccounted for by the model (RMSR = .052, CI = .050 - .052; WRMR =.046, CI = .044 - .047).

The one-factor model showed a clear common factor among the MRMS items. The loadings for the unidimensional FA model are recorded in Table 4 on the next page. The common factor had a substantial loading of .3 or higher on all the MRMS items. Item 19 had the lowest factor loading at .42, and item 13 had the highest common factor loading at .85. A total of 19 MRMS items had over half their variance accounted for by the common factor (loading ≥ .5). At this stage, no item presented as being particularly problematic, and all items appeared to be loading well on the same common construct.

Table 4

Factor Loadings and Item Communalities of the MRMS One-Factor and Bifactor Models.

Item no.	One-factor[a]		Bifactor[b]			
	λ	h^2	λ_{GEN}	λ_{F1}	λ_{F2}	λ_{F3}
6 (R)	.48	.23	.53	.53		
19 (R)	.42	.18	.41	.43		
1 (R)	.58	.34	.61	.42		
4	.72	.52	.52	.30	.73	
3	.69	.47	.50		.67	
8	.83	.69	.77		.61	-.32
5	.64	.41	.46		.59	
2	.62	.38	.42		.59	
11	.82	.67	.80		.49	
9	.81	.66	.71		.46	
7	.48	.23	.34		.41	
17	.53	.28	.32		.40	
20	.68	.46	.58		.31	
13	.85	.73	.83		.30	
14	.66	.44	.60			.53
16	.63	.39	.51			.40
18	.79	.62	.66			.36
15	.73	.53	.64			.31
12	.63	.40	.55			.31
22	.70	.50	.57			
21	.67	.45	.55			
10	.58	.33	.44			

Note. Loadings below .3 have been omitted. Reverse-scored items are denoted with an (R). h^2 = item communalities.

[a]One-factor model. Matrix = polychoric correlations. Factor estimation method = robust unweighted least squares (RULS). [b]Bifactor Model. Matrix = polychoric correlations. Factor estimation method = RULS. Rotation = Robust Promin. Bifactor modelling procedure = pure exploratory bifactor.

Dimensionality indices supported interpreting the MRMS data as unidimensional under the one-factor model. With a UniCo value greater than .95 (UniCo = .97, CI =

.96 - 0.99), an ECV value well above .85 (ECV = .91, CI = .91 - 0.92), and MIREAL below .3 (MIREAL = .17, CI = .15 - .18), the numbers looked outstanding.

Although strong preliminary evidence suggested that the MRMS was unidimensional, the evidence needed to be placed in further context. While MIREAL is model-independent, ECV and UniCo are model-dependent and limited by the model type. To get a more accurate sense of the dimensionality of the MRMS, it was necessary to compute a comparison model.

MRMS Bifactor Model. My factor retention indices of choice, the PA-MRFA and Hull-CAF procedures, both indicated only one dimension underlying the data. However, the MRMS data had three eigenvalues greater than 1, which provided grounds for specifying three group factors for the bifactor comparison model, based on Kaiser's criterion (Braeken & Van Assen, 2017). Kaiser's criterion has been the subject of much criticism but nonetheless continues to be commonly cited (Braeken & Van Assen, 2017).

Three disproportionately sized group factors loaded onto the MRMS items in the resulting bifactor solution. The pattern matrix for the bifactor solution is reported in Table 4 on the previous page. Group factor 1 (F1) had four items, group factor 2 (F2) had 11 items, and group factor 3 (F3) had six items. The bifactor model was an excellent fit to the data (GFI .998, CI = .998 - .998; AGFI = .997, CI = .997 - .997). The GFI index being so close to 1 indicated that the common variance in the data was almost entirely accounted for by the model.

General Factor. The general factor loaded substantially on all the items ($\lambda_{GEN} \geq .3$), providing evidence of a common factor underlying the MRMS items. However, six items had a general factor loading below .5. Nonetheless, the general factor was dominant relative to the group factors. The general factor accounted for roughly two-thirds of the common variance across all the MRMS items (ECV = .61, CI = .60 - .63), with the general factor saturation further confirmed by an OmegaH value of .58. Furthermore, the MRMS appeared to be tending towards unidimensional, with a UniCo value of .79 (CI = .77 - .81).

Table 5

MRMS inter-factor correlation matrix

Factor	Factor Correlations				Variance
	F1	F2	F3	GF	
1	1.00				0.83
2	-.08	1.00			2.85
3	.09	.68	1.00		0.83
GF	.00	.00	.00	1.00	7.28

Note. F1 = group factor 1. F2 = group factor 2. F3 = group factor 3. GF = general factor. Variance = variance explained by the rotated factors.

Group Factors. Although the MRMS was intended to be unidimensional, the bifactor model elucidated some multidimensionality. I explored the group factors and item-level variation in more detail to better understand where the multidimensionality was introduced.

All the group factors were constrained to be orthogonal to the general factor in the modelling stage, meaning the correlations between the group factors and general factor were set to zero as part of the modelling process. Therefore, factor correlations of interest are those between the group factors. The inter-factor correlations are captured below in Table 5. F2 and F3 were highly correlated, but F1 had close to zero correlation with the other two factors in the solution.

I argue that F1 presents as a methodology factor. F1 loaded on all three of the reverse-scored items in the MRMS, item 1 (IECV = .69, IUniCO = .91, IREAL = .38), item 6 (IECV = .50, IUniCO = .71, IREAL = .46) and item 19 (IECV = .48, IUniCO = .67, IREAL = .40). Item 1 is "It is a terrible experience for a man to be raped by a woman", Item 6 is "Most men who are raped by a woman are very upset by the incident," and item 19 is "Most men would not enjoy being raped by a woman". In addition to the reverse scoring setting these items apart from the rest of the scale, all three items are similarly worded and arguably content repeats. Out of all 22 MRMS items, only the reverse-scored items had an I-REAL value greater than .3, which confirmed that the items violated the assumption of local independence. A methodology factor based on scoring differences would also account for why F1 had close to zero correlation with F2 and F3.

However, item 4 was not reverse scored, yet also had a substantial F1 loading. As a counterpoint, item 4 had cross-loadings, with F2 having a much higher loading on the item by comparison. Therefore, I maintain that F1 presents as a methodology factor, with the F1 items grouping due to their reverse-scored phrasing.

Table 6

Group Factor 2 Item content and dimensionality indices

Item no. and Content	I-ECV	I-UniCo	I-REAL
13. Most men who are raped by a woman are somewhat to blame for not being more careful.	.87	.99	.05
11. Most men who are raped by a man are somewhat to blame for not escaping or fighting off the man.	.81	.93	.02
9. If a man engages in necking and petting and he lets things get out of hand, it is his own fault if his partner forces sex on him.	.74	.92	.00
8. Most men who are raped by a woman are somewhat to blame for not escaping or fighting off the woman.	.74	.84	.10
20. Men who parade around nude in a locker room are asking for trouble.	.70	.96	.14
7. Many men claim rape if they have consented to homosexual relations but have changed their minds afterwards.	.47	.57	.02
3. Any healthy man can successfully resist a rapist if he really wants to.	.46	.49	.15
5. A man can enjoy sex even if it is being forced upon him.	.43	.50	.21
4. If a man obtained an erection while being raped it probably means that he started to enjoy it.	.42	.44	.26
2. The extent of a man's resistance should be a major factor in determining if he was raped.	.41	.45	.04
17. Women who rape men are sexually frustrated individuals.	.29	.54	.18

Note. I-ECV = item level explained common variance. I-UniCo = item level unidimensional congruence. I-REAL = Item level residual absolute loadings.

Factor 2 (F2) was the largest group factor, consisting of half the items in the MRMS. F2 accounted for 24% of the common variance (see Table 6 above for item content). F2 had the highest loadings on rape myths related to victim-blaming (items 2, 3, 8 and 11) and rape myths that conflate consent with physiological arousal (items 4 and 5).

In terms of dimensionality indices, some noteworthy items were item 4, item 5 and item 17. Item 4 and item 5 both had an I-REAL value very close to .3, indicating that the items come close to violating the assumption of local independence. Item 17 had an I-ECV value substantially lower than the rest of the items in the scale, at .29. Item 17 also had one of the lowest general factor loadings of all the MRMS items ($_{Gen}$= .32). Six out of the 11 items on this factor had an IUniCo value well below .8, which suggests that most of the items in F2 are introducing item-level multidimensionality to the scale.

Table 7

Group Factor 3 Item content and dimensionality indices

Item no. and content	I-ECV	I-UniCo	I-REAL
8. Most men who are raped by a woman are somewhat to blame for not escaping or fighting off the woman.	.74	.84	.10
15. Most men who have been raped have a history of promiscuity.	.71	.97	.19
12. A man who has been raped has lost his manhood.	.68	.95	.20
18. A man who allows himself to be raped by another man is probably homosexual.	.64	.96	.26
14. If a man told me that he had been raped by another man, I would suspect that he is homosexual.	.61	.79	.27
16. No self-respecting man would admit to being raped.	.57	.86	.17

Note. I-ECV = item level explained common variance; I-UniCo = item level unidimensional congruence; I-REAL = item level residual absolute loadings.

Factor 3 (F3) loaded on a third of the MRMS items, however most of the loadings were small (see Table 7). Three items (item 8, item 12 and item 15) barely met the cut-off criteria of .3. However, the general factor loaded well on all F3 items, with general factor loadings ranging from .51 to .64. The F3 items had I-ECV values upwards of .57, and six F3 items had excellent I-UniCo values ranging from .79 to .97, which suggests the items introduce little multidimensionality to the scale, especially in comparison to F2.

Item Analysis. In this section, I cover items that would typically be flagged as complicating model interpretation for different reasons and were potentially worth removing. Specifically, I investigated items that presented with cross-loadings and items that did not load on a specific group factor, all of which have been captured in Table 8 on the next page. I was interested in whether any of these items should be removed from analysis when analysing the combined item pool of the IRMA and MRMS.

Only two items had cross-loadings, item 4 (I-ECV = .42, I-UniCo = .44, I-REAL= .26) and item 8 (I-ECV = .74, I-UNiCo= .84, I-REAL = .10). Items with cross-loadings are problematic because they violate the assumption of local independence, and can result in parameter distortion. However, the I-REAL did not flag either of the items as substantially violating the assumption of local independence, as reported I-REAL for both items was below .3. Furthermore, the dominant loadings were very high, and the cross-loadings were minor. Cross-loadings tend to be more problematic when they are quite similar.

Table 8

Content and dimensionality indices for items without a specific group factor

Item no. and content	I-ECV	I-UniCo	I-REAL
10. Male rape is usually committed by homosexual men.	.56	.94	.11
21. Male rape is more serious when the victim is heterosexual than when the victim is homosexual.	.63	.97	.15
22. I would find it difficult to believe a man who told me he was raped by a woman.	.63	.98	.00

Note. I-ECV = item level explained common variance; I-UniCo = item level unidimensional congruence; I-REAL = item level residual absolute loadings.

Three items did not load on a specific group factor, items 10, 21 and 22 (see Table 8). The general factor loadings were not particularly high on these items, however all three had good I-ECV values above .5, and excellent UniCo values above .9. Researchers interested in assessing the reliability and validity of subscales usually remove items that do not load on a specific factor from further analysis. However, the general construct is the primary focus of my research. Given that the general factor loadings for all three items were substantial, and all three items had acceptable I-ECV values, I opted to retain the items in further analyses.

MRMS: Model Comparison

The model comparison is based on the parameters presented in Table 9 on the next page. Starting with the factor analytic parameters, there was a large difference between the factor loadings in the one-factor model and the general factor loadings in the bifactor model. The relative bias in item loadings across both factor analytic models ranged from -10% to 66%, with 12 of the items falling outside the acceptable biasing range of -15% to 15%. I calculated the mean bias across the model items to be 21%.

Moving on to the IRT parameters, the differences in slopes between the unidimensional normal ogive model and the general dimension of the multidimensional IRT model were small. Across the items, the difference in slopes ranged from -.23 to .29. Item 6 and item 11 displayed the greatest parameter distortion, with their slopes across the two IRT models differing by .25 and .29 respectively. The three items with the most discriminating power under the unidimensional model showed an increase in discriminatory power under the bifactor IRT model.

The items with the most discriminating power in the bifactor model (i.e. >1), were among the least distorted in the unidimensional model, with the exception of item 11. Items 13, 8, 11, 9 and 18 showed very high discriminatory power under both the unidimensional IRT model as well as the bifactor IRT model. Out of those 5 items, item 18 was the most different in terms of item content, "A man who allows himself to be raped by another man is probably homosexual". The other 4 most-discriminating items were all characterised by victim-blaming.

Table 9

Comparison table of changes in item loadings and items slopes across the common dimensions of the one-factor and bifactor MRMS solutions.

	FA			IRT		
Item no.	λ_{1F}	λ_{GEN}	Relative Bias	α_{Uni}	α_{Gen}	Slope difference
13	.85	.83	3%	1.63	1.76	-0.13
8	.83	.77	8%	1.51	1.69	-0.19
11	.82	.80	3%	1.43	1.72	-0.29
9	.81	.71	15%	1.40	1.24	0.16
18	.79	.66	19%	1.27	1.17	0.09
15	.73	.64	14%	1.07	0.99	0.08
4	.72	.52	39%	1.03	0.86	0.17
22	.70	.57	23%	0.99	0.83	0.17
3	.69	.50	37%	0.94	0.74	0.20
20	.68	.58	18%	0.92	0.80	0.13
21	.67	.55	23%	0.91	0.75	0.16
14	.66	.60	11%	0.89	0.93	-0.04
5	.64	.46	39%	0.83	0.64	0.19
12	.63	.55	15%	0.81	0.74	0.07
16	.63	.51	22%	0.81	0.70	0.11
2	.62	.42	46%	0.79	0.57	0.22
1 (R)	.58	.61	-5%	0.72	0.91	-0.20
17	.53	.32	66%	0.63	0.40	0.23
10	.58	.44	31%	0.71	0.55	0.16
7	.48	.34	40%	0.55	0.40	0.15
6 (R)	.48	.53	-10%	0.54	0.80	-0.25
19 (R)	.42	.41	3%	0.47	0.52	-0.05

Note. λ_{1F} = item-level factor loadings, one-factor FA model; λ_{GEN} = item-level factor loadings, general factor of the bifactor FA model; α_{Uni} = item slope/discrimination, unidimensional IRT model; α_{Gen} = item slope/discrimination, general dimension of the bifactor IRT model. Items are ordered by slope size under the unidimensional IRT model, in descending order. Reverse-scored items are denoted with an (R).

Illinois Rape Myth Acceptance Scale Dimensionality Investigation Results

The following section contains the results of my dimensionality investigation into the IRMA. In this section I report the following: (a) traditional psychometrics for the

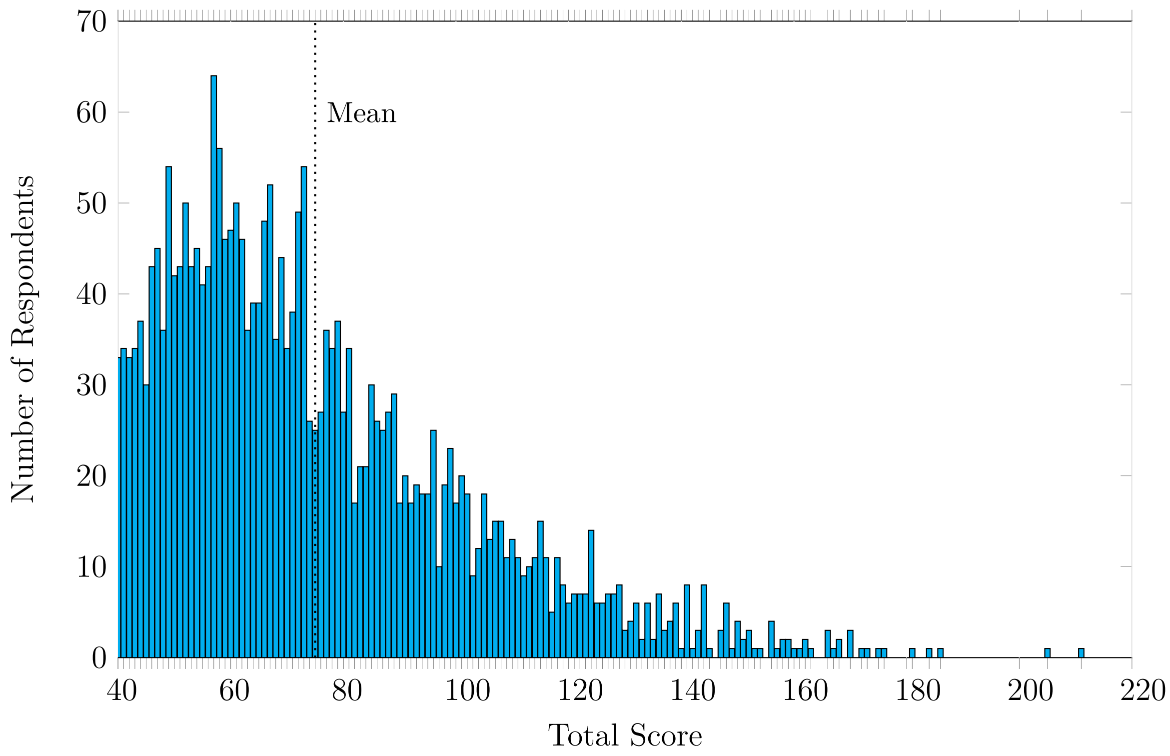
IRMA; (b) an exploratory factor analysis of the IRMA data, which includes a one-factor model and a bifactor model fitted to the data; and (c) I present the results of a model comparison between the one-factor and bifactor model as well as IRT transforms of each, to elucidate potential item-level parameter distortion.

IRMA: Traditional Psychometrics

Cronbach's coefficient alpha was excellent in the piloting stage (n = 41, α = .91), as well as for the 2,536 responses gathered in the main study (α = .94, GLB = .99, ω = .97). The IRMA data was non-normally distributed with a positive skew of 1.5 (SE = .05) and kurtosis of 1.30 (SE = .10). Figure 5 illustrates the distribution of the IRMA total scores.

Figure 5

Histogram showing the skewed distribution of IRMA Total Scores

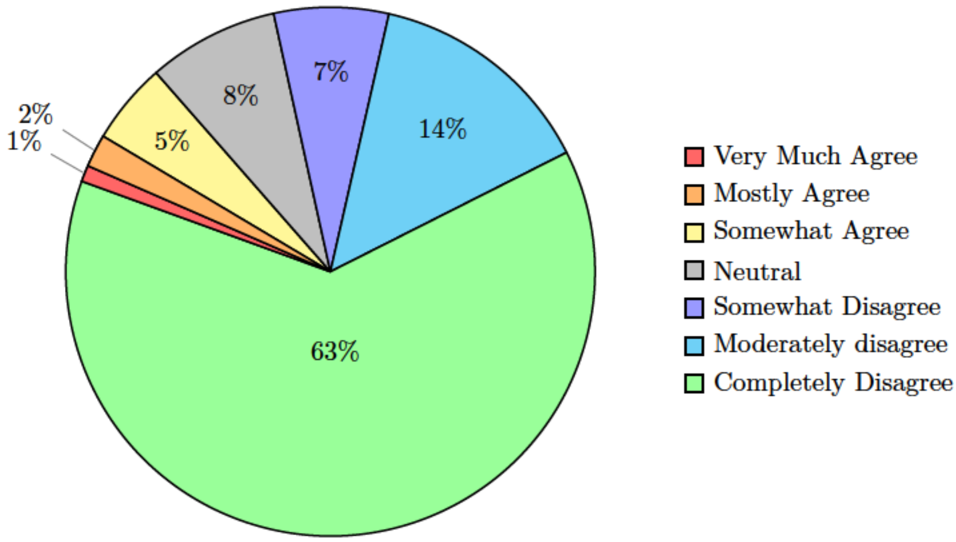


Note. N= 2536, Mean = 75.38 (SE = .53), SD = 26.87, Min = 40, Max = 211, Range = 171.

Item means ranged from 1.07 to 3.83, with an average of 1.88. Most of the items displayed a positive skew, with an average skew of 2.27 (range: -.10 to 8.13). However, there were three items with notably different distributions to the other IRMA items, with higher than average means, little to no skew (S), and negative kurtosis (K). The items that stood out at this stage in the analysis were item 3 (M = 3.38, SD = 1.97, S = 0.23, K = -1.18), item 5 (M = 3.83, SD = 1.56, S = -0.10, K = -0.54), and item 20 (M = 3.63, SD = 1.85, S = 0.06, K = -1.04). A table capturing item-level distribution statistics and response frequencies for all IRMA items is available in Appendix G.

Figure 6

IRMA Response Frequency by Category



Note. This pie chart was generated by summing the responses in each category across all 40 IRMA items. The data from the five filler items were removed prior to the summation.

Figure 6 shows the percentage of responses that fell into each response category. Overall, there was 8% agreement with the rape myth statements in the IRMA. Most of the responses indicated some level of disagreement with the items (84%). More responses were recorded in the neutral category (8.18%) than the remaining three categories combined (7.75%). At the item level, the first Likert category had the highest response frequency for 38 of the IRMA items, which indicated the strongest level of disagreement with rape myths. Two exceptions were item 5 and item 20, with response frequencies for both items peaking in the neutral category. Item 5 was "Women who are caught having an illicit affair sometimes claim it was rape", for which 32% of respondents selected the neutral response category. Item 20 was "Rapists are usually sexually frustrated individuals", with 24% of responses in the neutral category.

IRMA: Exploratory Factor Analysis

Preliminary Statistics. A polychoric correlation matrix formed the foundation for the exploratory factor analyses of the IRMA data. At both the scale and item level the correlation matrix demonstrated suitability for factor analysis. Bartlett's test of sphericity was significant ($X^2(780) = 29022.2$, $p = .00001$). The Kaiser-Meyer-Olkin measure of sampling adequacy result was marvellous (KMO = .97, CI = .93-.96). The items also individually demonstrated very good sampling adequacy, with all items recording an Item-MSA value of .91 or higher.

Table 10

IRMA eigenvalues greater than 1

Variable	Eigenvalue	Proportion of variance	Cumulative variance
1	18.24	.46	.46
2	2.21	.06	.51
3	2.08	.05	.56
4	1.49	.04	.60
5	1.27	.03	
6	1.04	.03	

Note. Only eigenvalues greater than 1 have been reported here.

The ratio of the first to second eigenvalue was 9:1. At well over the criterion of 3:1 recommended by Reise et al. (2011), the high eigenvalue ratio indicated a substantial general dimension worth continuing to investigate (see Table 10). The Hull-CAF procedure indicated only one common dimension underlying the IRMA. By contrast, the results of the PA-MRFA advised 3 dimensions for the IRMA. The Hull-CAF provided grounds for fitting a one-factor model to the data, and the PA-MRFA provided grounds for fitting a bifactor model with three group factors to the data.

IRMA One-Factor Model. The one-factor solution is captured in Table 11. Overall, the model appeared to be a good fit (CFI = .984, CI: .983 - .986; GFI = .978, CI: .976 - .980; RMSEA = .058, CI: .057 - .059; WRMR = .067, CI = .064 - .068). Most

Table 11

IRMA EFA Models

Item	One-factor[a]		Bifactor[b]			
	λ	h^2	λ_{GEN}	λ_{F1}	λ_{F2}	λ_{F3}
irma18	.65	.42	.64	.56		
irma02	.57	.33	.58	.49		
irma44	.63	.40	.58	.45		
irma08	.63	.39	.61	.40		
irma23	.71	.50	.66	.39		
irma16	.64	.40	.41	.36		
irma07	.65	.42	.41	.36	.30	
irma05	.46	.21		.35		
irma35	.63	.39	.50	-.35	.63	
irma13	.62	.39	.46		.62	
irma10	.70	.49	.53		.59	
irma14	.75	.56	.54		.58	
irma15	.77	.59	.53		.58	
irma28	.64	.41	.51		.58	
irma09	.52	.27	.38		.55	
irma22	.75	.56	.59		.54	
irma27	.62	.39	.55	-.31	.52	
irma41	.78	.61	.57		.49	
irma17	.74	.55	.62		.45	
irma01	.64	.41	.45		.44	
irma29	.79	.62	.64		.42	
irma25	.71	.50	.51		.41	.32
irma45	.78	.61	.69		.39	
irma19	.72	.52	.56		.39	
irma36	.79	.63	.68		.37	
irma12	.71	.50	.63		.35	
irma33	.77	.59	.60		.33	
irma31	.78	.60	.55	.32	.33	
irma43	.63	.39	.50		.33	
irma26	.62	.38	.58		.32	
irma38	.76	.57	.76		.30	
irma42	.45	.21	.57			.56
irma03	.36	.13	.46			.52
irma20	.42	.18	.47			.49
irma39	.62	.39	.63			.37
irma30	.52	.27	.56			.34
irma24	.77	.59	.78			-.32
irma34	.72	.52	.71			
irma37	.71	.50	.62			
irma04	.58	.34	.59			

Note. Loadings below .3 have been omitted. h^2 = unique item variance (uniqueness).

[a]One-factor model. Matrix = polychoric correlations. Factor estimation method = RULS. [b]Pure Exploratory Bifactor Model. Matrix = polychoric correlations. Factor estimation method = RULS. Rotation = Robust Promin.

of the items had good loadings in the one-factor model. Only 4 items had loadings lower than .5, and the remaining item loadings ranged from .52 to .79. Dimensionality indices computed based on the one-factor model strongly indicated the data was essentially unidimensional (UniCo = .956, CI: .946 - .976; ECV = .902, CI: .898 - .908; MIREAL = .166, CI: .15 - .17).

IRMA Exploratory Bifactor Model. The bifactor solution for the IRMA is reported in Table 11. The bifactor model demonstrated excellent fit (CFI = .996, CI: .996 - .997), and the common variance was almost entirely accounted for by the model (GFI = .995, CI: .995 - .995; AGFI = .993, CI: .993 - .994). The bifactor model showed a marked improvement over the one-factor model (RMSEA = .031, CI: .01 - .05; WRMR = .031, CI: .030 - .031). However, unidimensional congruence decreased under the bifactor model (UniCo = .798, CI: .872 - .817).

General Factor. The general factor accounted for over half of the common variance (ECV = .583, CI: .569 - .597, ω_h = .59). At the item level, a total of 32 IRMA items had a general factor loading of .5 or above. Item 5 was the only item that did not have a substantial loading on the general factor, with a loading of .24. From the items that had a substantial loading on the general factor, item 9 had the lowest general factor loading at .38, and item 24 had the highest at .78.

Group Factors. The group factors in the bifactor solution were very disproportionate. Group factor 1 (F1) had 11 substantial item loadings, group factor 2 (F2) had 24, and group factor 3 (F3) had 7. To facilitate discussion of the IRMA group factors, in the tables to follow I have included an indication of which subscale each item was classified under in the original development paper by Payne et al. (1999). The subscale indicators are intended to (a) highlight trends in item content, and (b) facilitate a comparison between how the items grouped under my model as opposed to their original subscales by Payne et al. (1999).

Table 12 contains the items of the first group factor. F1 was dominated by items from the She Wanted It (WI) subscale, and the She Lied (SL) subscale. F1 had negative loadings on only two items, item 27 and item 35, from the Rape is a Deviant

Table 12

Item content and item-level dimensionality indices for F1 items in the IRMA bifactor model

Subscale	Item no. and content	IECV	I-UniCo	I-REAL
WI	23. Some women prefer to have sex forced on them so they don't have to feel guilty about it.	.72	.95	.02
WI	8. Many women secretly desire to be raped.	.67	.91	.18
WI	44. Many women actually enjoy sex after the guy uses a little force.	.63	.86	.10
WI	2. Although most women wouldn't admit it, they generally find being physically forced into sex a real "turn on".	.63	.82	.09
WI	18. Many women find being forced to have sex very arousing.	.61	.79	.04
DE	27. Rape is unlikely to happen in the woman's own familiar neighbourhood.	.60	.73	.11
DE	35. Rape almost never happens in the woman's own home.	.45	.51	.12
LI	31. A lot of women lead a man on and then cry rape.	.42	.87	.17
LI	16. Rape accusations are often used as a way of getting back at men.	.32	.73	.08
LI	7. Many so-called rape victims are actually women who had sex and "changed their minds" afterwards.	.30	.71	.12
LI	5. Women who are caught having an illicit affair sometimes claim it was rape.	.16	.37	.21

Note. Subscale as per Payne et al. (1999). WI = She Wanted It; LI = She Lied; DE = Rape is a Deviant Event. I-ECV = item level explained common variance; I-UniCo = item level unidimensional congruence; I-REAL = item level residual absolute loadings. Items are arranged in descending order according to I-ECV.

Event (DE) subscale. I posit that F1, especially when compared to F2 and F3, appears to have loaded on items that invalidate victims in some way, especially the validity of experiences, and credibility of allegations. I will elaborate on these observations further in the discussion chapter.

An unexpected trend that emerged from the item-level dimensionality indices, was seeing the items group according to their subscales in Table 12 when ordered according to their I-ECV value. All of the items from the SL subscale had an I-ECV value below .5, indicating that less than half the items' common variance was explained by the general factor. Items 35 and item 5 had low I-UniCo values, indicating multidimensionality introduced at the item level by both items. Item 5 in particular had noticeably poor dimensionality indices, with the highest IREAL value of the F1 items, as well as the lower I-ECV and I-UniCo value by far.

Table 13

Select F2 items from the IRMA bifactor model

Subscale	Item no. and content	I-ECV	I-UniCO	I-REAL
NR	17. A rape probably didn't happen if the woman has no bruises or marks.	.65	.89	.21
DE	27. Rape is unlikely to happen in the woman's own familiar neighbourhood.	.60	.73	.11
DE	22. It is usually only women who dress suggestively that are raped.	.58	.77	.02
DE	28. In reality, women are almost never raped by their boyfriends.	.50	.60	.13
SA	41. A woman who dresses in skimpy clothes should not be surprised if a man tries to force her to have sex.	.49	.79	.14
DE	10. Usually, it is only women who do things like hang out in bars and sleep around that are raped.	.49	.62	.01
TE	14. Rape isn't as big a problem as some feminists would like people to think.	.47	.66	.20
DE	35. Rape almost never happens in the woman's own home.	.45	.51	.12
SA	15. When women go around wearing low-cut tops or short skirts, they're just asking for trouble.	.42	.63	.06
DE	13. Men from nice middle-class homes almost never rape.	.41	.47	.07
DE	9. Rape mainly occurs on the "bad" side of town.	.37	.42	.03

Note. Subscale as per Payne et al. (1999). SA = She Asked For It; DE = Rape is a Deviant Event. TE = Rape is Trivial Event. NR = It Wasn't Really Rape. I-ECV = item level explained common variance; I-UniCo = item level unidimensional congruence; I-REAL = item level residual absolute loadings. Items are arranged in descending order according to I-ECV. F2 had 24 items; for the sake of space, only items with an F2 group loading $\geq .45$ have been included here.

Moving on to group factor 2 (F2). F2 was the largest group factor, comprising over half the IRMA items (n = 24). Four of the subscales originally outlined by Payne et al. (1999) collapsed into F2: all seven items from the Rape is a Deviant Event (DE) subscale, 7 out of 8 items from the She Asked For It (SA) subscale, 4 of 5 items from the Rape is a Trivial Event (TE) subscale, as well as 3 out of 5 items from the It Wasn't Really Rape (NR) subscale.

Table 13 contains the item content and dimensionality indices for F2 items with a loading of .45 or higher. Interestingly, all seven items from the DE subscale featured in the subset of items with the highest loadings on F2 captured in Table 13. I posit that F2 appears to be characterised by rape myths that socially or spatially distance rape, and will elaborate on this point further in the discussion chapter.

Regarding item-level dimensionality indices, none of the F2 items had an IREAL over .3, meaning none of the items had substantial residual loadings. In all, 14 of the items that loaded on F2 had more than half their explained common variance accounted for by the general factor (I-ECV $\geq$ .5). However, most of the F2 items also had I-UniCo values below .8, suggesting that the items themselves may be multidimensional. Multidimensionality can be introduced at the item level when the same item is interpreted differently by respondents. Items with particularly low I-UniCo values were item 13 ("Men from nice middle-class homes almost never rape") and item 9 ("Rape mainly occurs on the “bad” side of town"). Interesting to note, both item 13 and item 9 link a rape myth to a particular socio-economic class.

Finally, group factor 3 (F3). F3 was the smallest group factor, with only 7 items (see Table 14). F3 was the only factor that closely aligned with a single subscale as intended by Payne et al. (1999), containing all five items from the “Mean to” subscale. However, F3 also included one item with a negative factor loading from the “not rape” subscale (item 24), as well as an item from the “she asked for it” subscale (item 25). All the items that loaded on F3 appeared to have a perpetrator focus, with the exception of item 25 ("When a woman is a sexual tease, eventually she is going to get into trouble").

Item 24 ("If the rapist doesn't have a weapon, you really can't call it rape") was

Table 14

Item content and item-level dimensionality indices for F3 items in the IRMA bifactor model

Subscale	Item no. and content	I-ECV	I-UniCo	IREAL
MT	3. When men rape, it is because of their strong desire for sex.	.42	.62	.58
MT	20. Rapists are usually sexually frustrated individuals.	.48	.67	.50
NR	24. If the rapist doesn't have a weapon, you really can't call it rape.	.77	.98	.37
SA	25. When a woman is a sexual tease, eventually she is going to get into trouble.	.45	.80	.21
MT	30. When a man is very sexually aroused, he may not even realise that a woman is resisting.	.71	.94	.39
MT	39. Men don't usually intend to force sex on a woman, but sometimes they get too sexually carried away.	.73	.95	.37
MT	42. Rape happens when a man's sex drive gets out of control.	.50	.72	.58

Note. Subscale as per Payne et al. (1999). MT = He Didn't Mean To. SA = She Asked For It. NR = It Wasn't Really Rape. I-ECV = item level explained common variance; I-UniCo = item level unidimensional congruence; I-REAL = item level residual absolute loadings. Items are arranged in descending order according to I-ECV.

the only item with a negative loading on F3. Item 24 also had the highest loading on the general factor of all the F3 items. The negative loading on F3 could be due to the fact that the rest of the perpetrator-focused items in F3 all confirm that rape can happen in the absence of a weapon, whereas item 24 contradicts those items by asserting that a weapon is a necessary condition for rape. Item 24 also had a near-perfect I-UniCo value of .98, which suggests that this item was particularly clearly worded to the respondents.

Many of the F3 items had excellent I-UniCo values (I-UniCo $\geq .8$). However, item 3, item 20 and item 25 had less than half their explained common variance accounted for by the general factor (I-ECV $< .5$). Furthermore, six of the F3 items had an I-REAL value well above .3, indicating substantial residuals that cannot be considered negligible. The high I-REAL values indicate that many of the items loaded on F3 would be highly susceptible to bias under a unidimensional FA solution. As a final note for this section, F3 showed close to no correlation with F1 and F2 (see Table 15).

Table 15

Inter factor correlation matrix for the IRMA bifactor solution

Factor	Factor Correlations				Variance
	F1	F2	F3	GEN	
1	1				2.04
2	.50	1			5.32
3	.11	.02	1		1.94
GF	0	0	0	1	12.97

Note. F1 = group factor 1. F2 = group factor 2. F3 = group factor 3. GEN = general factor. Variance = explained variance of the rotated factors.

Item-Level Model Complications. There were nine items that complicated the interpretation of the IRMA bifactor model. First, item 5 "Women who are caught having an illicit affair sometimes claim it was rape" did not have a substantial loading on the general factor (loading gen = .24). Item 5 was also flagged earlier for having notably poor item-level dimensionality indices (I-ECV = .37, CI: .19-.80; I-UniCO = .16, CI: .08-.39).

Second, three items did not load on a specific group factor. The item content and dimensionality indices for items that did not load on a specific group factor are captured in Table 16. Interesting to note, Item 4, item 34 and item 37 arguably all relate to narrow definitions of rape and consent, a point I will elaborate on further in the discussion chapter. All three items were close to perfectly unidimensional (I-UniCO > .99), had over 80% of the explained common variance accounted for by the general factor (I-ECV > .8), and had negligible residual loadings (I-REAL < .3).

Finally, 5 items had significant crossloadings in the bifactor solution. Item 7, item 35, item 27, item 25 and item 31 all had a loading of .3 or higher for more than one group factor. The cross-loadings were small, and only just met the criteria for a substantial factor loading, ranging from -.35 to .32. However, item 35 was the only item to have a sizeable dominant loading. In other words, for four of the items that had cross-loadings, the loadings on both factors were roughly the same size. While all the items with cross-loadings had a general factor loading of .5 or higher, most had an I-ECV value below .5, with item 27 (I-ECV = .60) as the only exception. The implication then is that most of the items with cross-loadings both (a) complicate the model and (b) do not present as particularly strong indicators of the general factor. I will discuss the implications further in Chapter 5.

Table 16

Flagged Items: IRMA Bifactor Model

Item no. and content	λ_{GEN}	I-ECV	I-UniCo	I-REAL
4. If a woman is willing to "make out" with a guy, then it's no big deal if he goes a little further and has sex.	.59	.88	.99	.12
34. If a woman doesn't physically resist sex – even when protesting verbally – it really can't be considered rape.	.71	.88	.99	.11
37. When women are raped, it's often because the way they said "no" was ambiguous.	.62	.77	.99	.05

Note. I-ECV = item level explained common variance; I-UniCo = item level unidimensional congruence; I-REAL = item level residual absolute loadings. Items are arranged in descending order according to I-ECV.

Model Comparison

The model comparison is based on the parameters presented in Table 17 on the next page. The difference between the item loadings on the general factor of the bifactor model and the single factor in the one-factor model was big. The relative bias calculation revealed that only 17 out of the 40 items fell within the acceptable biasing range of -15% to 15%. Furthermore, the relative bias ranged from -21% to 88%, with an average bias across all items of 20%. Item 5 had the greatest relative bias of all the items, with a relative bias of 88%.

By contrast, the differences in slopes between the unidimensional normal ogive model and the general dimension of the multidimensional model were small. Across the items, the difference in slopes ranged from -.54 to .30. Only two items had a concerning

Table 17

IRMA Model Comparison

	FA			IRT		
Item	λ_{1F}	λ_{GEN}	Relative Bias (%)	α Uni	α Gen	Slope difference
irma_01	.64	.45	42	.84	.61	.24
irma_02	.57	.58	-1	.70	.85	-.15
irma_03	.36	.46	-21	.39	.65	-.27
irma_04	.58	.59	-1	.72	.75	-.03
irma_05	.46	.24	88	.52	.31	.21
irma_07	.65	.41	58	.85	.62	.24
irma_08	.63	.61	3	.80	.91	-.11
irma_09	.52	.38	39	.61	.48	.13
irma_10	.70	.53	33	.98	.80	.18
irma_12	.71	.63	12	1.00	.92	.08
irma_13	.62	.46	37	.80	.65	.15
irma_14	.75	.54	38	1.12	.88	.25
irma_15	.77	.53	46	1.19	.90	.30
irma_16	.64	.41	56	.82	.59	.23
irma_17	.74	.62	19	1.10	.96	.14
irma_18	.65	.64	1	.84	1.10	-.25
irma_19	.72	.56	30	1.04	.81	.23
irma_20	.42	.47	-9	.47	.63	-.17
irma_22	.75	.59	26	1.13	.94	.19
irma_23	.71	.66	8	1.00	1.03	-.03
irma_24	.77	.78	-2	1.19	1.73	-.54
irma_25	.71	.51	40	.99	.77	.22
irma_26	.62	.58	7	.79	.78	.01
irma_27	.62	.55	13	.80	.79	.01
irma_28	.64	.51	25	.83	.73	.10
irma_29	.79	.64	24	1.29	1.07	.22
irma_30	.52	.56	-7	.60	.74	-.14
irma_31	.78	.55	42	1.23	1.03	.20
irma_33	.77	.60	27	1.20	.97	.23
irma_34	.72	.71	2	1.05	1.09	-.04
irma_35	.63	.50	26	.80	.74	.06
irma_36	.79	.68	17	1.30	1.11	.20
irma_37	.71	.62	13	1.00	.89	.12
irma_38	.76	.76	0	1.15	1.37	-.22
irma_39	.62	.63	-1	.79	.92	-.13
irma_41	.78	.57	37	1.26	.98	.27
irma_42	.45	.57	-21	.51	.98	-.47
irma_43	.63	.50	25	.81	.64	.16
irma_44	.63	.58	8	.81	.85	-.04
irma_45	.78	.69	13	1.25	1.19	.06

difference in IRT slopes, irma42 (Δ_α = -.47) and irma 24 (Δ_α = -.54). The average slope difference across all the items was .05 when negative differences were included as is, and .18 when all slope differences were scored in the same direction.

Combined Item Pool Dimensionality Investigation Results

In the following section, I explore the dimensionality of the combined item pool (CIP) of the IRMAS and MRMS, to determine if the items of the MRMS may be tapping into the same underlying construct as the items of the IRMA. The results of this exploration include the following: (a) correlation between the IRMAS and MRMS, and reliability indicators for the CIP, (b) a bifactor model fitted to the CIP data, with a focus on general factor strength and potential interpretation of the group factors and (c) I report the difference between the general factor of the PEBI and the common factor of an alternative one-factor model, with the compared models presented in both FA loadings and IRT slopes.

Combined Item Pool: Traditional Psychometrics

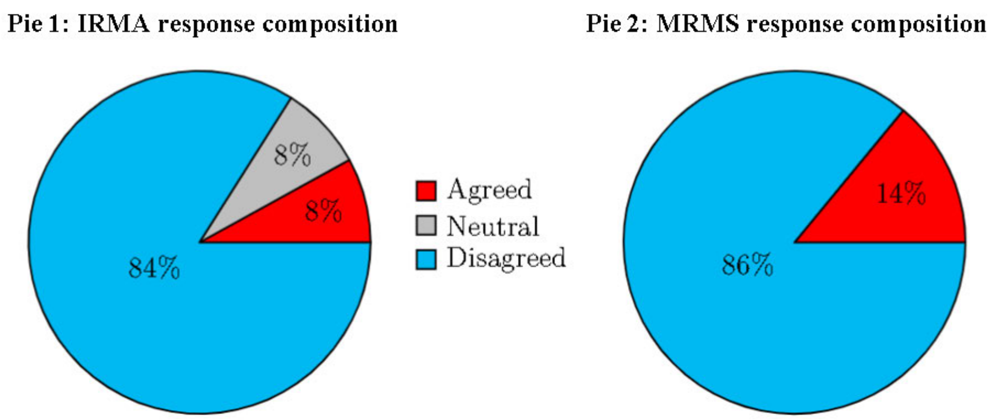
To determine the correlation between the MRMS and IRMA, I used Spearman's rho, calculated in R using the cor() function available as part of the *psych* package (Revelle, 2021). Spearman's rho correlation between the MRMS and IRMAS was ρ = 0.77, and ρ = 0.83 once corrected for attenuation. However, the cor() function of the *psych* package does not calculate p-values, therefore I also tested the correlation in PSPP. In PSPP, the only bivariate correlation available is the Pearson correlation, but the result was identical($r = .77$, $p < .001$).

A common reporting technique in MRMS literature is to split the 6-point rating scale into two halves and report the percentage of disagreement versus agreement with rape myths for each item. This technique reduces multiple ranked categories to a binary format that is easy to interpret. While reductionistic, looking at the data in this way can make it easier to spot trends worth exploring that might otherwise go unnoticed.

The pie charts in Figure 7 reveal a stark similarity in general response composition. Both the MRMS and IRMAS had roughly 85% of the responses within the three response categories that captured disagreement with rape myths. If the MRMS

Figure 7

Pie Charts: Simplified Response Compositions of the IRMAS and MRMS



Note. To generate these pie charts, responses across all agreement categories for each scale were summed, with the same process repeated for the disagreement categories. The summed number of responses was converted to percentages, and the final pie charts were formatted in OverLeaf.

had included a neutral category when it was administered, the response compositions may have looked even more similar. However, it is unknowable from this study whether a neutral category in the MRMS would have seen a decrease in agreement or disagreement with the rape myth statements.

Combined Item Pool: Exploratory Factor Analyses

Preliminary Statistics. Overall, the CIP demonstrated excellent reliability according to commonly cited reliability indices (n = 2536, α = .98, GLB = .99, ω = .98). At both the scale and item level the correlation matrix demonstrated suitability for factor analysis. Bartlett's test of sphericity was significant ($X^2(1891) = 28960.8$, $p <$ .001), and the Kaiser-Meyer-Olkin measure of sampling adequacy result was very good (KMO = .97). The items also individually demonstrated very good sampling adequacy, with item-MSA values ranging from .93 to .99.

Eigenvalues greater than 1 are captured in Table 18. The ratio of the first to second eigenvalue is 26.5:2.6, which rounds off to a ratio of 10:1. At well over the recommended criterion of 3:1 (Reise, Ventura, et al., 2011), the sizeable first eigenvalue is a good indicator that there is a general RMA construct underlying the rape myths.

The Hull-CAF indicated there was only one common factor underlying the combined RMA item pool. By contrast, the PA-MRFA analysis indicated there were at least 4 dimensions underlying the data. The comparison modelling approach made it possible to test the recommendations of both factor retention methods, with the one-factor model justified by the Hull-CAF, and a bifactor model with four group factors motivated by the PA-MRFA results.

Table 18

Eigenvalues for the combined item pool

Variable	Eigenvalue	Proportion of variance
1	26.51	.43
2	2.63	.04
3	2.61	.04
4	2.21	.03
5	1.57	.02
6	1.42	.02
7	1.14	.02
8	1.09	.02
9	1.07	.01

Combined Item Pool: Exploratory Bifactor Model. A pure exploratory bifactor model with four specific group factors was fitted to the data. The complete bifactor solution is available in Appendix H. Limited fit indices were available in the FACTOR output for this model, which I will discuss further in my limitations section in Chapter 5. However, the fit indices available in the output indicated excellent model fit (CFI = .997, GFI = .995, AGFI = .994).

The solution had a high UniCo value of .83, indicating it was close to unidimensional. The low MIREAL value of .16 indicated that the data did not display a substantial violation of the assumption of local independence. The general factor was strong relative to the group factors, and saturated most of the common item response variance (ECV = .58, omegaH = .60). Out of 62 items, 61 had a substantial general factor loading ($\lambda_{Gen} \geq .3$). However, the general factor did not load substantially on irmas05 ($\lambda_{Gen} = .27$), which also did not have a substantial group factor loading.

In addition to irmas05, there were 10 items that did not have a substantial group factor loading. Surprisingly, there were only two items that had cross-loadings, irmas24 and irmas29. The cross-loadings were quite similar in size; irmas29 had a loading of .32 on F1, and a loading of .33 on F4; irmas24 had a positive loading on F1 of .39 but a negative loading of -.47 on F2. Cross-loadings are to be expected in exploratory investigations. It was however surprising that out of such a large item pool, there were only two items with cross-loadings.

The four specific group factors were sized disproportionately to one another, with 11 items in the first group factor (F1), nine items in the second group factor (F2), 16 items in the third group factor (F3), and 17 items in the fourth group factor (F4). Despite being different sizes, the group factors were arguably interpretable and appeared to be substantively meaningful.

Table 19

Inter factor correlation matrix for the combined item pool bifactor solution

Factor	F1	F2	F3	F4	GEN
F1	1				
F2	.45	1			
F3	.18	.27	1		
F4	.50	.50	.27	1	
GEN	0	0	0	0	1

Note. F1 = group factor 1. F2 = group factor 2. F3 = group factor 3. F4 = group factor 4. GEN = general factor.

The inter-factor correlation matrix is captured in Table 19. Factor 3 correlated the least with each of the other group factors. While F1, F2 and F4 had correlations ranging from .45 to .50 with each other, correlations with F3 ranged from .18 to .27. It is important to reiterate here that the specific group factors are calculated based on shared variance among items over and above the common variance captured in the general factor (DeMars, 2013). Therefore, the inter-factor correlation matrix indicates that three of the specific group factors are tapping into related secondary constructs, whereas the specific construct tapped into by F3 may not be as closely related.

Group Factor 1. F1 entirely comprised of IRMAS items, and was dominated by items from the She Lied (LI) subscale and the She Wanted It (WI) subscale (Payne et al., 1999). The only two exceptions were item 29 from the Rape is a Trivial Event (TE) subscale, and item 24 from the It Wasn't Really Rape (NR) subscale (Payne et al., 1999). Irmas24 and irmas29 had the highest general factor loadings out of all the items in F1.

The following items had a higher F1 loading than a general factor loading: irmas08, irmas18, irmas2 and irmas44. All four items were originally grouped under the WI subscale by Payne et al. (1999). However, for the remaining items that loaded on F1, the general factor loading was higher than the F1 loading. All items in F1 had a general factor loading greater than .3. Furthermore, the general factor explained over half the item-level common variance for five of the items (I-ECV $>$ .5).

In terms of item content, F1 was characterised by rape myths that pertain to the validity of rape allegations and victim credibility, as well as rape myths that would signal a narrow understanding of rape and consent if endorsed by a respondent. For example, "Women tend to exaggerate how much rape affects them" (item 29) undermines the credibility of victims' experiences. Another example item is "Many women actually enjoy sex after the guy uses a little force" (item 44), which conflates rape with consensual rough sex. I will elaborate further on the factor interpretation in the discussion chapter.

Group Factor 2. Factor 2 consisted of nine items, and was the smallest of the group factors. F2 loaded on eight items from the IRMAS, and one item from the MRMS. Five of the items from the IRMAS came from the "He didn't mean to" subscale. The remaining three IRMAS items came from the "Not Rape" and "trivial event" subscales. However, all three items from the NR and TE subscales had negative loadings on F2. F2 loaded on IRMAS items with a perpetrator focus, as well as rape myths that relied on the assumption that rape is a sexually motivated crime.

All the items that F2 loaded on had a general factor loading greater than .5. The I-ECV values confirmed that the general factor explained at least half the common

Table 20

CIP Group Factor 1 Items

Subscale	Variable	Item content	λ_{F1}	λ_{GEN}	I-ECV
NR	irma24	If the rapist doesn't have a weapon, you really can't call it rape.	.39	.71	.68
TE	irma29	Women tend to exaggerate how much rape affects them.	.32	.66	.66
LI	irma33	A lot of times, women who claim they were raped just have emotional problems.	.37	.60	.59
WI	irma23	Some women prefer to have sex forced on them so they don't have to feel guilty about it.	.54	.59	.58
WI	irma08	Many women secretly desire to be raped.	.60	.53	.51
WI	irma18	Many women find being forced to have sex very arousing.	.72	.53	.42
LI	irma31	A lot of women lead a man on and then cry rape.	.37	.53	.40
WI	irma44	Many women actually enjoy sex after the guy uses a little force.	.56	.50	.46
WI	irma02	Although most women wouldn't admit it, they generally find being physically forced into sex a real "turn on".	.64	.48	.44
LI	irma07	Many so-called rape victims are actually women who had sex and "changed their minds" afterwards.	.38	.43	.36
LI	irma16	Rape accusations are often used as a way of getting back at men.	.38	.42	.37

Note. Subscale as per Payne et al. (1999). WI = She Wanted It subscale. LI = She Lied subscale. NR = It Wasn't Really Rape subscale. TE = Rape is a Trivial Event subscale. λ_{F1} = Item loadings on group factor 1. λ_{GEN} = item loadings on the general factor. I-ECV = Item-level explained common variance.

variance for all the F2 items (I-ECV $\geq$ .5). The single MRMS item that F2 loaded on, MRMS item 17 (mrms17), is almost identical in wording to item 20 from the IRMAS (irma20). Item 17 from the MRMS explicitly refers to female rapists, whereas item 20 from the IRMAS does not include a reference to perpetrator gender. However, it is worth noting that when item 20 was read in the context of the rest of the IRMA, most participants would likely have assumed the item was referring to male rapists. It is noteworthy that not only did the same group factor load mrms17 and irma20, the items also had very similar general and group factor loadings. I will discuss the implications further in Chapter 5.

Table 21

CIP Group Factor 2 Items

Subscale	Variable	Item content	λ_{F2}	λ_{GEN}	I-ECV
MT	irma3	When men rape, it is because of their strong desire for sex.	.61	.50	.50
MT	irma20	Rapists are usually sexually frustrated individuals.	.60	.55	.55
NR	irma24	If the rapist doesn't have a weapon, you really can't call it rape.	-.47	.71	.68
TE	irma26	Being raped isn't as bad as being mugged or beaten.	-.32	.60	.75
MT	irma30	When a man is very sexually aroused, he may not even realise that a woman is resisting.	.36	.51	.64
TE	irma38	If a woman isn't a virgin, then it shouldn't be a big deal if her date forces her to have sex.	-.35	.70	.75
MT	irma39	Men don't usually intend to force sex on a woman, but sometimes they get too sexually carried away.	.37	.60	.69
MT	irma42	Rape happens when a man's sex drive gets out of control.	.65	.59	.55
N/A	mrms17	Women who rape men are sexually frustrated individuals.	.61	.65	.62

Note. Subscale as per Payne et al. (1999). "MT" = He didn't mean to. "NR" = Not really rape. "TE" = Rape is a trivial event. "N/A" = Not applicable.

Group Factor 3. F3 loaded on 16 MRMS items, and none of the IRMAS items. The item content and factor loadings are available in Table 22. Eleven of the items in F3 had a general factor loading of .5 or higher. The general factor accounted for at least half the explained common variance for 13 of the items (I-ECV $\geq$ .5).

F3 presented as a potential methodology factor. It is possible that F3 loaded only on MRMS items because they were rated on a 6-point Likert scale, and were in the minority. However, worth noting is that F3 did not load on six of the MRMS items. Two of the MRMS items were loaded on by different specific group factors, and four of the MRMS items did not have any substantial group factor loadings. If F3 was indeed a methodology factor, it arguably should have loaded on all the MRMS items, not just most of them. The following MRMS items did not have specific group factor loadings: mrms07 "many men claim rape if they have consented to homosexual relations but have changed their minds afterwards"; mrms10 "Male rape is usually committed by

Table 22

CIP Group Factor 3 Items

Variable	Item content	λ_{F3}	λ_{GEN}	I-ECV
mrms18	A man who allows himself to be raped by another man is probably homosexual.	.31	.66	.70
mrms13	Most men who are raped by a woman are somewhat to blame for not being more careful.	.44	.63	.54
mrms9	If a man engages in necking and petting and he lets things get out of hand, it is his own fault if his partner forces sex on him.	.45	.59	.50
mrms11	Most men who are raped by a man are somewhat to blame for not escaping or fighting off the man.	.55	.59	.49
mrms21	Male rape is more serious when the victim is heterosexual than when the victim is homosexual.	.33	.57	.72
mrms4	If a man obtained an erection while being raped it probably means that he started to enjoy it.	.50	.57	.57
mrms8	Most men who are raped by a woman are somewhat to blame for not escaping or fighting off the woman.	.58	.57	.46
mrms5	A man can enjoy sex even if it is being forced upon him.	.31	.54	.65
mrms20	Men who parade around nude in a locker room are asking for trouble.	.30	.53	.56
mrms22	I would have a hard time believing a man who told me that he was raped by a woman.	.41	.53	.58
mrms3	Any healthy man can successfully resist a rapist if he really wants to.	.49	.51	.52
mrms12	A man who has been raped has lost his manhood.	.32	.49	.61
mrms1	It is a terrible experience for a man to be raped by a woman. (R)	.45	.48	.52
mrms2	The extent of a man's resistance should be a major factor in determining if he was raped.	.37	.47	.56
mrms6	Most men who are raped by a woman are very upset by the incident. (R)	.36	.41	.54
mrms19	Most men would not enjoy being raped by a woman. (R)	.36	.33	.41

homosexual men"; mrms14 "If a man told me that he had been raped by another man, I would suspect that he is homosexual."; and mrms16 "No self-respecting man would admit to being raped."

F3 items that alluded to homosexuality and manhood had the lowest group factor loadings, and many barely met the criteria for a substantial loading. F3 loaded the most highly on item 8 "most men who are raped by a woman are somewhat to blame for not fighting off the woman", and item 11 "most men who are rape by a man are somewhat to blame for not escaping the man". Item 8 and item 11 are worded identically apart from the specified perpetrator gender, and it is interesting to note that the items had very similar loadings and I-ECV values. To report on item content trends, I considered only items with a group loading greater than .4, and focused on keywords that repeated across the items. Keywords I noted across the items with an F3 loading greater than .4 were "raped by a woman", "erection/enjoy", "resist/fight off", and "to blame/own fault". For further discussion of these results, please see Chapter 5.

Group Factor 4. F4 was the biggest group factor, with 17 items. Upon review of the item content (see Table 23), it was clear that F4 loaded on rape myths characterised by victim blaming and othering. I will elaborate further on factor interpretation in Chapter 5. F4 loaded on one item from the MRMS and 16 items from the IRMA, particularly items from the She Asked For It (SA) subscale and the Rape is a Deviant Event (DE) subscale. General factor loadings ranged from .41 to .65, and I-ECV values ranged from .35 to .80. In total, only five of the F4 items had an I-ECV below .5.

Combined Item Pool: One-factor Model. A one-factor solution was generated as a point of comparison, to determine whether the combined item pool data could be classified as essentially unidimensional. The loadings for the one-factor model are available alongside the bifactor model loadings in Appendix H. The common factor loaded on all the CIP items, with a loading of .5 or higher for 56 of the items. The remaining six items that fell below the .5 mark still loaded substantially on the common factor. The common factor loaded below .5 on the following items: irmas03 (λ = .37), irmas20 (λ = .43), irmas42 (λ = .47), mrms06 (λ = .42), mrms07 (λ = .49), and mrms19 (λ =.37).

Measure-level dimensionality indices based on the one-factor model suggested the data was unidimensional (UniCo = .99, ECV = .92, MIREAL = .16). Furthermore, the one-factor model appeared to be a good fit to the data. The GFI and AGFI had

Table 23

CIP Group Factor 4 Items

Subscale	Variable	Item content	λ_{F4}	λ_{GEN}	I-ECV
SA	irma1	If a girl is raped while she is drunk, she is at least somewhat responsible for letting things get out of control.	.44	.41	.35
DE	irma9	Rape mainly occurs on the “bad” side of town.	.45	.50	.64
DE	irma10	Usually, it is only women who do things like hang out in bars and sleep around that are raped.	.48	.60	.64
DE	irma13	Men from nice middle-class homes almost never rape.	.50	.58	.67
TE	irma14	Rape isn’t as big a problem as some feminists would like people to think.	.51	.60	.58
SA	irma15	When women go around wearing low-cut tops or short skirts, they’re just asking for trouble.	.56	.51	.38
NR	irma17	A rape probably didn’t happen if the woman has no bruises or marks.	.35	.63	.69
SA	irma19	If a woman goes home with a man she doesn’t know, it is her own fault if she is raped.	.32	.51	.46
DE	irma22	It is usually only women who dress suggestively that are raped.	.44	.64	.67
SA	irma25	When a woman is a sexual tease, eventually she is going to get into trouble.	.35	.48	.39
DE	irma27	Rape is unlikely to happen in the woman’s own familiar neighbourhood.	.40	.65	.80
DE	irma28	In reality, women are almost never raped by their boyfriends.	.46	.61	.72
DE	irma35	Rape almost never happens in the woman’s own home.	.52	.62	.69
SA	irma36	A woman who “teases” men deserves anything that might happen.	.32	.62	.60
SA	irma41	A woman who dresses in skimpy clothes should not be surprised if a man tries to force her to have sex.	.44	.54	.42
N/A	mrms15	Most men who have been raped have a history of promiscuity.	.33	.60	.63

Note. Subscale as per Payne et al. (1999). “SA” = She asked for it. "DE" = Rape is a deviant event. “NR” = Not really rape. "TE" = Rape is a trivial event. "N/A" = Not applicable.

values close to 1 (GFI = .98, AGFI = .98), and the summary statistics for the fitted residuals (n = 1891, M = 0.0003, Variance = 0.0042) indicated that the residuals did not pose a concern.

Combined Item Pool: Parameter Bias Under Factor Analytic Parameters

In order to gauge the item-level parameter bias that occurred when the multidimensionality went unmodelled, I compared the common factor loadings in the one-factor model to the general factor loadings of the bifactor model. A table containing the relevant loadings from each model, as well as the calculated parameter bias for each item is available in Appendix I.

According to the results of the FA model comparison, most of the item loadings were overestimated in the one-factor model. Of the 61 items that formed part of the bias analysis, 36 fell outside the acceptable range of -15% to 15%. Parameter bias was not calculated for item 5 of the IRMAS as it did not load significantly on the general factor in the bifactor model. Some of the items only just fell outside the acceptable range, for example, 6 items displayed a 16% loading bias. However, 26 of the items were biased upwards of 20%.

Four items that fell within the acceptable range were items that did not load on a specific group factor. The bias captured for these items may be inaccurate due to the fact that they did not load on a specific group factor. In addition, it is worth noting that item 29 and item 24 of the IRMAS had cross-loadings on two group factors. It is therefore possible that the relative biasing effect may not be accurate for these two items, as they could have caused some distortion in item parameters across both models.

Combined Item Pool: Parameter Bias under IRT parameters

The slopes between the unidimensional and bifactor IRT models are reported for each item in Appendix I, along with the change in slope across the two IRT models. The table spans three pages and was too long to include here.

All the slopes on the general dimension were substantial and in the same direction. Based on interpretation guidelines by Baker and Kim (2017), 26 items were moderately discriminating (slope: .38-.79), 21 items were highly discriminating (slope: .80-.99), and 14 items had very high discriminatory power (slope: > 1). Irmas 5 was the only item with low discriminatory power ($\alpha = .33$).

I have included a frequency table below (see Table 24), which captures the total number of items I considered to have a very small, small, moderate, large and very large change in slope. As I noted in my methodology, I took the interpretation guidelines by Baker and Kim (2017, p. 26) for slopes considered to be low versus highly discriminating, and have used the ranges provided as an indication for what constitutes a small or large change in slope. Overall, the item slopes under the unidimensional IRT model were very similar to the item slopes for the general dimension of the bifactor IRT model. For 32 of the items, the slope difference was less than .1. For an additional 21 of the items, the slope difference was less than .21. A further seven items displayed a slope difference between .22 and .33. Only two items were of real concern: item 17 from the MRMS (Δ_α = -.50) and item 42 from the IRMAS (Δ_α = -.45). In both cases, the item discrimination parameter was notably underestimated by the unidimensional model.

Table 24

Summary of change in item slopes across the IRT models for the combined item pool

Interpretation of Δ_α	Change in slope (Δ_α)	No. of items
Very small	<0.21	52
Small	.21-.37	8
Moderate	.38-.79	2
Large	.80-.99	0
Very Large	$\geq$ 1.00	0

Note. The category ranges are based on recommendations for item slope interpretation made by Baker and Kim (2017, p. 26).

CHAPTER 5: DISCUSSION

A key component of my research was to explore whether the Illinois Rape Myth Acceptance Scale (IRMAS; Payne et al., 1999) and Male Rape Myth Scale (MRMS; Kerr Melanson, 1998) could be considered essentially unidimensional. In the following discussion, I review the evidence that suggests both the IRMAS and MRMS can be treated as essentially unidimensional under an item response theory (IRT) framework but not necessarily under a factor analytic (FA) framework. For the final component of my research, I explored whether the items of the IRMAS and MRMS were indicators of the same latent construct. Therefore, I also review evidence from my exploratory investigation suggesting that rape myths, whether they centre on female or male rape victims, are manifestations of the same general construct. Finally, I discuss what I uncovered about the multidimensionality inherent in each scale, insights from analysing the combined item pool, and the practical implications of my findings.

Dimensionality of the Illinois Rape Myth Acceptance Scale and the Male Rape Myth Scale: Evidence of Essential Unidimensionality

A key component of my research was exploring whether the MRMS and IRMAS could be considered essentially unidimensional. To test whether the scales could be considered unidimensional, I relied on conventional dimensionality indices and analysed item-level parameter differences across several EFA models. Although the MRMS and IRMAS data were investigated independently, I discuss both scales at the same time to highlight similarities and differences between their psychometric properties. In the following discussion, I will review the evidence suggesting that the IRMAS and MRMS can be treated as essentially unidimensional under an IRT framework but not an FA framework. The evidence discussed includes dimensionality indices, a strong common factor underlying the data, and an acceptable degree of parameter distortion when modelled as unidimensional.

Dimensionality Indices

The first indication of dimensionality was provided in the form of several model-independent dimensionality indices: parallel analysis based on minimum rank factor analysis (PA-MRFA), the Hull method based on common parts accounted for (Hull-CAF), eigenvalue ratios, and the mean of item residual absolute loadings (MIREAL). I also looked at several model-dependent dimensionality indices, namely unidimensional congruence (UniCo), explained common variance (ECV) and omega hierarchical (OmegaH).

The PA-MRFA and Hull-CFA are procedures that aim to identify the number of dimensions underlying data. The Hull-CAF procedure indicated that there was only one dimension underlying each scale. By contrast, the PA-MRFA procedure confirmed a single dimension for the MRMS but advised that three dimensions were underlying the IRMA. Timmerman et al. (2018) note that the PA-MRFA can overestimate the number of factors to extract when large samples are used. The Hull-CAF ignores minor factors and only focuses on major factors to minimise the risk of over-extraction (Lorenzo-Seva et al., 2011). Due to my comparison modelling approach, my research included models based on the recommendations of both factor retention indicators.

Although rooted in principal components theory, eigenvalues can provide valuable information when their limitations are appropriately acknowledged. If there is a large difference between the first and second eigenvalue, i.e., the ratio is greater than 3:1, this can serve as a preliminary indication that there is likely a general factor underlying the data that is substantial enough to warrant investigating (Reise et al., 2015). For the IRMAS, the first to second eigenvalue ratio was 9:1. For the MRMS, the ratio was 7.5:1. In both cases, the eigenvalue ratio indicated that both the MRMS and IRMAS were viable candidates for bifactor analysis, as they were likely to have a substantial general factor.

In my results, I reported several scale-level dimensionality indices: MIREAL, UniCo and ECV. Of those indices, only the MIREAL is model-independent and indicates departure from unidimensionality (Ferrando & Lorenzo-Seva, 2018). Both the IRMAS and MRMS had a recorded MIREAL of .17, which is considered low and did not flag a significant departure from unidimensionality.

For the IRMAS and MRMS, both the one-factor models and bifactor models were plausible upon initial inspection, and a good fit. To be clear, the one-factor models were not taken as definitive proof of common factors but did provide grounds to continue investigating a general factor for both the IRMAS and MRMS. Both the MRMS and IRMAS one-factor models had a UniCo value greater than .95; however, when recomputed for the bifactor models, the UniCo values dropped substantially, to .79 for the MRMS, and to .80 for the IRMAS. While these are still considered high UniCo values, the UniCo did not provide proof of essentially unidimensionality.

The reported ECV for the bifactor models was also notably lower than that reported for the one-factor models. Due to the decrease in ECV under the bifactor models, the ECV could not be cited as proof of essential unidimensionality. In terms of my study context, the decrease in ECV was expected, as under the bifactor model the general factor competes with the group factors to explain item variance (DeMars, 2013; Rodriguez et al., 2016). However, this difference illustrates why it is important to place dimensionality indices, and how they were calculated, in the appropriate context.

While the ECV did not provide evidence of essential unidimensionality when calculated based on the bifactor model, the ECV did indicate that the general factor underlying the MRMS and the general factor underlying the IRMAS were moderately strong, and dominated the response variation. The strength of the general factor in each scale was also confirmed by the OmegaH values calculated based on each bifactor model.

I posit that there is sufficient evidence of a moderately strong general factor underlying the MRMS, and there is also a moderately strong general factor underlying the IRMA. While the general factor did not completely saturate the item response variance for either scale, it did account for close to two-thirds of the common response variation in both the IRMAS and MRMS. Furthermore, all the MRMS items had substantial loadings on the general factor in the bifactor model, and the general factor explained over half the common variance at the item level for most items. The same was true for the IRMA, with the exception of item 5, which I will discuss in more detail in a later section.

Parameter Bias

The final, and arguably more strict criteria regarding item-level parameter bias was only fully uncovered in the model comparison stage. It was important to see whether the inherent multidimensionality in each scale resulted in biased item-level parameters under the one-factor model. The model comparison approach outlined by Reise et al. (2015) makes the impact of multidimensionality on item loadings salient.

For both the IRMAS and MRMS, the one-factor models showed excellent fit to the data and were plausible models. However, when compared to the general factor component of the alternative bifactor models, the one-factor models appeared less appropriate. Under an FA framework, there was considerable item-level parameter bias when the data was modelled as being explained by a single common factor. For both the IRMAS and MRMS, there was a large difference between the item loadings of the one-factor model, compared to the general factor of the bifactor model. The difference in FA parameters was calculated as relative bias percentage and revealed that too many items had a relative bias greater than 15%. The difference between the item loadings suggests that many loadings in the one-factor model were biased and appeared bigger than they really were. This parameter bias is something that occurs as the result of forcing multidimensional data into a unidimensional solution (Reise, Scheines, et al., 2013; Rodriguez et al., 2016).

However, under an IRT framework, I reached a very different conclusion. I argue that the differences in item slopes across the IRT models were small enough to justify treating both the IRMAS and MRMS as essentially unidimensional under an IRT framework. With that said, I think it would be important to acknowledge the models as essentially unidimensional, not absolutely unidimensional, and to acknowledge the bias in item parameters when interpreting the results.

Under a normal ogive IRT model, if an item has an item slope of .21 or less it is considered to have very low discriminatory power and is essentially not useful for IRT applications (Baker & Kim, 2017). When I compared the MRMS item discrimination under the unidimensional model to the item discrimination parameters for the general factor of the bifactor IRT model, the differences in item slopes were arguably very small. For 19 of the MRMS items, the difference in slope was less than .21. For the remaining three MRMS items, the slope difference was still small, coming in at under .3. I therefore posit that it would be justified to treat the MRMS as unidimensional for a wide range of IRT applications.

Worth noting is that the difference in slopes across the IRMAS models was more substantial than the MRMS models. The model comparison revealed that if modelled as unidimensional, the discriminatory power of many of the IRMAS items would be overestimated. However, the IRMAS bifactor model showed that the IRMAS items were in fact highly discriminating for the general dimension, and so I argue that the difference in slopes observed under the unidimensional model is small enough to justify modelling the IRMAS as unidimensional. With that said, I would limit treating the IRMAS as unidimensional only to IRT applications where small differences in slopes do not have major impacts.

It may not be appropriate to treat the IRMAS as unidimensional for the testing differential item functioning, but, it arguably could be treated as unidimensional for the purposes of adding the items to a broader rape myth acceptance item pool for computer adaptive testing. Most of the IRMAS items were moderately to very highly discriminating on the general dimension, which would be the construct of interest in computer adaptive testing. Take for example item 24 ("If the rapist doesn't have a weapon, you really can't call it rape"). Item 24 had a very high discrimination parameter under both the unidimensional and bifactor IRT models ($\alpha_{Uni} = 1.19$; $\alpha_{Uni} = 1.73$), and so would arguably be a good indicator for the general dimension in either model. However, it was also one of the most biased items ($\Delta_{\alpha} = -.54$), and so would arguably not be nearly as useful if the aim of fitting a unidimensional IRT model was to investigate differential item functioning.

When working with multidimensional models, the interpretation of item parameters and data structure can differ greatly across equivalent IRT and FA solutions (Reise, Moore, & Maydeu-Olivares, 2011; Reise et al., 2015). Although the models are transforms, the IRT item parameters factor in item communalities whereas FA parameters do not, which is one of the reasons why IRT parameter estimation is more robust in the presence of multidimensionality (Reise et al., 2015).

To conclude this section, I posit that neither the IRMAS nor the MRMS can be considered essentially unidimensional under a factor analytic framework. The loadings under the one-factor models for both scales were highly biased, with the majority appearing to be inflated and overestimated. This interpretation is further supported by the fact that neither model had an ECV or OmegaH value close to .8 (Rodriguez et al., 2016). However, both the IRMAS and MRMS could be considered essentially unidimensional under an IRT framework, particularly for IRT applications where the minor differences in slope are inconsequential. The differences in item slopes across the IRT models were arguably very small, likely because IRT parameters factor in item communalities, and therefore the unidimensional IRT models were relatively unbiased despite the presence of multidimensionality in both scales.

Dimensionality of the Combined Item Pool of the IRMAS and MRMS: Evidence of a Global RMA construct

For the final component of my research, I explored whether the items of the IRMAS and MRMS were indicators of the same latent construct. For this investigation, I combined the item pool of IRMAS and MRMS and analysed the dimensionality. In the discussion to follow, I will be presenting evidence from my exploratory investigation that suggests that rape myths, regardless of whether they centre on female or male rape victims, are effect indicators of the same causal construct. The pieces of evidence include a high correlation between the scales, similarities in response composition, a strong general factor underlying the combined data, and indications that the combined item pool could pass as essentially unidimensional under an IRT framework.

High Correlation and Similar Response Composition

The high correlation ($\rho = .77$; $r = .77$, $p < .0001$) between the IRMAS and MRMS was a preliminary indication of potential construct congruence. No scale measures a construct perfectly, there is a certain amount of noise and error that is inherent in every measure. Differences in noise and error would offer an explanation as to why the IRMAS and MRMS could be tapping into the same construct and yet not have a perfect correlation. In the context of this research investigation, the high correlation between the MRMS and IRMAS was anticipated and confirms the high correlation found in previous research studies (Davies et al., 2012; Walfield, 2018).

While high correlations are often sought out in literature focused on mapping out nomological networks, they can present data handling challenges if they are not detected and dealt with appropriately and should be dealt with early on if not the primary focus of a research investigation. For example, if two items have a correlation coefficient greater than .80 the researcher should consider removing one of the items from the analysis due to the possibility of item redundancy. When an entire scale is treated as a single variable, a correlation greater than 0.8 with another scale is considered a sign of serious multicollinearity. Multicollinearity can cause certain statistics, such as linear and multiple regression, to behave in unpredictable ways which is not desirable if a study aims to be replicable.

Therefore, the high correlation between the IRMAS and MRMS correlation has a practical relevance from the outset: it would be ill-advised to perform regression analyses that use both scales at the same time. This is something worth highlighting as these scales have been used in such analyses in literature in the past under the assumption that they measure related but different constructs (Walfield, 2018).

Both MRMS and IRMAS data were characterised by clear positive skews. Furthermore, when the response compositions were simplified into agreement and disagreement, the MRMS and IRMAS were very similar. It is my opinion that had the MRMS had a neutral category when it was administered, the response compositions would look even more similar. In addition, it is worth emphasising that for the IRMAS, more responses were recorded in the neutral category than all the agreement categories combined. I, therefore, posit that future research should consider including a neutral category for male rape myths too.

Strong General Factor

I used an exploratory bifactor modelling procedure to determine the extent of the common variance shared across all the items. Common variance shared between specific subsets of items was modelled over and above common variance already modelled by the general factor.

When the data was fitted to a bifactor model, the resulting general factor dominated the model and was more substantial than the specific group factors. If the items were measuring different constructs, the outcome should have looked quite different. Since there were double the number of IRMAS items, if the MRMS and IRMAS were tapping into different constructs, I would have expected all the MRMS items to either not load substantially on the general factor, or for all the MRMS items to have grouped at the lower end of the loading range, with the IRMAS items all grouping together with the highest loadings on the general factor. However, the results clearly showed that items from both the IRMAS and MRMS loaded substantially on the general factor, and furthermore that the general factor explained at least half the item-level common variance for most of the IRMAS and MRMS items.

Essentially unidimensional under IRT parameters

I explored whether the combined item pool could be characterised as essentially unidimensional under an IRT framework. The logic behind this investigation was that if the common variance between the items was substantial enough to result in an adequate unidimensional IRT model, it would provide a stronger level of evidence that the items tapped into the same causal construct.

Under IRT parameters, all item slopes for the general dimension were scored in the same direction under both the unidimensional model and for the general factor of the bifactor model. Furthermore, items that were the most discriminating for the general dimension came from both the MRMS and the IRMAS.

It was very interesting to see that the unidimensional IRT model for the combined item pool was even less biased than the unidimensional IRT models for the IRMAS and MRMS individually, which further supports the conclusion that the items likely tap into the same construct. An explanation for why the unidimensional model was less biased when both the IRMAS and MRMS items were included in the same analyses, was that the increase in the number of items resulted in an increase in the communality of the items. Communality is key to the calculation of IRT parameters (Reise et al., 2015). In this case, because the communality between the items increased, it appears the estimation of the unidimensional model was more accurate because there was more information for the modelling procedure to work with.

A unidimensional interpretation would not have been possible had the items of the IRMAS and MRMS been tapping into fundamentally different constructs. Therefore, based on the plausibility of the unidimensional IRT model, I posit that the items of the IRMAS and MRMS appear to share a common causal construct that influences their response variation.

However, it is important to note that interpretation under the factor analytic parameters leads to a very different conclusion. When the loadings of the general factor of the bifactor model were compared to loadings of the one-factor model, the relative bias for over half the items fell outside the acceptable range. For most of the items, the loadings in the one-factor model were overestimated. In other words, under the one-factor model, the item loadings on the common factor appeared much bigger than they really were. The results clearly showed that to retain a one-factor model for the data would be unacceptable under an FA framework, as the multidimensionality present in the data resulted in severely biased item loadings under the one-factor model. Therefore, under a factor analytic framework, the multidimensionality inherent in the data would need to be appropriately modelled, and the dataset could not be characterised as unidimensional.

Multidimensionality insights

While both the IRMAS and MRMS showed evidence of essential unidimensionality under an IRT framework, they also both had clear multidimensionality present. Through my explorative approach, I made several observations regarding the multidimensionality that presented across the three exploratory bifactor models I fitted to the IRMAS and MRMS data. While my research intention was to focus on unidimensionality, I would like to share my observations about the group factors that emerged, as well as some items that stood out for me.

Multidimensionality in the MRMS

From the outset, a key difference between the bifactor analysis of the IRMAS compared to the MRMS, was that the MRMS bifactor model was a theoretical exploration, that made use of bifactor modelling as a dissection tool. Neither the PA-MRFA nor Hull-CAF indicated multiple factors underlying the data. This is likely why although I reported 3 factors when I looked at the MRMS on its own- these same factors did not emerge under the combined item pool model.

The inter-factor correlation matrix for the MRMS solution indicated that, while F1 barely correlated with the other two factors, F2 and F3 were highly correlated, suggesting that F2 and F3 may have a substantive interpretation. Hogge and Wang (2022) recently revised the Male Rape Myth Scale by Kerr Melanson (1998) and investigated the underlying factor structure. The revised scale is notably different from the original, however the exploratory factor analysis by Hogge and Wang (2022, p. 426) revealed two correlated factors, which they named “Marginalisation” and “Victim Culpability” (p. 426). Relating back to my results, F2 is similar to the Victim Culpability factor identified by Hogge and Wang, as the factor loaded on most of the rape myths that lended themselves to victim blaming. F3 was also similar to the Marginalisation factor identified by Hogge and Wang, loading on rape myths that related rape to loss of manhood and homosexuality. However, at best, the factors that emerged in my study were only a rough approximation of the factors identified by Hogge and Wang.

I offer instead an alternative explanation for the factors that emerged in my MRMS bifactor model and posit that each factor elucidated variation potentially introduced by methodological factors. Of the 3 group factors elucidated by the bifactor model, F1 appeared to be a methodology factor caused by the reverse scoring of 3 MRMS items, F2 appeared to be a factor formed due to content repeats, and F3 consisted of items that barely met the cut-off criterion for a substantial loading. All three factors therefore elucidated potential avenues for multidimensionality to either be reduced in future, or to be taken into appropriate consideration by future researchers.

Multidimensionality in the IRMA

The IRMAS is divided into seven subscales, however, the PA-MRFA procedure identified only three factors underlying the IRMAS data. Under my bifactor model of the IRMAS data, the subscales outlined by Payne et al. (1999) appeared to merge, revealing (1) a factor that loaded predominantly on items from the She Wanted It and the She Lied subscales, (2) a factor that loaded predominantly on items from the Rape is a Deviant Event and the She Asked for It subscale, as well as (3) a factor that loaded predominantly on items from the He Didn't Mean To subscale.

According to Reise et al. (2018), it is common for group factors to constitute nuisance variation under a bifactor model, as once variation explained by the common factor has been accounted for, there is seldom enough reliable variation explained by specific group factors to warrant valid and reliable subscales. This is a phenomenon also known as factor collapse (Mansolf & Reise, 2016). Although it was not my intention to comment on the validity of the IRMAS subscales, in light of my exploratory investigation I would not recommend South African researchers plan a study that depends critically on utilising the subscale scores of the IRMAS.

With that said, there were very apparent similarities between the group factors that emerged under the IRMAS bifactor model and the factors that emerged under the bifactor model for the combined item pool of the IRMAS and MRMS. I posit that the specific factors that emerged captured variance from secondary constructs that future researchers may find relevant to consider, especially in the context of correlation and prediction studies. Due to the similarity in factors, I will unpack this further in my review of insights from my bifactor analysis of the combined item pool of the IRMAS and MRMS to avoid repetition.

One thing I will note before moving on, however, is that item 5 of the IRMAS had exceptionally poor psychometric properties. The distribution statistics for item 5 flagged it as an outlier from the beginning. It was the only item with a negative skew, and also had the highest mean of all the items, with 32% of the responses to this question falling in the neutral category. Furthermore, item 5 did not load substantially on the general factor when the IRMAS data were analysed independently, nor when the CIP data were analysed.

Multidimensionality in the Combined item pool

The exploratory bifactor model of the combined item pool revealed four group factors. Three of the group factors loaded predominantly on IRMAS items, with the factors appearing very similar to when the IRMAS data were analysed independently from the MRMS data. The fourth group factor loaded on 16 MRMS items, and none of the IRMAS items.

Methodology Factor. The factors that loaded mostly on IRMAS items appeared to be related secondary constructs as the factors were moderately correlated with each other. One specific factor, F3, consisted of only MRMS items. The specific construct that manifested as F3 barely correlated with the other three factors.

As noted by Howe et al. (2019), specific factors in a bifactor model can arise due to methodological differences among scales, such as differences in rating scales. The MRMS items were rated on a 6-point Likert scale, whereas the IRMAS items were rated on a 7-point Likert scale. Therefore, I believe that the difference in rating scales resulted in the emergence of a methodology factor, which loaded on the MRMS items because there were fewer items rated on a 6-point Likert scale. Furthermore, the MRMS factor had close to no correlation with the other factors in the solution. If it was tapping into a related construct, it should have been moderately correlated with the other three factors. I therefore note that the difference in rating scales was likely a major source of nuisance variation introduced in the combined item pool, and posed a significant limitation to my study.

The MRMS items may have grouped under different factors had the rating scales been applied consistently across my questionnaire. Furthermore, the MRMS items did all still load substantially on the general construct underlying the combined item pool, indicating male rape myth literature is relevant to the interpretation of the group factors. Therefore, in my interpretation of the remaining factors, which were almost exclusively loaded on IRMAS items, I will be drawing on both male rape myth and female rape myth literature in an attempt to make sense of the group factor loadings and provide meaningful insights for future researchers.

Invalidation, and differences in item interpretation. A clear source of multidimensionality that emerged from the IRMAS and CIP bifactor models, is the link specific rape myths have to invalidation - both in terms of victim's experiences and credibility. In both the combined item pool and IRMAS analysis, one of the specific factors loaded specifically on items from the She Wanted It (WI) and the She Lied (SL) subscales of the IRMAS (F1 in the IRMAS bifactor model, and F1 in the CIP bifactor model). Under both bifactor models, the factor had the highest loading on the following items "Many women find being forced to have sex very arousing" (irmas18) and "Although most women wouldn't admit it, they generally find being physically forced into sex a real 'turn on'" (irmas02).

I believe that what ties the rape myths from the SL and WI subscales together is that they invalidate the thoughts, sexual autonomy and experiences of victims, a feature of rape myths that has been well noted in RMA literature (Karimakwenda, 2021; Leverick, 2020; Shafer et al., 2018). What I would like to highlight, however, is that the grouping of items under this factor in the combined item pool analysis indicated that some items may have been interpreted differently by my South African student sample.

F1 loaded on item 29 of the IRMAS from the Trivial Event subscale, which is worded as follows "Women tend to exaggerate how much rape affects them". In my model, irmas29 grouped with items that imply that women are deceitful about having been raped, or that women secretly want to be raped. This suggests that participants interpreted this item not in terms of trivialising rape, but rather that women tend to lie about or embellish their experiences. Therefore, this item may have less to do with the perceived severity of rape, and more to do with the perceived validity of victims' experiences.

The other exception was irmas24, which originally fell under the "Not rape" scale and is worded as follows: "If the rapist doesn't have a weapon, you really can't call it rape". Read under the context of the original subscale, the implication was that if people agreed with the item, they supposedly held a narrow definition of rape. However, as was the case with irmas29, the grouping with the other items in this factor presents an interesting alternative interpretation: That women just "call" things rape, that are not really rape. In other words, the emphasis is perhaps less on whether the reader has a comprehensive understanding of what constitutes rape, and perhaps more on whether readers perceive victims as being able to reliably identify whether they have been raped.

Interesting to note is that when McMahon and Farmer (2011) revised the Illinois Rape Myth Acceptance scale, they completely eliminated the She Wanted It subscale after consulting with focus groups that deemed the items irrelevant. However, these items may still be relevant in the South African context. The items from the She Wanted It subscale had some of the highest I-ECV values and general factor loadings in the IRMAS bifactor solution. In particular, item 23 “Some women prefer to have sex forced on them so they don’t have to feel guilty about it.” was among the most highly discriminating items in the IRT models of the IRMAS. Therefore, while the She Wanted It subscale was removed by international researchers, my research suggests that the item content may still be relevant for RMA research in the South African context.

Gender diverse content range, othering, and rape as a sexually motivated crime. The second source of multidimensionality I would like to discuss, is the link some rape myths have to othering. In my literature I noted that many rape myths effectively distance rape, whether it be spatially, socially or behaviourally, and, furthermore, that rape myth acceptance has been linked to a multitude of oppressive belief systems such as sexism and racism (Aosved & Long, 2006; Obierefu & Ojedokun, 2017).

I argue that one of the specific factors that emerged under both the IRMAS bifactor model as well as the combined item pool analysis appeared to load on rape myths that related to othering rape, such as rape myths that relegated rapists to strangers, located rape as occurring far from home, and to victims who behave or dress in a certain way. The link between rape myths and othering in the South African context is well captured in the work of Dosekun (2013).

However, the presence of this specific factor suggests that many rape myths are introducing othering/marginalisation that is not entirely accounted for by rape myth acceptance. In particular, I would like to highlight irmas09 "Rape mainly occurs on the 'bad' side of town", and irmas13 "Men from nice middle-class homes almost never rape". Under the IRMAS bifactor model, both items had low I-ECV values, low loadings on the general factor, and a higher loading on the specific factor than the general factor. These are examples of items where, although they may be rape myths, the item-level variation was not predominately explained by the general rape myth acceptance factor.

However, I would also like to note that many items from the Deviant Event subscale appeared to have more variation explained by the specific group factor when modelled with only other IRMAS items. When the item pools were combined, many of the Deviant Event subscale items had increased general factor loadings and more substantial I-ECV values. This suggests that the introduction of MRMS items to the analysis increased the relevance of the IRMAS items characterised by othering in relation to the general dimension. Both Hogge and Wang (2022) and Hine et al. (2021) identified a form of othering/marginalisation as a key factor in their respective male rape myth measures.

When the items of the IRMAS and MRMS were analysed together, the items clustered together in slightly different ways compared to when the IRMAS and MRMS were looked at independently and revealed some interesting loadings. In particular, I would like to highlight the final factor in the CIP bifactor model, the factor that loaded on all the items from the He Didn't Mean To subscale of the IRMA.

The He Didn't Mean To subscale has persisted across multiple revisions of the Illinois Rape Myth Acceptance scale in various contexts (Fakunmoju et al., 2019; Johnson et al., 2023; McMahon & Farmer, 2011). However, I would like to draw attention to one item in particular that suggests the factor may not be as gendered as previously believed. The factor loaded on mrms17, “Women who rape men are sexually frustrated individuals”. The item mrms17 had a similar group factor loading and general factor loading to three IRMAS items in particular: irmas03, irmas20 and irmas42. Interesting to note is that all four of these items relate to the myth that rape is a sexually motivated crime. The implication then is that this factor may extend beyond justifying violence men perpetrate against women, and instead be better characterised as justifying rape, regardless of the gender of the perpetrator.

General discussion

My research supports the work of Urban and Porras Pyland (2022) and confirms that gendered rape myths are in fact highly likely to be manifestations of the same general rape myth acceptance construct. The one-factor and bifactor models, as well as the transformations to IRT parameters, elucidated that there is a general underlying factor/dimension that manifests in the items of the IRMA and the MRMS. I would therefore like to recommend researchers critically engage with the increasing likelihood of construct proliferation in rape myth acceptance research, and consider ways to overcome the associated limitations.

I would also like to relate my work to the observation made by Canan et al. (2023), who noted that gender-neutral items broaden the potential to place a greater focus on variables relevant to rape myths other than gender, such as the role of drugs and alcohol in rape myths, and rape myths that link into other oppressive belief systems such as classism and racism. It is true that the group factors that emerged in my study appeared to have less to do with gender, and more to do with (a) othering rape and its victims, (b) the invalidation of the experiences and credibility of victims and (c) denial and justification of rape motivations.

In addition, I would like to highlight that male rape myths that contained greater details about the perpetrator appeared to have notably high discrimination parameters in the IRT models. The suggestion, then, is to consider broadening the scope of rape myth items beyond stratifying measurement based on desires to adequately represent gender binaries at the item level. I posit instead considering the benefits of moving towards a wider view of rape myth acceptance that acknowledges the many various ways in which this particular variable manifests. Important to note is that my findings suggest that a gender-inclusive understanding of rape myths is highly relevant in the South African context.

It appears from the recent spate of literature on the role of gender in RMA measurement is to either (a) use gender-neutral language, (b) use variously gendered items or (c) use exclusively gendered items. I would like to suggest a radical alternative: to recognise that rape myth acceptance appears to be a general construct that can manifest in a wide variety of rape myths, both gendered and non. I do not think that gender should characterise every single item, nor do I think gender should be completely erased, as gender norms, roles and stereotypes do appear to interact with rape myth acceptance, and are a reality that should be recognised and accounted for. Further exploration of rape myths using item response theory may yield more insight into which variables are best manipulated at the item level, and to what extent those items are able to identify rape myth acceptance across different groups of people. I posit that given the sheer diversity of rape myths that exist, rape myth acceptance measurement could flourish under item response theory framework, and in particular could see benefit to creating a repository of rape myths for which the exact psychometric properties are known for each item, as is done for computer adaptive testing applications.

For any researchers interested in using item response theory, my research suggests that as a starting point, both MRMS and IRMA can be treated as essentially unidimensional, and combining the items from both scales in the same model would further reduce the biasing effect of the multidimensionality present in each scale due to the increased communality among items. As I noted in my literature review, when looked at individually, rape myths appear to have different functions and characterisations, such as some rape myths being characterised more by victim blaming, and other rape myths indicating narrow understandings of rape and consent. These differences in rape myths, the diversity of the manifestation of rape myth acceptance, is what makes item response theory such an attractive tool for dissecting this phenomenon.

Methodology insights

Comparison Modelling. It is worth highlighting that the one-factor model and accompanying model-dependent dimensionality indices were misleading under a factor analytic framework. For example, when the IRMA was fitted to a one-factor model, the fit indices and scale-level dimensionality indices supported the model and indicated the data was unidimensional. However, once the bifactor model was fitted, the scale level dimensionality indices did not indicate the data as being essentially unidimensional.

Therefore, my research supports the observations by Reise, Scheines, et al. (2013) and Ferrando and Lorenzo-Seva (2018) that model fit indices can look acceptable for inappropriate solutions. Comparison modelling elucidated the presence of multidimensionality inherent in the data that otherwise would have gone undetected had the one-factor model been looked at alone. Therefore, my recommendation is to use the comparison modelling approach outlined by Reise et al. (2015), and to consider using bifactor models to test the strength of the common factor in more depth before assuming data is unidimensional.

Number of Factors. The number of factors retained is an absolutely crucial step in the modelling process, and should always be well-motivated. Having seen first-hand how a different number of factors can impact on the loadings, I am now very wary of papers that do not specify this step. I would like to highlight that relying on Kaiser's criterion would have suggested a very different number of dimensions. According to Kaiser's criterion, the number of eigenvalues greater than 1 indicates how many components to specify. There were 6 eigenvalues that met this criterion for the IRMA and 3 for the MRMS. According to Taherdoost et al. (2014), Kaiser's criterion tends to overestimate the number of dimensions underlying data, possibly due to the fact that it is based on principal components theory and was not intended for factor analytic applications. Given the discrepancy in the number of recommended factors for both scales, I would like to caution future researchers against relying on Kaiser's criterion, and to rather consider the PA-MRFA and Hull-CAF procedures which are grounded in Factor Analytic theory.

Rescaling. The first thing I tried to do to overcome the differences in rating options was rescaling the responses of the IRMA and the MRMS to the same rating scale. This can be done through a simple conversion. I tried 3 separate rescalings: rescaling the MRMS to a 7-point, rescaling the IRMA to a 6-point, and rescaling both the IRMA and MRMS to an 8-point scale.

Although this can be an acceptable way to manipulate data, it was not feasible for this study. The polychoric correlation matrices passed Bartlett's test of sphericity but failed the KMO miserably, indicating that they were not an adequate foundation for further analysis. Even if the KMO had been slightly higher and closer to being acceptable, I would not have proceeded in this vein. A major concern was that the KMO responded erratically to the rescaling and the index varied greatly between rescaling trials. Given that the rescaling had made the KMO respond unpredictably, I was concerned that it might cause other analyses to behave erratically too, and so searched for an alternative way to combine the data.

It is possible that the rescaling was not successful because the IRMA had a neutral category, whereas the MRMS did not. Perhaps it may have been more successful had the one been rated on a 5-point Likert scale, and the other on a 7-point. A full investigation into rescaling was not the focus of this study, and so I decided to abandon this path and seek another solution. I have included my failed rescaling efforts in the hope that it can serve as a caution to future researchers considering such data manipulation.

Total Scores vs Item-Level Insights. The study sample did not contain many participants with extreme rape myth acceptance. Overall, 63% of the IRMA responses and 63% of the MRMS responses captured complete disagreement with the rape myths presented. Only 1% of IRMA responses indicated strong agreement with a rape myth, whereas 3% of MRMS responses indicated strong agreement.

Although the overall agreement rate was small by comparison, it is important to remember that global assessments can be reductionistic and misleading, which is a key reason why I have been investigating the potential for item response theory applications. There is a lot of information available at the item level that can get lost when a scale is only viewed in its entirety. One might be tempted to look at the above breakdown and conclude that rape myth acceptance was negligible in this sample, and not particularly concerning. However, an example of an item-level insight is the fact that 9% of the sample, 228 people, strongly agreed with the statement "The extent of a man's resistance should be a major factor in determining if he was raped". Even one person's RMA can have a serious impact on the lives of victims of sexual assault. I would therefore caution against looking to the majority when making conclusions about whether rape myth acceptance is a cause for concern.

Limitations

My sample, albeit relatively large, was drawn from a student population. Therefore, while my findings should be generalisable for student populations in South Africa, my findings cannot be generalised to the general South African population. However, my methodology should be applicable and usable for a general population sample.

I had a large sample of 2,536 participants. I needed a minimum of 500 participants to run my analyses, and the final sample was just over five times that. While fantastic for the factor analysis, large sample sizes do increase the chance of committing a type II error. A type II error is when there is a greater risk of failing to reject the null hypothesis (Foxcroft & Roodt, 2013). There are many statistics, such as the chi-square (X^2), that are sensitive to sample size (Mulaik, 2015). The large sample size also limited the type of factor estimation procedures available and meant that the more common ML procedure was not appropriate.

Although my sample was large, the response rate for this study was 12.76%. This is in line with postal response rates (Sinclair et al., 2012); however, I do think that there was room for improvement in the response rate. The majority of responses came from women (61.54%), and while women have been noted to dominate samples in rape myth acceptance literature (Walfield, 2018), I think it reflects a need to take a more proactive role in the recruitment process to ensure samples can be as representative as possible.

The questionnaire for this study was only available in English. The results showed that only half of the sample recorded English as their home language. The implication is that roughly 50% of the respondents completed the provided questionnaire in their second language. With that said, at Stellenbosch University the primary medium of instruction is English. Therefore, respondents were assumed to have an adequate grasp of English to complete the survey with proficiency.

Although the questionnaire was administered in South Africa, it is unclear what percentage of the sample identified as South African. All 11 of South Africa's official languages were recorded by the sample. Just over 90% of the respondents recorded either English, Afrikaans, or isiXhosa as their home language. All three languages are widely spoken in the Western Cape, the South African province in which Stellenbosch University is located. Shona and German were the only two languages to appear at a higher frequency than one or more of the official languages of South Africa. Shona and German are commonly spoken in Zimbabwe and Namibia respectively, and both countries border South Africa.

The data for both MRMS and IRMA were non-normally distributed, and positively skewed. Non-normal data distributions are typical in rape myth acceptance measurement, a criticism made particularly salient by Gerger et al. (2007). Due to the non-normal distribution of my data, I needed to ensure that the statistics and procedures I used did not rely on an assumption of normality.

The IRMAS and MRMS use different rating scales, which provided some challenges to data handling. First, I had to compute the polychoric correlation matrix in R, and could not use the raw data to analyse the combined item pool in FACTOR. This, then, had knock-on effects, which were not ideal. Due to running the analyses straight off the polychoric correlation matrix, the IRT parameter transform needed to be calculated manually, and FACTOR could not compute confidence intervals for the exploratory factor solutions.

While there were several significant limitations to my study, my hope is that future researchers can learn from and improve on my methods.

Recommendations

My first recommendation is to consider using similar rating scales if the MRMS and IRMA are administered in the same questionnaire. I received an email from one of the participants who was displeased that the IRMA items had a different rating scale to the items of the MRMS. They noted that the absence of the neutral category for the MRMS items meant that they felt forced to answer yes or no to statements that they neither disagreed nor agreed with. However, what was stressed by the respondent was not so much that there was no neutral category, but rather that one set of questions had the flexibility that comes with a neutral category and not the other. What strikes me is that the email assumed that the questions were trying to measure the same thing, which, as my exploration has uncovered, is likely the case.

Considering this response, and the fact that these indicators are likely tapping into the same construct, I would like to suggest that if the IRMA and MRMS are administered together, researchers consider standardising the rating scale. It is important to consider what kinds of biases are reinforced by the questions that we do and do not ask in surveys. Similarly, it is important to consider what may be implied when rating scales appear to differ by gender. While it was not known at the time whether these questionnaires were tapping into the same general construct or not, I do acknowledge that when exploring similarities between historically gendered measures it is important to not appear to be biasing measurement in the eyes of respondents.

My second recommendation is for future researchers to investigate partially completed responses. I received 250 incomplete responses. As all incomplete responses were taken as a sign of withdrawal, this study had a 10% withdrawal rate. Future researchers could perhaps have a closer look into incomplete response data to see at what point people stopped responding and the impact on sample composition.

My third recommendation is to use practical incentives to increase participation. I had a range of incentives, which included a bicycle, cash prizes and necklaces. I received many emails regarding the bicycle and believe it may have been the main reason why I had such a high entry rate into the lucky draws. The entry rate into the lucky draws was 97%, which strongly suggests that the incentives formed an important motivation for participation in my research, and the large sample size I had to work with.

While FA is particularly suited for dimensionality investigations, as it is easier to navigate and the terminology is more widely understood, I highly recommend researchers continue to investigate RMA data under an IRT framework. I think it would be well worthwhile to plot item characteristic curves for rape myths and consider research into computer adaptive testing applications. I think computer adaptive testing could be particularly well suited for South African rape myth acceptance research, as it would drastically reduce the number of items required to assess rape myth acceptance.

I also recommend exploring the extent to which differently gendered item content effectively discriminates between people with high and low rape myth acceptance, and to assess differential item bias. For example, rape myths that are male-centric may be more effective at indicating RMA among women than among men, and vice versa. Understanding more about the role gendered rape myth content plays in the measurement of rape myth acceptance may be necessary to capture the full extent of the way that this variable can manifest across different groups of people.

Conclusion

The strength of my research lies in the way I explored the data, and how I arrived at the models I looked at. The insights I provided in this discussion were based on a thorough understanding of the data structure for each data set I looked at, as well as a great deal of research into the strengths and limitations of each step of my methodology. I maximised the strengths of exploratory factor analysis, and utilised bifactor modelling as a tool for mapping the common variance between items. It is important to recognise that rape myth acceptance is a schema that does not operate within a vacuum. There is very likely interplay with other schemas, gender norms, masculine and feminine expectations, defensive attributions and othering. I hope that my research has provided a starting point in clarifying to what extent the IRMAS and

MRMS are measuring rape myth acceptance, as opposed to related secondary constructs.

Throughout my research, I have highlighted similarities between the rape myths used in the IRMAS, and the rape myths used in the MRMS. I believe that my exploration revealed similarities not only in literature findings and item wording, but also similarities in data structure, and evidence of a general rape myth acceptance variable that manifests in both male rape myths and female rape myths. I posit that modelling the data for both the IRMA and MRMS as unidimensional under an IRT framework would be justified, and actively encourage researchers to consider using item response theory to analyse rape myth acceptance data.

However, other researchers might look at the same slope differences, and the evidence I have provided for a general RMA construct and come to different conclusions. The value of my research does not necessarily lie in my conclusions, but rather in the method employed, and the transparency with which I have reported my results. My hope with my research was to provide a methodology that was replicable and to learn more about the general dimension that underlies RMA measures, particularly when used in the South African context, and I believe I have done both.

References

Abad, F. J., Garcia-Garzon, E., Garrido, L. E., & Barrada, J. R. (2017). Iteration of partially specified target matrices: Application to the bi-factor case. *Multivariate Behavioral Research*, *52*(4), 416 429. https://doi.org/10.1080/00273171.2017.1301244

Abrahams, N., Jewkes, R., & Mathews, S. (2013). Depressive symptoms after a sexual assault among women: Understanding victim-perpetrator relationships and the role of social perceptions. *African Journal of Psychiatry*, *16*(4), 288 293. https://doi.org/10.4314/ajpsy.v16i4.39

Adams-Clark, A. A., & Chrisler, J. (2018). What constitutes rape? The effect of marital status and type of sexual act on perceptions of rape scenarios. *Violence Against Women*, *24*(16), 1867 1886. https://doi.org/10.1177/1077801218755975

Adoch, C. (2022). Defining rape: The problem with consent. *Strathmore Law Journal*, *6*(1), 71 92. https://doi.org/10.52907/slj.v6i1.172

Adolfsson, K., Strömwall, L. A., & Landström, S. (2020). Blame attributions in multiple perpetrator rape cases: The impact of sympathy, consent, force, and beliefs. *Journal of Interpersonal Violence*, *35*(23), 5336 5364. https://doi.org/10.1177/0886260517721171

Aosved, A. C., & Long, P. J. (2006). Co-occurrence of rape myth acceptance, sexism, racism, homophobia, ageism, classism, and religious intolerance. *Sex Roles*, *55*(7), 481 492. https://doi.org/10.1007/s11199-006-9101-4

Baglin, J. (2014). Improving your exploratory factor analysis for ordinal data: A demonstration using FACTOR. *Practical Assessment, Research, and Evaluation*, *19*, Article 5. https://doi.org/10.7275/DSEP-4220

Baker, F. B., & Kim, S.-H. (2017). The basics of item response theory using R. Springer International Publishing.

Barnett, M., Hale, T., & Sligar, K. (2017). Masculinity, femininity, sexual dysfunctional beliefs, and rape myth acceptance among heterosexual college men and women. *Sexuality & Culture*, *21*(3), 741 753. https://doi.org/10.1007/s12119-017-9420-3

Bartlett, M. S. (1951). The effect of standardization on a chi square approximation in factor analysis. *Biometrika*, *38*, 337 344.

Bashonga, R., & Khuzwayo, Z. (2017). "This thing of the victim has to prove that the perp intended to assault is kak!": Social media responses to sexual violence on South African university campuses. *Agenda*, *31*(3), 35 49. https://doi.org/10.1080/10130950.2017.1394636

Beres, M. A., Treharne, G. J., & Stojanov, Z. (2019). A whole campus approach to sexual violence: The University of Otago model. *Journal of Higher Education Policy and Management*, *41*(6), 646 662. https://doi.org/10.1080/1360080X.2019.1613298

Bianchi, R. (2020). Do burnout and depressive symptoms form a single syndrome? Confirmatory factor analysis and exploratory structural equation modeling bifactor analysis. *Journal of Psychosomatic Research*, *131*, 109954. https://doi.org/10.1016/j.jpsychores.2020.109954

Bollen, K. A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology*, *53*(1), 605 634. https://doi.org/10.1146/annurev.psych.53.100901.135239

Bollen, K. A., & Bauldry, S. (2011). Three Cs in measurement models: Causal indicators, composite indicators, and covariates. *Psychological Methods*, *16*(3), 265 284. https://doi.org/10.1037/a0024448

Bollen, K. A., & Ting, K.-f. (2000). A tetrad test for causal indicators. *Psychological Methods*, *5*(1), 3 22. https://doi.org/10.I037//1082-9S9X.5.I.3

Bornovalova, M. A., Choate, A. M., Fatimah, H., Petersen, K. J., & Wiernik, B. M. (2020). Appropriate use of bifactor analysis in psychopathology research: Appreciating benefits and limitations. *Biological Psychiatry*, *88*(1), 18 27. https://doi.org/10.1016/j.biopsych.2020.01.013

Braeken, J., & Van Assen, M. A. L. M. (2017). An empirical Kaiser criterion. *Psychological Methods*, *22*(3), 450 466. https://doi.org/10.1037/met0000074

Burt, M. R. (1980). Cultural myths and supports for rape. *Journal of Personality and Social Psychology*, *38*(2), 217 230. https://doi.org/10.1037/0022-3514.38.2.217

Cai, L., & Moustaki, I. (2018). Estimation methods in latent variable models for categorical outcome variables. In P. Irwing, T. Booth, & D. J. Hughes (Eds.), *The Wiley handbook of psychometric testing: A multidisciplinary reference on survey, scale and test development* (pp. 253 277, Vol. 1). John Wiley & Sons Ltd.

Canan, S. N., Cozzolino, L., Myers, J. L., & Jozkowski, K. N. (2023). Does gender inclusive language affect psychometric properties of the Illinois Rape Myth Acceptance Scale-Short Form? A two-sample validation study. *Journal of Interpersonal Violence*, *38*(3), 3373 3394. https://doi.org/10.1177/08862605221106144

Chapleau, K. M., Oswald, D. L., & Russell, B. L. (2007). How ambivalent sexism toward women and men support rape myth acceptance. *Sex Roles*, *57*(1), 131 136. https://doi.org/10.1007/s11199-007-9196-2

Chapleau, K. M., Oswald, D. L., & Russell, B. L. (2008). Male rape myths: The role of gender, violence, and sexism. *Journal of Interpersonal Violence*, *23*(5), 600 615. https://doi.org/10.1177/0886260507313529

Charly, D., & Reddy, K. (2019). Understanding blame attributions in rape among legal professionals. *International Journal of Criminal Justice Sciences*, *14*(2), 222 238. https://doi.org/10.5281/zenodo.3719281

Chen, F. F., Jing, Y., Hayes, A., & Lee, J. M. (2013). Two concepts or two approaches? A bifactor analysis of psychological and subjective well-being. *Journal of Happiness Studies*, *14*(3), 1033 1068. https://doi.org/10.1007/s10902-012-9367-x

Cherniawsky, S., & Morrison, M. (2022). “You should have known better”: The social ramifications of victimization-focused sexual assault prevention tips. *Journal of Interpersonal Violence*, *37*(1), NP125 NP146. https://doi.org/10.1177/0886260520913650

Dartnall, E., & Jewkes, R. (2013). Sexual violence against women: The scope of the problem. *Best Practice & Research Clinical Obstetrics & Gynaecology*, *27*(1), 3 13. https://doi.org/10.1016/j.bpobgyn.2012.08.002

Davies, M., Gilston, J., & Rogers, P. (2012). Examining the relationship between male rape myth acceptance, female rape myth acceptance, victim blame, homophobia, gender roles, and ambivalent sexism. *Journal of Interpersonal Violence*, *27*(14), 2807 2823. https://doi.org/10.1177/0886260512438281

Davies, M., & Rogers, P. (2006). Perceptions of male victims in depicted sexual assaults: A review of the literature. *Aggression and Violent Behavior*, *11*(4), 367 377. https://doi.org/10.1016/j.avb.2006.01.002

Davies, M., Walker, J., Archer, J., & Pollard, P. (2013). The scripting of male and female rape. *Journal of Aggression, Conflict and Peace Research*, *5*(2), 68 76. https://doi.org/10.1108/17596591311313663

DeMars, C. E. (2013). A tutorial on interpreting bifactor model scores. *International Journal of Testing*, *13*(4), 354 378. https://doi.org/10.1080/15305058.2013.799067

Dewald, S., & Lorenz, K. (2022). Lying about sexual assault: A qualitative study of detective perspectives on false reporting. *Policing and Society*, *32*(2), 179 199. https://doi.org/10.1080/10439463.2021.1893725

Dima, A. L. (2018). Scale validation in applied health research: Tutorial for a 6-step R-based psychometrics protocol. *Health Psychology and Behavioral Medicine*, *6*(1), 136 161. https://doi.org/10.1080/21642850.2018.1472602

DiStefano, C., Liu, J., Jiang, N., & Shi, D. (2018). Examination of the weighted root mean square residual: Evidence for trustworthiness? *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(3), 453 466. https://doi.org/10.1080/10705511.2017.1390394

DiStefano, C., McDaniel, H. L., Zhang, L., Shi, D., & Jiang, Z. (2019). Fitting large factor analysis models with ordinal data. *Educational and Psychological Measurement*, *79*(3), 417 436. https://doi.org/10.1177/0013164418818242

Dosekun, S. (2013). 'Rape is a huge issue in this country': Discursive constructions of the rape crisis in South Africa. *Feminism & Psychology*, *23*(4), 517 535. https://doi.org/10.1177/0959353513493614

Dowling, N. M., Raykov, T., & Marcoulides, G. A. (2020). A method for examining the equating of psychometric scales and tests: An application using dementia screening test batteries. *Educational and Psychological Measurement*, *80*(1), 199 209. https://doi.org/10.1177/0013164418775785

Du Plessis, N., Kagee, A., & Maw, A. (2009). Women's experiences of reporting rape to the police: A qualitative study. *Social Work*, *45*(3), 275 289. https://doi.org/10.15270/45-3-204

Egan, R., & Wilson, J. C. (2012). Rape victims' attitudes to rape myth acceptance. *Psychiatry, Psychology and Law*, *19*(3), 345 357. https://doi.org/10.1080/13218719.2011.585128

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Lawrence Erlbaum Associates, Inc.

Emezue, C., & Udmuangpia, T. (2022). Authentic empathy and the role of victim service providers in (de)stigmatizing male sexual victimization. *Journal of Interpersonal Violence*, *37*(7), NP3832 NP3855. https://doi.org/10.1177/0886260520948150

Eyssel, F., & Bohner, G. (2011). Schema effects of rape myth acceptance on judgments of guilt and blame in rape cases: The role of perceived entitlement to judge. *Journal of Interpersonal Violence*, *26*(8), 1579 1605. https://doi.org/10.1177/0886260510370593

Fakunmoju, S., Abrefa-Gyan, T., & Maphosa, N. (2019). Confirmatory factor analysis and gender invariance of the revised IRMA scale in Nigeria. *Affilia*, *34*(1), 83 98. https://doi.org/10.1177/0886109918803645

Fernando, G. (2019). Rape and murder of teenager sparks protests across South Africa. *New Zealand Herald*.

https://www.nzherald.co.nz/world/rape-and-murder-of-teenager-sparks-protests-across-south-africa/6KCIEX4UOKEB3ZYAWLQWM3S2RQ/

Ferrando, P. J., & Lorenzo-Seva, U. (2018). Assessing the quality and appropriateness of factor solutions and factor score estimates in exploratory item factor analysis. *Educational and Psychological Measurement*, *78*(5), 762 780. https://doi.org/10.1177/0013164417719308

Finch, W. H. (2020). Using fit statistic differences to determine the optimal number of factors to retain in an exploratory factor analysis. *Educational and Psychological Measurement*, *80*(2), 217 241. https://doi.org/10.1177/0013164419865769

Finchilescu, G., & Dugard, J. (2021). Experiences of gender-based violence at a South African university: Prevalence and effect on rape myth acceptance. *Journal of Interpersonal Violence*, *36*(5), NP2749 NP2772. https://doi.org/10.1177/0886260518769352

Fine, M. A. (1992). On the distinction between two tests as measures of the same construct and as interchangeable. *American Psychologist*, *47*(9), 1146. https://doi.org/10.1037/0003-066X.47.9.1146.a

Fisher, N. L., & Pina, A. (2013). An overview of the literature on female-perpetrated adult male sexual victimization. *Aggression and Violent Behavior*, *18*(1), 54 61. https://doi.org/10.1016/j.avb.2012.10.001

Fleming, K., & Kruger, L.-M. (2013). “She keeps his secrets”: A gendered analysis of the impact of shame on the non-disclosure of sexual violence in one low-income South African community. *African Safety Promotion Journal*, *11*(2), 107 124. https://www.ajol.info/index.php/asp/article/view/136104

Flora, D. B. (2020). Your coefficient alpha is probably wrong, but which coefficient omega is right? A tutorial on using R to obtain better reliability estimates. *Advances in Methods and Practices in Psychological Science*, *3*(4), 484 501. https://doi.org/10.1177/2515245920951747

Foxcroft, C., & Roodt, G. (Eds.). (2013). *Introduction to psychological assessment in the South African context* (4th ed.). Oxford University Press.

Francke, R.-L. (2019). Thousands protest in South Africa over rising violence against women. *The Guardian*. https://www.theguardian.com/world/2019/sep/05/thousands-protest-in-south-africa-over-rising-violence-against-women

Free Software Foundation. (2018). *GNU PSPP* (Version 1.2.0-g0fb4db) [Statistical Analysis Software]. Boston, MA. https://www.gnu.org/software/pspp/

Furr, R. M., & Bacharach, V. R. (2014). *Psychometrics: An introduction* (2nd ed.). SAGE Publications.

Gamerman, D., Gonçalves, F. B., & Soares, T. M. (2018). Differential item functioning. In W. J. v. d. Linden (Ed.), *Handbook of item response theory. volume 3: Applications* (pp. 67 86). CRC Press, Taylor & Francis Group.

Garcia-Garzon, E., Abad, F. J., & Garrido, L. E. (2021). On omega hierarchical estimation: A comparison of exploratory bi-factor analysis algorithms. *Multivariate Behavioral Research*, *56*(1), 101 119. https://doi.org/10.1080/00273171.2020.1736977

George, D., & Mallery, P. (2016). *IBM SPSS statistics 23 step by step: A simple guide and reference* (Fourteenth edition). Routledge, Taylor & Francis Group.

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. https://doi.org/10.1002/ab.20195

Gouws, A. (2018). #EndRapeCulture campaign in South Africa: Resisting sexual violence through protest and the politics of experience. *Politikon*, *45*(1), 3 15. https://doi.org/10.1080/02589346.2018.1418201

Grubb, A., & Turner, E. (2012). Attribution of blame in rape cases: A review of the impact of rape myth acceptance, gender role conformity and substance use on victim blaming. *Aggression and Violent Behavior*, *17*(5), 443 452. https://doi.org/10.1016/j.avb.2012.06.002

Hartmann, L. (2019). *"Los my poes af"* the fine line between being radical enough and being too radical. *Agenda*, *33*(2), 74 83. https://doi.org/10.1080/10130950.2019.1602950

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods*, *7*(2), 191 205. https://doi.org/10.1177/1094428104263675

Hine, B. A., & Murphy, A. D. (2019). The influence of 'high' vs. 'low' rape myth acceptance on police officers' judgements of victim and perpetrator responsibility, and rape authenticity. *Journal of Criminal Justice*, *60*, 100 107. https://doi.org/10.1016/j.jcrimjus.2018.08.001

Hine, B. A., Murphy, A. D., & Churchyard, J. S. (2021). Development and validation of the Male Rape Myth Acceptance Scale (MRMAS). *Heliyon*, *7*(6), e07421. https://doi.org/10.1016/j.heliyon.2021.e07421

Hogge, I., & Wang, Y.-W. (2022). Revising the Male Rape Myths Scale. *Psychology of Men & Masculinities*, *23*(4), 422 423. https://doi.org/10.1037/men0000405

Howe, G. W., Dagne, G. A., Brown, C. H., Brincks, A. M., Beardslee, W., Perrino, T., & Pantin, H. (2019). Evaluating construct equivalence of youth depression measures across multiple measures and multiple studies. *Psychological Assessment*, *31*(9), 1154 1167. https://doi.org/10.1037/pas0000737

Hughes, D. J. (2018). Psychometric validity: Establishing the accuracy and appropriateness of psychometric measures. In P. Irwing, T. Booth, & D. J. Hughes (Eds.), *The Wiley handbook of psychometric testing: A multidisciplinary reference on survey, scale and test development* (1st ed., pp. 751 780, Vol. 2). John Wiley & Sons Ltd.

Ip, E. H., & Chen, S.-H. (2015). Using projected locally dependent unidimensional models to measure multidimensional response data. In S. P. Reise & D. A. Revicki (Eds.), *Handbook of item response theory modeling: Applications to typical performance assessment* (1st, pp. 226 251). Routledge, Taylor & Francis Group.

Javaid, A. (2017). Making the invisible visible: (un)meeting male rape victims' needs in the third sector. *Journal of Aggression, Conflict and Peace Research*, *9*(2), 106 115. https://doi.org/10.1108/JACPR-08-2016-0248

Javaid, A. (2018). Male rape, masculinities, and sexualities. *International Journal of Law, Crime and Justice*, *52*, 199 210. https://doi.org/10.1016/j.ijlcj.2017.12.003

Javaid, A. (2019). The invisible, the alien and the marginal: Social and cultural constructions of male rape in voluntary agencies. *International Review of Victimology*, *25*(1), 107 123. https://doi.org/10.1177/0269758017745614

Jennrich, R. I. (2018). Rotation. In P. Irwing, T. Booth, & D. J. Hughes (Eds.), *The Wiley handbook of psychometric testing: A multidisciplinary reference on survey, scale and test development* (1st ed., pp. 278 304, Vols. 2, Vol. 1). John Wiley & Sons Ltd.

Jennrich, R. I., & Bentler, P. M. (2011). Exploratory bi-factor analysis. *Psychometrika*, *76*(4), 537 549. https://doi.org/10.1007/s11336-011-9218-4

Jennrich, R. I., & Bentler, P. M. (2012). Exploratory bi-factor analysis: The oblique case. *Psychometrika*, *77*(3), 442 454. https://doi.org/10.1007/s11336-012-9269-1

Jewkes, R., Sikweyiya, Y., Dunkle, K., & Morrell, R. (2015). Relationship between single and multiple perpetrator rape perpetration in South Africa: A comparison of risk factors in a population-based sample. *BMC Public Health*, *15*(1), 616. https://doi.org/10.1186/s12889-015-1889-9

Jewkes, R., Sikweyiya, Y., Morrell, R., & Dunkle, K. (2010). Why, when and how men rape. *SA Crime Quarterly*, *34*, 23 31. https://doi.org/10.17159/2413-3108/2010/v0i34a874

Jewkes, R., Sikweyiya, Y., Morrell, R., & Dunkle, K. (2011). Gender inequitable masculinity and sexual entitlement in rape perpetration South Africa: Findings of a cross-sectional study. *PLOS One*, *6*(12), e29590. https://doi.org/10.1371/journal.pone.0029590

Jina, R., Jewkes, R., Christofides, N., & Loots, L. (2013). Knowledge and confidence of South African health care providers regarding post-rape care: A cross-sectional study. *BMC Health Services Research*, *13*(257). https://doi.org/10.1186/1472-6963-13-257

Jina, R., Machisa, M., Labuschagne, G., Vetten, L., Loots, L., & Jewkes, R. (2020). Unspoken victims: A national study of male rape incidents and police investigations in South Africa. *South African Medical Journal*, *110*(9), 926 931. https://doi.org/10.7196/SAMJ.2020.v110i9.14525

Johnson, N. L., Lipp, N. S., & Stone, H. K. (2023). Initial evaluation of a gender-inclusive version of the Illinois Rape Myth Acceptance Scale. *Psychology of Sexual Orientation and Gender Diversity*, *10*(2), 206 216. https://doi.org/10.1037/sgd0000536

Joreskog, K. G. (2003). Factor analysis by MINRES [Technical Report]. Retrieved July 8, 2022, from https://www.ssicentral.com/wp-content/uploads/2021/04/lis_minres.pdf

Kahn, A., Rodgers, K., Martin, C., Malick, K., Claytor, J., Gandolfo, M., Seay, R., McMillan, J., & Webne, E. (2011). Gender versus gender role in attributions of blame for a sexual assault. *Journal of Applied Social Psychology*, *41*(2), 239 251. https://doi.org/10.1111/j.1559-1816.2010.00711.x

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31 36. https://doi.org/10.1007/BF02291575

Kalaf, J., Vilete, L. M. P., Volchan, E., Fiszman, A., Coutinho, E. S. F., Andreoli, S. B., Quintana, M. I., De Jesus Mari, J., & Figueira, I. (2015). Peritraumatic tonic immobility in a large representative sample of the general population: Association with posttraumatic stress disorder and female gender. *Comprehensive Psychiatry*, *60*, 68 72. https://doi.org/10.1016/j.comppsych.2015.04.001

Kalichman, S., Simbayi, L., Cain, D., Cherry, C., Henda, N., & Cloete, A. (2007). Sexual assault, sexual risks and gender attitudes in a community sample of South African men. *AIDS Care*, *19*(1), 20 27. https://doi.org/10.1080/09540120600984003

Kalichman, S., Simbayi, L., Kaufman, M., Cain, D., Cherry, C., Jooste, S., & Mathiti, V. (2005). Gender attitudes, sexual violence, and HIV/AIDS risks among men and women in Cape Town, South Africa. *Journal of Sex Research*, *42*(4), 299 305. https://doi.org/10.1080/00224490509552285

Karimakwenda, N. (2021). The language of erasure: Neutralization techniques in contemporary South African marital rape judgments. *Violence Against Women*, *27*(3), 378 398. https://doi.org/10.1177/1077801219897842

Kassing, L. R., Beesley, D., & Frey, L. L. (2005). Gender role conflict, homophobia, age, and education as predictors of male rape myth acceptance. *Journal of Mental Health Counseling*, *27*(4), 311 328. https://doi.org/10.17744/mehc.27.4.9wfm24f52kqgav37

Kassing, L. R., & Prieto, L. R. (2003). The rape myth and blame-based beliefs of counselors-in-training toward male victims of rape. *Journal of Counseling & Development*, *81*(4), 455 461. https://doi.org/10.1002/j.1556-6678.2003.tb00272.x

Keehn, E., Stemple, L., Sanger, C., & Peacock, D. (2014). Uneven and still insufficient: South African Police Services' station-level compliance with sexual offences laws. *Feminist Criminology*, *9*(2), 87 112. https://doi.org/10.1177/1557085113501220

Kerr Melanson, P. S. (1998, August). *Belief in male rape myths: A test of two competing theories* [Doctoral dissertation]. Queen's University.

Koss, M. P. (1985). The hidden rape victim: Personality, attitudinal, and situational characteristics. *Psychology of Women Quarterly*, *9*(2), 193 212. https://doi.org/10.1111/j.1471-6402.1985.tb00872.x

Landström, S., Strömwall, L., & Alfredsson, H. (2016). Blame attributions in sexual crimes: Effects of belief in a just world and victim behavior. *Nordic Psychology*, *68*(1), 2 11. https://doi.org/10.1080/19012276.2015.1026921

Lankster, N. M. (2019). Chap Chat: Gender relations and perceptions of rape amongst adolescent males in South Africa. *Journal of Feminist Family Therapy*, *31*(2), 129 143. https://doi.org/10.1080/08952833.2019.1574491

Lee, J., Kim, J., & Lim, H. (2010). Rape myth acceptance among Korean college students: The roles of gender, attitudes toward women, and sexual double standard. *Journal of Interpersonal Violence*, *25*(7), 1200 1223. https://doi.org/10.1177/0886260509340536

Leverick, F. (2020). What do we know about rape myths and juror decision making? *International Journal of Evidence & Proof*, *24*(3), 255 279. https://doi.org/10.1177/1365712720923157

Levin, R. J., & Van Berlo, W. (2004). Sexual arousal and orgasm in subjects who experience forced or non-consensual sexual stimulation   a review. *Journal of Clinical Forensic Medicine*, *11*(2), 82 88. https://doi.org/10.1016/j.jcfm.2003.10.008

Lisak, D., Gardinier, L., Nicksa, S. C., & Cote, A. M. (2010). False allegations of sexual assault: An analysis of ten years of reported cases. *Violence Against Women*, *16*(12), 1318 1334. https://doi.org/10.1177/1077801210387747

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

Lorenzo-Seva, U., & Ferrando, P. J. (2019a). A general approach for fitting pure exploratory bifactor models. *Multivariate Behavioral Research*, *54*(1), 15 30. https://doi.org/10.1080/00273171.2018.1484339

Lorenzo-Seva, U., & Ferrando, P. J. (2019b). Robust promin: A method for diagonally weighted factor rotation. *Liberabit: Revista Peruana de Psicología*, *25*(1), 99 106. https://doi.org/10.24265/liberabit.2019.v25n1.08

Lorenzo-Seva, U., & Ferrando, P. J. (2021). *FACTOR* [Computer software]. Universitat Rovira i Virgili, Tarragona, Spain. https://psico.fcep.urv.cat/utilitats/factor/index.html

Lorenzo-Seva, U., Timmerman, M. E., & Kiers, H. A. L. (2011). The Hull method for selecting the number of common factors. *Multivariate Behavioral Research*, *46*(2), 340 364. https://doi.org/10.1080/00273171.2011.564527

Lowe, M., & Rogers, P. (2017). The scope of male rape: A selective review of research, policy and practice. *Aggression and Violent Behavior*, *35*(2), 38 43. https://doi.org/10.1016/j.avb.2017.06.007

Macleod, C., Böhmke, W., Mavuso, J., Barker, K., & Chiweshe, M. (2018). Contesting sexual violence policies in higher education: The case of Rhodes University. *Journal of Aggression, Conflict and Peace Research*, *10*(2), 83 92. https://doi.org/10.1108/JACPR-05-2017-0295

Makongoza, M., & Nduna, M. (2017). Awareness and rejection accounts of intimate partner violence by young women in Soweto, Johannesburg, South Africa. *Journal of Interpersonal Violence*, *36*(1), 7 32. https://doi.org/10.1177/0886260517726413

Mansolf, M., & Reise, S. P. (2016). Exploratory bifactor analysis: The Schmid-Leiman orthogonalization and Jennrich-Bentler analytic rotations. *Multivariate Behavioral Research*, *51*(5), 698 717. https://doi.org/10.1080/00273171.2016.1215898

Maxwell, L., & Scott, G. (2014). A review of the role of radical feminist theories in the understanding of rape myth acceptance. *Journal of Sexual Aggression*, *20*(1), 40 54. https://doi.org/10.1080/13552600.2013.773384

Mayeza, E. (2022). Rape culture: Sexual intimidation and partner rape among youth in sexually diverse relationships [Advance online publication]. *Sexualities*. https://doi.org/10.1177/13634607221144618

McLean, I. A. (2013). The male victim of sexual assault. *Best Practice & Research Clinical Obstetrics & Gynaecology*, *27*(1), 39 46. https://doi.org/10.1016/j.bpobgyn.2012.08.006

McMahon, S., & Farmer, G. L. (2011). An updated measure for assessing subtle rape myths. *Social Work Research*, *35*(2), 71 81. https://doi.org/10.1093/swr/35.2.71

Mgolozeli, S., & Duma, S. (2019). "As I was walking down the street, four strange guys came and took me under the bridge, where they all raped me": An interpretative phenomenological analysis of the types of rape experienced by men in South

Africa. *American Journal of Men's Health*, *13*(6), Article 1557988319882589. https://doi.org/10.1177/1557988319882589

Mgolozeli, S. E., & Duma, S. E. (2020a). "They all laughed and asked me if I enjoyed having sex with those guys": Exploring men's lived experiences when reporting rape to police in South Africa. *PLOS One*, *15*(8), e0235044. https://doi.org/10.1371/journal.pone.0235044

Mgolozeli, S. E., & Duma, S. E. (2020b). "They destroyed my life because I do not feel like a man anymore": An interpretative phenomenological analysis of men's lived experiences of rape victimization. *Heliyon*, *6*(5), E0381. https://doi.org/https://doi.org/10.1016/j.heliyon.2020.e03818

Mkhize, S., & Sibanyoni, E. K. (2019). Police turn blind eye to homosexual victims of crime in Durban of South Africa. *Journal of Humanities and Social Sciences Studies*, *1*(6), 123 132. https://media.neliti.com/media/publications/351603-police-turn-blind-eye-to-homosexual-vict-ea062535.pdf

Modri, J. (2014). The rhetoric of rape: An extended note on apologism, depoliticisation and the male gaze in Ndou V S. *South African Journal on Human Rights*, *30*(1), 134 158. https://doi.org/10.1080/19962126.2014.11865101

Morrell, R., Jewkes, R., & Lindegger, G. (2012). Hegemonic masculinity/masculinities in South Africa: Culture, power, and gender politics. *Men and Masculinities*, *15*(1), 11 30. https://doi.org/10.1177/1097184X12438001

Mulaik, S. A. (2015). Fundamentals of common factor analysis. In S. P. Reise & D. A. Revicki (Eds.), *Handbook of item response theory modeling: Applications to typical performance assessment* (1st, pp. 211 254). Routledge, Taylor & Francis Group.

Naylor, N. (2008). The politics of a definition. In L. Artz & D. Smythe (Eds.), *Should we consent? Rape law reform in South Africa* (pp. 22 51). Juta & Company Ltd.

Ngubane, L. B., Nöthling, J., Moletsane, R., Wilkinson, A., & Qulu, L. (2022). Why men rape: Perspectives from incarcerated rapists in a KwaZulu-Natal prison, South Africa. *Frontiers in Psychology*, *13*, 805289. https://doi.org/10.3389/fpsyg.2022.805289

Ngubane-Mokiwa, S., & Chisale, S. S. (2019). Male rape or consensual sex: Hidden hegemonic masculinities by Zulu speaking men with disabilities. *Ubuntu Special Issue*, *8*(2), 107 124. https://hdl.handle.net/10520/EJC-19f635815e

Norton, R., & Grant, T. (2008). Rape myth in true and false rape allegations. *Psychology, Crime & Law*, *14*(4), 275 285. https://doi.org/10.1080/10683160701770286

Nyúl, B., Kende, A., Engyel, M., & Szabó, M. (2018). Perception of a perpetrator as a successful person predicts decreased moral judgment of a rape case and labeling it as rape. *Frontiers in Psychology*, *9*, Article 2555. https://doi.org/10.3389/fpsyg.2018.02555

Obierefu, P. O., & Ojedokun, O. (2017). Hostile sexism as predictor of rape supportive attitude among men. *African Journal for the Psychological Study of Social Issues*, *20*(3), 163 176. Retrieved May 19, 2019, from https://www.ajol.info/index.php/ajpssi/article/view/176170

Orth, Z., Andipatin, M., & van Wyk, B. (2020). "These women are making a statement against rape and yet the only thing y'all can focus on is 'eww they're naked'": Exploring rape culture on Facebook in South Africa. *Gender Issues*, *38*, 243 259. https://doi.org/10.1007/s12147-020-09268-x

Parratt, K. A., & Pina, A. (2017). From "real rape" to real justice: A systematic review of police officers' rape myth beliefs. *Aggression and Violent Behavior*, *34*, 68 83. https://doi.org/10.1016/j.avb.2017.03.005

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

Pearson, J., & Barker, D. (2018). Male rape: What we know, don't know and need to find out a critical review. *Crime Psychology Review*, *4*(1), 72 94. https://doi.org/10.1080/23744006.2019.1591757

Peterson, Z. D., & Muehlenhard, C. L. (2004). Was it rape? The function of women's rape myth acceptance and definitions of sex in labeling their own experiences. *Sex Roles*, *51*(3), 129 144. https://doi.org/10.1023/B:SERS.0000037758.95376.00

Pretorius, G. (2009). The male rape survivor: Possible meanings in the context of feminism and patriarchy. *Journal of Psychology in Africa*, *19*(4), 575 580. https://doi.org/10.1080/14330237.2009.10820332

R Core Team. (2020). *R: A language and environment for statistical computing* [Computer Software]. Vienna, Austria, R Foundation for Statistical Computing. https://www.R-project.org/

Reed, R., Pamlanye, J., Truex, H., Murphy-Neilson, M., Kunaniec, K., Newins, A., & Wilson, L. (2020). Higher rates of unacknowledged rape among men: The role of rape myth acceptance. *Psychology of Men and Masculinity*, *21*(1), 162 167. https://doi.org/10.1037/men0000230

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, *47*(5), 667 696. https://doi.org/10.1080/00273171.2012.715555

Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment*, *95*(2), 129 140. https://doi.org/10.1080/00223891.2012.725437

Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2018). Bifactor modelling and the evaluation of scale scores. In P. Irwing, T. Booth, & D. J. Hughes (Eds.), *The Wiley handbook of psychometric testing: A multidisciplinary reference on survey, scale and test development* (pp. 677 707, Vol. 2). John Wiley & Sons Ltd. https://doi.org/10.1002/9781118489772.ch22

Reise, S. P., Cook, K. F., & Moore, T. M. (2015). Evaluating the impact of multidimensionality on unidimensional item response theory model parameters. In S. P. Reise & D. A. Revicki (Eds.), *Handbook of item response theory modeling: Applications to typical performance assessment* (1st, pp. 13 40). Routledge, Taylor & Francis Group.

Reise, S. P., Kim, D. S., Mansolf, M., & Widaman, K. F. (2016). Is the bifactor model a better model or is it just better at modeling implausible responses? Application of iteratively reweighted least squares to the Rosenberg Self-Esteem Scale. *Multivariate Behavioral Research*, *51*(6), 818 838. https://doi.org/10.1080/00273171.2016.1243461

Reise, S. P., Moore, T., & Maydeu-Olivares, A. (2011). Target rotations and assessing the impact of model violations on the parameters of unidimensional item response theory models. *Educational and Psychological Measurement*, *71*(4), 684 711. https://doi.org/10.1177/0013164410378690

Reise, S. P., Scheines, R., Widaman, K. F., & Haviland, M. G. (2013). Multidimensionality and structural coefficient bias in structural equation modeling: A bifactor perspective. *Educational and Psychological Measurement*, *73*(1), 5 26. https://doi.org/10.1177/0013164412449831

Reise, S. P., Ventura, J., Keefe, R. S. E., Baade, L. E., Gold, J. M., Green, M. F., Kern, R. S., Mesholam-Gately, R., Nuechterlein, K. H., Seidman, L. J., & Bilder, R. (2011). Bifactor and item response theory analyses of interviewer report scales of cognitive impairment in schizophrenia. *Psychological Assessment*, *23*(1), 245 261. https://doi.org/10.1037/a0021501

Revelle, W. (2021). *Psych: Procedures for psychological, psychometric, and personality research* (Version 2.1.6) [R package]. Evanston, Illinois, Northwestern University. https://CRAN.R-project.org/package=psych

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, *21*(2), 137 150. https://doi.org/10.1037/met0000045

Rosenstein, J., & Carroll, M. (2015). Male rape myths, female rape myths, and intent to intervene as a bystander. *Violence and Gender*, *2*(4), 204 208. https://doi.org/10.1089/vio.2015.0027

Russell, K. J., & Hand, C. J. (2017). Rape myth acceptance, victim blame attribution and just world beliefs: A rapid evidence assessment. *Aggression and Violent Behavior*, *37*, 153 160. https://doi.org/10.1016/j.avb.2017.10.008

Selepe, M., Lindegger, G., & Govender, K. (2021). Discourses in accounts of rape by sex offenders in Limpopo province, South Africa. *South African Journal of Psychology*, *51*(4), 521 532. https://doi.org/10.1177/0081246320960551

Setia, A., Marks, M., & An, S. (2020). Double standards in perceived traits of women labeled victims versus survivors. *Sexuality and Culture*, *24*(5), 1562 1578. https://doi.org/10.1007/s12119-020-09712-w

Shafer, A., Ortiz, R. R., Thompson, B., & Huemmer, J. (2018). The role of hypermasculinity, token resistance, rape myth, and assertive sexual consent communication among college men. *Journal of Adolescent Health*, *62*(3), S44 S50. https://doi.org/10.1016/j.jadohealth.2017.10.015

Sinclair, M., O'Toole, J., Malawaraarachchi, M., & Leder, K. (2012). Comparison of response rates and cost-effectiveness for a community-based survey: Postal, internet and telephone modes with generic or personalised recruitment approaches. *BMC Medical Research Methodology*, *12*(1), 132. https://doi.org/10.1186/1471-2288-12-132

South Africa Police Service. (2021a). *Fourth quarter crime statistics 2021/2022* [Data set]. SAPS & Statistics South Africa. https://www.saps.gov.za/services/older_crimestats.php

South Africa Police Service. (2021b). *Third quarter crime statistics 2021/2022* [Data set]. https://www.saps.gov.za/services/older_crimestats.php

South Africa Police Service. (2022a). *First quarter crime statistics 2022/2023* [Dataset]. https://www.saps.gov.za/services/crimestats.php

South Africa Police Service. (2022b). *Second quarter crime statistics 2022/2023* [Data set]. https://www.saps.gov.za/services/crimestats.php

Spruin, E. (2018). An exploration into the acceptance of male rape myths within the UK. *Journal of Forensic Sciences & Criminal Investigation*, *9*(3), 555763. https://doi.org/10.19080/JFSCI.2018.09.555763

Stabile, B., Grant, A., Purohit, H., & Rama, M. (2019). "She lied": Social construction, rape myth prevalence in social media, and sexual assault policy. *Sexuality, Gender & Policy*, *2*(2), 80 96. https://doi.org/10.1002/sgp2.12011

Stern, E., Cooper, D., & Greenbaum, B. (2015). The relationship between hegemonic norms of masculinity and men's conceptualization of sexually coercive acts by women in South Africa. *Journal of Interpersonal Violence*, *30*(5), 796 817. https://doi.org/10.1177/0886260514536275

Stucky, B. D., Thissen, D., & Orlando Edelen, M. (2013). Using logistic approximations of marginal trace lines to develop short assessments. *Applied Psychological Measurement*, *37*(1), 41 57. https://doi.org/10.1177/0146621612462759

Suarez, E., & Gadalla, T. M. (2010). Stop blaming the victim: A meta-analysis on rape myths. *Journal of Interpersonal Violence*, *25*(11), 2010 2035. https://doi.org/10.1177/0886260509354503

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. https://doi.org/10.1177/0886260512475317

Swemmer, S. (2019). Ethical loneliness and the development of a victim-focused approach to rape cases in South Africa. *Aggression and Violent Behavior*, *47*, 298 306. https://doi.org/10.1016/j.avb.2019.01.014

Taherdoost, H., Sahibuddin, S., & Jalaliyoon, N. (2014). Exploratory factor analysis; concepts and theory. *Advances in Applied and Pure Mathematics*, *27*, 375 382. https://hal.science/hal-02557344

Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods*, *16*(2), 209 220. https://doi.org/10.1037/a0023353

Timmerman, M. E., Lorenzo-Seva, U., & Ceulemans, E. (2018). The number of factors problem. In P. Irwing, T. Booth, & D. J. Hughes (Eds.), *The Wiley handbook of psychometric testing: A multidisciplinary reference on survey, scale and test development* (pp. 305 345, Vols. 2, Vol. 1). John Wiley & Sons Ltd.

Turchik, J., & Edwards, K. (2012). Myths about male rape: A literature review. *Psychology of Men and Masculinity*, *13*(2), 211 226. https://doi.org/10.1037/a0023207

Urban, R. E., & Porras Pyland, C. (2022). Development and preliminary validation of the gender inclusive rape myth acceptance scale. *Journal of Interpersonal Violence*, *37*(21), NP20630 NP20652. https://doi.org/10.1177/08862605211055076

van der Bruggen, M., & Grubb, A. (2014). A review of the literature relating to rape victim blaming: An analysis of the impact of observer and victim characteristics on attribution of blame in rape cases. *Aggression and Violent Behavior*, *19*(5), 523 531. https://doi.org/10.1016/j.avb.2014.07.008

Vetten, L. (2017). Daughters of the revolution: Legal narrativity in S v Zuma. *South African Journal on Human Rights*, *33*(2), 236 257. https://doi.org/10.1080/02587203.2017.1354423

Vonderhaar, R. L., & Carmody, D. C. (2015). There are no “innocent victims”: The influence of just world beliefs and prior victimization on rape myth acceptance. *Journal of Interpersonal Violence*, *30*(10), 1615 1632. https://doi.org/10.1177/0886260514549196

Walfield, S. M. (2018). “Men cannot be raped”: Correlates of male rape myth acceptance. *Journal of Interpersonal Violence*, *36*(13), 1 27. https://doi.org/10.1177/0886260518817777

Waller, N. G. (2018). Direct Schmid Leiman transformations and rank-deficient loadings matrices. *Psychometrika*, *83*(4), 858 870. https://doi.org/10.1007/s11336-017-9599-0

Wilson, L. C., & Miller, K. E. (2016). Meta-analysis of the prevalence of unacknowledged rape. *Trauma, Violence, & Abuse*, *17*(2), 149 159. https://doi.org/10.1177/1524838015576391

Wilson, L. C., Newins, A. R., & White, S. W. (2018). The impact of rape acknowledgment on survivor outcomes: The moderating effects of rape myth acceptance. *Journal of Clinical Psychology*, *74*(6), 926 939. https://doi.org/10.1002/jclp.22556

Wilson, L. C., & Newins, A. (2019). Rape acknowledgment and sexual minority identity: The indirect effect of rape myth acceptance. *Psychology of Sexual Orientation and Gender Diversity*, *6*(1), 113 119. https://doi.org/10.1037/sgd0000304

Woods, C. M. (2015). Estimating the latent density in unidimensional IRT to permit non-normality. In S. P. Reise & D. A. Revicki (Eds.), *Handbook of item response theory modeling: Applications to typical performance assessment* (1st, pp. 60 84). Routledge, Taylor & Francis Group.

Yapp, E. J., & Quayle, E. (2018). A systematic review of the association between rape myth acceptance and male-on-female sexual violence. *Aggression and Violent Behavior*, *41*, 1 19. https://doi.org/10.1016/j.avb.2018.05.002

Young, S. M., Pruett, J. A., & Colvin, M. L. (2016). Comparing help-seeking behavior of male and female survivors of sexual assault: A content analysis of a hotline. *Sexual Abuse*, *30*(4), 454 474. https://doi.org/10.1177/1079063216677785

Appendix A

Recruitment Email: Main Study

From: Demi Geldenhuis
Sent: Monday, 22 March 2021 18:15
Subject: Win a bike and other prizes: Invitation to complete a questionnaire

Dear fellow student

My name is Demi Geldenhuis, and I am a student in the Department of Psychology in the Faculty of Humanities. I would like to invite you to take part in a study by filling out a questionnaire, the results of which will contribute to a research project in order to complete my MA thesis in Psychology.

The purpose of this study is to investigate how attitudes towards rape are measured, and whether they can be measured more precisely. You can make a difference by adding your honest response set to this study. Every set of responses will play an important role in measuring this construct accurately.

The questionnaire will take approximately 15 minutes to complete and will require you to select the extent to which you agree/disagree with roughly 60 statements about the causes, context and consequences of sexual assault.

The survey is completed on an **anonymous basis**, and all responses will remain confidential. Your participation is entirely voluntary, and you are free to decline to participate. You are also free to withdraw from the study at any point, in which case your data and responses will be deleted. However, once you have submitted the survey it will no longer be possible to withdraw, as your answers will be completely anonymous and there will be no way for me to find your answer set in order to delete it.

At the end of the survey, you will be taken to a **new page** where you may submit your email address to enter **the lucky draws**; this will not be linked to your answers in any way. The following prizes are up for grabs (see photographs below):

- **x1 bicycle** (with custom colouring, valued at **R2500**)
- **x6 R500 cash prizes**
- **x10 moonstone necklaces**

This questionnaire deals with sensitive content that some students may find distressing or triggering. Should you experience any distress as a result of this survey, please contact one of the following for confidential and professional service:

1. Welgevallen Community Psychology Clinic
   - Tel: 021 808 2696
   - Email: WCPC@sun.ac.za
2. Centre for Student Counselling and Development (CSCD)
   - 021 808 4994 (office hours) (Stellenbosch campus)
   - 021 938 9590 (office hours) (Tygerberg campus)
   - Or request an appointment via email: supportus@sun.ac.za
   - For emergency services, please contact ER24 at 010 205 3032 (day and night).

Please also take note of the following 24-hour crisis helplines:

- TEARS Foundation: *134*7355# (SMS service)
- Rape Crisis Centre: 021-447-9762

Click here to take the survey

Thank you for your time!

Kind regards,

Demi Louise Geldenhuis
BA Hons (Psychology) | BA (Law)
MA Candidate: Department of Psychology
Faculty of Arts and Social Sciences



Appendix B

Recruitment email: Pilot Study

From: Geldenhuis, D, Me
Subject: Invitation to complete a survey

Dear fellow student

My name is Demi Geldenhuis, and I am a student in the Department of Psychology in the Faculty of Humanities. I would like to invite you to contribute to a research project I am conducting in order to complete my MA thesis in Psychology. You can do this in two ways: 1) by **completing the survey** linked to this email, and/or 2) **forwarding this email** to a fellow undergraduate student.

The purpose of this study is to investigate how attitudes towards rape are measured, and whether they can be measured more precisely.

I only require 40 responses for this phase of data collection, and every response is invaluable. The questionnaire will take approximately 15 minutes to complete and will require you to select the extent to which you agree/disagree with roughly 60 statements about the causes, context and consequences of sexual assault.

The survey is completed on an **anonymous basis**, and all responses will remain confidential. Your participation is entirely voluntary, and you are free to decline to participate. You are also free to withdraw from the study at any point, in which case your data and responses will be deleted. However, once you have submitted the survey it will no longer be possible to withdraw, as your answers will be completely anonymous and there will be no way for me to find your answer set in order to delete it.

As a token of appreciation for your time, you may enter a lucky draw to stand a chance to **win R200**. At the end of the survey you will be taken to a **new page** where you may submit your email address to enter the lucky draw; this will not be linked to your answers in any way.

This questionnaire deals with sensitive content that some students may find distressing or triggering. Should you experience any distress as a result of this survey, please contact one of the following for confidential and professional service:

1. Welgevallen Community Psychology Clinic

- Tel: 021 808 2696
- Email: WCPC@sun.ac.za

2. Centre for Student Counselling and Development (CSCD)

- 021 808 4994 (office hours)
  (Stellenbosch campus)
- 021 938 9590 (office hours)
  (Tygerberg campus)
- Or request an appointment via email: supportus@sun.ac.za
- For emergency services, please contact ER24 at 010 205 3032 (day and night).

Please also take note of the following 24-hour crisis helplines:

- TEARS Foundation: *134*7355# (SMS service)
- Rape Crisis Centre: 021-447-9762

Click here to take the survey

Appendix C

STUDY DETAILS AND BRIEF

Dear fellow student

Thank you for your interest! Please take some time to read the following information, which will explain the details of the project.

The purpose of this study is to investigate how attitudes towards rape are measured, and whether these can be measured more precisely. The questionnaire will take approximately **15 minutes** to complete and will require you to select the extent to which you agree/disagree with roughly 60 statements about the causes, context and consequences of sexual assault.

The survey is completed on an **anonymous** basis, and all responses will remain confidential. I am looking at **response patterns**, and therefore ask that you ***please* answer as honestly as possible.** A grey progress bar is shown at the top of each page. Should you enter the **lucky draw**, your details will not be attached to your responses in any way.

You will need to fully complete a section before continuing. Your answers will only be submitted once you finish the survey, therefore if you exit at any point before the end none of your answers will be recorded. Once you have submitted the survey it will no longer be possible to withdraw, as your answers will be completely anonymous and there will be no way for me to find your answer set in order to delete it.

Please note: This questionnaire deals with sensitive content that some students may find distressing or triggering. Should you experience any distress as a result of this survey, please contact one of the following for confidential and professional service:

- Welgevallen Community Psychology Clinic
  - Tel: 021 808 2696
  - Email: WCPC@sun.ac.za
- Centre for Student Counselling and Development (CSCD)
  - 021 808 4994 (office hours) - Stellenbosch campus
  - 021 938 9590 (office hours) - Tygerberg campus
  - Or request an appointment via email: supportus@sun.ac.za
  - For emergency services, please contact ER24 at 010 205 3032 (day and night).

Please also take note of the following 24-hour crisis helplines:

- TEARS Foundation: *134*7355# (SMS service)
- Rape Crisis Centre: 021-447-9762

RIGHTS OF RESEARCH PARTICIPANTS:
Your participation is entirely voluntary, and you are free to decline to participate. You have the right to decline answering any questions and you can exit the survey at any time without giving a reason. You are not waiving any legal claims, rights or remedies because of your participation in this research study.

If you have questions regarding your rights as a research participant, contact Mrs Maléne Fouché [mfouche@sun.ac.za; 021 808 4622] at the Division for Research Development.

If you have any questions or concerns about the research, please feel free to contact the researcher Demi Geldenhuis at 20233515@sun.ac.za, or the Supervisor, Dr Kafaar at zkafaar@sun.ac.za.

To save a copy of this text, you may contact the researcher at any time and a copy will be forwarded to you.

*Participant consent

In order to proceed with the survey, please tick "YES" to the following:

	Yes
I confirm that I have read and understood the information provided for the current study.	○
I agree to take part in this survey.	○
I am over 18 years old	○

Appendix D

RMA Questionnaire

Demographic Questions

Page 2

***Age:**

***Gender:**

- Female
- Male
- Non-binary

***Home language:**

- English
- Afrikaans
- isiXhosa
- Other:

***Currently studying towards a(n):**

- Undergraduate degree
- Postgraduate degree
- Other:

IRMA items

Page 3

*Illinois Rape Myth Acceptance Scale

Please indicate to what extent you agree/disagree with the following statements:

	Not at all agree 1	2	3	Neutral 4	5	6	Very much agree 7
If a girl is raped while she is drunk, she is at least somewhat responsible for letting things get out of control.							
Although most women wouldn't admit it, they generally find being physically forced into sex a real "turn on".							
When men rape, it is because of their strong desire for sex.							
If a woman is willing to "make out" with a guy, then it's no big deal if he goes a little further and has sex.							
Women who are caught having an illicit affair sometimes claim it was rape.							
Newspapers should not release the name of a rape victim to the public.							
Many so-called rape victims are actually women who had sex and "changed their minds" afterwards.							
Many women secretly desire to be raped.							
Rape mainly occurs on the "bad" side of town.							
Usually, it is only women who do things like hang out in bars and sleep around that are raped.							
Most rapists are not caught by the police.							
If a woman doesn't physically fight back, you can't really say that it was rape.							
Men from nice middle-class homes almost never rape.							
Rape isn't as big a problem as some feminists would like people to think.							
When women go around wearing low-cut tops or short skirts, they're just asking for trouble.							

Page 4

***Illinois Rape Myth Acceptance Scale (continued)**

Please indicate to what extent you agree/disagree with the following statements:

	Not at all agree 1	2	3	Neutral 4	5	6	Very much agree 7
Rape accusations are often used as a way of getting back at men.							
A rape probably didn't happen if the woman has no bruises or marks.							
Many women find being forced to have sex very arousing.							
If a woman goes home with a man she doesn't know, it is her own fault if she is raped.							
Rapists are usually sexually frustrated individuals.							
All women should have access to self-defence classes.							
It is usually only women who dress suggestively that are raped.							
Some women prefer to have sex forced on them so they don't have to feel guilty about it.							
If the rapist doesn't have a weapon, you really can't call it rape.							
When a woman is a sexual tease, eventually she is going to get into trouble.							
Being raped isn't as bad as being mugged or beaten.							
Rape is unlikely to happen in the woman's own familiar neighbourhood.							
In reality, women are almost never raped by their boyfriends.							
Women tend to exaggerate how much rape affects them.							
When a man is very sexually aroused, he may not even realise that a woman is resisting.							

Page 5

Illinois Rape Myth Acceptance Scale (continued)

Please indicate to what extent you agree/disagree with the following statements:

	Not at all agree 1	2	3	Neutral 4	5	6	Very much agree 7
A lot of women lead a man on and then cry rape.	○	○	○	○	○	○	○
It is preferable that a female police officer conduct the questioning when a woman reports a rape.	○	○	○	○	○	○	○
A lot of times, women who claim they were raped just have emotional problems.	○	○	○	○	○	○	○
If a woman doesn't physically resist sex – even when protesting verbally – it really can't be considered rape.	○	○	○	○	○	○	○
Rape almost never happens in the woman's own home.	○	○	○	○	○	○	○
A woman who "teases" men deserves anything that might happen.	○	○	○	○	○	○	○
When women are raped, it's often because the way they said "no" was ambiguous.	○	○	○	○	○	○	○
If a woman isn't a virgin, then it shouldn't be a big deal if her date forces her to have sex.	○	○	○	○	○	○	○
Men don't usually intend to force sex on a woman, but sometimes they get too sexually carried away.	○	○	○	○	○	○	○
This society should devote more effort to preventing rape.	○	○	○	○	○	○	○
A woman who dresses in skimpy clothes should not be surprised if a man tries to force her to have sex.	○	○	○	○	○	○	○
Rape happens when a man's sex drive gets out of control.	○	○	○	○	○	○	○
A woman who goes to the home or apartment of a man on the first date is implying that she wants to have sex.	○	○	○	○	○	○	○
Many women actually enjoy sex after the guy uses a little force.	○	○	○	○	○	○	○
If a woman claims to have been raped but has no bruises or scrapes, she probably shouldn't be taken too seriously.	○	○	○	○	○	○	○

MRMS items

Page 6

*Male Rape Myth Scale

Please indicate the extent to which you agree/disagree with the following statements:

	Strongly Disagree 1	2	Neutral 3	4	5	Strongly Agree 6
It is a terrible experience for a man to be raped by a woman.						
The extent of a man's resistance should be a major factor in determining if he was raped.						
Any healthy man can successfully resist a rapist if he really wants to.						
If a man obtained an erection while being raped it probably means that he started to enjoy it.						
A man can enjoy sex even if it is being forced upon him.						
Most men who are raped by women are very upset by the incident.						
Many men claim rape if they consented to sex with another guy but then changed their mind afterwards.						
Most men who are raped by a woman are somewhat to blame for not escaping or fighting off the woman						
If a man engages in kissing and foreplay and allows things to get out of hand, it is his own fault if his partner forces sex on him.						
Male rape is usually committed by homosexual men						
Most men who are raped by a man are somewhat to blame for not escaping or fighting off the man.						

Page 7

*Male Rape Myth Scale (continued)

Please indicate the extent to which you agree/disagree with the following statements:

	Strongly Disagree 1	2	Neutral 3	4	5	Strongly Agree 6
A man who has been raped has lost his manhood.	○	○	○	○	○	○
Most men who are raped by a woman are somewhat to blame for not being more careful.	○	○	○	○	○	○
If a man told me that he had been raped by another man, I would suspect that he is homosexual.	○	○	○	○	○	○
Most men who have been raped have a history of sleeping around.	○	○	○	○	○	○
No self-respecting man would admit to being raped.	○	○	○	○	○	○
Women who rape men are sexually frustrated individuals.	○	○	○	○	○	○
A man who allows himself to be raped by another man is probably not straight.	○	○	○	○	○	○
Most men would not enjoy being raped by a woman.	○	○	○	○	○	○
Men who parade around nude in a locker room are asking for trouble.	○	○	○	○	○	○
Male rape is more serious when the victim is heterosexual than when the victim is homosexual.	○	○	○	○	○	○
I would find it difficult to believe a man who told me that he was raped by a woman.	○	○	○	○	○	○

Final Page

The last page of the survey differed, depending on which path respondents took to get there. For path 1, if respondents had completed the entire questionnaire, they were presented with the opportunity to enter into a series of lucky draws. For path 2, respondents who did not pass the screening question were redirected straight to the end of the survey and were not given the opportunity to enter into the lucky draws.

Final page: Path 1

Thank you for taking the survey! Your responses are greatly appreciated.

You now have the opportunity to enter a series of lucky draws to win one of the following: a bicycle from Reventon with custom colouring, R500, or a moonstone necklace.

If you click on the link below you will be taken to a new page where you can leave your email address; this will ensure that your contact information is not linked to your responses. Your answers will remain anonymous and confidential.

Click here to enter

Final page: Path 2

Thank you for your interest in this research project!

Unfortunately, I am only collecting responses from undergraduate students.

If you received an invite to this survey by mistake, please let me know at

Appendix E

R Code

R Code for Entire Project

```r
setwd("C:/Users/User/Documents/MASTERS/1.Data")
library(psych)
library(CTT)
library(GPArotation)
library(ggplot2)
library(lavaan)

## Data Prep ##
rma_dat<-read.csv("MA_Dataset_DG.csv", header = TRUE, sep = "
    ,") # import data
str(rma_dat)
irma_dat <- rma_dat[1:2536, paste0("irma_", 1:45)] # select
    the IRMA data
irma_dat <- irma_dat[,-c(6,11,21,32,40)] # remove IRMA
    fillers
str(irma_dat)
mrms_dat <-rma_dat[1:2536, paste0("mrms_", 1:22)] # select
    the MRMS data
mrms_dat[, paste0("mrms_", c(1, 6, 19))]<- 7 - mrms_dat[,
    paste0("mrms_", c(1, 6, 19))] # reverse score
str(mrms_dat)
global_dat<-(cbind(irma_dat,mrms_dat)) # rejoin the data sets
str(global_dat)

# alpha and item stats
irma_item_stats <- psych::alpha(irma_dat)
```

```
irma_alpha <- psych::alpha(irma_dat)$total$std.alpha
mrms_item_stats <- psych::alpha(mrms_dat)
mrms_alpha <- psych::alpha(mrms_dat)$total$std.alpha
global_item_stats <- psych::alpha(global_dat)
global_alpha <- psych::alpha(global_dat)$total$std.alpha

# correlate raw summed scores of IRMA and MRMS
{mrms_totals <- rowSums(mrms_dat)
  irma_totals <- rowSums(irma_dat)
  irma_mrms_cor <- cor(irma_totals,mrms_totals, method = "
    spearman")
  y <- c(irma_alpha, mrms_alpha)}
irma_mrms_discor <- disattenuated.cor(irma_mrms_cor,y)
irma_mrms_discor

## polychoric correlation matrix combined item pool ###
poly_global <- polychoric(x= global_dat, smooth=TRUE, global=
    FALSE, delete=FALSE, max.cat=7, correct=FALSE)
poly_global <- poly_global$rho
options(scipen=999) # ensures p values are returned in right
    format
KMO(poly_global)
cortest.bartlett(poly_global, n = 2536)
write.csv(poly_global, "poly_global_full.csv") # export
    matrix as a .csv file

## omega hierarchical combined item pool
global_comm<-read.csv("global_comm.csv", header = TRUE, sep =
     ",", row.names = 1) # import data, make sure was saved as
```

```
a .csv
global_PM<-read.csv("global_PM.csv", header = TRUE, sep = ",", row.names = 1) # import data, code assumes the first data line is row headings
genload <- global_PM[, 1] # note, code assumes the first line of loadings are the gen factor loadings
genload
grpload <- global_PM[, 2:5]
grpload
sum_genload_sq <- (sum(genload))^2
sum_grpload_sq <- (sum(grpload))^2
uniq <- sum(global_comm[, "u2"])
Tot_Var <- (sum_genload_sq + sum_grpload_sq + uniq)
omegah_global = sum_genload_sq/Tot_Var
omegah_global

## omega hierarchical mrms
mrms_comm<-read.csv("mrms_comm.csv", header = TRUE, sep = ",", row.names = 1) # import data
mrms_PM<-read.csv("mrms_PM.csv", header = TRUE, sep = ",", row.names = 1) # import data
genload <- mrms_PM[, 1]
genload
grpload <- mrms_PM[, 2:4]
grpload
sum_genload_sq <- (sum(genload))^2
sum_grpload_sq <- (sum(grpload))^2
uniq <- sum(mrms_comm[, "u2"])
Tot_Var <- (sum_genload_sq + sum_grpload_sq + uniq)
```

```
omegah_mrms = sum_genload_sq/Tot_Var
omegah_mrms
## omega hierarchical irma
```

MRMS OmegaH Output

Below is the output for the calculation of omegaH for my MRMS bifactor model. I have included the output here to provide a more detailed insight into the components that formed part of the calculation.

```
# MRMS omegaH Output #
> genload <- mrms_PM[, 1]
> genload
 mrms_1  mrms_2  mrms_3  mrms_4  mrms_5  mrms_6  mrms_7  mrms_8  mrms_9 mrms_10 mrms_11 mrms_12 mrms_13 mrms_14 mrms_15 mrms_16
  0.613   0.424   0.499   0.516   0.457   0.530   0.344   0.769   0.707   0.442   0.797   0.549   0.825   0.597   0.641   0.512
mrms_17 mrms_18 mrms_19 mrms_20 mrms_21 mrms_22
  0.321   0.659   0.413   0.576   0.546   0.573
> grpload <- mrms_PM[, 2:4]
> grpload
            F1     F2     F3
mrms_1   0.420  0.101 -0.038
mrms_2   0.114  0.594 -0.136
mrms_3   0.176  0.666 -0.232
mrms_4   0.301  0.730 -0.288
mrms_5   0.281  0.592 -0.192
mrms_6   0.529  0.018 -0.027
mrms_7   0.031  0.413 -0.070
```

```
mrms_8    0.036  0.608 -0.324
mrms_9    0.030  0.464 -0.073
mrms_10   0.012  0.258  0.168
mrms_11  -0.063  0.493 -0.245
mrms_12  -0.056  0.090  0.309
mrms_13  -0.052  0.302  0.008
mrms_14  -0.061 -0.081  0.530
mrms_15  -0.046  0.135  0.307
mrms_16   0.046  0.063  0.399
mrms_17  -0.012  0.398  0.146
mrms_18  -0.076  0.182  0.355
mrms_19   0.434  0.178 -0.122
mrms_20  -0.063  0.306  0.090
mrms_21  -0.027  0.261  0.191
mrms_22   0.152  0.229  0.220
> sum_genload_sq <- (sum(genload))^2
> sum_grpload_sq <- (sum(grpload))^2
> uniq <- sum(mrms_comm[, "u2"])
> Tot_Var <- (sum_genload_sq + sum_grpload_sq + uniq)
> omegah_mrms = sum_genload_sq/Tot_Var
> omegah_mrms
[1] 0.5815862
```

Appendix F

MRMS Item Statistics

Table F1

MRMS Item-level distribution statistics and response frequency by category

	Distribution				Response Frequency by Category					
Item no	M	SD	Skewness	Kurtosis	1	2	3	4	5	6
1.00	1.49	1.08	2.59	6.44	0.76	0.11	0.05	0.04	0.01	0.02
2.00	2.43	1.74	0.86	-0.69	0.49	0.14	0.09	0.11	0.08	0.09
3.00	2.09	1.45	1.27	0.59	0.52	0.19	0.11	0.09	0.04	0.05
4.00	1.48	0.98	2.35	5.39	0.75	0.12	0.07	0.04	0.01	0.01
5.00	1.90	1.28	1.40	1.12	0.57	0.17	0.11	0.09	0.03	0.02
6.00	2.02	1.4	1.42	1.11	0.52	0.22	0.11	0.07	0.04	0.04
7.00	2.37	1.28	0.58	-0.41	0.35	0.19	0.27	0.14	0.04	0.02
8.00	1.65	1.22	1.98	3.14	0.7	0.12	0.07	0.06	0.03	0.02
9.00	1.72	1.22	1.72	2.09	0.66	0.14	0.08	0.07	0.04	0.01
10.00	2.52	1.47	0.56	-0.76	0.37	0.15	0.21	0.16	0.07	0.03
11.00	1.35	0.9	3.04	9.52	0.82	0.08	0.04	0.03	0.01	0.01
12.00	1.59	1.23	2.18	3.77	0.76	0.08	0.05	0.05	0.03	0.02
13.00	1.45	0.95	2.35	5.25	0.76	0.12	0.06	0.04	0.01	0.01
14.00	1.53	1.06	2.24	4.58	0.74	0.12	0.06	0.05	0.02	0.01
15.00	1.44	0.86	2.07	3.90	0.74	0.13	0.09	0.03	0.01	0
16.00	1.82	1.35	1.55	1.29	0.66	0.1	0.09	0.08	0.04	0.02
17.00	2.77	1.63	0.45	-1.02	0.33	0.16	0.16	0.17	0.11	0.07
18.00	1.44	1.04	2.67	6.78	0.79	0.1	0.05	0.03	0.02	0.01
19.00	2.44	1.8	0.94	-0.60	0.49	0.16	0.08	0.09	0.05	0.12
20.00	1.45	0.98	2.48	5.97	0.77	0.11	0.06	0.04	0.01	0.01
21.00	1.59	1.19	2.00	3.05	0.76	0.06	0.08	0.07	0.02	0.02
22.00	2.22	1.54	0.99	-0.32	0.51	0.15	0.1	0.11	0.08	0.04

Appendix G

IRMA Item Statistics

Table G1

IRMA Item-level distributions and response frequencies

	Distribution				Response Frequency by Category						
Item	M	SD	Skew	Kurtosis	1	2	3	4	5	6	7
irma01	1.74	1.30	1.93	3.00	.66	.17	.06	.04	.06	.01	.01
irma02	1.67	1.22	1.97	3.34	.69	.14	.05	.08	.03	.01	.01
irma03	3.38	1.97	.23	-1.18	.28	.11	.12	.17	.16	.09	.07
irma04	1.41	.98	2.81	8.03	.80	.10	.04	.04	.01	.01	.00
irma05	3.83	1.56	-.10	-.54	.09	.14	.12	.32	.21	.08	.05
irma07	2.30	1.38	.89	.02	.39	.25	.13	.16	.05	.01	.01
irma08	1.21	.72	4.49	23.17	.89	.06	.01	.02	.01	.00	.00
irma09	2.30	1.64	1.05	-.02	.50	.15	.12	.09	.10	.03	.02
irma10	1.50	1.04	2.45	5.96	.74	.13	.05	.04	.02	.01	.00
irma12	1.33	.91	3.44	12.68	.84	.08	.03	.02	.01	.01	.00
irma13	1.67	1.17	2.08	4.31	.66	.17	.08	.05	.02	.01	.01
irma14	1.33	.87	3.32	12.21	.83	.09	.04	.03	.01	.00	.00
irma15	1.65	1.26	2.15	4.03	.72	.12	.06	.05	.04	.01	.01
irma16	2.59	1.51	.79	-.19	.30	.27	.15	.15	.09	.02	.02
irma17	1.26	.71	3.51	14.24	.84	.10	.03	.02	.01	.00	.00
irma18	1.73	1.16	1.71	2.41	.62	.19	.06	.09	.02	.00	.00
irma19	2.06	1.51	1.40	1.13	.55	.17	.10	.09	.06	.02	.02
irma20	3.63	1.85	.06	-1.04	.19	.13	.10	.24	.16	.10	.07
irma22	1.54	1.04	2.22	4.81	.71	.15	.07	.04	.02	.01	.00
irma23	1.69	1.15	1.65	1.91	.66	.15	.06	.11	.02	.00	.00
irma24	1.07	.39	8.13	82.61	.96	.03	.01	.01	.00	.00	.00
irma25	2.43	1.65	.91	-.29	.45	.17	.12	.12	.10	.03	.02
irma26	1.16	.64	5.18	31.13	.92	.04	.01	.02	.00	.00	.00

irma27	1.54	1.12	2.47	6.11	.74	.13	.05	.05	.02	.01	.01
irma28	1.80	1.20	1.68	2.58	.59	.20	.11	.06	.02	.01	.00
irma29	1.38	.92	2.93	9.20	.80	.10	.03	.04	.01	.00	.00
irma30	2.51	1.85	.96	-.33	.48	.14	.08	.11	.09	.05	.04
irma31	2.18	1.40	1.13	.54	.45	.24	.13	.11	.06	.01	.01
irma33	1.72	1.13	1.92	3.91	.60	.22	.08	.07	.02	.00	.01
irma34	1.34	.93	3.24	11.03	.84	.07	.03	.03	.01	.01	.00
irma35	1.61	1.07	2.03	3.94	.67	.18	.07	.05	.02	.01	.00
irma36	1.31	.82	3.36	12.46	.83	.09	.04	.02	.01	.00	.00
irma37	1.58	1.10	2.17	4.39	.71	.14	.06	.05	.02	.01	.00
irma38	1.10	.51	7.50	66.21	.95	.03	.01	.01	.00	.00	.00
irma39	2.30	1.58	1.01	-.05	.47	.18	.10	.13	.08	.03	.01
irma41	1.80	1.35	1.80	2.50	.64	.15	.08	.05	.05	.01	.01
irma42	2.92	1.88	.53	-.94	.36	.14	.10	.16	.13	.06	.05
irma43	2.10	1.54	1.27	.55	.56	.14	.09	.10	.06	.03	.01
irma44	2.49	1.58	.69	-.64	.41	.18	.08	.21	.08	.02	.01
irma45	1.25	.73	3.90	17.88	.86	.08	.03	.02	.01	.00	.00

Appendix H

Exploratory Factor Analytic Models for the Combined Item Pool

	Bifactor Model					One-factor model	
Variable	**GF**	**F1**	**F2**	**F3**	**F4**	**Loading**	**h2**
irma18	.53	.72				.61	.38
irma02	.48	.64				.56	.31
irma08	.53	.60				.60	.36
irma44	.50	.56				.60	.36
irma23	.59	.54				.69	.47
irma24	.71	.39	-.47			.76	.57
irma16	.42	.38				.61	.37
irma07	.43	.38				.62	.39
irma31	.53	.37				.75	.56
irma33	.60	.37				.75	.56
irma29	.66	.32			.33	.76	.58
irma42	.59		.65			.47	.22
mrms17	.65		.61			.56	.31
irma03	.50		.61			.38	.14
irma20	.55		.60			.43	.19
irma39	.60		.37			.62	.38
irma30	.51		.36			.52	.27
irma26	.60		-.32			.60	.36
irma38	.70		-.35			.75	.57
mrms08	.57			.58		.74	.55
mrms11	.59			.55		.74	.55
mrms04	.57			.50		.63	.40
mrms03	.51			.49		.60	.36
mrms01	.48			.45		.50	.25
mrms09	.59			.45		.78	.61
mrms13	.63			.44		.82	.67
mrms22	.53			.41		.64	.41
mrms02	.47			.37		.57	.33
mrms06	.41			.36		.42	.18
mrms19	.33			.36		.37	.13
mrms21	.57			.33		.63	.40
mrms12	.49			.32		.59	.35
mrms05	.54			.31		.62	.39
mrms18	.66			.31		.77	.59
mrms20	.53			.30		.69	.47
irma15	.51				.56	.75	.56
irma35	.62				.52	.62	.39
irma14	.60				.51	.72	.52
irma13	.58				.50	.62	.38

Table H1 continued from previous page

	Bifactor Model		One-factor model	
irma10	.60	.48	.70	.49
irma28	.61	.46	.63	.40
irma09	.50	.45	.52	.27
irma22	.64	.44	.75	.56
irma01	.41	.44	.63	.40
irma41	.54	.44	.77	.60
irma27	.65	.40	.62	.38
irma17	.63	.35	.72	.52
irma25	.48	.35	.69	.48
mrms15	.60	.33	.73	.54
irma36	.62	.32	.79	.62
irma19	.51	.32	.72	.52
irma04	.56		.59	.35
irma05	.27		.44	.19
irma12	.61		.71	.51
irma34	.66		.73	.53
irma37	.60		.69	.48
irma43	.49		.63	.40
irma45	.68		.77	.59
mrms07	.36		.49	.24
mrms10	.49		.56	.31
mrms14	.55		.63	.40
mrms16	.54		.62	.38

Appendix I

Model Bias for Combined Item Pool

Table I1

Relative Loading Bias and Slope difference between Combined Item Pool Models

	FA			IRT		
Item	1F	Gen	Relative Bias	a Uni	a Gen	Slope difference
irma01	0.63	0.41	53%	0.81	0.57	0.24
irma02	0.56	0.48	15%	0.67	0.70	-0.03
irma03	0.38	0.50	-25%	0.41	0.71	-0.30
irma04	0.59	0.56	5%	0.73	0.72	0.01
irma05	0.44	0.27	61%	0.48	0.33	0.15
irma07	0.62	0.43	44%	0.79	0.62	0.17
irma08	0.60	0.53	13%	0.75	0.80	-0.05
irma09	0.52	0.50	4%	0.61	0.64	-0.03
irma10	0.70	0.60	16%	0.97	0.91	0.06
irma12	0.71	0.61	16%	1.02	0.89	0.13
irma13	0.62	0.58	8%	0.79	0.81	-0.02
irma14	0.72	0.60	20%	1.04	0.98	0.06
irma15	0.75	0.51	46%	1.14	0.93	0.21
irma16	0.61	0.42	45%	0.77	0.59	0.18
irma17	0.72	0.63	14%	1.05	0.98	0.07
irma18	0.61	0.53	16%	0.78	0.91	-0.14
irma19	0.72	0.51	41%	1.04	0.78	0.26
irma20	0.43	0.55	-21%	0.48	0.81	-0.33
irma22	0.75	0.64	18%	1.14	1.02	0.12
irma23	0.69	0.59	17%	0.95	0.94	0.01
irma24	0.76	0.71	6%	1.15	1.43	-0.27
irma25	0.69	0.48	44%	0.95	0.75	0.20
irma26	0.60	0.60	0%	0.75	0.83	-0.08
irma27	0.62	0.65	-5%	0.79	0.96	-0.17
irma28	0.63	0.61	3%	0.82	0.89	-0.07
irma29	0.76	0.66	15%	1.18	1.14	0.04
irma30	0.52	0.51	1%	0.60	0.66	-0.06
irma31	0.75	0.53	40%	1.13	1.00	0.13
irma33	0.75	0.60	25%	1.13	0.96	0.17
irma34	0.73	0.66	11%	1.07	1.00	0.07
irma35	0.62	0.62	1%	0.79	0.92	-0.13

irma36	0.79	0.62	28%	1.28	1.02	0.26
irma37	0.69	0.60	17%	0.96	0.84	0.12
irma38	0.75	0.70	8%	1.14	1.19	-0.05
irma39	0.62	0.60	3%	0.78	0.85	-0.07
irma41	0.77	0.54	43%	1.21	0.98	0.24
irma42	0.47	0.59	-20%	0.53	0.98	-0.45
irma43	0.63	0.49	29%	0.81	0.63	0.17
irma44	0.60	0.50	20%	0.76	0.75	0.01
irma45	0.77	0.68	13%	1.21	1.16	0.05
mrms01	0.50	0.48	4%	0.58	0.65	-0.07
mrms02	0.57	0.47	21%	0.70	0.61	0.09
mrms03	0.60	0.51	16%	0.74	0.73	0.01
mrms04	0.63	0.57	12%	0.81	0.85	-0.04
mrms05	0.62	0.54	16%	0.79	0.72	0.08
mrms06	0.42	0.41	2%	0.46	0.49	-0.03
mrms07	0.49	0.36	36%	0.56	0.43	0.13
mrms08	0.74	0.57	30%	1.11	1.09	0.02
mrms09	0.78	0.59	33%	1.26	1.07	0.19
mrms10	0.56	0.49	14%	0.67	0.60	0.07
mrms11	0.74	0.59	26%	1.11	1.08	0.03
mrms12	0.59	0.49	21%	0.73	0.62	0.11
mrms13	0.82	0.63	29%	1.43	1.24	0.18
mrms14	0.63	0.55	16%	0.81	0.72	0.09
mrms15	0.73	0.60	23%	1.07	0.90	0.17
mrms16	0.62	0.54	15%	0.78	0.69	0.09
mrms17	0.56	0.65	-15%	0.67	1.17	-0.50
mrms18	0.77	0.66	16%	1.20	1.08	0.12
mrms19	0.37	0.33	12%	0.39	0.38	0.01
mrms20	0.69	0.53	29%	0.94	0.75	0.19
mrms21	0.63	0.57	11%	0.81	0.77	0.04
mrms22	0.64	0.53	20%	0.82	0.74	0.08