

IMPLICIT RELATIONAL ASSESSMENT PROCEDURE (IRAP): CAN SO-CALLED IMPLICIT RAPE-SUPPORTIVE BELIEFS BE RESTRUCTURED AND DO THEY PREDICT BEHAVIOUR?

Anna Brown BSc, MSc

A thesis submitted in partial fulfilment of the requirements of the University of Lincoln for the degree of Doctor of Clinical Psychology

2014

Thesis Abstract

Research into rape-supportive cognition has been hampered by methodological problems such as social desirability. The study of so-called implicit cognition has been dominated by explanations that the effects produced on implicit/indirect measures are mediated by associations held in memory. The Implicit Relational Assessment Procedure (IRAP; Barnes-Holmes et al., 2006) offers an alternative perspective. Derived from Relational Frame Theory (RFT, Hayes, Barnes-Holmes & Roche, 2001), here it is applied for the first time to study the malleability of implicit rape-supportive cognition, following a cognitive-restructuring task in university males. The relationship between implicit (IRAP) and explicit measures (the Acceptance of Modern Myths about Sexual Aggression scale (AMMSA; Gerger, Kley, Bohner & Siebler, 2007), and behaviour was explored (measured by a forced-choice task and ratings of the researcher). IRAP scores did not change following the intervention. Behaviour on the forced-choice measure was predicted by IRAP scores but not by scores on the AMMSA.

Additional analyses into the predictive ability of the AMMSA and IRAP measure on behavioural measures (charity-box overall giving score and Researcher Rating Scale) were conducted. Findings from hierarchical multiple regression analyses indicated that the AMMSA predicted none of the variance whereas the IRAP predicted 12.4% and 11.5% of the variance respectively.

Acknowledgements

I extend significant gratitude to Dr David Dawson and Dr Nima Golanji-Moghaddam for their constant support throughout the research process. I would like to thank Dr Mark Gresswell for his advice when designing the behavioural outcome measures. I also would like to thank my partner Darren for his significant support and patience throughout the research process.

Statement of Contribution

1. Project design:

The first author was primarily responsible for project design. Appreciation is extended to the second and third authors for their advice.

2. Applying for ethical approval:

The first author was responsible for applying for ethical approval.

3. Writing the literature review:

The first author was responsible for writing the literature review (regarding the systematic literature review, journal paper, and extended paper).

4. Recruiting participants:

The first author was responsible for the recruitment of participants from both university sites.

5. Data collection:

The first author collected the data with regards to both phases of the study.

6. Analysis:

The first author carried out the statistical analysis. Appreciation is extended to the second and third authors for their statistical advice and support.

7. Write-up:

The first author was responsible for writing up the research.

Contents

Thesis Abstract	1
Acknowledgements	2
Statement of Contribution.....	3
List of Tables.....	7
List of Figures	8
List of Appendices	9
Systematic Review	10
1. Abstract	11
2. Introduction.....	12
3. Method	16
3.1 Searching	16
3.2 Selection	17
3.3 Data Abstraction.....	19
4. Results	21
4.1 General Characteristics and Key Findings.....	21
4.2 Methodological Quality	41
5. Discussion	51
5.1 Summary	51
5.2 Limitations.....	52
5.3 Future Research	53
5.4 Conclusions.....	55
References	56
Journal Paper	62
Abstract	63
Background.....	64
Method	72
Participants	72
Materials and Apparatus.....	73
Procedure.....	76
Results	82

Direct Measure.....	82
AMMSA Differences Between Groups.....	83
Indirect Measure: Scoring the IRAP	83
IRAP Results.....	84
Relationship between Indirect and Direct Measures	85
Malleability of Implicit Beliefs.....	85
Relationship between the Direct Measure and Behavioural Outcomes..	91
Discussion.....	91
References	96
Extended Paper	108
1 Extended Background	109
2 Extended Method	116
2.1 Participants.....	116
2.2 AMMSA.....	117
2.3 Charity-box Task	118
2.4 Researcher Rating Scale	120
2.5 Intervention and Control Conditions	121
2.6 IRAP Stimulus Set.....	125
2.7 Procedure	126
2.8 Ethics.....	129
3 Extended Results	132
3.1 Test Assumptions: Mann-Whitney U Test	132
3.2 Test Assumptions: Pearson’s r.....	132
3.3 Managing Outliers.....	133
3.4 Normality	134
3.5 Skewness and Kurtosis.....	134
3.6 Calculating the Shape of the Distribution Using Z-Scores	136
3.7 Shapiro-Wilk Test	138
3.8 Linearity and Homoscedasticity	140
3.9 Modifications to Analyses	140
3.10 Test Assumptions: Independent-Samples T-Test.....	142
3.11 Test Assumptions: Mixed Between-Within Analysis of Variance (ANOVA)	142

3.12 Further Analyses.....	142
3.13 Type II Errors.....	145
4 Extended Discussion	145
5 Critical Reflection.....	152
Extended References.....	156
Appendices.....	165
Appendix a	166
Appendix b.....	167
Appendix c	168
Appendix d.....	172
Appendix e.....	176
Appendix f.....	178
Appendix g.....	179
Appendix h.....	180
Appendix i.....	182
Appendix j.....	183

List of Tables

- Table 1. General Characteristics and Key Findings
- Table 2. Methodological Characteristics of Quantitative Studies
- Table 3. The Demographic Characteristics (Age) of Participants
- Table 4. The IRAP Stimulus Set
- Table 5. D-Implicit Relational Assessment Procedure Scores (D-IRAP):
Individual Trial-types (Pre-Experimental/Control Group)
- Table 6. Mixed Between-Within Analysis of Variance by IRAP Trial-type
- Table 7. Significant Correlations between Mean IRAP and Measure of
Overall Giving
- Table 8. Significant Correlations between Mean IRAP and Researcher Rating
Scale
- Table 9. Procedure for Participants in the Intervention and Control
Conditions
- Table 10. Skewness and Kurtosis
- Table 11. Values of ZSkewness and ZKurtosis
- Table 12. Shapiro-Wilk Test of Normality

List of Figures

Figure 1. Selection Procedure

Figure 2. Examples of the Four Implicit Relational Assessment Procedure
Trial-types

Figure 3. Rape-Supportive Verbal Relations Bias Pre- and Post-
Intervention/Control (Overall Mean D-IRAP Effect Score)

Figure 4. A Boxplot Highlighting an Extreme Outlier

List of Appendices

Appendix a	Search terms
Appendix b	Researcher Rating Scale (RRS)
Appendix c	Research crib sheets: Intervention group
Appendix d	Research crib sheets: Control group
Appendix e	Participant information sheet
Appendix f	Consent form: Phase one
Appendix g	Consent form: Phase two
Appendix h	Participant debrief sheet
Appendix i	Ethics approval: University of Lincoln
Appendix j	Ethics approval: University of Nottingham

Systematic Review

The effectiveness of rape prevention interventions for non-convicted males: A systematic review of the last 10 years*

Anna Brown¹, David Dawson², Nima Golanji-Moghaddam²

¹ Trent Doctorate in Clinical Psychology, University of Lincoln

² Research Tutor, Trent Doctorate in Clinical Psychology, University of Lincoln, UK

1. Abstract

Background: A wealth of rape prevention interventions for non-convicted males exist however, empirical evaluation of effectiveness is currently limited. **Aims:** To systematically review the methodology of the literature base in order to: reduce bias in the interpretation of findings regarding treatment effectiveness; identify gaps in knowledge; and draw conclusions from the research. **Method:** A systematic search of five databases was conducted. Inclusion and exclusion criteria were applied to relevant studies, resulting in 18 studies being included for review. **Results:** Studies predominantly demonstrated reductions in rape-supportive attitudes; however behavioural outcome assessment was limited. **Conclusions:** Future interventions need to link to theory, account for individuals' risk status and include more robust measures of behavioural outcomes.

Key Words: Rape, prevention, sexual assault, intervention, programme.

* Prepared for submission to the Journal of Sexual Aggression

2. Introduction

Sexual assault is a significant health and social problem worldwide leading to wide-ranging problems for both individuals and society. It is defined, for the purposes of this review, to include any sexual act in which the victim is threatened, coerced or forced to engage in without consent. The long-term personal effects and economic burden of sexual violence is well known and, as such, prevention programmes have been developed in order to address this with the aim of reducing sexual violence.

Programmes are available for convicted males in prisons and secure health settings however, the majority of sexual assaults are not reported due to the difficulties in disclosure and lack of confidence in conviction rates (Bedard-Gilligan, Jaeger, Echiverri-Cohen, & Zoellner, 2012; Tjaden & Theonnes, 2006), with many sexual aggressors living in the community, never entering the criminal justice system. Research suggests that 12% of female undergraduates in American universities have been raped (Kilpatrick, Resnick, Rugiero, Conoscenti, & McCauley, 2007). The need for interventions to reduce sexual violence with non-convicted males is clear.

The majority of sexual assault prevention programmes for non-convicted males especially college students are found in the United States (US) and are increasing in frequency (Anderson & Whiston, 2005). However, the theoretical basis for programmes is predominantly weak, with many programmes including largely intuitive components (Anderson & Whiston, 2005; Morrison, Hardison, Mathew & O'Neil, 2004). Theoretical approaches which have shown some utility with regards attitudinal and behavioural

change more generally (Paul & Gray, 2011) have been applied to a handful of interventions. Specifically, social norms theory proposes that individuals behave in accordance with their perceptions of expected behaviour of others. Therefore, perceptions that others hold rape-supportive attitudes are thought to lead to the development of cognitive distortions (excuses) to justify sexually aggressive behaviour. Interventions based on this theory expose men to information regarding the accurate social norms of others, with reference to sexual attitudes (Gidycz, Orchowski & Berkowitz, 2011; Hillenbrand-Gunn, Heppner, Mauch & Park, 2010).

Belief systems theory (Rokeach, 1968) proposes that beliefs which are structured centrally form part of the personality and have influence over peripherally held beliefs, and that in order to produce long-term attitudinal change interventions must take into consideration individuals' existing belief systems (Grube, Mayton, & Ball-Rokeach, 1994). Interventions based on this theory (Foubert, Newberry & Tatum, 2007; Langhinrichssen-Rohling, Foubert, Brasfield, Hill & Shelley-Tremblay, 2011) address men as potential helpers of victims of rape, rather than as potential perpetrators, with the intention they will integrate the messages from the programme more fully with their central beliefs (as a potential helper). Research evaluating the effectiveness of sexual assault interventions is increasing but is often methodologically flawed, with few studies using an experimental design and limited follow-up testing.

The concept of effectiveness is widely debated, with attitudinal change currently being the most widely used measure of programme effectiveness. Outcome evaluation consists predominantly of self-reported responses to questionnaires typically measuring attitudes towards women, rape, and empathy for victims of sexual violence. Social desirability bias inherent in self-report measures means

evidence supporting programme effectiveness is weak (Morrison et al., 2004). Generally, research points to improvements in reported attitudes associated with sexual violence immediately following completion of interventions. However, the links between attitudinal and behavioural change both generally and specifically regarding sexual violence are under-researched (Anderson & Whiston, 2005), and therefore conclusions that can be drawn from this are limited. The ethnocentricity of current research efforts means little is known about the effectiveness of these programmes with regards minority groups. Evaluation of the effectiveness of sexual assault prevention programmes is therefore currently hampered by a lack of methodological quality. The need for systematic reviews of sexual assault prevention programmes is clear in order to reduce bias in the collection and interpretation of findings, by providing a rigorous evaluation of the methodological quality of studies. This will enable a more robust understanding of what works, identify gaps in knowledge and make recommendations for future research.

The wide variety of outcome measures used to evaluate attitude change in the sexual assault prevention literature makes comparisons between studies difficult; nevertheless, reviews have been undertaken. Morrison et al. (2004) conducted a systematic review of sexual assault intervention programmes using studies published between 1990 and 2003 and found that 14% of studies reported positive intervention effects at post-test and follow up, and 80% reported mixed results. They concluded that attitudinal changes often found post-intervention were not maintained over time and noted a lack of behavioural outcome measures in order to evaluate effectiveness in terms of a reduction in sexual violence.

Anderson and Whiston (2005) conducted a meta-analysis and found significant effect sizes for rape-related attitudes, rape

knowledge, behavioural intent and incidence of sexual assault. They found that rape empathy and rape awareness behavioural outcomes were not affected by intervention and longer intervention programmes were more effective than shorter ones. Anderson & Whiston (2005) concluded that the content of the interventions, facilitator type and gender of group members may be associated with increased effectiveness.

More recently two further reviews have been conducted, however they were not systematic (Garrity, 2011; Lonsway et al., 2009). Lonsway et al. (2009) completed a review of the literature to inform practitioners and provide concrete guidance for interventions, they also noted that longer interventions appeared to be more effective than shorter interventions and that repeated exposure to content was more effective. However, the review did not provide a systematic review or evaluation of methodological quality within the studies they discussed.

Garrity (2011) conducted a review of seven studies (qualitative, $n = 2$; quantitative $n = 5$) between 2000 and 2007, and concluded that a decrease in adherence to rape myths was found post-intervention within several studies. They also found that some participants reported an increased understanding of the legal definition of rape. Garrity (2011) did not provide a comprehensive systematic review, and several studies were not included that met the inclusion criteria (Banyard, Moynihan & Plante, 2007; Foubert et al. 2007; Stephens & George, 2004; 2009). Whilst data from the studies was abstracted and reported, no systematic assessment of methodological quality was carried out. Therefore, the present study aims to provide a systematic review of the literature with analysis of methodological quality as the central component to the review. The review will evaluate studies meeting the inclusion criteria (see method) since

2003 in order to capture all relevant studies conducted after the previous systematic review (Anderson & Whiston, 2005). The overall aims of the systematic review are:

1. To investigate the types of interventions employed to reduce sexual assault in non-convicted males
2. To explore the effectiveness of these interventions
3. To evaluate the methodological quality of studies in order to reduce bias in interpretation of findings
4. To discover and highlight gaps in knowledge

3. Method

3.1 Searching

Initial studies were identified through systematic searching of the following five databases in August 2012: CINAHL, AMED, Academic Search Elite, MEDLINE, and PsycINFO. Databases were selected for relevance to the review question and include journals relating to the behavioural sciences, mental health, medicine, allied healthcare, social sciences, alternative treatments, and nursing.

Multiple combinations of words were used to identify relevant articles, consisting of a variety of key words corresponding to various combinations including the intervention, the nature of the problem (i.e. rape) the common outcome measure (i.e. rape myths), and attitudinal change (see Appendix a for search terms). A variety of different words were used to describe each component to insure relevant articles were not ignored, such as those with a variation in

the spelling of a key word. A truncation mark (*) was inserted at the end of words such as *attitud** to include: *attitudinal*; *attitude*, *attitudes*. A truncation mark (*) was also used to search for plurals of the word (eg. *Rape myth**). Google Scholar was searched using the search terms (*Rape prevention*) AND (*programme OR program*), limiting the search to the years 2003-2012, and reviewing the first 100 results. Finally, reference lists of review articles were screened to identify relevant articles.

3.2 Selection

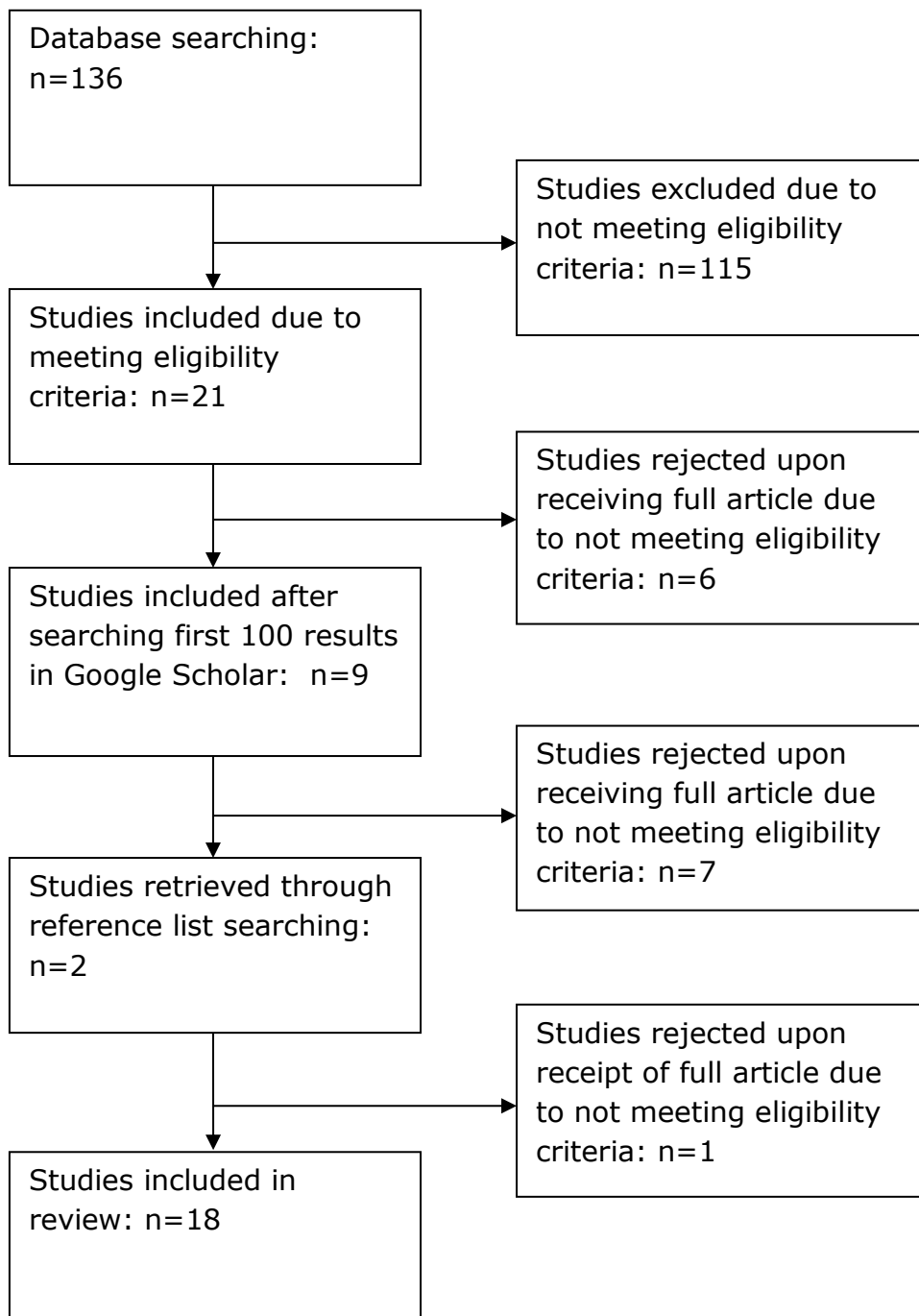
Articles were included in the review if they included: male participants; a control group; a pre-post quantitative experimental design; an intervention which sought to change attitudes or behaviour relating to sexual violence; articles written in English, peer reviewed (to increase quality); and articles that reported original data between the years 2003-2012. Articles were excluded from the review if they studied: convicted offenders; females only; victims only; attitude change regarding sexual violence without comprising an intervention; and media priming.

Database searching revealed 136 studies within which articles were reviewed for eligibility. Following exclusion of articles that did not meet eligibility criteria ($n = 115$), 21 articles remained. Six studies were rejected upon receipt of the full articles due to not meeting eligibility criteria. Nine additional articles were identified as a result of searching the Google Scholar website (first 100 results), seven of which were later excluded due to not meeting eligibility criteria. Two further studies were identified through scanning review articles references, one of which was later rejected due to not meeting eligibility criteria. This selection procedure resulted in the inclusion of

18 studies for systematic review (see Figure 1 for a flow chart depicting the identification of articles for inclusion). Two articles reported data generated from the same study (Foshee, Bauman, Ennett, Linder, Benefield, Suchindran, 2004; 2005), however they were included as they each analysed different components of the data.

Figure 1.

Selection Procedure



3.3 Data Abstraction

Each article was reviewed and data was abstracted under the following two categories: General characteristics and key findings; and methodological quality. Within the first category, the intervention type, length, gender of the participants in the intervention group, outcome measures, follow up time, sample characteristics, summary points and key findings were abstracted. The decision to focus on these factors was taken with reference to pertinent factors in the literature relating to intervention studies and other distinguishing variables.

In order to abstract relevant information with regards to the methodological characteristics of each study, a quality assessment strategy was developed. It was developed to include components of existing quality assessment tools (Critical Appraisal Skills Programme, 2005; Newcastle-Ottawa Scale, Wells et al., 2010). The quality assessment strategy also includes important factors relating to the study of rape intervention programmes, such as the use of behavioural outcome measures in improving the quality of the study (Anderson & Whiston, 2005).

In order to evaluate the methodological quality of each study the following information was assessed: participant demographics, standardisation of measures, inclusion of behavioural outcome measures, quality of statistics, level of deception, length of post-intervention follow up for administering outcome measures, and other sources of potential bias. As the inclusion criteria required only studies with a control group to be included in the review, the inclusion of a control group did not form part of the quality assessment strategy.

3.3.1 Coding Frame. Regarding participant demographics, studies were given a score of 0 if the level of information given was not adequate, 1 if the information was partially reported, and 2 if the information was clearly reported. In relation to the standardisation of measures, studies were awarded 0 if appropriate and standardised measures were not used, 1 if appropriate but modified measures were used alone or alongside standardised measures, or a limited number of standardised measures were used, and 2 if appropriate and standardised measures were used.

The behavioural outcome refers to the quality of measures of behavioural intent or action. Studies were awarded a 0 if they did not include any behavioural assessment following the completion of the intervention. Studies were given a score of 1 if limited or partially relevant behavioural assessment was included (such as an assessment of bystander behaviour or future behavioural intent, as opposed to a measure of actual reported sexual aggression). Studies were awarded a 2 if appropriate behavioural assessment was included that measured reported sexually aggressive behaviour specifically.

With regards to the assessment of statistical quality, studies were given a score of 0 if the incorrect statistical analysis was performed on the data (e.g. parametric tests on nominal data), and 1 if the appropriate test was used but effect sizes were not reported, or if correctional analyses for multiple analyses were not conducted. Studies received a score of 2 if effect sizes were reported and appropriate correctional analyses were employed.

To rate the level of deception used regarding the true aims of the study, (blinding to the position of the participant in terms of experimental or control group is not possible with intervention studies) studies were given a scores of 0 if no attempts were made to

shield the participants or the experimenters from the true aims of the study. A score of 1 was awarded if the participants were partially aware of the aims of the study (e.g. participation in a bystander group that aimed to reduce perpetration) and 2 to studies where both participants and experimenters were unaware of the true aims of the study. This variable was included as a measure of quality as it reduces bias in the form of experimenter effects and socially desirable responding to outcome measures.

Regarding the assessment of follow-up quality, studies received a score of 0 if the follow-up occurred immediately following completion of the intervention, 1 if the follow up was weeks or months later and 2 if the follow up included testing at least one year later. Greater scores were awarded to studies that included longer follow-up periods; previous research suggests drift in attitudes occurs over time following interventions; therefore an important outcome variable in assessing the effectiveness of interventions is the sustainability of cognitive change over time. However, this is only a pertinent measure of effectiveness if behavioural change results from attitudinal change. A further quality indicator was included; namely, 'other sources of potential bias,' to include specific factors which might serve to bias the findings for individual studies.

4. Results

4.1 General Characteristics and Key Findings

The abstracted data documenting the general characteristics of each study and summarising key findings is presented in Table 1.

Table 1: General Characteristics and Key Findings (for key, refer to end of table)

Author(s) and Date	Intervention Type	Intervention Length	Group Gender	Outcome Measures	Follow-up	Sample Characteristics	Summary Points and Key Findings
Johansson-Love & Geer 2003	Video discussing the effects of rape on the victim & pamphlets	22 minutes	Single sex: Male	RMAS*	Immediately and two weeks later	151 US undergraduate males. Experimental condition (n=78). Control condition (n=73). 18-39 years (M = 20.06 SD = 2.27) 84% White, 8.7% African-American, 6.7% Other Minorities	Rape myth attitudes were lower at both the immediate and the subsequent (2 weeks) assessments. Lowered adherence to rape myths was unrelated to previously held rape myth level
Langhinrichsen-Rohling et al. (2011)	Course: The Men's Programme. Based on belief system theory it works to appeal to men as helpers of victims of sexual assault and bystanders rather than perpetrators	1 hour	Single sex: Male	BES,* BWHS,* IRMA-SF*	Immediately after course	179 US college students, 17-32 years (M = 18.88 SD = 2.14)	The Men's Programme significantly increased self-reported willingness to help as a bystander and perceived efficacy as a bystander and significantly reduced adherence to rape myths in comparison to control group

Table 1: General Characteristics and Key Findings

Author(s) and Date	Intervention Type	Intervention Length	Group Gender	Outcome Measures	Follow-up	Sample Characteristics	Summary Points and Key Findings
O'Donohue et al. (2003)	Video : 3 components consisting of victim empathy, rape myth acceptance and outcome expectancies	45 minutes	Single sex: Male	Pilot study: EMC,* RMAS,* Probability Questionnaire (of sexual aggression). Main study: HS,* SES,* Motivation ratings (to decrease sexual coercion), RMAS,* AIVs,* ASBs,*ASAs,* Res,* SER, credibility ratings.	Immediately after video	102 male US undergraduates (M = 19.7 SD = 2.4) White 77.5%.	Experimental group demonstrated reductions at post-testing across all measures relating to rape. High Risk participants (i.e. prior assault history) showed significant change in expected direction pre to post across six measures (RMAs*, AIVs*, ASBs*, ASAs*, REs*, SER*). Low Risk participants (i.e. no prior assault history) showed significant change in predicted direction across two measures (ASBs* and REs*)
Rau et al. (2010)	Lecture (57 slides), 3 minute discussions x2, 3 audio dramatisations and 25 minute film. Main focus: acquaintance	Unclear, at least 31 minutes plus audio material and lecture	Single sex: Male	RMAS,* RKS,* RMS,* RES,* SES*	Immediately after programme	1,505 male, U.S. Navy Personnel. 786 in experimental group (410 with pre-testing, 376 without), 719 in control group	The Sexual Assault Intervention Training (SAIT) programme increased rape knowledge, reduced adherence to rape myths, increased empathy for victims of rape regardless of previous sexual aggression or the

Table 1: General Characteristics and Key Findings

Author(s) and Date	Intervention Type	Intervention Length	Group Gender	Outcome Measures	Follow-up	Sample Characteristics	Summary Points and Key Findings
	sexual assault and military regulations concerning the consequences, examined rape myth and misreading sexual cues, consent, sexual behaviour, aiding victims and effects of peer pressure on sexual aggression.					(427 with pre-testing, 292 without). The majority were single (89%). 17-37 years (M= 20.00 SD = 2.90)	<p>effect of pre-testing. However those with previous sexual aggression reported lower knowledge, higher acceptance of rape myths and lower empathy towards victims prior to the programme.</p> <p>Men who had completed a pre-test displayed less acceptance of rape myths and greater victim empathy at post-test.</p> <p>For men with a history of sexual coercion, rape empathy scores post-test were higher for men that completed a pre-test than for those that only did a post-test.</p> <p>Men with a history of sexual coercion exhibited greater acceptance of rape myths and lower empathy than those without a history.</p>

Table 1: General Characteristics and Key Findings

Author(s) and Date	Intervention Type	Intervention Length	Group Gender	Outcome Measures	Follow-up	Sample Characteristics	Summary Points and Key Findings
Stephens & George (2009)	Video (50 minutes), question relating to convincing another man not to be sexually aggressive, and listing three ways to help a survivor	1 hour	Single sex: Male	M-SES,* BIDR,* RMS,* IRMAS-SF,* RES,* SRAES,*BI,* ASA,* ELMQ*	Average of 11.5 days (SD 8.2) after pre-test and five week follow up occurred on average 49.2 days (mode= 33, median = 41, SD = 32.2) after post-test sessions	146 male White US undergraduates. 18-29 years (M = 19.3 SD = 1.8) Heterosexual (93.2%) 1st experience of consensual sexual intercourse = 17 years (SD 1.5) High risk males (45.2%) as measured by M-SES*	At pre-test high risk men held more rape-supportive views than low risk men across all dependent measures. Post intervention reductions are found for rape myth acceptance, attraction to sexual aggression and behavioural intentions to rape. Victim empathy scores increased post-intervention. At five week follow up, only rape myth acceptance and victim empathy effects are sustained.
Foubert & Newberry (2006)	Workshop with added bystander training module, and workshop with added module on defining consent in situations involving alcohol	1 hour plus module length (not reported)	Single sex: Male	IRMAS,* LRS,* RES*	Immediately after workshop	261 male college students (seniors = 29%, juniors = 34%, sophomores = 37%)	Post-intervention participants' empathy towards victims of sexual assault increased significantly, and rape myth adherence, likelihood of raping (LOR) and likelihood of committing sexual assault (LOCSA) significantly decreased. Those in the bystander group demonstrated significant reduction in scores post-test with regards to LOR

Table 1: General Characteristics and Key Findings

Author(s) and Date	Intervention Type	Intervention Length	Group Gender	Outcome Measures	Follow-up	Sample Characteristics	Summary Points and Key Findings
Foubert et al. (2007)	Course: The Men's Programme. Based on belief system theory it works to appeal to men as	1 hour	Single sex: Male	IRMAS,* SES*	Immediately after course and seven months later	565 male undergraduates	<p>[low effect size (ES)], LOCSA (low ES), rape myth acceptance (medium-high ES) and a significant increase in victim empathy (medium ES). Consent group experienced similar results to the bystander group apart from a medium effect size for rape myth adherence and a low effect size for increases in victim empathy.</p> <p>Between groups: no significant difference in LOR (post-test) to control. Consent and bystander groups had significantly less LOCSA. Bystander group had significantly lower RMA than controls and significantly more empathy.</p> <p>Men that had completed the intervention were significantly less likely to commit a sexually coercive act than controls during the first seven months at university.</p> <p>Long-term attitude change</p>

Table 1: General Characteristics and Key Findings

Author(s) and Date	Intervention Type	Intervention Length	Group Gender	Outcome Measures	Follow-up	Sample Characteristics	Summary Points and Key Findings
	helpers of victims of sexual assault and bystanders rather than perpetrators						(reduction in rape myth adherence) was associated with programme completion as measured seven months later. Participants that completed a pre-test produced significantly lower scores for rape myth acceptance on immediate (but not follow up) testing regardless of programme or control condition
Stephens & George (2004)	Video concerning acquaintance rape: Interviews of rape victims, rape myth challenging, discussion of links between sex and violence, cultural causes of rape, discussion around alcohol and rape	28 minutes	Single sex: Male	MSES,* RMAS,* ATWS,* SRAES,*	Immediately after video	45 US male undergraduates. 18-25 years. White 71.1%, Asian, 17.8%, Other 6.7%, African American 2.2% or Native American 2.2%.	Investigated individual differences in moderating programme effects. Effects were moderated by past coerciveness. The video reduced rape myth acceptance and sex-related alcohol expectancy scores in non-coercives, however in coercives no such effects were found.

Table 1: General Characteristics and Key Findings

Author(s) and Date	Intervention Type	Intervention Length	Group Gender	Outcome Measures	Follow-up	Sample Characteristics	Summary Points and Key Findings
Banyard et al. (2007)	One and three session programme and booster session two months later	One session + booster = 2 hours. Three sessions + booster = 5 hours	Single sex: Male	IRMAS-SF,* CDRAS,* BAS,* BB,* BES,*DBS,*	Immediately after programme, at 2 month, 4 month and 12-month follow-up.	389 US undergraduates (217 females and 172 males) at pre-test. 363 at post-test (165 males and 198 females). 284 at 2-month follow up (121 males, 162 females). 140 at 4-month follow up (62 males and 78 females). 83 at 12 month follow up (26 males and 57 females).	Compared doses of programming. Large effect sizes for the intervention were found for males. Participants in both intervention groups demonstrated improvement across outcome measures post-test compared to controls. Significant increases in pro-social attitudes regarding bystanders, and increases in self-reported bystander behaviours and perceptions of bystander efficacy were found to be significant within intervention groups, particularly so for the longest intervention group. Most programme effects were consistently present at 4 and 12-month follow ups. Some effects declined over 4 and 12 month follow ups. Nevertheless, outcomes measuring efficacy,

Table 1: General Characteristics and Key Findings

Author(s) and Date	Intervention Type	Intervention Length	Group Gender	Outcome Measures	Follow-up	Sample Characteristics	Summary Points and Key Findings
Gidycz et al. (2011)	Programme covering empathy, a norms correction component, a discussion concerning consent and a bystander intervention component. Based on the integrated model of sexual assault in which attitudes, beliefs, socialisation and peer group relationships determine conditions for sexual assault.	1.5 hour programme and 1 hour booster	Single sex	IRMAS-SF,* Hypergender Ideology Scale, Differential Reinforcement subscale of Social Norms Measure (SNM), Bystander Intervention subscale of Sexual Social Norms Inventory (SSNI), Association with Aggressive Peers subscale of SNM,	4 months and 7 months following programme completion	635 undergraduate males USA. 18-19 years (98%), unmarried (98.7%), heterosexual (98.1%) White (91.8%), African American (5%), Hispanic or Latino (2.5%), Asian (1.7%), Native Hawaiian or Pacific Islander (0.2%), American Indian or Alaska Native (0.3%), Other (0.9%)	knowledge and attitudes remained significant. Bystander behaviour significantly improved post-test and at 2 months but this did not last longer term. Reduction in sexual aggression found in programme completers at 4-months (1.5% of intervention group, to 6.7% controls). Sexually aggressive (SA) men (intervention group) - less reinforcement (than non-coercives) for SA behaviour (at 4 month not 7 months). Intervention group - fewer associations with coercives (regardless of past sexual aggression), less viewing of SA media. However, controls reported same. Intervention group - association with coercive peers higher for SA than non- across time. No differences in rape myths over time as a function of group. Intervention group

Table 1: General Characteristics and Key Findings

Author(s) and Date	Intervention Type	Intervention Length	Group Gender	Outcome Measures	Follow-up	Sample Characteristics	Summary Points and Key Findings
	A booster session is also included.			Modelling subscale of SNM, Overall Reinforcement subscale of SNM, Bystander Intervention subscale of SSNI. Rape scenarios SES,*MCSDS*			viewed other men more likely to act as bystanders compared to controls (at 4 and 7-months). Intervention group SA- increased perceptions of other mens' willingness to be bystanders than baseline and SA controls at 4 months. Mens' likelihood to intervene did not change as a function of time or group.
Foshee et al. (2004)	Programme (first stage) including theatre production, 10 curriculum sessions and a poster contest then booster (second stage) which consisted of an 11 page newsletter posted to their home and a telephone	Theatre production (unreported length) Ten 45 minute curriculum sessions and a poster contest (unreported length). Booster (unreported).	Mixed	Sexual violence was defined by the sum of a subset of 2 acts of sexual aggression	1 month (wave 2) and 1 year (wave 3) post intervention (first stage). Further analyses were conducted on those that consented at 2 years (wave 4) post intervention, then following completion of	Adolescents in the 8th grade in the autumn of 1994 in 10 public American schools. Analysis is limited to those who completed baseline (wave 1) and wave 4 and 6 questionnaires (n=460). Control group comprised of 201 participants and of those that	Race nor gender moderated programme effects. Safe Dates effect on sexual perpetration: participants (male and female combined) reported perpetrating significantly less sexual violence at 4-year follow up than controls. The booster did not improve the effectiveness of Safe Dates in preventing sexual dating violence.

Table 1: General Characteristics and Key Findings

Author(s) and Date	Intervention Type	Intervention Length	Group Gender	Outcome Measures	Follow-up	Sample Characteristics	Summary Points and Key Findings
	call for those that consented to second stage. Focus is on social norms, reducing cognitive barriers to increase likelihood of taking preventative action, community support, and conflict management strategies.	length) 4 weeks after the newsletter was posted			booster or non-booster control (second stage) further follow up tests were completed at 3 years (wave 5) and 4 years (wave 6) post completion of intervention.	received the intervention, 124 were in the group that received only 'Safe Dates' intervention and 135 made up the group that received Safe Dates and the booster.	There were no significant differences between booster group and control in follow-up sexual dating violence perpetration.
Foshee et al. (2005)	Programme (first stage) including theatre production, 10 curriculum sessions and a poster contest then booster (second stage) which consisted of an 11 page	Theatre production (unreported length) 10 45 minute curriculum sessions and a poster contest (unreported length).	Mixed	Sexual violence was defined by the sum of a subset of 2 acts of sexual aggression	1 month (wave 2) and 1 year (wave 3) post intervention (first stage). Further analyses were conducted on those that consented at 2 years (wave 4)	Adolescents in the 8th or 9th grade in the autumn of 1994 in 14 public American schools. Analysis is limited to those who completed wave 1,2,3,4 and wave5 (n=1566). White	Race nor gender moderated programme effects. Safe Dates effect on sexual perpetration: participants (male and female combined) reported perpetrating significantly less sexual violence at 1 month, 1 year, 2 years and 3 years following programme than controls.

Table 1: General Characteristics and Key Findings

Author(s) and Date	Intervention Type	Intervention Length	Group Gender	Outcome Measures	Follow-up	Sample Characteristics	Summary Points and Key Findings
	newsletter posted to their home and a telephone call for those that consented to second stage. Focus is on social norms, reducing cognitive barriers to increase likelihood of taking preventative action, community support, and conflict management strategies	Booster (unreported length) 4 weeks after the newsletter was posted.			post intervention, then following completion of booster or non-booster control (second stage) further follow up tests were completed at 3 years (wave 5) and 4 years (wave 6) post completion of intervention.	(72.2%, male (46.8%). Mean age at baseline was 13.9 years. Control group consisted of 930 participants and intervention group 636 participants.	The effects of the intervention were mediated by dating violence norms, gender-role norms, and awareness of community services.
Currier & Carlson (2009)	1 course covering women and violence, (students read materials from sociology, psychology,	One semester	Mixed	RMAS,* ATRVS,*MCSD S*	Immediately following course completion	Pre-test (n = 214, 77 in women and violence course, 56 in gender course and 81 in sociology course).	With regards to those in the 'Women and Violence' course, significant reductions were observed regarding negative attitudes towards rape victims and changes in attitudes regarding date rape.

Table 1: General Characteristics and Key Findings

Author(s) and Date	Intervention Type	Intervention Length	Group Gender	Outcome Measures	Follow-up	Sample Characteristics	Summary Points and Key Findings
	<p>women’s studies, and political science about the social, cultural, and legal aspects of violence against women and engage in discussions/presentations). 1 course concerning gender. 1 Sociology course</p>					<p>Post-test (n = 137, 66 in women and violence course, 48 in gender course and 23 in sociology course). Participant ages ranged from 18-35 years (M = 20.23, SD = 1.82). Women (77%). White (86%), African American (5%), hispanci/Latino/Latina (3%), Asian/Asian American (3%), Native American (1%). Freshman (22%), Sopomores (19%), Juniors (20%), Seniors (36%), Those beyond 8th semester at university (3%).</p>	<p>Rape myth acceptance was un-changed. May be due to low base rate of rape myth acceptance.</p> <p>Males had higher rape myth acceptance than females at pre-and post-testing.</p>

Table 1: General Characteristics and Key Findings

Author(s) and Date	Intervention Type	Intervention Length	Group Gender	Outcome Measures	Follow-up	Sample Characteristics	Summary Points and Key Findings
Kress et al. (2006)	Programme consisting of information sharing, peer theatre, peer group facilitation and large group discussion. Role of alcohol, rape myth challenging by peers, and personalising sexual assault prevention for individuals in the group was also included.	2.5 hours	Mixed	IRMAS-SF,*	Immediately following programme completion	234 college freshman USA. (174 fully completed measures and so used in analysis). Female (66%), Caucasian (90%), African American (4.3%). 17-19 years (97%), 18 years (78%).	Regardless of gender, age and race, rape myth attitudes were significantly reduced following the programme. Females endorsed rape myths significantly less than males at pre-test.
Fay & Medway (2006)	Programme: encouraging critical thinking, reflection and discussion rather than lecture. Learning about rape and cultural influences,	2 hours	Mixed	RMAS,* Attitudes Towards Dating Violence.	Immediately following programme completion and between 5 and 7 months later for the next post-test.	American high-school students with re-test and initial post-test data (n = 154, males = 67, females = 85) control group (n = 78) treatment group (n = 76).	Rape myths significantly decreased following intervention compared to control. Attitudes towards dating violence were unchanged. Pre-programme, males=more rape supportive attitudes than females. Difference remained at post-

Table 1: General Characteristics and Key Findings

Author(s) and Date	Intervention Type	Intervention Length	Group Gender	Outcome Measures	Follow-up	Sample Characteristics	Summary Points and Key Findings
	exploring feelings, learning about mixed messages in communication, identifying rape prevention strategies and local victim support services.					Groups did not differ by gender but did so by race; more African Americans were in the intervention (57%) group than control (40%).	testing no main effect observed for race. Males held attitudes more supportive of dating violence than females pre- and post- programme. This correlated with high scores on rape myth scales. Intervention (not gender) accounted for significant reduction in immediate post-test scores regarding RMA. Acquaintance rape attitudes: more pro-social attitudes were held within programme participants at post-test and delayed follow up. RMA (not treatment or gender) significantly predicted attitudes towards dating violence as post-test. Adherence to rape myths significantly decreased over all testing points and results were not affected by gender. No such change was seen within controls.

Table 1: General Characteristics and Key Findings

Author(s) and Date	Intervention Type	Intervention Length	Group Gender	Outcome Measures	Follow-up	Sample Characteristics	Summary Points and Key Findings
Bradley et al. (2009)	Programme. Overhead projector: topics covered included rape myths and facts, risk factors/perception, response strategies, Video covered: Victim empathy and outcome expectancies	50 minutes	Mixed	RMAS,* AIVs,* ASBs,* Adjective Checklist, Rape Outcome Expectancy Scale, Programme Information Quiz, Consumer Satisfaction Survey	Immediately after programme and 2 week follow up	309 undergraduate males (n = 113) and females (n = 196) USA from 11 classes. Control group (n = 132), experimental group (n = 177). Mean age = 23.2 (SD = 6.0).	Male victim empathy, distress and sadness increased following programme. Adherence to rape myths decreased significantly between pre-test and 2-week follow up. No significant changes in scores on AIVs* or ASBs* at 2-week follow up. No significant differences between outcome expectancy as a function of group or time. Knowledge of sexual assault increased between pre-test and 2-week follow up for experimental group.
Moynihan et al. (2010)	Bringing in the Bystander program informs about sexual and intimate partner violence and uses skill building practise to encourage safe	4.5 hours	Single sex (single sex female groups also included in	IRMAS-SF,* BES,* Bystander Intention to Help-Short Form, BBs,* Post-programme bystander	Immediately following programme and 2-month follow up	Participants whose data could be used in analysis of three attitudinal outcome measures (n=98). Experimental group (n=36) Control group (n=62). For bystander behaviour	Significant reductions in rape supportive attitudes reported over time however differences were observed between males and females. No significant differences were found pre-test to 2-month follow up regarding bystander behaviour.

Table 1: General Characteristics and Key Findings

Author(s) and Date	Intervention Type	Intervention Length	Group Gender	Outcome Measures	Follow-up	Sample Characteristics	Summary Points and Key Findings
	intervening in risk situations		study)	behaviour questions		measures (n=93) Experimental group (n=35), Control group (n = 58).	
Hillenbrand-Gunn et al. (2010)	Programme based on the men as allies philosophy, and theoretical framework of social norms. Discussion and a music video by a male rap artist, male role-models, helping victims	3 sessions each lasting 45 minutes	Mixed	IRMAS-SF,* WWYD,* Self-Protective Behaviours Measure, WWYD-T,* DSS-R,* DSS-R-T,* MCSD-C*	Immediately following the programme and 4-week follow up	212 participants completed all assessments at three points of testing. Experimental group (n = 124, male = 78, female = 46) Control group (n = 88, male = 50, female = 38). Caucasian (83.0%), African American (5.2%), Hispanic, Latina/o, or Chicana/o (1.9%), Asian American or Pacific Islander (1.4%), Other (2.4%), no response	Males demonstrated significant decrease in adherence to rape myths both at post-test and 4-week follow up. A significant difference between control and experimental group at post-test regarding willingness to avoid engaging in sexually coercive behaviour was found for males (experimental group remained the same as pre-test; control group reported less willingness to avoid engaging in sexual aggression post-test). Male student rated their peers as more rape supportive than the participants rated

Table 1: General Characteristics and Key Findings

Author(s) and Date	Intervention Type	Intervention Length	Group Gender	Outcome Measures	Follow-up	Sample Characteristics	Summary Points and Key Findings
						(4.2%). Mean age = 16.58 (SD = .73).	themselves at pre-test, however, this bias reduced following intervention.

*Key: AIVs = Acceptance of Interpersonal Violence scale, ASBs = Adversarial Sexual Beliefs scale, RMAs = Rape Myth Acceptance scale, ASAs = Attraction to Sexual Aggression scale, ATWS = Attitudes Towards Women Survey, BAS = Bystander Attitudes Scale, BB = Bystander Behaviours, BES = Bystander Efficacy Scale, BI = Behavioural Intentions, BWHS = Bystander Willingness to Help Scale, CDRAS = College Date Rape Attitude Survey, DBS = Decisional Balance Scale, ELMQ = Elaboration Likelihood Model Questionnaire, EMC = Empathy Manipulation Check, HS = Hypermasculinity Scale, IRMAS-SF = Illinois Rape Myth Acceptance Scale-Short Form, LRS = Likelihood of Raping Scale, MCSD = Marlowe-Crowne Social Desirability Scale, MCSD-C = Marlowe-Crowne Social Desirability Scale-Form C, MSES = Modified Sexual Experiences Scale, RMAS = Rape Myth Acceptance Scale, RMS = Rape Myth Scale, REs = Rape Empathy scale, RKS = Rape Knowledge Scale, SER = Self-Efficacy Ratings, SES = Sexual Experiences Survey, SBS = Sexual Beliefs Scale, SRAES = Sex-Related Alcohol Expectancies Scale, WWYD = What Would You Do? WWYD-T = What Would You Do-Typical Guy, DSS-R = Discomfort With Sexist Situations-Revised Scale, DSS-R-T = Discomfort With Sexist Situations-Revised Scale-Typical Guy

4.1.1 Intervention type. Out of the 18 studies: four interventions were video based (Johansson-Love & Geer, 2003; O'Donohue, et al., 2003; Stephens & George, 2004; 2009); one consisted of a lecture, film and discussion (Rau et al., 2010); two interventions were workshops (Banyard et al., 2007; Foubert & Newberry, 2006); and ten were programmes (Bradley et al., 2009; Currier & Carlson, 2009; Fay & Medway, 2006; Foubert et al., 2007; Foshee, 2004; 2005; Gidycz et al., 2011; Hillenbrand-Gunn et al., 2010; Kress et al., 2006; Langhinrichsen-Rohling et al., 2011; Moynihan et al., 2010). See Table 1 for more information regarding the content of each intervention.

4.1.2 Intervention Length. A wide range of different length interventions were reviewed, ranging from 22 minutes to one semester in length. Two interventions lasted less than one hour (Johansson-Love & Geer, 2003; Stephens & George, 2004). The majority ($n = 7$) lasted approximately one hour (Bradley, 2009; Foubert & Newberry, 2006; Foubert et al., 2007; Langhinrichsen-Rohling et al., 2011; O'Donohue et al., 2003; Rau et al., 2010; Stephens & George, 2009). Two interventions lasted approximately two hours (Fay & Medway, 2006; Gidycz, 2011; Kress, 2006). One intervention lasted between two and five hours with the inclusion of a booster programme (Banyard et al., 2007). One intervention lasted for 4.5 hours (Moynihan, 2010) and others consisted of a course of sessions; for example, Foshee (2004; 2005) studied an intervention which involved participants attending ten 45-minute sessions. The Men as Allies programme (Hillenbrand-Gunn, 2010) consisted of three 45-minute sessions. The longest course lasted one semester (Currier & Carlson, 2009).

4.1.3 Group Gender. The majority of groups consisted of males ($n = 10$), one programme was designed for both males and females but consisted of single sex groups (Moynihan et al., 2010) and seven groups were mixed (Bradley et al., 2009; Currier & Carlson, 2009; Fay & Medway, 2006; Foshee et al., 2004; 2005; Hillenbrand-Gunn et al., 2010; Kress et al., 2006).

4.1.4 Outcome Measures. The majority of studies used both attitudinal and behavioural outcome measures ($n = 10$), some used attitudinal measures only ($n = 5$) and others used behavioural measures only ($n = 2$). The most commonly used instruments included: the Rape Myth Acceptance Scale (Burt, 1980); the Illinois Rape Myth Acceptance Scale – Short Form (Payne, Lonsway & Fitzgerald, 1999); the Sexual Experiences Survey (Koss, 1985); the Acceptance of Interpersonal Violence scale (Burt, 1980); and the Adversarial Sexual Beliefs scale (Burt, 1980). Several studies ($n = 3$) also included measures designed by the authors.

4.1.5 Follow-up Period. The majority of studies only completed post-intervention testing immediately following completion of the programme ($n = 7$). Many included follow-up testing months later in addition to testing immediately following completion of the programme ($n = 6$). A limited number of studies included follow up testing weeks after completion of the programme ($n = 2$). Other studies held several follow-up tests over a period of four years ($n = 3$).

4.1.6 Sample Characteristics. The majority of studies sampled male undergraduates ($n = 13$). Others sampled males in the final year of school ($n = 4$) and one study sampled males in the US Navy. The majority of studies consisted of sample sizes between 100-200 participants ($n = 6$). Two studies sampled less than 100 participants and two studies sampled more than 1,500. Other studies sampled between 300-700 participants ($n = 8$). Participants were predominantly aged 18-20 years ($n = 9$), with some studies sampling participants aged 16-18 years ($n = 4$) and others including participants up to the age of 37, however these studies predominantly sampled 20 years old males ($n = 5$). Most studies sampled White participants as a significant majority ($n = 16$), however participants of Asian origin accounted for 17.8% of the total sample in one study (Stephens & George, 2004), and those of African American origin accounted for 57% of the intervention group in another (Fay & Medway, 2006).

4.2 Methodological Quality

Data regarding the methodological quality of studies was abstracted, evaluated and reported in Table 2.

Table 2: Methodological Characteristics of Quantitative Studies (for Key, refer to end of table)

Study	Participant Demographics	Measures: Standardisation	Measures: Behavioural	Statistics	Deception	Follow-up	Other Sources of Potential Bias
Johansson-Love & Geer 2003	1	1	0	2	0	1	Rape Myth Acceptance Scale (Burt, 1980) outdated, and norms not generated on same sample in study
Langhinrichsen-Rohling et al. (2011)	1	2	0	2	1	0	Sample selection bias (e.g. motivated men participated). Controls may have also participated in experimental group.
O'Donohue et al. (2003)	1	1	1	1	1	0	Proxy variables measured rather than actual behavioural change
Rau et al. (2010)	2	1	0	2	1	0	Potential for socially desirable responding in relation to past sexually assaultive behaviour prior to completing the programme
Stephens & George (2009)	2	1	1	2	1	1	Vast majority of participants were white and so sample lacks generalisability
Foubert & Newberry (2006)	1	2	1	2	0	0	Poor sample representativeness (predominantly white). Floor effects (very low likelihood of raping scores in controls) precluded a significant finding between control and intervention groups on likelihood of raping

Table 2: Methodological Characteristics of Quantitative Studies

Study	Participant Demographics	Measures: Standardisation	Measures: Behavioural	Statistics	Deception	Follow-up	Other Sources of Potential Bias
Foubert et al. (2007)	0	1	2	1	1	1	1 Experimenters returned to residences to collect post-test measures. This might have influenced participant responding, thinking they may be identified more easily.
Stephens & George (2004)	1	1	0	2	2	0	0 No pre-post testing of attitudes. The randomised pre-test post-test design allowed inference of change; however it did not permit determination of change.
Banyard et al. (2007)	2	1	1	2	1	2	2 Data at 4- and 12-month test points had much smaller sample sizes. Lack of standardised measures for bystander attitudes and behaviour.
Gidycz et al. (2011)	1	1	2	2	1	1	1 The intervention group had higher exposure to sexually explicit material and more sexually aggressive experiences than controls at the start of the programme. Reduction in self-reported sexual aggression for intervention group- however lack of long term follow up means low base rate and potential for type II error.
Foshee et al. (2004)	0	0	1	1	0	2	2 High level of attrition rate due to length of study.

Table 2: Methodological Characteristics of Quantitative Studies

Study	Participant Demographics	Measures: Standardisation	Measures: Behavioural	Statistics	Deception	Follow-up	Other Sources of Potential Bias
Foshee et al. (2005)	1	0	1	1	0	2	Threats to internal validity imposed by differential attrition and/or differential predictors of attrition by treatment condition are not controlled by design. Potential explanation for the favourable effects not controlled by design is that adolescents in the treatment group as compared to those in the control group provided more socially desirable responses to the behavioural measures. Multiple imputation procedures were used and research is limited in examining how variations in the missingness equation can influence study conclusions.
Currier & Carlson (2009)	1	2	0	2	0	0	More women than men in the intervention groups but equal numbers in the control (sociology) group. Participants were younger in the control group. Reduced power to calculate control group as lower sample size.
Kress et al. (2006)	1	1	0	2	0	0	Males in the study tended to be older and age was not controlled for. Social desirability effects due to limited blinding and only one measure of outcome, potentially increasing awareness of required response following intervention.

Table 2: Methodological Characteristics of Quantitative Studies

Study	Participant Demographics	Measures: Standardisation	Measures: Behavioural	Statistics	Deception	Follow-up	Other Sources of Potential Bias
Fay & Medway (2006)	1	1	0	2	0	1	Loss of much of the delayed post-test data reduces ability to make firm conclusions of longer term effects of the programme
Bradley et al. (2009)	1	2	0	2	0	1	Ceiling and floor effects in outcome measures, reducing their sensitivity. No random assignment of participants to experimental or control groups.
Moynihan et al. (2010)	2	2	1	2	1	1	45% of participants had already received rape prevention training prior to the current study. Possibility of sharing learning from experimental to control group through conversation in the 2 month follow up period. Small follow-up time for behavioural assessment potentially causing type II errors due to low base rate.
Hillenbrand-Gunn et al. (2010)	2	1	1	2	1	1	Non-randomised assignment to control or experimental groups leading to threats to internal validity (selection effects). Two of the outcome measures were constructed by the researcher and so need to be validated.

NB. Key

Participant Demographics

- 0 Participant demographics not adequately reported
- 1 Participant demographics partially reported
- 2 Participant demographics clearly reported

Measures: Standardisation

- 0 No appropriate standardised measures are used
- 1 Appropriate but modified measures are used alone or with standardised measures or limited standardised measures are used
- 2 Appropriate and standardised measures are used

Measures: Behavioural

- 0 No assessment of behaviour included
- 1 Limited/moderately relevant behavioural assessment included (eg bystander behaviour, or likelihood of future behaviour rather than reported sexual aggression)
- 2 Appropriate behavioural assessment measures were included with regards to sexually aggressive behaviour

Statistics

- 0 wrong statistical test used
- 1 no effect sizes reported and/or correction tests used
- 2 effect sizes reported and/or appropriate post-hoc tests used

Deception

- 0 Participants and experimenters were aware of the true aims of the study
- 1 Participants were partially aware of the aims of the study (eg a bystander group), experimenters were not
- 2 Participant and experimenters were both unaware of the true aims of the study

Follow-up

- 0 immediately after intervention
- 1 Weeks or months after intervention
- 2 Years after intervention

The following factors are pertinent in assessing methodological quality of intervention studies.

4.2.1 Participant demographics. Demographic variables are important in assessing the quality of rape intervention studies; research points to gender and age differences in: adherence to rape myths; sexually aggressive behaviour; and intervention effectiveness (Anderson & Whiston, 2005; Morrison et al., 2004). The participant demographics were partially reported in the majority of studies ($n = 11$), fully reported in some studies ($n = 5$) and under-reported in a handful of studies ($n = 2$). All studies sampled participants from the US. Adequate participant demographic information enables bias to be limited when generalising findings.

4.2.1 Standardisation of measures. Outcome measures were rated regarding the degree to which they were standardised. The majority of studies included standardised measures alongside modified measures and, in a few cases, author-designed measures ($n = 14$). In some studies a wide range of standardised measures were used ($n = 5$) and in others only author-designed measures were used ($n = 2$) such as a two-item questionnaire to assess participants' reported sexual violence following the completion of the intervention. The use of standardised measures enables more valid comparisons to be made between studies, normative data to be established and for outcome data to be synthesised (e.g. for use in meta-analysis).

4.2.2 Behavioural outcome measures. The inclusion of measures of behaviour is pertinent when studying the effectiveness

of interventions designed primarily to change behaviour. Until recently, however, this aspect of evaluating effectiveness has been largely ignored, predominantly due to a lack of effective behavioural measures. Of the studies under review, eight did not include any form of behavioural assessment as an outcome measure; and eight included a limited or moderately relevant behavioural measure, such as, assessing behaviour indirectly associated with sexual violence (i.e. bystander behaviour).

One study comprising two articles (Foshee et al., 2004; 2005) used a behavioural measure alone to evaluate effectiveness; however this was a self-report measure which is subject to social desirability bias. There is a need for the development of further measures of behavioural outcomes in relation to sexual violence in order to provide a more reliable evaluation of programme effectiveness.

4.2.3 Statistics. Studies were deemed to be of high quality statistically if: the appropriate analysis was carried out; effect sizes were reported; and a correctional test was employed (if multiple analyses were performed). The number of studies meeting the criteria for high quality statistical analysis was high ($n = 14$). The remaining studies ($n = 4$) failed to report effect sizes within the article, meaning limited conclusions can be drawn from the data and it is unclear as to the magnitude of any given significant result (Foshee et al., 2004; 2005; Foubert et al., 2007; O'Donohue et al., 2003).

Many studies carried out multiple comparisons of the data, thus increasing the likelihood of making a type I error (incorrectly rejecting the null hypothesis), however in order to correct this appropriate post-hoc tests were carried out (e.g. Bonferroni

correction). As a result of applying corrections to the data several significant results were found to be no longer significant. Finally, multiple imputation procedures were employed in one study (Foshee et al., 2005) and research is limited in examining how variations in the missingness equation can influence study conclusions.

4.2.4 Deception. As it is largely unfeasible to employ a double blind experimental design within intervention studies, another method of reducing bias is to examine the level of deception used. When participants and experimenters are unaware of the true aims of the study they are less biased toward socially desirable responding or influenced by experimenter effects.

The studies included in this review included partially deceiving participants with regard to the true aims of the study (n = 9), and many studies employed limited deception strategies (n = 7). For example, interventions aiming to reduce the perpetration of sexual violence by recruiting participants to a bystander awareness programme fell into this category. One study did not inform the individual responsible for recruiting participants, of participants' previous history of sexual violence, and did not inform the participants that the measure of past sexual behaviour was linked to the study (Stephens & George, 2004). All participants were fully debriefed following completion of the study.

4.2.5 Post-intervention follow-up. In order to evaluate the effectiveness of an intervention programme, the design of the study needs to be able to capture the sustainability of any attitudinal and behavioural change. Therefore the length of follow-up period

following the completion of the study is an important factor in the methodological evaluation of intervention studies. Eight studies in the review employed a design that tested outcome measures several weeks or months after the intervention. Seven studies employed no follow up testing other than immediately following the completion of the intervention. Other studies tested participants one year ($n = 1$) and at one, two, three and four years ($n = 2$) after completion of the initial intervention.

4.2.6 Other sources of potential bias. One study failed to randomise assignment of participants to the control or experimental groups, which leads to threats to internal validity (Hillenbrand-Gunn et al. 2010). Many studies used the Rape Myth Acceptance Scale (Burt, 1980) as an attitudinal outcome measure, however this measure has been criticised for being outdated and the norms were generated on males approximately 20 years older than those included in the majority of studies in this review. As younger males have been found to hold higher levels of rape-supportive beliefs than older males (Anderson & Wiston 2005), it is likely that the norms and qualitative meaning, assigned to scores within this scale, under-represent the strength of beliefs held within younger males.

Some studies employed behavioural outcome measures of sexual violence post-intervention, however due to the limited follow up time and low base rate of sexual violence, the likelihood of making a Type I error in statistical analysis is high (incorrect rejection of the null hypothesis). Attrition rates were high in studies with long-term follow up periods which resulted in diminished sample sizes over time, and made comparisons over time less reliable.

5. Discussion

5.1 Summary

The majority of interventions found a significant reduction in rape related attitudes as a function of time and group ($n = 14$). Of those studies that investigated the effects of risk (as measured by past self-reported sexual aggression) on outcome measures, three studies found no change in rape myth acceptance over time for high risk males within intervention groups (Gidycz, 2011; Stephens & George, 2004; 2009). In contrast, two studies found decreases in rape myth acceptance for high risk males in intervention groups (O'Donohue et al., 2003; Rau et al., 2010).

Following an intervention group, one study found high risk males to have lower perception of reinforcement for sexually aggressive behaviour than low risk males, however this was not sustained at four-month follow-up (Gidycz, 2011). Pro-social attitude change was greatest for low risk males (Stephens & George, 2004; 2009). The majority of studies that included a follow-up component found that attitude change was sustained over time (Banyard et al., 2007; Currier & Carlson, 2009; Foubert et al., 2007; Gidycz et al., 2011; Hillenbrand-Gunn et al., 2010; Johannson-Love & Geer, 2003; Stephens & George, 2009), with only one finding changes were not sustained (Fay & Medway, 2006). However, as the majority of studies employed only limited follow-up periods, conclusions regarding the effectiveness of interventions in creating sustained attitudinal change are tenuous.

Booster groups had a positive impact on reported bystander behaviours and attitudes following the groups, however, these effects were not sustained (Banyard et al., 2007).

In relation to behavioural change, all studies investigating this outcome found significant decreases in self-reported sexual violence following the intervention compared to control groups (Foshee, 2004; 2005; Foubert et al., 2007; Gidycz, 2011). However, due to the method of data collection (self-report) social desirability bias is significant.

Booster groups did not further reduce reported sexual violence at follow-up testing (Foshee, 2004; 2005). Behavioural intent to commit sexual assault was reduced following intervention compared to controls (Foubert & Newberry, 2006; Stephens & George, 2009) but this effect did not remain at five-week follow-up (Stephens & George, 2009).

Regarding victim empathy, five studies reported increases as a function of time and group (Bradley et al., 2009; Foubert et al., 2007; O'Donahue et al., 2003; Rau et al., 2010; Stephens & George, 2009). Little is known about the effects of risk status on changes in victim empathy following intervention. Therefore, whilst victim empathy components warrant inclusion intuitively, little research evidence exists to support its inclusion. Research with convicted sex-offenders has found that having empathy for victims of sexual assault has no effect on risk of sexual recidivism (Brown, Harkins & Beech, 2011). Further research is warranted with non-convicted males to explore this in order to assess effective components for interventions.

5.2 Limitations

Dissertations were excluded from this review, resulting in a bias toward publications, meaning that studies reporting significant results are more likely to have been submitted and published. Therefore this

review is likely to contain a degree of bias in favour of studies which found significant results.

Although attention was given to systematic and objective methodological coding of the data, it is likely that the coding strategy contained bias due to the difficulty in ensuring ratings were applied objectively.

Studies often failed to report in sufficient detail, the content of the interventions in question. This lack of detail means that treatment integrity cannot be assessed; meaning evaluation of treatment is difficult due to the high level of confounding variables. In light of this, it is important for future studies to include information regarding content and integrity.

5.3 Future Research

There is a need for further development of behavioural outcome measures to assess sexual violence. The limited reliance on behavioural measures in this review restricted the conclusions that could be drawn in relation to programme effectiveness. One option for future research could include surveying the population after the intervention has finished, assessing behavioural change through better use of local university statistics regarding surveys of sexual violence.

Better consideration needs to be given to developing theoretically driven treatment programmes, so as to enable theoretically linked hypotheses to be proposed, regarding individual components of treatment. This will enable hypotheses to be tested empirically. For example, future programmes might benefit from including components drawn from theories such as the hypocrisy salience

theory (Aronson, 1999) which has received attention with regards to its application to sexual assault prevention programmes (Paul & Gray, 2011).

Hypocrisy salience theory states that incoming information that is incongruent with own beliefs is retained less effectively than consistent information, and leads to cognitive dissonance (Festinger, 1957). Cognitive dissonance is defined as an uncomfortable feeling resulting from holding two conflicting beliefs simultaneously. Behaviour that is incongruent with beliefs is also proposed to lead to cognitive dissonance (Stone & Cooper, 2001). Individuals seek to reduce cognitive dissonance by changing discordant factors, one of which includes behavioural change if past behaviour is at odds with current stance (Festinger, 1957). Interventions basing themselves on this theory induce hypocrisy and subsequent cognitive dissonance in individuals by giving them a task in which they are required to take a stance which conflicts with their past sexually aggressive attitudes or behaviour. This is thought to induce subsequent motivation to reduce cognitive dissonance. There is a need for rape-prevention interventions to utilise theoretically-driven components such as these in order to provide a basis within which hypotheses can be empirically tested.

Studies ought to use updated standardised measures of rape-supportive beliefs and longer-term follow up procedures are warranted. Future measures should seek to improve validity by reducing socially desirable responding, through the use of indirect/implicit measures of beliefs.

As the majority of studies sample American college students, there is a need to develop future research with other groups in order to increase the sample representativeness.

5.4 Conclusions

A review of the literature regarding sexual assault interventions with non-convicted males has shown that whilst attitudinal change appears to occur following interventions, this has limitations in terms of evaluating programme effectiveness. Behavioural outcome assessment is limited and the links between attitudinal and behavioural change regarding sexual violence remains unclear.

In relation to clinical practice, attention should be given to the risk status of men undertaking the programmes as treatment effectiveness has often been found to be limited for high risk groups. This will enable more effective treatment programmes to be developed for those in greatest need. Interventions need to develop a theoretical basis in order to allow empirical studies to be conducted into the effectiveness of individual components and hypotheses to be tested.

References

- Anderson, L. & Whiston, S. (2005). Sexual assault education programmes. A meta-analytic examination of their effectiveness. *Psychology of Women Quarterly, 29*, 374-388.
- Aronson, E. (1999). Dissonance, hypocrisy, and the self-concept. *Readings About the Social Animal, 227-224*.
- Banyard, V., Moynihan, M., & Plante, E. (2007). Sexual violence prevention through bystander education: An experimental evaluation. *Journal of Community Psychology, 35*(4), 463-481.
- Bedard-Gilligan, M., Jaeger, J., Echiverri-Cohen, A., & Zoellner, L. (2012). Individual differences in trauma disclosure. *Journal of Behaviour Therapy and Experimental Psychiatry, 43*(2), 716-723.
- Bradley, A. R., Yeater, E. A., & O'Donohue, W. (2009). An evaluation of a mixed-gender sexual assault prevention program. *The Journal of Primary Prevention, 30*(6), 697-715.
- Brown, S., Harkins, L., & Beech, A. (2011). General and victim-specific empathy: Associations with actuarial risk, treatment outcomes and sexual recidivism. *Sex Abuse, 24*(5), 411-430.
- Burt, M.R. (1980). Cultural myths and supports for rape. *Journal of Personality and Social Psychology, 38*, 217-230.
- Critical Appraisal Skills Programme (CASP; 2005). *CASP Checklists (URL used)*. Oxford: CASP.

- Currier, D. M., & Carlson, J. H. (2009). Creating attitudinal change through teaching: How a course on "women and violence" changes students' attitudes about violence against women. *Journal of Interpersonal Violence, 24*(10), 1735-1754.
- Fay, K. E., & Medway, F. J. (2006). An acquaintance rape education program for students transitioning to high school. *Sex Education, 6*(3), 223-236.
- Festinger, L. (1957). *Theory of cognitive dissonance*. Evanston, IL: Row, Peterson.
- Foubert, J. D., & Newberry, J. T. (2006). Effects of two versions of an empathy-based rape prevention program on fraternity men's survivor empathy, attitudes, and behavioral intent to commit rape or sexual assault. *Journal of College Student Development, 47*(2), 133-148.
- Foshee, V., Bauman, K., Ennett, S., Linder, F., Benefield, T., & Suchindran, C. (2004). Assessing the long-term effects of the Safe Dates programme and a booster in preventing and reducing adolescent dating violence victimisation and perpetration. *American Journal of Public Health, 94*(4), 619-624.
- Foshee, V., Bauman, K., Ennett, S., Linder, F., Benefield, T., & Suchindran, C. (2005). Assessing the effects of the dating violence prevention programme "Safe Dates" using random coefficient regression modelling. *Prevention Science, 6*(3), 245-258.
- Foubert, J., Newberry, J., & Tatum, J. (2007). Behaviour differences seven months later: Effects of a rape prevention programme on first year men who join fraternities. *NASPA Journal, 44*, 728-749.

- Garrity, S. (2011). Sexual assault prevention programs for college-aged men: A critical evaluation. *Journal of Forensic Nursing, 7*, 40-48.
- Gidycz, C., Orchowski, L., & Berkowitz, A. (2011) Preventing sexual aggression among college men: An evaluation of a social norms and bystander intervention programme. *Violence Against Women, 17*(16), 720-742.
- Grube, J. W., Mayton, D. M., & Ball-Rokeach, S. J. (1994). Inducing change in values, attitudes and behaviours: Belief system theory and the method of value self-confrontation. *Journal of Social Issues, 50*, 153-173.
- Hillenbrand-Gunn, T., Heppner, M. J., Mauch, P. A., & Park, H. (2010). Men as allies: The efficacy of a high school rape prevention intervention. *Journal of Counseling & Development, 88*(1), 43-51.
- Johansson-Love, J., & Geer, J. H. (2003). Investigation of attitude change in a rape prevention program. *Journal of Interpersonal Violence, 18*(1), 84-99.
- Kilpatrick, D., Resnick, H., Rugiero, K., Conoscenti, L., & McCauley, J. (2007). *Drug-facilitated, Incapacitated, and Forcible Rape: A National Study*. Charleston, SC: Medical University of South Carolina and National Crime Victims Research and Treatment Center.
- Koss, M.P., & Gidycz, C.A. (1985). Sexual experiences survey: Reliability and validity. *Journal of Consulting Clinical Psychology, 53*, 422-423.

- Kress, V. E., Shepherd, J. B., Anderson, R. I., Petuch, A. J., Nolan, J. M., & Thiemke, D. (2006). Evaluation of the impact of a coeducational sexual assault prevention program on college students' rape myth attitudes. *Journal of College Counseling, 9*(2), 148-157.
- Langhinrichsen-Rohling, J., Foubert, J. D., Brasfield, H. M., Hill, B., & Shelley-Tremblay, S. (2011). The Men's program: Does it impact college men's self-reported bystander efficacy and willingness to intervene? *Violence Against Women, 17*(6), 743-759.
- Lonsway, K., Banyard, V., Berkowitz, A., Gidycz, C., Katz, J., Koss., Schewe, P., & Ullman, S. (2009). Rape prevention and risk reduction: Review of the research literature for practitioners. Special report. Washington, DC: National Institute of Justice and the Centers for Disease Control and Prevention.
- Morrison, S., Hardison, J., Mathew, A., & O'Neil, J. (2004). *An evidence-based review of sexual assault preventive intervention programmes*. Special report. Washington, DC: National Institute of Justice and the Centers for Disease Control and Prevention.
- Moynihan, M. M., Banyard, V. L., Arnold, J. S., Eckstein, R. P., & Stapleton, J. G. (2010). Engaging intercollegiate athletes in preventing and intervening in sexual and intimate partner violence. *Journal of American College Health, 59*(3), 197-204.
- O'Donohue, W., Yeater, E. A., & Fanetti, M. (2003). Rape prevention with college males: The roles of rape myth acceptance, victim empathy, and outcome expectancies. *Journal of Interpersonal Violence, 18*(5), 513-531.

- Paul, L., & Gray, M. (2011). Sexual assault programming on college campuses: Using social psychological belief and behaviour change principles to improve outcomes. *Trauma, Violence & Abuse, 12*(2), 99-109.
- 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 (IRMA) Scale. *Journal of Research in Personality, 33*, 27-68.
- Rau, T. J., Merrill, L. L., McWhorter, S. K., Stander, V. A., Thomsen, C. J., Dyslin, C. W., & Milner, J. S. (2010). Evaluation of a sexual assault education/prevention program for male U.S. navy personnel. *Military Medicine, 175*(6), 429-434.
- Rokeach, M. (1968). *Beliefs, attitudes and values: A theory of organisation and change*. San Francisco: Jossey-Bass.
- Stephens, K. A., & George, W. H. (2004). Effects of anti-rape video content on sexually coercive and noncoercive college men's attitudes and alcohol expectancies. *Journal of Applied Social Psychology, 34*(2), 402-416.
- Stephens, K. A., & George, W. H. (2009). Rape prevention with college men: Evaluating risk status. *Journal of Interpersonal Violence, 24*(6), 996-1013.
- Stone, J., & Cooper, J. (2001). A self-standards model of cognitive dissonance. *Journal of Experimental Social Psychology, 37*, 228-243.
- Tjaden, P., & Thoennes, N. (2006). *Extent, nature and consequences of rape victimization: Findings from the national violence against*

women survey. Special report. Washington, DC: National Institute of Justice and the Centers for Disease Control and Prevention.

Wells, G. A., Shea, B. O'Connell, D. Peterson, J. Welch, V., Losos, M., & Tugwell, P. (2010). The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomised studies in meta-analyses. Retrieved August 12, 2012, from http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

Journal Paper

Implicit Relational Assessment Procedure (IRAP): Can So-Called Implicit Rape-Supportive Beliefs be Restructured and Do They Predict Behaviour?*

Anna Brown¹, David Dawson², Nima Golanji-Moghaddam²

¹ Trent Doctorate in Clinical Psychology, University of Lincoln

² Research Tutor, Trent Doctorate in Clinical Psychology, University of Lincoln, UK

Abstract

Rape-supportive attitudes have been linked to sexual aggression, both theoretically and empirically, and form a central focus of most sex-offender interventions. Research into rape-supportive cognition has been hampered by an over-reliance on self-report measures, and associated socially-desirable responding. The Implicit Relational Assessment Procedure offers an alternative approach (IRAP; Barnes-Holmes et al., 2006). Derived from Relational Frame Theory (RFT, Hayes, Barnes-Holmes & Roche, 2001), here it is applied for the first time to study the malleability of implicit rape-supportive cognition, following a cognitive-restructuring task in university males. This sample forms a valid analogue, given the prevalence of sexual offending in this population. The relationship between implicit (IRAP) and explicit measures and behaviour was explored. IRAP scores did not change following the intervention. Behaviour within the forced-choice measure was predicted by IRAP scores but not by scores on the AMMSA. Findings are discussed in relation to future research, theoretical and clinical implications.

Keywords: implicit beliefs; sexual aggression; rape myths; Implicit Relational Assessment Procedure (IRAP); malleability.

* Prepared for submission to Archives of Sexual Behaviour

Background

Evidence provided within Home Office statistics suggests that approximately¹ 404,000 sexual offences² were committed against women aged 16 to 59 between 2012 and 2013. Such estimates were based on averages of reported crimes from 2009 to 2012, with full-time students being one of the highest at-risk groups: accounting for 6.8% of prevalence rates (Ministry of Justice, 2013). True base rates of sexual offending are notoriously difficult to ascertain due to the implications for victims reporting these crimes, such as re-traumatisation and disbelief (Yamawaki, Darby & Queiroz, 2007).

A wealth of literature points to the prevalence of sexually aggressive behaviour perpetrated by male undergraduates and college students in the United States (US, for reviews, see Murnen, Wright & Kaluzny, 2002; Schewe, 2002). However, there is a significant lack of research within this population in the UK. Stenning, Mitra-Kahn and Gunby, (2013) studied sexual violence against females in a UK university using an online survey ($N = 580$), focus groups and interviews. They found that since being at the university, 8% of women reported being the victims of sexual assault. Abbey, McAuslan, Zawacki, Clinton, and Buck (2001) reported in a US study that 33% of college males ($N = 342$) admitted to having perpetrated some form of sexual assault. Given the concerns around the social desirability of reporting sexual aggression (Malamuth, 1989), the possibility of minimisation (Langton et al., 2008) and cognitive distortions around past actions, these figures are likely to under-estimate prevalence.

¹ Based on figures of police reported crime and the Crime Survey for England and Wales (CSEW) in the last 12 months.

² Sexual Offences here include rape, attempted rape and sexual assault. Rape is defined as penetration by a penis of the vagina, anus or mouth of another person without consent. Sexual assault is an act of physical, psychological and emotional violation, in the form of a sexual act, which is inflicted on someone without consent.

Rape-supportive attitudes, often referred to as *rape-myths*, are defined in the literature as beliefs that serve to divert responsibility by, blaming the victim and exonerating the perpetrator (Burt, 1980; Gerger et al., 2007). [See Extended Background 1.0]. Suarez and Gadalla, (2010) reviewed 37 studies relating to the analysis of rape-myths and confirmed a strong positive correlation with sexual aggression. Although causality cannot be inferred from this, evidence has suggested the prevalence of rape-supportive beliefs in the lead up to the offence within a sample of convicted rapists (Polascheck & Gannon, 2004). However, data was drawn from retrospective self-reports and likely to be confounded by issues, such as, the reliability of memory. Helmus, Hanson, Babchishin and Mann (2013) conducted a meta-analysis on 46 sex-offender studies ($N = 13,782$), exploring the degree to which sexual offence-supportive attitudes predicted recidivism, and found a small but consistent relationship (Cohen's $d = .22$; Cohen, 1988). This indicates that sexual attitudes are associated with sexual aggression. Rape-myth acceptance (RMA) has been found to be higher within perpetrator samples, than those who have not been convicted or self-disclosed sexually aggressive behaviour (Abbey & Jacques-Tiura, 2011; DeGue & DeLillo, 2004; Field, 1978; Malamuth 1986; Murphy, Coleman & Haynes, 1986). Rape-supportive attitudes or *cognitive distortions* form a key component in models of sexual offending (Hall & Hirschman, 1991; Knight & Sims-Knight, 2003; Malamuth; 2003; Marshall & Barbaree, 1990; O'Ciardha & Ward, 2013; Polaschek & Gannon, 2004; Polaschek & Ward, 2002; Ward, 2000; Ward & Beech, 2006; Ward & Casey, 2010). Rape-supportive attitudes are also central to structured sexual offender risk-assessment tools (Thornton, 2002).

Difficulties with self-report measures, such as socially-desirable responding (Nosek, 2007), has led to an increased interest in so-called *implicit/indirect* measures. Such measures propose to measure automatic responses to the environment with reduced awareness and control (Gawronski & Payne, 2010). Examples include the Affective Priming Task (Fazio, Jackson, Dunton & Williams, 1995), the Implicit Association Test (IAT, Greenwald, McGhee & Schwartz, 1998) the Go/No-Go Association

Task (GNAT, Nosek & Banaji, 2001) and Extrinsic Affective Simon Task (EAST; De Houwer, 2003). Instead of relying on self-reports, and introspective accessibility (De Houwer, 2006) these measures compare the fluency (response times) in which individuals associate specific pairs of stimuli in relation to other pairs, and comparative fluency is purported to indicate the relative strength of the belief. [See Extended Background 1.1].

The IAT has been the most popular *indirect* measure, particularly within domains where social desirability is problematic, such as prejudice (Greenwald et al., 2002; Nosek, Greenwald & Banaji, 2005), and sexual offending (Brown, Gray & Snowden, 2009; Mihailides, Devilly, & Ward, 2004; Nunes, Firestones, & Baldwin, 2007). The IAT has been applied across many domains (Greenwald, Nosek, Banaji & Klauer, 2005). Nunes, Hermann and Ratcliffe (2013) used the IAT to examine whether biases in the speed in which individuals paired words such as *rape - good* and *not rape - bad* were related to self-reported sexual aggression in a sample of male university students. They found that implicit beliefs were significantly associated with greater levels of self-reported past sexual aggression; however, the IAT only provides a relative measure of strength of an association rather than the purported direction of the association. For example, the IAT does not determine whether responses indicate biases towards seeing rape as good, or seeing rape as neutral (but still better than 'not rape') it simply notes that there was a difference. [See Extended Background 1.2].

There is general consensus regarding the validity of implicit measures, and much of this evidence draws from IAT studies (Nosek, Hawkins & Frazier, 2011), as the IAT is the most commonly used measure in the area. However, a critical assumption borne out of the dominant stream of research into implicit attitudes is that psychological structures serve to moderate the outcome effects within the measurement procedure, and that the strength of that moderation provides an index of the strength of the attitude (Fazio, 2007; Greenwald et al., 2002; Rydell & McConnell, 2006; Strack & Deutsch, 2004). For example, faster responding on the IAT to

pairs of stimuli (e.g. *black* - negative) is assumed to mean they must be more easily categorised and thus strongly associated in memory than pairs that are responded to more slowly (e.g. *black* - positive). This is known as the *associative* paradigm. Hughes, Barnes-Holmes and De Houwer (2011) argue that this assumption is a theoretical one rather than an “immutable truism,” (p. 472) and note that this has led to the development of research procedures [e.g. GNAT; IAT; Implicit Association Procedure (IAP) Schnabel, Banse & Asendorpf, 2006] examining *how* associative structures operate in memory without testing the hypothesis that associative structures exist. [See Extended Background 1.3].

Research within the associative paradigm appeared to demonstrate that explicit attitudes were more malleable than implicit attitudes (Gawronski & Strack, 2004; Nosek, Banaji & Greenwald, 2002), and therefore assumed to be available for intervention (Bargh, 1999). This lack of malleability within implicit attitudes established the concept that implicit attitudes were highly stable associations across differing temporal and contextual parameters. However, as further research has been carried out using the IAT, examples of the sensitivity of implicit attitudes to the context in which they are measured has emerged (Feroni & Mayr, 2005), challenging the idea that implicit attitudes are highly stable associations in memory that are hard to change. [See Extended Background 1.4]. Associative researchers have argued that the malleability of implicit attitudes is actually a result of problems with the validity of measures themselves rather than a problem with the idea of stability in memory (Han, Czellar, Olson, & Fazio, 2010; Olson & Fazio, 2004). [See Extended Background 1.5].

The lack of specificity of beliefs indicated on the IAT poses problems in interpreting how attitudes have changed following an intervention, and prevents further understanding of the malleability of beliefs. For example, consider an IAT regarding attitudes towards women that shows changes in response-bias from pre- to post-intervention, in the direction of holding more positive views towards women. It would not be clear if attitudes towards women were initially neutral and became positive following the

intervention, or if the attitudes were initially negative and became neutral following the intervention.

An alternative theoretical approach to the conceptualisation and measurement of implicit cognition has emerged from the functional contextual paradigm,³ specifically from Relational Frame Theory (RFT; Hayes et al., 2001). RFT is a behavioural approach to the study of language and cognition and proposes that all verbal behaviour (cognition) is relational (Hayes et al., 2001). From RFT, a novel measurement procedure has been generated, the Implicit Relational Assessment Procedure (IRAP; Barnes-Holmes et al., 2006). The IRAP offers a relational assessment and so provides greater specificity of attitudes. For example, it provides information regarding the exact nature and direction of the belief, which is missing from previous methodologies, such as the IAT.

Related to RFT is the Relational Elaboration and Coherence (REC) model. The REC model offers a purely functional model of implicit cognition, completely replacing any notions of mental constructs or associations (see Barnes-Holmes et al., 2010). [See Extended Background 1.6]. From the REC perspective, attitudes are conceptualised as involving the acquisition of positive or negative evaluative functions based on arbitrarily applicable relational responding, or more simply, the behavioural consequences of the environment and behaviour interacting. Implicit and explicit responding is conceptualised as a single process and as such, behaviour is viewed on a continuum. The way in which behaviours diverge on indirect and direct measures is thought to be a function of time and accuracy, rather than relating to a dual process. Pressure to respond accurately under time-constraints, such as within indirect procedures, means brief implicit responses are more probable. Alternatively, within direct measures, which are thought to allow time for relational networks to unfold, elaborated explicit responses are more probable.

³ Functional contextualism is a philosophy of science, rooted in pragmatism and contextualism. It is actively applied in the field of behaviour analysis and emphasises the importance of predicting and influencing psychological events using empirical concepts and rules. Knowledge constructed within this paradigm is general, abstract and spatiotemporally unrestricted (Fox, 2006).

The REC model explains divergence between implicit (indirect) and explicit (direct) measures by stating that brief, immediate evaluative responses may or may not correspond with relational responding. When they correspond, indirect and direct measures converge and when they do not, they diverge; it is thought individuals discount their brief relational responses if they do not correspond with elaborate and extended relational responding (Barnes-Holmes et al., 2010). Regarding areas of social sensitivity, the REC model predicts constraints of time-pressure will lead to divergence, as individuals have less time to contact elaborated responses.

The REC model is developing and there are difficulties in distinguishing it from dual-process explanations on the basis of current evidence. Nevertheless, the REC model offers a coherent account of the convergence and divergence of behavioural patterns that can be observed within both direct and indirect measures, and an explanation regarding how they predict different types of behavioural outcomes (Barnes-Holmes et al., 2010). In contrast to associative paradigms, the REC model views behaviour on indirect measures, as being highly related to the context in which it is obtained. The model assumes that the IRAP effect, produced under time-pressure, is driven predominantly by immediate and relatively brief relational responses, whereas explicit measures reflect extended and coherent relational networks (Barnes-Holmes et al., 2010). Specifically, the IRAP is thought to capture brief implicit responses and requires that individuals relate stimuli (words) directly, by quickly and accurately confirming or denying relations between a stimulus and a target word. The output consists of the speed in which individuals relate pairs of stimuli. For example, within some trials, participants are instructed to respond to stimuli in a way that is consistent with their prior history of relating to similar stimuli (i.e. a relation of co-ordination) and on other trials, they have to respond in way that is inconsistent with their past history of relating such stimuli. Response time is considered to be made up of two distinct parts; the brief and immediate relational response, and time-taken to press the response key. Under time-pressure brief implicit responses become objectively measurable. Although, the degree of time-pressure considered

sufficient for brief relational responding to occur is debated (Barnes-Holmes et al., 2010). The difference between time taken to respond to consistent and inconsistent trials is defined as the IRAP effect and is hypothesised to indicate the strength of relational responses (Barnes-Holmes et al., 2010). It is expected that individual responses should be completed more quickly on consistent trials than inconsistent trials because; during consistent trials they are thought to be responding in-line with their most probable relational responses.

Current treatment programmes for individuals convicted of rape (Gannon, Collie, Ward, & Thakker, 2008; Hanson, Bourgon, Helmus & Hodgson, 2009) and rape-prevention programmes within universities (for review see Anderson & Whiston, 2005) predominantly adopt cognitive-behavioural techniques such as cognitive-restructuring to effect attitudinal change, with the intention that the change in attitudes will correspond with changes in behaviour. [See Extended Background 1.7]. However, outcome measures predominantly rely on self-report questionnaires (elaborated explicit responses); little is known about the influence of these techniques on brief implicit responses and the degree to which rape-supportive brief implicit responses can predict different types of behaviour, in comparison with elaborated explicit responses. This study aims to begin to develop greater awareness of these important areas and crucially, the IRAP is not easily faked (McKenna, Barnes-Holmes, Barnes-Holmes & Stewart, 2007), which is pertinent in socially-sensitive areas such as sexual aggression.

Importantly, Cullen, Barnes-Holmes, Barnes-Holmes and Stewart (2009) demonstrated the utility of the IRAP as a measure of attitude change following exemplar training to reverse anti-old bias. Similarly, Hussey and Barnes-Holmes (2012) demonstrated the ability of the IRAP to detect change in depressive emotional reactions following a sad mood-induction procedure, in individuals with mild/moderate depressive symptoms.

Only one study has used the IRAP to investigate the effects of clinical treatment techniques on responding. Hooper, Villate, Neofotistou, and McHugh (2010) investigated the effects of a mindfulness intervention on

acceptance of negative emotions, within university students. They randomly assigned participants to either a mindfulness intervention group or a thought-suppression (TS) group. The mindfulness group were found to have a reduction in experiential avoidance from pre- to post-intervention in comparison to those in the TS group on the IRAP, but not the explicit measure. No studies to-date have used the IRAP to examine rape-supportive verbal relational responding in this important domain. Notably, no studies have explored the utility of indirect measures in predicting behaviour directly (within the domain of rape-supportive attitudes), as opposed to self-reported behaviour (see Nunes et al., 2013), which is itself under contextual control such as demand characteristics and social desirability. Given that brief implicit responses have predictive utility regarding clinical behaviour such as relapse (Steinberg, Karpinski, & Alloy, 2007) and attempts at suicide (Nock et al., 2010), it is pertinent to investigate the malleability of brief implicit responses within this clinically-relevant domain, as it may have important implications for the shaping of intervention programmes and dynamic risk-assessment tools.

The primary aim of this study was to therefore, investigate the malleability of rape-supportive brief implicit responses, by measuring them using the IRAP before and after a brief cognitive-restructuring task. A control condition was included (see Materials). It was hypothesised that brief implicit responses would be sensitive to the intervention and so rape-supportive bias, as measured by the IRAP, would be reduced at post-intervention testing.

The second aim was to explore the predictive utility of brief implicit responses in comparison to elaborated explicit responses, in relation to overt behaviour (charity donation and ratings of the researcher). The measure used to capture, what is considered to be elaborated explicit responses, was the Acceptance of Modern Myths About Sexual Aggression Scale (AMMSA). It was hypothesised that the baseline IRAP would predict behaviour (as measured by charity donation), for example, strong rape-supportive bias was thought to predict fewer tokens being placed in a

female-victim charity-box relative to the male-victim charity-box. Another hypothesis was that the explicit measure, (the AMMSA) would predict charity-box behaviour but to a lesser degree than the implicit measure. The AMMSA would be potentially tapping into elaborated explicit responses. However, the charity-box measure was considered to be tapping into behaviour falling more closely towards the brief end of a continuum, with brief responses on one end and elaborate responses on the other. Furthermore, it was hypothesised that the baseline IRAP would predict ratings of the female researcher; it was thought that strong rape-supportive bias would predict negative ratings of the researcher. It was hypothesised that the AMMSA would show some degree of predictive ability but to a lesser degree than the IRAP. [See Extended Background 1.8].

Method

Participants

The study was comprised of two phases. Phase one was conducted online to obtain a normal population baseline (AMMSA scores) with which to compare with the group that participated in the main study. Phase two consisted of the experimental intervention design, and participants for this phase were recruited through expression of interest in phase one.

Phase one. Participants ($N = 143$) were recruited from two UK university sites using email networks and digital recruitment methods at the universities to both staff and students (daily alerts, twitter, and targeted emails to departments). Participants were invited to take part in an online questionnaire investigating attitudes towards sex. The inclusion criteria were that participants were male, heterosexual, with English as a first language and normal or corrected-to-normal vision. Table 3 presents the demographic information.

Table 3

The Demographic Characteristics (Age) of Participants

Phase	Group	Mean Age (<i>Years</i>)	<i>SD</i>	Range (<i>Years</i>)
One (<i>N</i> = 143)	–	27.48	9.85	18 – 65
Two (<i>N</i> = 37)	Control (<i>n</i> = 17)	32.12	13.45	19 – 61
	Intervention (<i>n</i> = 20)	28.40	8.65	19 – 48

Phase two. Participants who agreed to complete phase two of the study comprised *N* = 39. The data from two participants were removed due to distractions occurring during the IRAP trials, rendering their responses invalid. [See Extended Method 2.1].

Materials and Apparatus

The Acceptance of Modern Myths of Sexual Aggression (AMMSA). The AMMSA (Gerger et al., 2007) was provided online for participants to complete. The AMMSA is a 30-item self-report scale, measuring adherence to myths relating to sexual aggression. Participants rate their agreement with statements such as: “When a man urges his female partner to have sex, this cannot be called rape,” using a seven point Likert scale ranging from 1 (totally disagree) to 7 (totally agree). High-scores indicate adherence to myths relating to sexual aggression.

The English version of the AMMSA has good internal consistency, with a Cronbach alpha coefficient reported of .92 (Gerger et al., 2007). In the present study, the Cronbach alpha coefficient was .93. The AMMSA has been shown to have high internal consistency (.92; Gerger et al., 2007) and good construct validity (.80 convergent validity with Illinois Rape Myth Acceptance Scale; Payne, Lonsway & Fitzgerald, 1999). [See Extended Method 2.2].

Charity-box task. The charities consisted of one which provided support for female victims of domestic violence including sexual and physical abuse (Women's Refuge), one which provided support for male victims of sexual abuse (The Blue Silence Foundation), and two charities which were deemed unrelated to interpersonal abuse (The Wildlife Foundation, and The Natural World Conservation Society). [See Extended Method 2.3].

Researcher rating scale. The Researcher Rating Scale (RRS) consisted of a Likert-type rating scale with six questions, designed to elicit participants' verbal behaviour relating to their judgement of several attributes of the female researcher: Knowledge; friendliness; competence; intelligence; approachability; and warmth (see Appendix b). Scores ranged from 1 (not at all) to 5 (very much). [See Extended Method 2.4].

Intervention and control conditions. Evidence suggests that men that have not been convicted of sexual offences are more likely to be receptive to cognitive-restructuring around rape-supportive cognition if they are not addressed as potential perpetrators of rape, as this can increase defensiveness and reduce engagement (Langhinrichsen-Rohling, Foubert, Brasfield, Hill, & Shelley-Tremblay, 2011; Scheel, Johnson, Schneider & Smith, 2001). The cognitive-restructuring intervention was designed to place participants in the role of helper rather than potential perpetrator. The intervention condition consisted of participants listening alone to four computerised audio-clips, lasting ten minutes in total, depicting a male student verbalising rape-supportive attitudes at different points during a fictional night out. The script was designed to map onto the rape-supportive attitudes within the AMMSA. Participants were to write down alternative statements that would challenge the student's beliefs (see Procedure and Appendix c). [See Extended Method 2.5]. From a REC perspective, the cognitive-restructuring task was deemed to be combining networks of

relations together in an additive manner for the purposes of this event; namely, the investigation of malleability.

The control condition involved participants listening alone to four computerised audio-clips, lasting a total of ten minutes, of the same actor verbalising his thoughts regarding a forthcoming job interview. Participants were required to write down alternative statements which would challenge the interviewee's beliefs about the job interview (see Appendix d). The written statements provided by the participants were not included in the analysis, other than to check if participants engaged in the task. Both the control group task and the intervention task lasted approximately 20 minutes each (please contact the authors for further details).

IRAP. The IRAP computer package was completed on a separate computer. The IRAP stimulus set was developed to reflect the rape-myths targeted within the AMMSA by the first author and was validated by a clinical psychologist and researcher working within the field. As it is a first study, stimuli were chosen to reflect a broad range of rape-myths. The final choice of stimuli were based on pilot testing and effect sizes from previous research using the IRAP (e.g. controlling for word length is unnecessary; each target word acts as its own control). Two category labels *Women are* and *Women are not* were used with two series of target stimuli, one series of terms that describe women as decent (e.g., *Honest*), and a series of semantically opposite terms (e.g., *Deceitful*). (See Table 4).

The IRAP included equal blocks of consistent and inconsistent trials. During what is termed a *consistent* trial, participants were required to confirm that women are, effectively, good people (e.g., *Honest; Faithful*) and during *inconsistent* trials confirm women are not (e.g., *Deceitful; Slutty*). Category labels in the IRAP act as controls for one another. Differences between response latencies generated from consistent and inconsistent trials make up the *D-IRAP* score (see Results, Indirect Measure: Scoring the IRAP).

The IRAP has enabled discrimination between known groups, albeit to varying degrees (Vahey, Barnes-Holmes, Barnes-Holmes & Stewart, 2009; Dawson, et al., 2009; Drake et al, 2010; Hussey & Barnes-Holmes, 2012) and is comparable with other measures of implicit cognition such as the IAT (Snowden et al., 2011). The IRAP has demonstrated a reasonable split-half reliability, with a Spearman-Brown coefficient of .72 (Barnes-Holmes, Murtagh, Barnes-Holmes, & Stewart, 2010). In the current study, the Spearman-Brown coefficient for the pre-intervention IRAP was .70 and for the post-intervention IRAP .71. [See Golijani-Moghaddam, Hart and Dawson, (2013), for a recent review]. [See Extended Method 2.6].

Table 4

The Implicit Relational Assessment Procedure Stimulus Set

Sample 1: Women are	Sample 2: Women are not
Relational Term 1: True	Relational Term 2: False
<i>Target words consistent with a positive view of women</i>	<i>Target words consistent with a negative view of women</i>
Honest	Deceitful
Equal	Subordinate
Faithful	Slutty
Direct	Teasers
Truthful	Liars
Victims in rape	Deserving of rape
Damaged by rape	Glad to be raped
Strong	Weak

Procedure

Pilots. Expert consensus was sought from a clinical psychologist and researcher in the field to check the relevance of the IRAP stimulus set in relation to rape-myths and how closely the stimuli mapped onto the AMMSA. A pilot was conducted in order to test the utility of the IRAP

stimulus set, and the IRAP instructions. Feedback indicated the stimuli were understood, and that the length and complexity of target words were feasible in an IRAP with a response time cut off equalling 2500ms and an accuracy target of at least 75%. The intervention and control conditions were then piloted to ensure instructions were sufficient, and the tasks could be completed as specified. Piloting of the charity-box task indicated that tokens were disproportionately decanted into the charity-box representing support for victims of natural disasters. The charity type was subsequently changed to represent a less emotive cause.

Phase one. Ethical approval for both phases of the study was granted from the University ethics boards. Participants were invited to participate in a survey examining attitudes to sex accessed via a link that led participants to an information sheet explaining: the details of the study, their right to withdraw, anonymity, and that they would be entered into a prize draw to win fifty pounds (see Appendix e). Participants indicated consent by ticking an online box (see Appendix f) before completing the AMMSA (Gerger et al., 2007). Participants generated a unique identifier code. Once the AMMSA had been completed (approximately 15 minutes) participants followed a link to a webpage to leave their email address in order to be contacted with potential prize money and to enable phase two to be arranged.

Phase two. Participants that replied to an email from the researcher, stating they were interested in completing phase two, were randomly allocated to the control or intervention condition (using software from *random.org*). Participants were given a consent form for phase two to sign before completing the first IRAP (hereafter referred to as *Pre-IRAP*), explaining their right to withdraw, and anonymity (see Appendix g). Phase two consisted of five stages: Pre-IRAP; intervention/control condition; second IRAP (hereafter referred to as *Post-IRAP*); researcher rating scale; and charity-box task. Each participant completed phase two alone.

Pre-IRAP. All participants completed the IRAP first and were given standardised instructions verbally from the researcher, describing the screen layout, how to complete the task, whilst highlighting the importance of speed and accuracy. Participants were informed they would be required to respond either in accordance or opposition with their beliefs. In order to reduce the likelihood of random responding and to confirm participants understood the task, participants each completed up to four practice blocks to ensure mean response time was less than 2500ms, with an accuracy rating above 75%. If participants did not achieve the specified criteria, feedback was presented on the screen at the end of the practice-block, highlighting speed, accuracy and instructing participants to meet the criteria. The researcher was seated next to the participant during practice-blocks but participants completed the test-blocks alone.

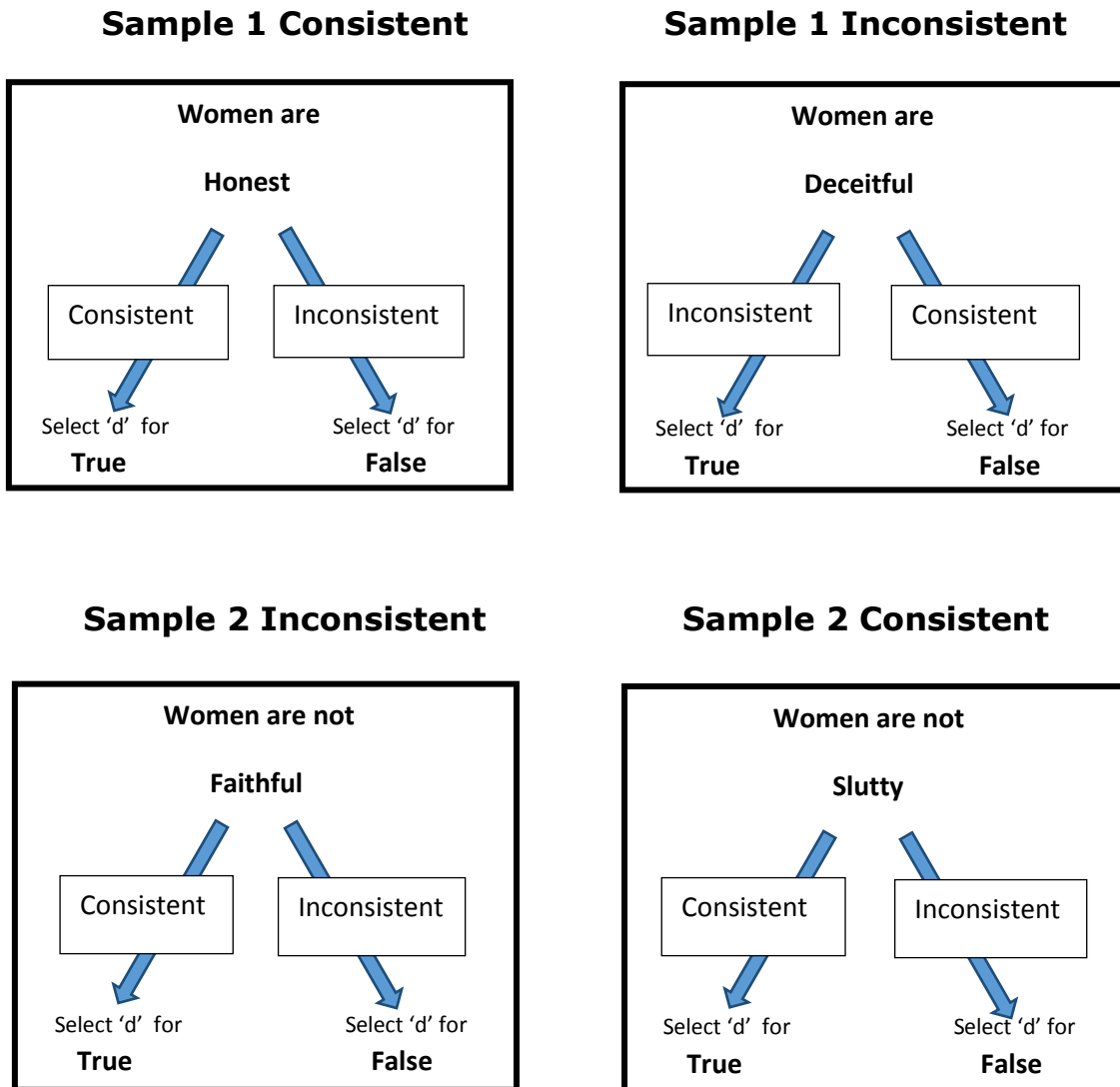
The IRAP consisted of six blocks of trials, with 24 trials per block. Simultaneously, within each trial a category label such as *Women are* appeared at the top of the screen, with 1 of 12 target words underneath (e.g., *Deceitful*), and the response terms *True* and *False* in each bottom corner (see Figure 2). Participants were to select the *D* key on the keyboard for the relational term *True* and the *K* key for *False*. If the relational term selected was incorrect, a red *X* appeared in the middle of the screen and was only removed once the correct response was selected. Once a correct response had been selected the stimuli disappeared from the screen for 400ms before a new trial was presented. Participants had to respond to the stimuli (a combination of category label and target word) in line with social norms during consistent blocks (e.g., *Women are - Honest - True*; *Women are not - Honest - False*) and in opposition to social norms during inconsistent blocks (e.g., *Women are - Honest - False*; *Women are not - Honest - True*). Table 4 displays which combinations were consistent and inconsistent.

Test-blocks (of 24 trials each) alternated between consistent and inconsistent, starting with a consistent block, resulting in a total of three consistent and three inconsistent blocks. During each block every target

word was presented twice with each category label; the pairings of which were randomly assigned. On-screen instructions at the start of each block informed participants how to respond (i.e., whether it was a consistent or inconsistent block). At the end of the IRAP, instructions appeared on a blue screen informing participants to report to the researcher. The pre-IRAP lasted approximately ten minutes.

Figure 2

Examples of the Four Implicit Relational Assessment Procedure Trial-types



Note: The arrows and text boxes denoting *Consistent* and *Inconsistent* were not presented to participants but denote for the purposes of this illustration responses consistent or inconsistent with social norms. Correct responses varied depending on whether the block was consistent or inconsistent.

Intervention and control conditions. Participants were randomly allocated to one of two conditions: intervention and control (see *Materials* for a full description). Audio-clips were generated using *Psycho-Py* software (Peirce, 2009) and played from a separate laptop. Participants were given

four-page hand-outs, and the researcher gave verbal instructions relating to the task. Participants were instructed to imagine the actor was their friend and if they did not intervene, their friend might commit rape later that night. Participants were invited to listen to each of the four clips and then write down alternative statements after each audio clip on the hand-out, in order to challenge their friend's cognitions, with the intention of changing his rape-supportive attitudes.

The control condition involved participants listening to four audio-clips lasting a total of ten minutes, of the same actor verbalising his thoughts regarding a forthcoming job interview. Participants were required to listen to one audio-clip and write down alternative statements, designed to challenge the actor's beliefs about the job interview, before starting the next audio-clip. Participants were informed that if no advice was given, their *friend* would not get the job. The written advice provided by participants was not included in the analysis, other than to check if participants engaged in the task. Both the control and intervention task lasted twenty-minutes. The researcher was not present.

Post-IRAP. The second IRAP was administered immediately after the intervention/control condition, using the same stimulus set and procedure as used in the pre-IRAP, including practice blocks. Participants were told the study had ended and received five pounds reimbursement for their time and two research credits, if required.

RRS. The RRS was given to participants to complete following the post-IRAP. Participants were informed the purpose of the questionnaire was to evaluate the researcher for use at an annual performance review. An envelope was given to participants and they were informed that the researcher would not see it; they were asked to put the completed RRS inside and seal it before returning it. The researcher remained in the room,

so it did not appear to be part of the experiment, but stood far enough away to enable privacy. Completion time was approximately one minute.

Charity-box task. Participants were informed the research was sponsored by a social-enterprise body and that, as part of the sponsorship, they were given three tokens to place in any of the four charity boxes located on another table in the room. Three tokens were given, instead of one or four, in order to increase the base-rate of responses in any particular box to make detection of statistically significant difference in responding more likely. Thus reducing the likelihood of making a Type II error. Participants were informed that the social enterprise would match the tokens each charity received in monetary value. The researcher remained in the room, but stood away from the participant to reduce likelihood of socially-desirable responding.

Debrief. Following completion of the study, participants were provided with a sealed envelope containing a debrief sheet explaining the true nature of the study, the elements of deception (i.e., regarding the RRS and the charity-box task), contact details of the researchers and support agencies (see Appendix h). [See Extended Method 2.7 and 2.8].

Results

Direct Measure

To assess a normal population baseline, in phase one, the AMMSA was scored by averaging the 30-items. Higher scores indicated greater adherence to rape-myths. For example, a high score denoted greater agreement with a statement such as: "When a woman starts a relationship

with a man, she must be aware that the man will assert his right to have sex." The group overall produced a mean score ($N = 143$, $M = 2.97$, $SD = .90$) which is lower than the mean reported for males in the previous study (Gerger et al., 2007) using the English version of the scale, $N = 148$, $M = 3.60$, $SD = .98$.

AMMSA Differences Between Groups

A Mann-Whitney U test was conducted to compare the AMMSA scores for those that completed phase one only with those that completed phase two, as the distribution of data was not normally distributed in the phase two sample. This was completed to check the comparability of the sample in terms of bias in reported adherence to myths of sexual aggression with the wider university male population. There was a significant difference in scores, with those completing phase one producing higher scores ($Md = 3.15$, $n = 104$) than those completing phase two, ($Md = 2.47$, $n = 39$), $U = 1481.5$, $z = -2.48$, $p = .01$, $r = -.21$. This effect size is considered to be small to medium (Cohen, 1988). Therefore, participants who attended phase two reported significantly less adherence to rape-myths than those completing phase one only, meaning those completing phase two were less biased towards denying or trivialising sexual aggression towards females. [See Extended Results 3.1].

Indirect Measure: Scoring the IRAP

The main data produced by the IRAP programme are raw latency scores consisting of elapsed time (milliseconds) between presentation of the stimulus in the IRAP trial and the accurate response elicited by the participant. Raw latency scores were transformed into standardised difference scores, (D -IRAP scores), using an adaptation of Greenwald, Nosek, and Banaji's (2003) D -algorithm. [See Extended Results]. This minimises the effects of confounding factors such as cognitive ability and motor skills (Barnes-Holmes, Waldron, Barnes-Holmes & Stewart, 2009;

Greenwald et al., 2003), in accordance with research in the field of implicit measurement (Barnes-Holmes et al., 2010; Cullen et al., 2009; Hussey & Barnes-Holmes, 2012).

Raw latency scores were transformed into five *D*-IRAP scores; one for each of the trial-types (e.g. *women are honest*, *women are deceitful*; *women are not honest*; *women are not deceitful*) and an overall *D*-IRAP score (the mean of the scores from the four trial-types). Boxplots were generated to check for outliers and one extreme score was transformed to three standard deviations above the mean (Field, 2009), rather than removing the data, in order to preserve statistical power.

A larger *D*-IRAP score signifies a greater difference in response latencies between consistent (with social norms, i.e., *women are honest*, *women are not deceitful*) and inconsistent trials (with social norms, i.e., *women are deceitful*, *women are not honest*). A positive *D*-IRAP score indicates participants are faster at confirming rather than denying women are honest, faithful, and direct, for example. In contrast, a negative *D*-IRAP score denotes participants are faster at confirming as opposed to denying women are deceitful, slutty, and teasers. Scores around zero signify there is no differentiation between scores produced from consistent and inconsistent trials.

IRAP Results

The *D*-IRAP scores for the group as a whole ($N = 37$), calculated from the Pre-IRAP, were in a positive direction ($M = .25$, $SD = .26$), indicating a general bias in the expected direction towards viewing women as honest and faithful. The results for three of the four individual trial-types indicated a similar bias (see Table 5), however very little differentiation was observed for the trial-type *women are deceitful* ($M = .01$, $SD = .38$), meaning participants neither confirmed nor denied that women were deceitful.

Table 5

D-Implicit Relational Assessment Procedure Scores (D-IRAP): Individual Trial-types (Pre-Experimental/Control Group)

Trial-type	Mean <i>D</i> -IRAP (<i>N</i> = 37)
Women are honest	.51 (<i>SD</i> .40)
Women are deceitful	.01 (<i>SD</i> .38)
Women are not honest	.25 (<i>SD</i> .38)
Women are not deceitful	.24 (<i>SD</i> .43)

Relationship between Indirect and Direct Measures

The relationship between implicit and explicit rape-supportive beliefs (as measured by the pre-experimental/control *D*-IRAP scores and AMMSA scores respectively) was investigated using Pearson product-moment correlation coefficient. Slight divergence was noted between implicit and explicit measures but did not reach significance, $r = -.12$, $n = 37$, $p = .48$, bootstrapped 95% CI: $-.38$ to $-.18$. Individual Pre-IRAP trials, again, highlighted some relationship with the explicit measure but this did not reach significance: Women are honest, $r = -.19$, $n = 37$, $p = .25$, bootstrapped 95% CI: $-.48$ to $.19$; Women are deceitful, $r = -.07$, $n = 37$, $p = .68$, bootstrapped 95% CI: $-.38$ to $.23$; Women are not honest, $r = -.27$, $n = 37$, $p = .11$, bootstrapped 95% CI: $-.52$ to $.03$; and Women are not deceitful, $r = .19$, $n = 37$, $p = .27$, bootstrapped 95% CI: $-.06$ to $.46$. [See Extended Results 3.2].

Malleability of Implicit Beliefs

To check randomisation had been successful, an independent-samples *t*-test was conducted to compare the mean Pre-*D*-IRAP scores between those in the intervention group and the control group. There was no significant

difference in scores between those in the intervention group ($M = .21$, $SD = .26$) and the control group, $M = .30$, $SD = .26$; $t(35) = 1.07$, $p = .29$ (two-tailed), indicating randomisation had been successful. The magnitude of the differences in the means (mean difference = .09, 95% CI: -.08 to .27) was small (Cohen, 1988; eta squared = .03). [See Extended Results 3.3 - 3.10].

Engagement with task. Checking of the written responses indicated participants demonstrated engagement with the task, albeit to differing degrees. For example, within the control condition participants' responses ranged from: "*You are as nervous as anyone would be, you shouldn't worry as much;*" to "*Focus on times where you were nervous before but were still successful.*" Within the intervention condition responses ranged from: "*Having sex against a woman's will, can harm them in ways we don't understand,*" to "*She will be traumatised for the rest of her life. Your behaviour will also destroy the lives of those around her and perpetuate women's distrust of men.*"

Mixed between-within analyses of variance (ANOVAs) were conducted to assess the impact of the group type (intervention versus control) on participants' implicit beliefs (mean *D*-IRAP scores) across two time points (pre- and post-intervention). Assumptions for parametric testing were met, specifically homogeneity of inter-correlations, measurement at ratio level, normal distribution, equality of error variance, and independence of observations. [See Extended Results: 3.11 - 3.12]. Models were equivalent in finding no significant main effects for group ($F = .13 - .75$, $p = .39 - .72$, $\eta_p^2 = .004 - .02$) or time ($F = 2.96 - .09$, $p = .09 - .76$, $\eta_p^2 = .003 - .08$) and no significant interaction effects for group*time ($F = .07 - 3.12$, $p = .09 - .80$, $\eta_p^2 = .002 - .80$), suggesting no difference in the effectiveness of the intervention in comparison to the control group (see Table 6 & Figure 3).

Table 6

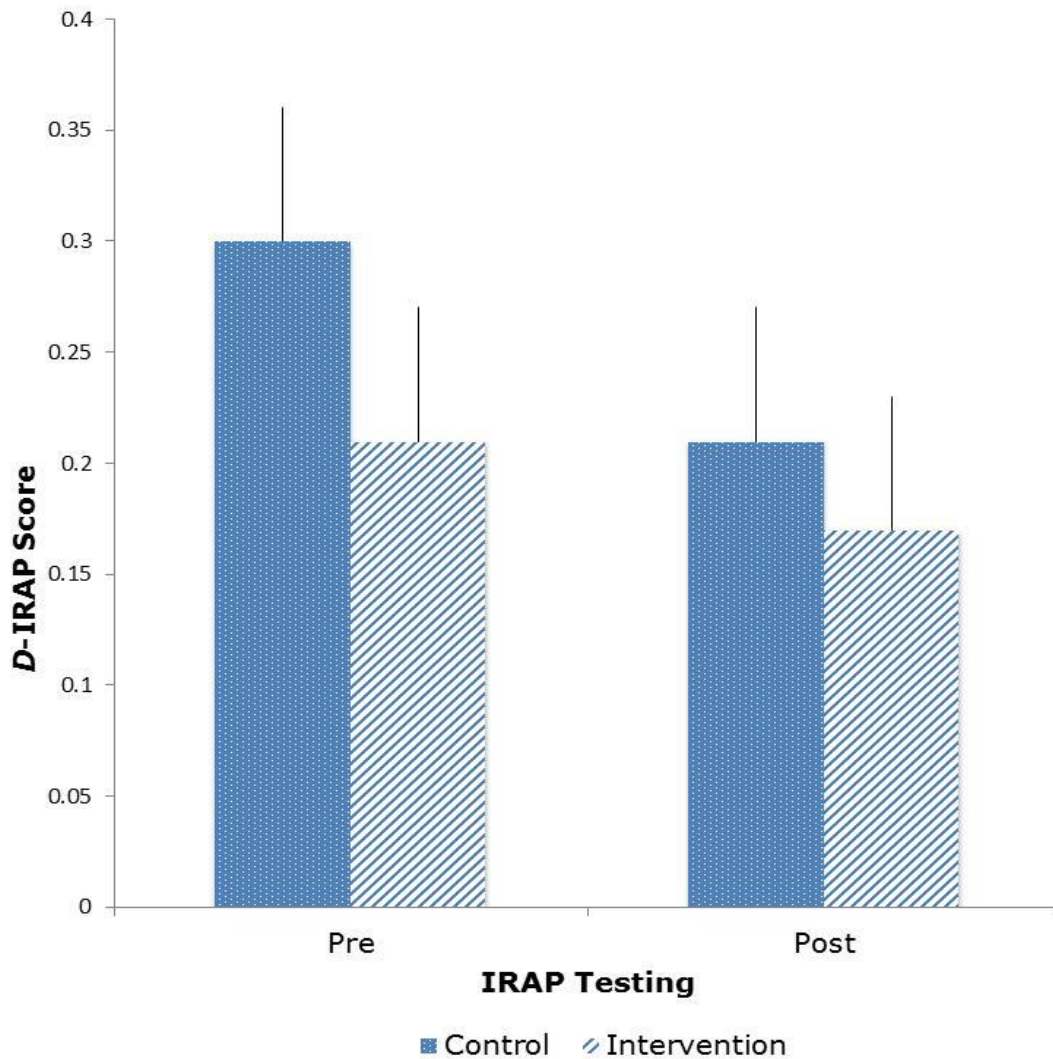
Mixed Between-Within Analysis of Variance by IRAP Trial-type

Trial-type	Λ	F	P	η_p^2
Women are honest (Consistent)				
Interaction effect: group*time	1.00	.07	.80	.002
Main effect: time	.95	1.74	.20	.05
Main effect: group		.75	.39	.02
Women are deceitful (Inconsistent)				
Interaction effect: group*time	.92	3.12	.09	.08
Main effect: time	.95	1.75	.19	.05
Main effect: group		.45	.51	.01
Women are not honest (Inconsistent)				
Interaction effect: group*time	1.00	.11	.74	.003
Main effect: time	.92	2.96	.09	.08
Main effect: group		.13	.72	.004
Women are not deceitful (Consistent)				
Interaction effect: group*time	1.00	.07	.79	.002
Main effect: time	1.00	.09	.76	.003
Main effect: group		.20	.66	.006

Note. Λ = Wilks Lambda.

Figure 3

Rape-Supportive Verbal Relations Bias Pre- and Post- Intervention/Control
(Overall Mean D-IRAP Effect Score)



Bars represent standard errors

Relationship between the Indirect Measure and Behaviour

Scoring the Researcher Rating Scale (RRS) consisted of totalling six individual-item scores and obtaining an overall mean score for each participant. Scoring the charity-box task consisted of summing the tokens in the *Blue Silence Foundation* box and the *Women's Refuge* box separately to

obtain absolute values of giving to each charity. In order to develop a single index of giving relative to *The Blue Silence Foundation* and *Women’s Refuge*, tokens donated to the former were subtracted from the latter. The relationship between implicit rape-supportive beliefs (as measured by the IRAP) and behaviour (measured by the RRS and charity-box task) was investigated using Pearson product-moment correlation coefficient. Bootstrapping was applied to the data to ensure robustness given the uneven distribution of the behavioural measures (Field, Miles, & Field, 2012; Wright, London & Field, 2011). [See Extended Results 3.9].

Measure of overall giving. A significant correlation was found between mean Pre-D-IRAP scores and overall *relative giving* score, and between the Women are honest trial-type and overall giving score (see Table 7), meaning that the more the participant endorsed implicit rape myths, the more likely participants were to give a greater proportion of tokens to *The Blue Silence Foundation* in comparison to *Women’s Refuge*.

Table 7

Significant Correlations Between Mean IRAP and Measure of Overall Giving

Trial-type	<i>R</i>	<i>P</i>	bootstrapped CI
Pre-D-IRAP	.33	.04	.07 to .54
Woman are not honest (Inconsistent)	.45	≤.01	.16 to .70

RRS. A medium negative correlation was found between Women are not honest and participant’s ratings of the researcher’s degree of approachability (RRS: question four) and both Women are deceitful and Women are not honest and RRS: question four, meaning the more rape-

supportive bias, the more approachable they perceived the researcher (See Table 8).

Table 8

Significant Correlations between Mean IRAP and Researcher Rating Scale

Researcher rating scale			
Trial-type	<i>R</i>	<i>P</i>	bootstrapped 95% CI
Women are not honest (Inconsistent)	-.38	.02	-.64 to -.15
Woman are deceitful (Inconsistent)	-.33	.04	-.65 to -.03

The pre-IRAP data were used to investigate the strength of relationship between implicit beliefs and behaviour as opposed to the post-IRAP data, as the intervention condition was deemed to create noise in the data, thus confounding the post-IRAP data. Cullen et al. (2009) reported test-retest reliability of .49, whilst comparable with the IAT (.56, Nosek, Greenwald, & Banaji, 2007), the findings were subject to the effects of active intervention and so caution must be applied. The test re-test reliability in the current study was $r = .51$, $n = 37$, $p \leq .01$ (.53 intervention group; .49 control group). Golijani-Moghaddam et al. (2013) reviewed the emerging reliability and validity data for the IRAP and concluded that when comparing test re-test IRAP data, intervention effects were likely to lead to underestimation of stability across testing. Given the lack of reliability of the IRAP at post-measurement following intervention (Golijani-Moghaddam et al, 2013), all significant relationships between pre-IRAP scores and behavioural measures dropped out at post-IRAP due to noise in the data and potentially the intervention not working.

Relationship between the Direct Measure and Behavioural Outcomes

In contrast to some of the findings relating to the indirect measure, the AMMSA did not significantly correlate with any of the behavioural outcome measures.

Discussion

Against predictions, rape-supportive brief implicit responses were not significantly reduced following intervention. In fact, there was a general trend in both groups to respond on the post-IRAP in accordance with an increased rape-supportive bias, although this increase was less pronounced for the intervention group, albeit not reaching statistical significance. From a REC perspective, given that brief implicit responses are viewed as highly sensitive to context-manipulation, it appears that the cognitive-restructuring intervention did not change these relational responses as hoped for, but potentially changed elaborated explicit responses, which are unlikely to be captured by the IRAP. However, we cannot infer this from the current design. Brief implicit responses are considered to be shaped by elaborate explicit responses and vice-versa, so it could be argued that change would be expected in brief implicit responses but to a lesser degree than the elaborated explicit responses. A limitation of this study was not including a post-intervention measure of the AMMSA. This might have captured any change in elaborated explicit responses, however, this was not an initial aim of the study (and administration was deemed to increase participant fatigue).

Similarly, the intervention may have changed rape-supportive brief implicit responses but the IRAP did not demonstrate sufficient reliability over time to capture this. The test-retest reliability of the current study was .51. Whilst responsivity to change is desirable in treatment-outcome measures,

high reliability is desired in order to reduce error-variance when interpreting treatment outcome. A difficulty with a sensitive measure is that it is less reliable at test-retest. Only one other IRAP study has reported test re-test reliability which was found to be .49 (Cullen et al., 2009). More research is needed to establish greater confidence in its test re-test reliability. However, intervention studies manipulate the context and so this may underestimate test re-test reliability, as change between pre- and post-intervention is likely due to contextual manipulation, so it will be measuring that effect rather than the reliability of the measure (Golijani-Moghaddam et al., 2013). Completing the IRAP changes the context and so implicit responding is likely to change. Future research might be improved by shortening the response-latency criteria still further.

Findings might also mean that the intervention did not work as designed; whilst engagement with the task was demonstrated by examining participant responses, the degree of engagement is likely to have varied between participants. One explanation might be that more biased individuals engaged less with the task in the intervention group (due to agreement with the actor's rape-supportive statements) than individuals that disagreed completely with the actor's views. The task involved perspective taking skills, which may have varied across individuals, thus confounding the results. Whilst cognitive-restructuring is commonly applied by asking clients to consider what they would say to a friend that was having a particular (unhelpful) thought, there may be a lack of effect following intervention as the rape-supportive beliefs being targeted were not specific to those endorsed by each participant.

Lai et al. (2014) studied the effectiveness of 17 interventions in reducing implicit racial preferences and found that interventions that engaged participants with other's perspectives were ineffective. They found that providing participants with counter-stereotypical exemplars, putting the participant into the story over longer periods including heightened vividness, and using evaluative conditioning methods were most effective in

reducing explicit racial preferences. This provides useful ideas for future research.

The use of the same target words across individuals was limited in terms of capturing salient rape-supportive attitudes for each individual and could be seen to reflect a more structuralist approach. This may have meant that change was undetected by the IRAP, if an individual did not endorse the beliefs targeted by the stimulus set. Nevertheless, standardisation enabled analysis at the group-level. Adopting a functional-contextual approach to the selection of target words would prove an interesting avenue for research. This could be achieved by completing repertory grids, taken from Personal Construct Theory (PCT; Kelly, 1955; 1969), to form target words.

In-line with hypotheses, the mean pre-*D*-IRAP scores predicted the relative proportion of tokens donated to charity, meaning individuals who exhibited greater bias toward rape-supportive brief implicit responses were more likely to place a greater number of tokens in the male victim charity-box relative to the tokens placed in the female charity-box. This finding was also replicated for the trial-type Women are not honest. The AMMSA did not predict charity-box behaviour, so whilst in-line with hypotheses, in that it was less predictive than the IRAP, it did not demonstrate any predictive ability. This possibly indicates that the charity-box task produced the context for brief relational responding, as it was administered when participants were leaving. However, without timing responses this cannot be tested.

In contrast to hypotheses, the pre-IRAP predicted⁴ ratings of the female researcher's approachability (Q4, RRS). Greater rape-supportive bias was significantly correlated with perceiving the researcher to be more approachable. This might reflect a wider belief relating to perceived dominance over women and subsequent right to approach. It could also reflect wider entitlement-type beliefs relating to ability to approach women, which has been linked to sexual aggression (Jewkes, Sikweyiya, Morrell, Dunkle, 2011; Pemberton & Wakeling, 2009; Polaschek & Gannon, 2004;

⁴ Correlations between variables are being used here to make inferences about predictions.

Thornton, 2002). From an REC perspective this would mean that brief implicit responses, were better able to predict similar behaviour that is less deliberative and more spontaneous. Studying the time taken to rate the researcher on each item would have shed light on this. Further research is warranted to unpick this. [See Extended Discussion 4.1].

Against hypotheses, the AMMSA did not predict ratings of the researcher. Future research could measure the time taken to respond to questions to assess the degree to which the RRS involved elaborate responding. Future research might also consider observational approaches with independent judges to rate behaviour towards the researcher. The ratings could be regressed onto indirect/direct responses as potential predictors.

Whilst limitations have been noted, this study is the first to investigate rape-supportive implicit cognition from a REC perspective using the IRAP. Whilst this was a relatively new approach, it offered greater specificity in the measurement of exact verbal relations (cognitions) and provided information regarding the direction of attitude differences (*D*-IRAP effect). [See Extended Discussion 4.2 & 4.3].

The clinical implications of these findings highlight that taking a functional contextual approach, specifically an REC perspective to the measurement and explanation of rape-supportive cognition; can be informative in gathering a comprehensive picture of the specific verbal relations that might be relevant in predicting specific behaviours. The behavioural outcome measure cannot be generalised to sexually aggressive behaviour. Nevertheless, the promising initial findings of this study might prompt further research, in order to tease out specific behaviours that can be predicted using the IRAP. Further research might explore how dynamic risk-factors (Thornton, 2002) such as impulsivity may be related to brief implicit responses. For example, exploring how well rape-relevant brief implicit responses can predict different types of sexual offending, such as, carefully planned/deliberated to impulsive/opportunistic offending; it is hypothesised that brief implicit responses would offer better predictive utility regarding spontaneous/opportunistic offences than elaborated explicit responses,

which might better predict elaborate plans. However, a large-scale prospective cohort study or targeted follow-up of offender recidivism would be required.

Future research is proposed to employ longitudinal designs to explore brief implicit responses from a developmental perspective. As the current study was a cross-sectional study it cannot answer questions regarding causality in relation to implicit beliefs and sexual aggression. Expanding research in this way would enhance the development of treatment approaches to reduce the prevalence of sexual aggression. Further exploration of the types of brief implicit responses that predict specific spontaneous behaviour and elaborated explicit responses that determine deliberative action might also enhance current sex-offender risk-assessment tools. [See Extended Discussion 4.4].

References

- Abbey, A., McAuslan, P., Zawacki, T., Clinton, A. M., & Buck, P. O. (2001). Attitudinal, experiential, and situational predictors of sexual assault perpetration. *Journal of Interpersonal Violence, 16*(8), 784-807.
- Abbey, A., & Jacques-Tiura, A. (2011). Sexual assault perpetrators' tactics: Associations with their personal characteristics and aspects of the incident. *Journal of Interpersonal Violence, 26*(14), 2866-2889. doi:10.1177/0886260510390955
- Anderson, L.A., & Whiston, S.C. (2005). Sexual assault education programs: A meta-analytic examination of their effectiveness. *Psychology of Women Quarterly, 29*(4), 374-388. doi:10.1111/j.1471-6402.2005.00237.x
- Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken, & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 361-382). New York, NY US: Guilford Press.
- Barnes-Holmes, D., Barnes-Holmes, Y., Power, P., Hayden, E., Milne, R., & Stewart, I. (2006). Do you really know what you believe? Developing the Implicit Relational Assessment Procedure (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist, 32*(7), 169-177.
- Barnes-Holmes, D., Murphy, A., Barnes-Holmes, Y., & Stewart, I. (2010). The implicit relational assessment procedure: Exploring the impact of private versus public contexts and the response latency criterion on pro-white and anti-black stereotyping among white Irish individuals. *The Psychological Record, 60*, 57-80.
- Barnes-Holmes, D., Murtagh, L., Barnes-Holmes, Y., & Stewart, I. (2010). Using the implicit association test and the implicit relational assessment

procedure to measure attitudes toward meat and vegetables in vegetarians and meat-eaters. *Psychological Record*, 60(2), 287-306.

Barnes-Holmes, D., Barnes-Holmes, Y., Stewart, I., & Boles, S. (2010). A sketch of the implicit relational assessment procedure (IRAP) and the relational elaboration and coherence (REC) model. *Psychological Record*, 60(3), 527-542.

Barnes-Holmes, D., Waldron, D., Barnes-Holmes, Y., & Stewart, I. (2009). Testing the validity of the implicit relational assessment procedure and the implicit association test: Measuring attitudes toward Dublin and country life in Ireland. *Psychological Record*, 59(3), 389-406.

Brown, A. S., Gray, N. S., & Snowden, R.J. (2009). Implicit measurement of sexual associations in child sex abusers: Role of victim type and denial. *Sexual Abuse: Journal of Research and Treatment*, 21(2), 166-180. doi:10.1177/1079063209332234

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

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. Hillsdale, NJ: Lawrence Erlbaum.

Cullen, C., Barnes-Holmes, D., Barnes-Holmes, Y., & Stewart, I. (2009). The implicit relational assessment procedure (IRAP) and the malleability of ageist attitudes. *Psychological Record*, 59(4), 591-620.

DeGue, S., & DiLillo, D. (2004). "You would if you loved me": Toward an improved conceptual and etiological understanding of nonphysical male sexual coercion. *Aggression and Violent Behavior*, 10, 513-532. doi:10.1016/j.avb.2004.09.001

De Houwer, J. (2003). The Extrinsic Affective Simon Task, *Experimental Psychology*, 50(2), 77-85.

- De Houwer, J. (2006). What are implicit measures and why are we using them? In R. W. Wiers & A. W. Stacy (Eds.). *The handbook of implicit cognition and addiction* (pp. 11–28). Thousand Oaks, CA: Sage.
- Drake, C. E., Kellum, K. K., Wilson, K. G., Luoma, J. B., Weintstein, J. H., & Adams, C. H. (2010). Examining the Implicit Relational Assessment Procedure: four preliminary studies. *The Psychological Record, 60*, 81-100.
- Fazio, R. H. (2007). Attitudes as object-evaluation associations of varying strength. *Social Cognition, 25*(5), 603-637.
- Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fida pipeline? *Journal of Personality and Social Psychology, 69*(6), 1013-1027.
- Field, A. P., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. London: SAGE.
- Field, H. S. (1978). Attitudes toward rape: A comparative analysis of police, rapists, crisis counselors, and citizens. *Journal of Personality and Social Psychology, 36*(2), 156-179.
- Froni, F., & Mayr, U. (2005). The power of a story: New, automatic associations from a single reading of a short scenario. *Psychonomic Bulletin and Review, 12*(1), 139-144.
- Fox, E. J. (2006). Constructing a pragmatic science of learning and instruction with functional contextualism. *Educational Technology Research & Development, 54*(1), 5-36.
- Gannon, T. A., Collie, R. M., Ward, T., & Thakker, J. (2008). Rape: Psychopathology, theory and treatment. *Clinical Psychology Review, 28*(6), 982-1008. doi:10.1016/j.cpr.2008.02.005

- Gawronski, B., & Payne, B. K. (Eds.). (2010). *Handbook of implicit social cognition: Measurement, theory, and applications*. New York, NY: Guilford Press.
- Gawronski, B., & Strack, F. (2004). On the propositional nature of cognitive consistency: Dissonance changes explicit, but not implicit attitudes. *Journal of Experimental Social Psychology, 40*(4), 535-542. doi:10.1016/j.jesp.2003.10.005
- 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. doi:10.1002/ab.20195
- Golijani-Moghaddam, N., Hart, A., & Dawson, D. L. (2013). The implicit relational assessment procedure: Emerging reliability and validity data. *Journal of Contextual Behavioral Science*. Advance online publication. doi:10.1016/j.jcbs.2013.05.002
- Greenwald, A.G., Banaji, M.R., Nosek, B.A., & Klauer, K.C. (2005). Validity of the salience asymmetry interpretation of the implicit association test: Comment on Rothermund and Wentura (2004). *Journal of Experimental Psychology: General, 134*(3), 420-425. doi:10.1037/0096-3445.134.3.420
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. K. L. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74*, 1464 -1480.
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*(2), 197-216. doi:10.1037/0022-3514.85.2.197

- Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, *109*(1), 3-25. doi:10.1037/0033-295X.109.1.3
- Hall, G. C. N. & Hirschman, R. (1991). Towards a theory of sexual aggression: a quadripartite model. *Journal of Consulting and Clinical Psychology*, *59*(5), 662-669.
- Han, H. A., Czellar, S., Olson, M. A., & Fazio, R. H. (2010). Malleability of attitudes or malleability of the IAT? *Journal of Experimental Social Psychology*, *46*, 286-298. doi:10.1016/j.jesp.2009.11.011
- Hanson, R. K., Bourgon, G., Helmus, L., & Hodgson, S. (2009). The principles of effective correctional treatment also apply to sexual offenders: A meta-analysis. *Criminal Justice and Behavior*, *36*(9), 865-891. doi:10.1177/0093854809338545.
- Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. New York London: Kluwer Academic/Plenum Publishers.
- Helmus, L., Hanson, R., K., Babchishin, K., M., & Mann, R., E. (2013). Attitudes supportive of sexual offending predict recidivism: A meta-analysis. *Trauma, Violence & Abuse*, *14*(1), 34-53. doi:10.1177/1524838012462244
- Hooper, N., Villatte, M., Neofotistou, E., & McHugh, L. (2010). The effects of mindfulness versus thought suppression on implicit and explicit measures of experiential avoidance. *International Journal of Behavioral Consultation and Therapy*, *6*(3), 233-244.
- Hughes, S., Barnes-Holmes, D., & De Houwer, J. (2011). The dominance of associative theorizing in implicit attitude research: Propositional and behavioral alternatives. *Psychological Record*, *61*(3), 465-496.

- Hughes, S., Barnes-Holmes, D., & Vahey, N. (2012). Holding on to our functional roots when exploring new intellectual islands: A voyage through implicit cognition research. *Journal of Contextual Behavioral Science, 1*(1-2), 17-38. doi:10.1016/j.jcb.2012.09.003
- Hussey, I., & Barnes-Holmes, D. (2012). The implicit relational assessment procedure as a measure of implicit depression and the role of psychological flexibility. *Cognitive and Behavioral Practice, 19*(4), 573-582. doi:10.1016/j.cbpra.2012.03.002
- 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) doi:10.1371/journal.pone.0029590
- Kelly, G. (1955). *The psychology of personal constructs*. New York: Norton.
- Kelly, G. (1969). Humanistic methodology in psychological research. In B. Maher (Ed.) *Clinical psychology and personality: The selected papers of George Kelly*. Chichester: Wiley & Sons.
- Knight, R. A., & Sims-Knight, J. E. (2003). The developmental antecedents of sexual coercion against women in adolescents: Testing alternative hypotheses with structural equation modelling. *Annals of the New York Academy of Sciences, 989*, 72–85.
- Langhinrichsen-Rohling, J., Foubert, J., D., Brasfield, H., M., Hill, B., & Shelley-Tremblay, S. (2011). The Men's program: Does it impact college Men's self-reported bystander efficacy and willingness to intervene? *Violence Against Women, 17*(6), 743-759. doi:10.1177/1077801211409728
- Langton, C. M., Barbaree, H. E., Harkins, L., Arenovich, T., McNamee, J., Peacock, E. J., Dalton, A., Hansen, K., Luong, D., & Marcon, H. (2008). Denial and Minimization Among Sexual Offenders: Post-treatment

Presentation and Association With Sexual Recidivism. *Criminal Justice and Behavior*, 35, 69-98. doi:10.1177/0093854807309287

Malamuth, N. M. (1986). Predictors of naturalistic sexual aggression. *Journal of Personality and Social Psychology*, 50(5), 953-962.

Malamuth, N. M. (1989). The attraction to sexual aggression scale: Part one. *Journal of Sex Research*, 26(1), 1989.

Malamuth, N. M. (2003). Criminal and noncriminal sexual aggressors: Integrating psychopathy in a hierarchical-mediational confluence model. *Annals of the New York Academy of Sciences*, 989, 33-58.

Marshall, W. L. & Barbaree, H. E. (1990). An integrated theory of the etiology of sexual offending. In W. L. Marshall, D. R. Laws & H. E. Barbaree (Eds), *Handbook of sexual assault: Issues, theories, and treatment of the offender* (pp. 257-275). New York: Plenum Press.

McKenna, I. M., Barnes-Holmes, D., Barnes-Holmes, Y., & Stewart, I. (2007). Testing the fake-ability of the Implicit Relational Assessment Procedure (IRAP): The first study. *International Journal of Psychology and Psychological Therapy*, 7, 123-138.

Mihailides, S., Devilly, G.J., & Ward, T.. (2004). Implicit cognitive distortions and sexual offending. *Sexual Abuse: Journal of Research and Treatment*, 16(4), 333-350.
doi:10.1023/B:SEBU.0000043327.10168.5b

Ministry of Justice, Home Office & the Office for National Statistics (2013). *An overview of sexual offending in England and Wales*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/214970/sexual-offending-overview-jan-2013.pdf

Murnen, S. K., Wright, C., & Kaluzny, G. (2002). If "boys will be boys," then girls will be victims? A meta-analytic review of the research that relates masculine ideology to sexual aggression. *Sex Roles*, 46(11), 359-375.

- Murphy, W. D., Coleman, E. M. & Haynes, M. R. (1986). Factors related to coercive sexual behaviour in a nonclinical sample of males. *Violence and Victims, 1*(4), 255-278.
- Nock, M. K., Park, J. M., Finn, C. T., Deliberto, T. L., Dour, H. J., & Banaji, M. R. (2010). Measuring the suicidal mind: Implicit cognition predicts suicidal behavior. *Psychological Science, 21*(4), 511-517. doi:10.2307/41062239
- Nosek, B.A. (2007). Implicit-explicit relations. *Current Directions in Psychological Science, 16*(2), 65-69. doi:10.1111/j.1467-8721.2007.00477.x
- Nosek, B. A. & Banaji, M. R. (2001). The Go/No-Go Association Task. *Social Cognition, 19*(6), 161-176.
- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math \neq me. *Journal of Personality and Social Psychology, 83*(1), 44-59. doi:10.1037/0022-3514.83.1.44
- Nosek, B.A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and using the implicit association test: II. Method variables and construct validity. *Personality and Social Psychology Bulletin, 31*(2), 166-180. doi:10.1177/0146167204271418
- Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The implicit association test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Automatic processes in social thinking and behavior* (pp. 265-292). New York, NY US: Psychology Press.
- Nosek, B. A., Hawkins, C. B., & Frazier, R. S. (2011). Implicit social cognition: From measures to mechanisms. *Trends in Cognitive Sciences, 15*(4), 152-159. Nunes, K.L., Firestone, P., & Baldwin, M. W. (2007). Indirect assessment of cognitions of child sexual abusers with

- the implicit association test. *Criminal Justice and Behavior*, 34(4), 454-475. doi:10.1177/0093854806291703
- Nunes, K. L., Hermann, C. A., & Ratcliffe, K. (2013). Implicit and explicit attitudes toward rape are associated with sexual aggression. *Journal of Interpersonal Violence*, 28(13), 2657-2675. doi:10.1177/0886260513487995
- Ó Ciardha, C., & Ward, T. (2013). Theories of cognitive distortions in sexual offending: What the current research tells us. *Trauma, Violence, and Abuse*, 14(1), 5-21. doi:10.1177/1524838012467856
- Olson, M.A., & Fazio, R. H. (2004). Reducing the influence of extra-personal associations on the implicit association test: Personalizing the IAT. *Journal of Personality and Social Psychology*, 86(5), 653-667. doi:10.1037/0022-3514.86.5.653
- 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. doi:10.1006/jrpe.1998.2238
- Pemberton, A. E., & Wakeling, H. C. (2009). Entitled to sex: Attitudes of sexual offenders. *Journal of Sexual Aggression*, 15(3), 289-303. doi:10.1080/13552600903097212
- Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, 2, 1-10. doi: 10.3389/neuro.11.010.2008
- Polaschek, D.L.L., & Gannon, T. A. (2004). The implicit theories of rapists: What convicted offenders tell us. *Sexual Abuse: Journal of Research and Treatment*, 16(4), 299-314. doi:10.1023/B:SEBU.0000043325.94302.40

- Polaschek, D. L. L., & Ward, T. (2002). The implicit theories of potential rapists: What our questionnaires tell us. *Aggression and Violent Behavior, 7*(4), 385-406. doi:10.1016/S1359-1789(01)00063-5
- Rydell, R.J., & McConnell, A.R. (2006). Understanding implicit and explicit attitude change: A systems of reasoning analysis. *Journal of Personality and Social Psychology, 91*(6), 995-1008. doi:10.1037/0022-3514.91.6.995
- Scheel, E. D., Johnson, E. J., Schneider, M., & Smith, B. (2001). Making rape education meaningful for men: The case for eliminating the emphasis on men as perpetrators, protectors, or victims. *Sociological Practice: A Journal of Clinical and Applied Sociology, 3*(4), 257-278.
- Schewe, P. A. (2002). Guidelines for developing rape prevention and risk reduction interventions. In P. A. Schewe (Ed.), *Preventing violence in relationships: Interventions across the life span* (pp. 107-136). Washington DC, US: American Psychological Association. doi:10.1037/10455-005
- Schnabel, K., Asendorpf, J., & Banse, R. (2006). Employing automatic approach and avoidance tendencies for the assessment of implicit personality self-concept the implicit association procedure (IAP). *Experimental Psychology, 53*(1), 69-76. doi:10.1027/1618-3169.53.1.69
- Snowden, R. J., Craig, R. L., & Gray, N. S. (2011). Indirect behavioral measures of cognition among sexual offenders. *Journal of Sex Research, 48*(2), 192-217. doi:10.1080/00224499.2011.557750
- Steinberg, J. A., Karpinski, A., & Alloy, L. B. (2007). The exploration of implicit aspect of self-esteem in vulnerability - stress models of depression. *Self and Identity, 6*(3), 101-117. doi:10.1080/15298860601118884

- Stenning, P., Mitra-Kahn, T., & Gunby, C. (2013). Sexual violence against female university students in the U.K.: A case study. (Italian). *Rivista Di Criminologia, Vittimologia E Sicurezza*, 7(2), 100.
- Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior, *Personality & Social Psychology Review*, 8(3), 220-247.
- Suarez, E., & Gadalla, T. M. (2010). Stop blaming the victim: A meta-analysis on rape myths. *Journal of Interpersonal Violence*, 25(11), 2010-2035. doi:10.1177/0886260509354503
- Thornton, D. (2002). Constructing and testing a framework for dynamic risk assessment. *Sexual Abuse: A Journal of Research & Treatment*, 14(2), 139-153.
- Vahey, N.A., Barnes-Holmes, D., Barnes-Holmes, Y., & Stewart, I. (2009). A first test of the implicit relational assessment procedure as a measure of self-esteem: Irish prisoner groups and university students. *Psychological Record*, 59(3), 371-388.
- Ward, T. (2000). Sexual offenders' cognitive distortions as implicit theories. *Aggression and Violent Behavior*, 5(5), 491-507. doi:10.1016/S1359-1789(98)00036-6
- Ward, T., & Beech, A. (2006). An integrated theory of sexual offending. *Aggression and Violent Behavior*, 11, 44-63. doi:10.1016/j.avb.2005.05.002
- Ward, T., & Casey, A. (2010). Extending the mind into the world: A new theory of cognitive distortions in sex offenders. *Aggression and Violent Behavior*, 15, 49-58. doi:10.1016/j.avb.2009.08.002
- Wright, D. B., London, K., & Field, A. P. (2011). Using bootstrap estimation and the plug-in principle for clinical psychology data. *Journal of Experimental Psychopathology*, 2(2), 252-270. doi:10.5127/jep.013611

Yamawaki, N., Darby, R., & Queiroz, A. (2007). The moderating role of ambivalent sexism: The influence of power status on perception of rape victim and rapist. *Journal of Social Psychology, 147*, 41-56.

Extended Paper

1 Extended Background

1.0 Rape myths are defined by Lonsway & Fitzgerald (1994) as being attitudes pertaining to rape which are generally false. However, Gerger, Kley, Bohner & Siebler, (2007) critique this definition, noting that it is almost impossible to test empirically as it is hard to falsify, especially if the rape-myth relates to secretly held beliefs. Gerger et al. (2007) propose a definition of rape myths as ethically wrong rather than false. This side-steps the issues relating to empirical falsification.

Burt (1980) developed a measure of rape myths (the Rape Myth Acceptance Scale), which was later developed into the RAPE scale by Bumby (1996). Existing rape-myth acceptance scales, however, are somewhat outdated given ever-shifting societal attitudes towards sex and gender (Gerger et al., 2007). In response to these concerns, Gerger et al. (2007) developed a measure which aimed to capture the subtleties of rape-supportive attitudes so as to reduce problems with social desirability.

1.1 Research on implicit measures has been plagued by conceptual confusion related to defining the types of beliefs being measured (Machado & Silva, 2007), definitions of the properties of measurement procedures and, the theoretical underpinnings of such cognition (De Houwer, 2006). De Houwer (2006) provided a comprehensive analysis and suggested terms such as *direct* and *indirect* should refer to the features of the measurement procedure and *implicit* and *explicit* to define the features of the to-be-assessed attributes.

1.2 To illustrate this further, a bias in responding on the IAT to *thin* versus *fat* could mean a positive attitude towards *thin* and a neutral attitude toward *fat* or a neutral attitude towards *thin* and a negative attitude towards *fat*.

The IAT has also been used to predict future behaviour, for example, McConnell and Leibold (2001) explored the predictive validity of the IAT (in comparison to an explicit measure of prejudice) in relation to inter-group discrimination in an American sample of university students. They found that negative racial IAT performance predicted negative social interactions with black individuals as measured explicitly by amount of time smiling in interactions with black individuals, however, explicit measures did not demonstrate this predictive validity. Nock et al. (2010) found that implicit cognitions relating to suicide within an IAT predicted attempts at suicide significantly better than explicit measures. Similarly, Friese, Hofmann and Wänke (2008) found that IAT performance predicted impulsive but not self-controlled future eating behaviour and the latter was better predicted by explicit measurement. More recently, Hauben, Havermans and Weirs, (2010) found that an evaluative conditioning intervention produced changes in implicit beliefs regarding alcohol and a corresponding reduction in drinking behavior. Greenwald, Poehlman, Uhlmann, Banaji (2009) conducted a meta-analysis of the predictive validity of the IAT and found that within the 32 socially-sensitive studies, the predictive validity of the IAT was significantly better than that of the explicit measures. Interestingly, in other studies greater convergence between explicit and implicit measures has been found (Gawronski, Geschke & Banse, 2003; Payne, 2001), most notably when responses on explicit measures are based on fast, intuitive responses rather than spending time deliberating over the answers (Gawronski & LeBel, 2008).

Sequential priming tasks have also been used to research implicit cognition (Dovidio, Evans, & Tyler, 1986); however, they tended to produce small effect sizes and limited reliability (Payne & Gawronski, 2010). They can be conceptualised as based in an associative paradigm and therefore are limited in terms of making assumptions about cognitive structures existing in memory without being able to empirically prove this.

1.3 This idea is captured by the Motivation and Opportunity as Determinants (MODE) model (Fazio, 1990; Fazio, 2007). The MODE model proposes that attitudes are stable associations within memory between an object and summary evaluation. Strong associations are considered to trigger automatic evaluations that are supposedly outside of awareness or executive control (Fazio, 2007). If opportunity and motivation to respond is high, then the impact of automatic associative processes to influence behaviour is low and responses are deemed to reflect deliberative processing. In contrast, if the opportunity and motivation is low then automatic associations are proposed to govern behaviour to a greater degree. For example, Widman and Olson (2013) used an evaluative priming task to assess the predictive validity of implicit rape-supportive cognition in a community male sample and undergraduate males and found that it predicted self-reported past sexual aggression which was not predicted by explicit measures of rape-supportive attitudes. They explained this effect using the MODE model, attributing it to the mediation of cognitive associations in memory automatically guiding sexual behaviour.

Similarly, Blake and Gannon (2010) employed a priming paradigm (Lexical Decision Task) to investigate the relationship between implicit beliefs and explicit beliefs on a proclivity to rape measure in a community male sample. They sought to test Polascheck and Ward's (2002) implicit theories hypothesis, whereby it was proposed that convicted rapists held specific beliefs about themselves, others and the world, (such as *male sex drive is uncontrollable*), which was proposed to be causal in sexual offending. Blake and Gannon (2010) found that only the explicit measure of rape-supportive cognition predicted self-reported proclivity to rape. They concluded that Ward's (2000) theory, (which explains implicit cognition as shaping behaviour through the process of associations in memory mediating information processing), may need to be reconsidered in light of their evidence. Others, however, have proposed that such beliefs may be post-hoc distortions to rationalise their offending behaviour in order to absolve themselves of responsibility (Maruna & Mann, 2006), which may offer an explanation for Blake and Gannon's findings. Blake and Gannon (2012)

carried out a similar study using a different indirect measure, an interpretative bias task. However, whilst finding support for the implicit theory *women are sex objects* (Polaschek & Ward, 2002), as measured by the interpretative bias task, again Blake and Gannon (2012) found the explicit measure better predicted self-reported proclivity to rape, and conceptualised attitudes as schemata, forming mental associations in memory. Leibold and McConnell (2004) used a sequential priming paradigm to investigate the predictive validity of implicit cognition relating to women and sex/hostility, and they found that this type of cognition predicted self-reported past sexually aggressive behaviour. However, in-line with common explanations of the effect produced on indirect measurement procedures; they attributed this finding to sexually aggressive men holding stronger associations in memory between women and sex/hostility, than those with a less sexually aggressive history.

1.4 For example, Foroni and Mayr (2005) found implicit bias against insects could be reversed following reading a story about the positive attributes of insects within a post-nuclear world. Blair, Ma and Lenton (2001) found that taking five minutes to imagine a strong woman reduced previous negative gender stereotyping.

1.5 Han, Czellar, Olson and Fazio, (2010) argued that imprecise response labels used in the IAT allow for plasticity and thus explain findings of malleability. Similarly, Olson and Fazio (2004) argued that responses can be due to multiple interpretations influencing associations so that extra-personal rather than personal associations are measured.

1.6 RFT views cognition as verbal behaviour and applies behavioural principles to the study of it. Central to the theory is the notion of arbitrarily applicable relational responding which accounts for the way in which human

cognition (verbal behaviour) operates. Arbitrarily applicable relational responding refers to a way of relating, which is governed by contextual cues that inform which relation is to be inferred, rather than the relation being determined by the specific physical properties of the object (Hayes, Barnes-Holmes & Roche, 2001). For example, Törneke (2010) explains this, by asking you to imagine that you are told that the symbol # is larger than the symbol @ (thus establishing a new relation) despite the fact that # is physically smaller than @. You are then asked to imagine that you are informed that @ is equal to one hundred pounds. When asked to choose one of the symbols, it is almost certain that you would pick #, as you would think that it represents a greater monetary value. Therefore, you would be responding to the newly learned arbitrary relation (e.g. that # is larger than @) rather than responding to its physical properties (e.g. that @ is larger than #). This approach to implicit cognition proposes that stimulus relations range from low to high complexity and can relate to other stimulus relations in complex relational networks. These are proposed to subsequently relate to other relational networks (Hughes, Barnes-Holmes & Vahey, 2012). Furthermore, the REC model explains responses to direct or indirect procedures as being determined by the previous learning histories governing similar behaviour in the past (Hughes et al., 2012).

The REC model is not a dual process model in which associative and propositional processes are deemed to make up responding on implicit and explicit measures (as can be seen with the Associative-Propositional Evaluation (APE) model; Gawronski & Bodenhausen, 2007). Instead, it is based upon a single process of arbitrarily applicable relational responding, as explained by Relational Frame Theory. This means that divergence of implicit and explicit attitudes is explained by the degree to which relational responses are elaborated and cohere, rather than by the workings of associative and propositional processes (Barnes-Holmes, et al., (2010). It should be noted that the REC model includes also, the behavioural processes of respondent conditioning and stimulus generalisation within its conceptualisation of the formation of language and cognition. Arbitrarily applicable relational responding is proposed to involve different patterns of

behaviour of varying degrees of complexity and derivation (Barnes-Holmes, et al., (2010). For example, Hughes et al., (2012) argue that relational responding can be separated into four categories consisting of: low relational complexity, low derivation; high relational complexity, low derivation; low relational complexity, high derivation; and high relational complexity with high derivation. Brief and immediate relational responses are considered to involve short time-frames, low complexity and low derivation. In contrast, elaborate and extended relational responses are thought to involve greater time, high complexity and high derivation. Therefore, the REC model posits that as behaviours take time to unfold, complex responses take longer than simpler counterparts and are less accurate. The more often a response has been previously derived is considered to lead to greater speed and accuracy in responding.

Therefore, in relation to the current study, consider participants that respond with greater speed and accuracy to relational frames of co-ordination between the stimuli 'women' and 'slutty' on the IRAP. From an REC perspective, this would indicate that they have a prior learning history which involves the frequent derivation of this relational frame, thus increasing the probability of them emitting the same relational response with speed and accuracy in the study. According to the REC model, for individuals exhibiting greater adherence to rape-supportive attitudes, such response patterns would likely emerge from exposure to some of the verbal and non-verbal behavioural contingencies that function for university males living in England. These may be related to media and societal narratives concerning women and rape. Difficulties exist with the REC model, for instance, the presence of IRAP effects for non-associative relations implies that a purely associative account may constrain understanding of implicit responses. Measures such as the IAT, when viewed from an REC perspective can be conceptualised as tapping into relations of co-ordination, and so lack the ability to measure a wide variety of relations. Alternatively, brief implicit responses as captured by the IRAP, can offer analysis with any type of relation such as a relation of opposition, hierarchy, spatial, temporal or comparative relations. This distinguishes the IRAP as a useful measure in

the study of implicit cognition. Previous research using indirect measures to explore implicit cognition within the domain of sexual offending, has been dominated by an associative approach to measurement and interpretation (Snowden, Craig & Gray, 2011). We argue that taking a functional contextual approach to this area offers unique specificity and insight into implicit rape-supportive cognition.

1.7 Despite the prevalence of rape-myths and their link with sexual aggression, there has been limited consideration regarding how rape-prevention programmes can address this. A wide variety of outcome measures are used to evaluate attitude change in the rape-prevention literature within community samples. This heterogeneity of measures makes comparisons between studies difficult; nevertheless, reviews have been undertaken. For example, Morrison, Hardison, Mathew & O'Neill (2004) conducted a systematic review of sexual assault intervention programmes using studies published between 1990 and 2003. They found that 14% of studies reported positive intervention effects at post-test and follow up, and 80% reported mixed results. Morrison et al. concluded that attitudinal changes, often found post-intervention, were not maintained over time and noted a lack of behavioural outcome measures in order to evaluate effectiveness in terms of a reduction in sexual violence.

The inclusion of measures of behaviour is pertinent when studying the effectiveness of interventions designed primarily to change behaviour. Until recently, however, this aspect of evaluating effectiveness has been largely ignored, predominantly due to a lack of effective behavioural measures.

Research evaluating the effectiveness of interventions aimed at changing beliefs supportive of sexual aggression within college males is limited. A meta-analysis was conducted by Flores and Hartlaub (1998) into the effectiveness of interventions in reducing rape-myth acceptance in male college students as measured by explicit (questionnaire) measures. They concluded that interventions such as sexuality courses, workshops and

video interventions were equally effective, with no evidence that length of intervention mediated effectiveness. Other studies have also demonstrated the effectiveness of interventions aiming to reduce adherence to rape-supportive cognition in university males (Milhausen et al., 2006; Schewe & O'Donohue, 1996). However, findings should be interpreted with caution as explicit measures are subject to social desirability biases, especially with regard to socially-sensitive topics such as sexual aggression.

2 Extended Method

2.1 Participants

The sample size calculation using G*Power 3.0 software (Faul, Erdfelder, Buchner, & Lang, 2009) indicated that a total of 42 participants would be required for phase two of the study (effect size $f = .39$; alpha $.05$; levels = 2, power = $.80$; correlation among repeated measures $r = .49$). Both the effect size and the correlation among repeated measures were drawn from previous research examining the malleability of implicit cognition (Cullen et al., 2009). Attrition rate in previous IRAP research has been found to be 13% (Drake et al., 2010). Therefore, in the present study a sample size of 48 participants was aimed for. This sample size was considered to be feasible given recruitment was to take place from two university samples. As the main body of the study was to be completed in one session including pre- and post-measures, drop out from the research was deemed to be less likely than if it was conducted over two testing periods, as participants would not need to have their return to the study reinforced.

Participants were initially recruited easily at the beginning of the data collection period. However, as potential participants typically moved back to their original home during holiday periods, recruitment became significantly more difficult during the latter stages. As a result, the data collection

process was stopped after collecting data from 39 participants. Initial data analysis was carried out to check if power was sufficient. As power was found to be sufficient, further participants were not recruited.

Demographic information was only collected with regards to participants' age and sexuality. This was to ensure that only data necessary for the present study was collected, thus preserving anonymity.

2.2 AMMSA

Please refer to Gerger et al. (2007) for the questions included in the AMMSA questionnaire. In order to comply with copyright law, this was not provided in the current study. The AMMSA provides a modern measure of rape-myths as previous measures were developed over thirty years ago (Burt, 1980; Field, 1978), meaning subtle changes in language used and societal attitudes could impact on the utility of older measures. Older measures have been shown to produce positively skewed distributions of data (Gerger et al., 2007), rendering parametric analysis less valid. The AMMSA has been shown to produce data that is normally distributed (Gerger et al., 2007).

Gerger et al. (2007) explain that the 30 items that make up the AMMSA consist of statements which relate to a number of themes including: denial of the scope of the problem; antagonism towards victims' demands; lack of support for policies designed to help alleviate the effects of sexual violence; beliefs that male coercion form a natural part of sexual relationships; and beliefs that exonerate male perpetrators by blaming the victim or circumstances. Examples of questions within the AMMSA that reflect these main themes respectively are as follows: "Many women tend to misinterpret a well-meant gesture as sexual assault;" "Although the victims of armed robbery have to fear for their lives, they receive far less psychological support than do rape victims;" "After a rape, women nowadays receive ample support;" "When a woman starts a relationship with a man, she must be aware that the man will assert his right to have sex;" and "Nowadays, a

large proportion of rapes is partly caused by the depiction of sexuality in the media as this raises the sex drive of potential perpetrators.”

2.3 Charity-box Task

The charity-box task was developed to capture behaviour that was deemed to be conceptually relevant whilst ensuring it was not too obvious that it was related to the study. Fictional charities were created rather than using pre-existing charities in order that any pre-existing verbal relations relating to real charities were less likely to influence responding. Of course, verbal relational networks are highly complex and far reaching and so it is still likely that pre-existing relational networks were contacted even though the charities were fictional, through the process of stimulus generalisation (Törneke, 2010).

Two charities were designed to reflect conceptually neutral areas. Initially one charity was designed which aimed to raise money for victims of natural disasters. However, initial piloting proved that this attracted a far greater proportion of all tokens, thus reducing the base rate of tokens being placed into the other conceptually relevant charities. Therefore this charity was changed to a wildlife charity to reflect a less emotive area. Piloting highlighted that having the four final charities (women’s refuge; the natural world conservation society; the wildlife foundation; and the blue silence foundation) enabled more even distribution of tokens.

The two charities ‘women’s refuge,’ and ‘the blue silence foundation,’ were designed to reflect conceptually relevant areas for the study. The principle of correspondence (Ajzen & Fishbein, 1977) proposes that the more a behavioural measure shares conceptually relevant attributes with the construct measure, the greater the relationship. However, whilst this has emerged from an associative paradigm (attitudes are mediated by structures in memory), the REC model would also support this, but would stipulate that the crucial feature of correspondence would be the type of relational responding (i.e. brief and immediate or extended and elaborate).

Instead of referring to the 'strength of the relationship,' the REC model would refer to the probability that a particular behaviour occurs in the presence of particular relational responding.

The women's refuge charity was designed to provide support to female victims of domestic abuse, including sexual and physical violence. This was considered to tap into rape-myth constructs, as found in the AMMSA, which related to: denial of the scope of the problem; beliefs that male coercion forms a natural part of sexual relationships (as it was a domestic abuse charity and not for victims of stranger rape); and not wanting to support policies designed to help victims of sexual violence.

Similarly, the blue silence foundation was designed to support male victims of sexual abuse. This was developed in order to provide participants with the option of donating tokens to male victims of a similar crime. This was included in the range of charities because it was thought to provide a measure of degree of support for male victims of sexual abuse.

Research suggests that when studying the prediction of behaviour from implicit measures, if the behavioural task is similarly relational in nature (to the relational implicit measure) then prediction ability is enhanced (Perugini, Richetin & Zogmaister, 2010). For example, Greenwald et al. (2009) conducted a meta-analysis into the predictive validity of the IAT and prediction of behaviour was greater when complementary categories were used as the behavioural measure. These two charities were designed with the aim (at the point of analysis), of combining the number of tokens donated to both, into one score, denoting relative giving to female compared to male victims. This was carried out by taking the number of tokens donated to the male charity from the number of tokens donated to the female charity, to provide a relative measure of giving to both. Positive scores indicated greater giving (and negative scores indicated fewer giving) to the female charity relative to the male charity.

2.4 Researcher Rating Scale

The researcher rating scale was developed specifically for the current research (see Appendix b). The aim was to provide a measure of the participants' ratings of the female researcher, and they were told that their ratings would be shown only to her supervisor for the purpose of evaluating the researcher at her annual review. Participants were informed that the researcher would not see their ratings, and they were to put their response sheet into a sealed envelope. This was to ensure that participants would be less constrained by impression management and social desirability bias that could impact on their responses. For example, if participants thought that the researcher might see their responses they might have been more likely to score her favourably.

During the development process, expert consensus was sought to seek appropriate attributes to measure that related to the area of rape myths and sexual aggression. Literature regarding how women are perceived by sexual aggressors was drawn upon. For example, Malamuth and Brown (1994) used videotaped scenarios in which a woman's reactions to male advances were varied in different conditions, to test hypotheses relating to findings that sexually aggressive men interpret women's communications with them differently to less sexually aggressive men. They tested three hypotheses, namely: sexually aggressive men are less competent at decoding women's negative emotions; aggressors fail to distinguish between women's friendliness and seductiveness and between their expression of assertiveness and hostility; and that aggressors are suspicious of women's intentions and so interpret their communications as untrustworthy. Malamuth and Brown found support for the third hypothesis.

Similarly, Polaschek and Gannon (2004) coded descriptions of convicted rapists' accounts of their offences. They reported that rapists view heterosexual encounters as adversarial, with women seeking to deceive men. Based on this research, questions were developed in the RRS to capture an interpretation of women as being cold, hostile and hurtful. For example, participants were asked to rate the researcher in terms of her

warmth, friendliness and how approachable they found her. Questions were posed to ask about the positive qualities as opposed to the negative aspects of the same construct, in order to make the task more believable (as an evaluation form for the researcher's annual review).

Some questions were designed to capture how intelligent and knowledgeable the researcher was deemed to be. These questions were derived from the literature that points to an association between sexual aggression and a perception that women are sexual objects as opposed to having qualities such as intelligence and knowledge (Polaschek & Gannon, 2004; Polaschek & Ward, 2002). As the RRS was given to participants as a supposed measure of formal evaluation, the questions were structured around the concept of intelligence rather than sexual objectification. When developing the scale, it was important to strike a balance between capturing the relevant attributes for the purpose of the research, whilst ensuring the researcher was not harmed by inviting personal comments from participants.

2.5 Intervention and Control Conditions

Cognitive-restructuring is the term used for a broad range of techniques that are commonly used within cognitive behavioural therapy (Beck, 1976). Initially, cognitive restructuring typically involves identifying the unhelpful cognitions that are deemed important to change, before learning how to challenge and dispute them, with the view that the change will subsequently impact upon behaviour (Beck, 1976). Often comprised of a number of techniques, cognitive restructuring can involve methods such as: gathering evidence that refutes the particular belief; consequential analysis, which involves the weighing up the costs and gains of holding on to a particular belief; and generating alternatives which involves reconceptualising the situation in order that a more adaptive representation is created (Clark, 2013).

Cognitive restructuring is the predominant modality of working with group members' offence-supportive attitudes within modern sexual offender treatment programmes (Marshall & Laws, 2003). The effectiveness of such sex offending programmes is typically measured at an individual level through a structured assessment of risk and treatment need, (Thornton, 2002) and at a group level by the rate and nature of reconviction (Hanson, Bourgon, Helmus and Hodgson, 2009). However, this proves to be very problematic for a number of reasons. For instance, often the base rate of recidivism (re-conviction) is extremely low, especially in offenders assessed as being of low risk, using actuarial risk assessment tools such as the Risk Matrix 2000 (Hanson & Thornton, 2000). Therefore, follow up periods need to be long, often many years, in order to detect the rate of recidivism, which proves to be difficult when researching interventions as the research can be expensive and takes many years to complete.

Another difficulty relates to the difference between recidivism rates and true re-offending rates. Reconviction data is often used to evaluate effectiveness of such programmes. However, the true rate of re-offending is thought to be far higher, as often victims do not come forward to the police and even if the offence makes it to court, the prosecution rates are very low due to difficulties with obtaining reliable evidence. These factors make accurate evaluation of interventions difficult to achieve. The available evidence suggests that sex offender treatment programmes, do reduce rate of recidivism. For example, Losel and Schmucker (2005) completed a meta-analysis of 69 studies which included a sample of 9512 sex offenders that had completed treatment, and a sample of 12,669 untreated sexual offenders. They found that sexual offender treatment; particularly cognitive behavioural interventions had a positive treatment effect in relation to sexual and general recidivism. In contrast, Kenworthy, Adams, Brooks-Gordon and Fenton (2004) concluded from their review that it is questionable how effective treatment is when applied outside of a well-designed study. Analysis of the effectiveness of sex offender interventions is marred by poor quality studies and a lack of random assignment. Few

studies have employed robust designs to study interventions deemed to be of a good standard (Hanson et al., 2009).

The rating of study quality is often variable amongst reviews. Hanson et al., (2009) looked to address this by employing the guidelines of the Collaborative Outcome Data Committee (CODC) to determine the quality of studies in their review of the effectiveness of sexual offender treatment. These guidelines were developed specifically for the evaluation of sexual offender outcome studies in the context of meta-analysis. The definition of a good study in the literature in this area is something researchers have struggled to agree on. The principles of treatment effectiveness, within the general offending population, which are most likely to produce reductions in recidivism, are considered to be those of risk, need and responsivity (RNR; Bonta & Andrews, 2007). For example, interventions are most effective if they: target offenders deemed to be of medium or high risk of recidivism; seek to work on criminogenic needs (such as criminal attitudes); and if they are responsive to individual learning abilities.

Hanson et al., (2009) sought to identify if the same principles of effective treatment in the general offending literature could be applied to sex offender treatment. They included 23 recidivism outcome studies; the majority of which were based on Canadian samples (12), five were based on American samples, three studies from the United Kingdom, two from New Zealand, and one from Holland. The majority of studies included in the meta-analysis were investigating treatment effectiveness for adult male sex offenders. However, it is worth noting that four studies focussed on adolescent sex offenders and three studies included females, amounting to less than ten percent of the total samples respectively (Hanson, et al., 2009). A recent meta-analysis has found recidivism rates for female sexual offenders to be very low; less than 3% over an average follow-up period of 6.5 years (Cortoni, Hanson & Coache, 2010). Therefore, caution must be taken when interpreting the results with respect to treatment effectiveness in adult males. The treatment programmes studied consisted of 10 that were delivered in institutions, 11 in the community and

2 in both institutions and the community. Of the 23 studies, 19 investigated sex-offender treatment programmes designed specifically for this population, however 4 studies related to the outcomes for sex offenders attending general offending behaviour programmes. In 10 of the studies, recidivism was defined as reconviction and in 12 studies it was defined as re-arrest; one study did not specify how they defined reconviction (Hanson et al., 2009). The median follow-up period used when measuring recidivism was 4.7 years, with a range from 1 to 21 years. Results indicated that the recidivism rates of the 'treated' sex offender group (10.9%, $n = 3121$) were lower than that of 'untreated' groups (19.2%, $n = 3625$).

However, given the limitations of the study noted above, it is clear that this is far from conclusive and highlights the need for further good quality randomised research designs.

The intervention condition in the current study was developed as a cognitive restructuring task. Whilst it is clear that it does not accurately replicate the methods used in treatment programmes, and the study focusses on university males rather than convicted offenders, it is recognised that there are benefits to studying this under-researched area. For example, in the present study, using a novel assessment procedure (IRAP) in controlled conditions with men that have been shown to have the capacity to sexually offend is the first step in a process; using the IRAP in a sample of convicted sex-offenders to assess offence-related beliefs is a future possibility but extensive research is warranted first, to answer many questions relating to its utility and predictive validity.

As the primary aim within the current study was to investigate the malleability of implicit rape-supportive cognition, it was decided to use an approach which evidence suggests may be more effective in this population. For example, it was decided to approach the participants in the intervention as potential helpers instead of potential perpetrators, as research has found that a sample of college men were more responsive to intervention if they are not made to feel defensive and ashamed (Langhinrichsen-Rohling et al. 2011). Therefore the intervention condition was structured in this way,

asking participants to generate alternative more adaptive statements (Clark, 2013), to try to change the rape-supportive beliefs of their 'friend,' as a means of vicariously restructuring their own attitudes. Or from an RFT and REC perspective, participants were engaging in a single event, which manipulated relations in this context, by temporarily combining relational networks in an additive manner (bringing in other relations) around a response.

The intervention condition consisted of four short audio clips with questions written on a hand-out for participants to complete in-between each clip, which asked them to write down advice to help their friend change his beliefs (see Appendix c).

The control condition was designed to replicate the structure of the intervention condition. For example, the audio clips were designed to be administered in four short clips, leaving the same amount of time between them to write down responses to the questions. The questions within the control condition were constructed in order to replicate the style of questions in the intervention condition. For example, participants were asked to write down what they would say to their 'friend' in order to change his beliefs (see Appendix d).

2.6 IRAP Stimulus Set

The IRAP stimulus set was designed to target rape-supportive verbal relations and aimed to link to similar constructs within the AMMSA more specifically. As it was a first study, stimuli were designed to cover a broad range of rape-supportive beliefs, with a view to narrowing focus in subsequent studies. For example, the stimuli, *Women are deserving of rape*, *Women are glad to be raped*, and *Women are not damaged by rape*, were developed to target the broad concept of 'denial of the scope of the problem' within the AMMSA. The stimuli *Women are deceitful*, *Women are not honest*, *Women are liars*, and *Women are not truthful*, were developed in order to tap into the related AMMSA construct 'exonerating the

perpetrator by blaming the victim or the circumstances.’ For example, by perceiving women as deceitful links to the question in the AMMSA that relates to women making false accusations of rape. The stimuli *Women are slutty*, and *Women are not faithful*, also relates to the construct ‘exonerating the perpetrator by blaming the victim or the circumstances’ because viewing women in sexually provocative and promiscuous terms enables men to place greater responsibility on women for ‘inviting’ sexual aggression.

The stimuli *Women are weak*, *Women are not strong*, *Women are subordinate*, and *Women are not equal*, link to the following construct in the AMMSA that when it comes to sexual contact, women expect men to take the lead as they are submissive to men. The stimuli *Women are subordinate*, and *Women are not equal*, also tap into the AMMSA construct that ‘male coercion forms a natural part of sexual relationships.’ The stimuli, *Women are not direct*, and *Women are teasers*, relate to the constructs within the AMMSA that ‘women like to play coy, but it does not mean they do not want sex’ and ‘women invite men in for coffee, meaning instead they really want sex.’

Stimuli were validated by clinical psychologists that have completed research in the field. Following this process, the stimuli were pilot tested on two male volunteers to check that they could be both understood and responded to within the response latency provided. This process enabled confirmation that the stimuli were understood, and that the length and complexity of target words were feasible to use in an IRAP programme with a response time cut off equalling 2500ms and an accuracy target of at least 75%.

2.7 Procedure

Table 9 shows a detailed account of the procedure for participants in both the intervention group and the control group from the beginning of the study to completion.

Table 9

Procedure for Participants in the Intervention and Control Conditions

<ol style="list-style-type: none"> 1. All participants complete the AMMSA online (phase one), which includes a participant information sheet (see Appendix e) and online consent form (see Appendix f) 2. Participants that opt in leave an email address on a separate web page which enables the researcher to contact them to arrange meeting for phase two 3. Participants that come to meet the researcher at the university for phase two sign a consent sheet which signifies their consent to participate in phase two (see Appendix g). 4. Participants are given an overview of the three tasks they will be completing before being briefed on the IRAP. Participants then start the practise blocks of the Pre-IRAP with the researcher present to check they are able to complete it and to see if they have any questions 5. The researcher leaves the room 6. Participants complete the main test blocks of the Pre-IRAP and once finished, they notify the researcher they have completed that part of the study 	
Intervention condition	Control condition
<ol style="list-style-type: none"> 7. Participants that have been randomly assigned to the intervention condition are then given another laptop with the audio files relating to sexual aggression ready to play. 8. Participants are given instructions that there are four audio clips, each one lasting less than a minute, and they are to imagine that the actor in the clips is their friend. They are then instructed that their task is to write down what they would say to their friend to change his beliefs following each clip. At this point 	<ol style="list-style-type: none"> 7. Participants that have been randomly assigned to the control condition are then given another laptop with the audio files relating to the job interview ready to play. 8. Participants are given instructions that there are four audio clips, each one lasting less than a minute, and they are to imagine that the actor in the clips is their friend. They are then instructed that their task is to write down what they would say to their friend to change his beliefs following each clip. At this point

the researcher points to each question in the hand-out booklet (see Appendix c) and shows how many questions there are and what they are asking. They are then asked if they have any questions relating to the task before the researcher leaves the room.

the researcher points to each question in the hand-out booklet (see Appendix d) and shows how many questions there are and what they are asking. They are then asked if they have any questions relating to the task before the researcher leaves the room.

9. Once they have completed the task they contact the researcher again who then comes back to set up the Post-IRAP on the same laptop they completed the Pre-IRAP on.
10. Participants then complete the Post-IRAP (which is exactly the same stimulus set as in the Pre-IRAP). They are given practice blocks to complete with the researcher present again first.
11. The researcher then leaves the room and participants move on to complete the Post-IRAP main test blocks.
12. Once completed participants notify the researcher who comes back into the room.
13. Participants are then informed it is the end of the study and are given £5 and a receipt to sign to say they have received it.
14. Participants are asked if they would mind completing a short evaluation form of the researcher, to be used at her annual review. They are informed that the researcher will not see what they write as they are to put it into a sealed envelope following completion.
15. The researcher starts to pack away materials in a different part of the room so as to give participants privacy when completing it to reduce bias.
16. Participants are then informed that the study has been funded by a social enterprise fund and that as part of this funding, it has been agreed that each participant is has the opportunity to donate three tokens to the charities on the boxes on the table near the door. They are informed that the social enterprise has stated that they will match the tokens donated with a monetary contribution and they are free to donate any of their tokens to any of the boxes in front of them. The researcher then briefly reads the labels on front of each box (in the same order each time) so the participant is aware

- of what each charity-box stands for.
17. The participant then places their three tokens into any of the four charity boxes. At this point the researcher moves to pack materials away again on the other side of the room so as to reduce biasing their responses.
 18. The researcher then gives the participant a debrief form (see Appendix h), whilst verbally explaining what the study was about.
 19. The researcher then checks out with the participant that they are ok and answers any questions they have following completion in the study.
-

2.7. 1. Ordering of implicit and explicit measures. Previous research using the IAT has found a relation between the IAT and explicit measures in relation to prejudice (McConnell & Leibold, 2001). This effect has been attributed by McConnell and Leibold (2001) to the transparency of the IAT sensitising participants to the purpose of the research and therefore completion of the explicit measure is deemed to be influenced more by social desirability concerns than if it was completed before the IAT. It was considered that the same bias might also be applicable to the IRAP. In an attempt to reduce this potential bias, the AMMSA was completed first prior to completion of phase two. Whilst it is of course possible that this may have influenced their responses on the IRAP, it was considered the least biased option. Further research is needed to investigate the effects of ordering implicit and explicit measures when conducting research using the IRAP.

2.8 Ethics

Ethical approval was granted by the ethics committees at both University sites (see Appendix i and Appendix j). With regards to informed consent, an information sheet containing information about the nature of the study, how the data will be stored, processed and used, and information regarding their right to withdraw, was provided online during phase one of the study (see Appendix f), and provided in written form during phase two (see Appendix g). Informed consent was required for participation in the research; during

phase 1 of the research, consent was obtained by completing online tick-boxes. During Phase 2, participants were provided with further written information and asked to sign a consent form. The researcher's details were provided on the consent form in case participants had any questions about the study prior to giving their consent.

Participants were informed of their right to withdraw their data at any point in the week following participation in the study. Participants were informed that withdrawal of their data would not result in the removal of payment with regards to compensation of their time or any other adverse effects. Regarding the storage of data and confidentiality, participants generated their own unique code from the first three letters of their mother's maiden name and the numerical form of their birth month (eg. March was coded as 03). This unique identifier code was used to link participants' data together across testing conditions and phases of the research. The data was only available to the researcher and research supervisors. Online consent forms did not contain personally identifiable information that linked to participants' data in any way. Participants were made aware that if they decided to withdraw from the study they were able to request that any personal data be destroyed. Data was stored in a locked filing cabinet at Lincoln University. Electronic data was password protected. Data was to be stored for seven years in archives following the completion of the study. Participants were able to give their preferred email address to be contacted on if they won the prize draw, by following a link at the end of the online questionnaire which took them to a separate web page in order keep their personal information separate from their data.

Elements of deception were used in the study. For example, participants were informed that the study was investigating attitudes towards sex and they were not told about the true aims of the study (until they were debriefed). It was not felt that this degree of deception would have an adverse effect on participants, as the audio-clips within the intervention condition were not deemed to be too dissimilar to scripts they might be exposed to on television. Other elements of deception related to the RRS

and the charity-box task. For example, participants were falsely informed that the experiment had finished prior to completion of the RRS and charity-box tasks. This deception was considered necessary in order to gain a less biased assessment of behaviour and was not envisaged to cause harm to potential participants. In order to minimise the impact of deception, participants were informed at the earliest opportunity as to the true nature of the research and were fully debriefed following completion of the study.

The debriefing following completion of the study, involved informing participants that the aim of the study was actually to investigate the malleability of implicit beliefs and to see whether implicit beliefs, as measured by the IRAP, predicted behaviour (eg. RRS score and charity donation). The debriefing sheet included contacts for support such as university counselling services in case they were required (see Appendix h). Participants were also informed that they would be able to receive feedback regarding the results of the study, if they wished to do so.

A risk assessment was conducted prior to the start of the study. The risk of physical harm to participants was considered to be low as it was deemed unlikely that participation would present any increase in risk of physical harm when compared to the risk generally at the university. Risk of psychological harm to participants was also deemed to be low. Whilst there was the potential for participants to find some questions embarrassing, they were made aware of this when told that the study was about attitudes towards sex, prior to obtaining informed consent. Participants were not informed of their AMMSA or IRAP scores in order to protect them from any potentially distressing results.

3 Extended Results

3.1 Test Assumptions: Mann-Whitney U Test

Non-parametric tests have fewer required assumptions than their parametric counterparts. The Mann-Whitney U test was chosen to analyse the differences between scores on the AMMSA in both groups because the scores were not normally distributed,⁵ which is often the case in social science research (Field, 2009). Non-parametric tests, however, are less sensitive than parametric tests and so more open to Type 2 errors. Stevens (1996) reported that once a sample size reaches 100, power is not a concern. As the sample size was relatively large ($N = 143$) in this part of the study, it is likely to have been powered sufficiently and in support of this claim, significant results were obtained. An assumption when using non-parametric tests is that the data is comprised of independent observations (Pallant, 2007). This means that each measurement must not be influenced by any other measurement. The measurements were independent of each other in the current study; therefore, the data met this assumption.

3.2 Test Assumptions: Pearson's r

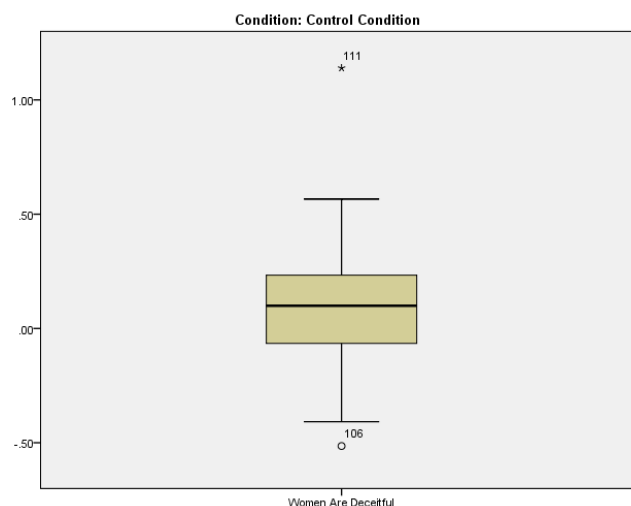
Pearson's r requires that the data are at least interval for it to be an accurate measure of relationships (Field, 2009). Interval data means that data is measured on a scale in which the intervals are equal. This assumption was met in the current study. In order to test for significance further assumptions are required; we have to check to see if the sample is normally distributed (Field, 2009).

⁵ Whilst the scores were only marginally divergent from a normal distribution, a non-parametric alternative was available and so used in this case to answer this particular question.

3.3 Managing Outliers

It is necessary to check for outliers when applying correlation analyses to small samples (Pallant, 2007). Boxplots are useful to compare the distribution of scores visually (Pallant, 2007) and so they were generated to check each variable for outliers in the current study. Outliers denote scores that are 1.5 times the interquartile range (the middle 50% of data) away from the median. Extreme scores are scores that are more than three times the interquartile range away from the median. As recommended by Field (2009), extreme outliers can be converted to a score which is three times the standard deviation added to the mean, so as to reduce skew in the distribution. In the current study, extreme scores were deemed informative to the aims of the study and so the decision was made to keep the extreme outlier in the data. It was not removed altogether because there were no reasons to assume that the score had been produced in error, such as under conditions which posed threats to validity (eg. distractions). Therefore, the extreme outlier was converted to a score of three times the standard deviation added to the mean (see Figure 4).

Figure 4. A Boxplot Highlighting an Extreme Outlier



3.4 Normality

Histograms were generated in order to visually check for the normality of distribution, as recommended by Field (2009).

Whilst Histograms enabled visual checking of the data, they are subjective and so for further assurance, I moved on to quantify the shape of the distribution by exploring skewness, kurtosis and by running the Kolmogorov-Smirnov test.

3.5 Skewness and Kurtosis

In normally distributed data, the value of skewness and kurtosis should be zero (Field, 2009). When interpreting values of skewness, positive values suggest over-representation of scores to the left of the distribution and negative values suggest overrepresentation to the right. If Kurtosis scores are positive it means the distribution is heavy tailed and pointed whereas negative scores indicate a light tailed, flat distribution (Field, 2009). Table 10 indicates that some scores did not appear to be normally distributed. For example, the distribution of scores on the RRS appeared to be heavily skewed to the right of the distribution and heavy tailed. Although it is useful to examine the skewness and kurtosis values, I decided to convert the scores to Z scores by dividing the scores by their standard error. This was in order to see how the scores compared to each other using different measures and to estimate how likely the values of skewness and kurtosis were to occur due to chance.

Table 10

Values of Skewness and Kurtosis

Variable	Skewness	SE	Kurtosis	SE
AMMSA Phase One	-.22	.24	-.60	.47
AMMSA Phase Two	.40	.39	.99	.76
Mean Pre- <i>D</i> -IRAP	-.07	.39	.03	.76
Mean Post- <i>D</i> -IRAP	.52	.39	-.42	.76
Women are honest ^a	-.37	.39	-.49	.76
Women are deceitful ^a	-.11	.39	.08	.76
Women are not honest ^a	-.05	.39	-.43	.76
Women are not deceitful ^a	-.22	.39	.86	.76
Women are honest ^b	.64	.39	.24	.76
Women are deceitful ^b	.35	.39	-.50	.76
Women are not honest ^b	.47	.39	-.68	.76
Women are not deceitful ^b	.02	.39	-.39	.76
Mean RRS	-2.46	.39	5.85	.76
Blue Silence Foundation	.07	.39	-.20	.76
Women's Refuge	.52	.39	.54	.76
Overall Giving Score	.74	.39	1.01	.76

Note. AMMSA = Acceptance of Modern Myths of Sexual Aggression Scale; AMMSA Phase One ($N = 104$); AMMSA Phase Two ($N = 37$) RRS = Researcher Rating Scale.

^a = Individual Trial-type Pre-*D*-IRAP Mean Score; ^b = Individual Trial-type Post-*D*-IRAP Mean Score.

3.6 Calculating the Shape of the Distribution Using Z-Scores

A Zskewness or Zkurtosis value of greater than 1.96 is not expected by chance and is significant at the $p \leq .05$ level. Likewise, a value above 2.58 is significant at the $p \leq .01$ level and a value above 3.29 is significant at the $p \leq .001$ level. The Z scores representing skew and kurtosis for variables in the current study are presented in Table 11. It is very clear from looking at the Zskewness and Zkurtosis values for the RRS (-6.31 and 7.70 respectively), that the data is significantly skewed, towards the right of the distribution, and is pointy with a heavy tailed distribution.

Table 11

Values of ZSkewness and ZKurtosis

Variable	ZSkewness	ZKurtosis
AMMSA Phase One	-.92	-1.28
AMMSA Phase Two	1.03	1.30
Mean Pre- <i>D</i> -IRAP	-.18	.04
Mean Post- <i>D</i> -IRAP	1.33	-.55
Women are honest ^a	-.95	-.64
Women are deceitful ^a	-.28	.11
Women are not honest ^a	-.13	-.57
Women are not deceitful ^a	-.56	1.13
Women are honest ^b	1.64	.32
Women are deceitful ^b	.90	-.66
Women are not honest ^b	1.21	.89
Women are not deceitful ^b	.05	-.51
Mean RRS	-6.31***	7.70***
Blue Silence Foundation	.18	-.26
Women's Refuge	1.33	.71
Overall Giving Score	1.90	1.33

Note. AMMSA = Acceptance of Modern Myths of Sexual Aggression Scale; AMMSA Phase One ($N = 104$); AMMSA Phase One & Two ($N = 37$) RRS = Researcher Rating Scale.

^a = Individual Trial-type Pre-*D*-IRAP Mean Score; ^b = Individual Trial-type Post-*D*-IRAP Mean Score.

* $p \leq .05$. ** $p \leq .01$. *** $p \leq .001$

3.7 Shapiro-Wilk Test

The Shapiro-Wilk test was performed on the data in order to explore whether the distribution for each variable as a whole deviated from a comparable normal distribution. Test statistics and significance values are provided for each variable in Table 12. The Shapiro-Wilk statistic was significant for the following variables: the AMMSA sample ($N = 37$); mean RRS, Blue Silence Foundation, Women's Refuge and Overall Giving. This means that the data from these variables was not normally distributed. All the *D-IRAP* scores were normally distributed. A limitation of the Shapiro-Wilk test is that if the sample is large it is easy to obtain significant results from very small deviations. The Kolmogorov-Smirnov test was carried out on the data from the large AMMSA sample ($N = 104$) as this test is deemed more appropriate for large samples than the Shapiro-Wilk test which has more power (Field, 2009). The Kolmogorov-Smirnov test statistic for the large AMMSA sample $D(104) = .05$, $p = .20$, was not significant meaning the data were normally distributed.

Table 12

Shapiro-Wilk Test of Normality

Variable	Shapiro-Wilk Statistic	Sig.
AMMSA Phase Two	.94	.04*
Mean Pre- <i>D</i> -IRAP	.98	.86
Mean Post- <i>D</i> -IRAP	.95	.12
Women are honest ^a	.96	.24
Women are deceitful ^a	.97	.51
Women are not honest ^a	.99	.94
Women are not deceitful ^a	.98	.76
Women are honest ^b	.96	.21
Women are deceitful ^b	.98	.56
Women are not honest ^b	.96	.16
Women are not deceitful ^b	.98	.71
Mean RRS	.52	.00*
Blue Silence Foundation	.76	.00*
Women's Refuge	.82	.00*
Overall Giving Score	.86	.00*

Note. AMMSA = Acceptance of Modern Myths of Sexual Aggression Scale; AMMSA Phase One ($N = 104$); AMMSA Phase One & Two ($N = 37$) RRS = Researcher Rating Scale.

^a = Individual Trial-type Pre-*D*-IRAP Mean Score; ^b = Individual Trial-type Post-*D*-IRAP Mean Score.

* $p \leq .05$.

3.8 Linearity and Homoscedasticity

The relationship between the variables being analysed with Pearson's r should be linear in order for the results to be valid (Pallant, 2007). Linearity of the data can be assessed using scatterplots to see if the data forms a straight line rather than a curve. Homoscedasticity refers to the variability in scores and in order for this assumption to be met, the variability for one variable in the relationship (x) should be similar at all values of the other variable (y). Again, this can be assessed using scatterplots to check that the data form a fairly even cigar shape (Pallant, 2007). Inspection of the scatterplots indicates that the relationships between variables were linear and that homoscedasticity was met. An exception to this was found in the relationship between the Pre-*D*-IRAP score and Women's Refuge donation, where it appeared that the relationship between scores may not be linear. The modifications made to analyses are reported in the next section.

3.9 Modifications to Analyses

As the data were not normally distributed within the variable AMMSA Phase One, a non-parametric test, the Mann-Whitney U test was used to analyse the differences between the AMMSA phase one sample and the AMMSA phase two sample. With regards to the other variables that were not normally distributed, namely: Mean RRS; Blue Silence Foundation; Women's Refuge; and Overall Giving, a procedure called the bootstrap was applied when carrying out Pearson's r statistical analysis (Efron & Tibshirani, 1998). Field, Miles and Field (2013) recommend this approach for Pearson's r analysis, when data is not normally distributed. Pearson's correlation coefficient has more statistical power than Spearman's correlation coefficient and so it is preferable to use if assumptions can be met or overcome (Field et al. 2013). The problem with not having a normally distributed sample is that the shape of the sampling distribution is unknown and therefore the probability of a test statistic occurring is also unknown (Field, 2009). Bootstrapping offers a unique method of estimating the

sampling distribution. For example, bootstrapping uses the sample data to obtain estimates of the sampling distribution by treating the sample data as a population from which smaller samples (named bootstrap samples) are taken and the mean calculated from each (Efron & Tibshirani, 1998; Field, 2009). Once many samples are taken and replaced, the sampling distribution can be estimated. Confidence intervals and significance tests can be carried out using the standard error, which is estimated from calculating the standard deviation of the sampling distribution (Efron & Tibshirani, 1998; Field, 2009).

Whilst bootstrapping sidesteps the issue of distributional assumptions, further support for using Pearson's r in samples with non-normally distributed data comes from evidence that modelling has shown that Pearson's r is highly robust to non-normality (Havlicek & Peterson, 1977). For example, Havlicek and Peterson (1977) sought to study empirically, the effects of violating the assumption of normality on the Pearson product-moment correlation co-efficient. They generated populations of data using Monte Carlo procedures⁶ with varying distributions: normal; positively skewed; negatively skewed; and leptokurtic (positive kurtosis). Various sized samples were randomly selected from these generated populations and distributions of r were calculated on 5000 sets of samples with $n = 5$ or $n = 15$, and 3000 sets of samples where $n = 30$ or $n = 60$ (Havlicek & Peterson, 1977). Havlicek and Peterson (1977) reported, "for the 216 distributions of r computed in this study, there were no significant deviations from the theoretical expected proportions of r at the .005, .01, .025, or .05 levels of significance" (p.376). They concluded that their results indicated that Pearson's r is insensitive to extreme violations of assumptions of normality.

⁶ Monte Carlo procedures are a broad class of computational algorithms that rely on repeated random sampling to obtain numerical results.

3.10 Test Assumptions: Independent-Samples T-Test

In order to check for randomisation, an independent-samples t-test was used to compare the means of the pre-*D*-IRAP scores in the control and intervention conditions. In order to carry out this test it is important to check the homogeneity of variance in addition to the previous assumptions (level of measurement, independence of observations, normal distribution) as it assumes the variability is similar for each group (Pallant, 2007). Levene's test for equality of variances was performed in order to test this assumption in these variables. A non-significant score indicates that variance for the group is equal. With reference to the variability across both groups for the Pre-*D*-IRAP score, Levene's test for equality of variances proved the variability was not statistically significant ($F = .03, p = .88$).

3.11 Test Assumptions: Mixed Between-Within Analysis of Variance (ANOVA)

With regards to the mixed between-within ANOVA the same general assumptions apply as is required for the t-test. However, an additional assumption is made, namely, the homogeneity of inter-correlations. This means that the pattern of inter-correlations amongst the within-subjects variable (e.g. time) should be the same at each level of the between subjects variable (e.g. condition; Pallant, 2007). Box's M statistic tests this assumption and was applied to all analyses using the mixed between-within ANOVA. All tests of homogeneity of inter-correlations proved to be insignificant, meaning this assumption was met.

3.12 Further Analyses

3.12.1 Multiple regression. To determine the relative contributions of responses on the IRAP (the implicit measure) and the AMMSA (the explicit measure) in predicting overall giving scores and behavioural ratings of the researcher (RRS: Question 4), two separate hierarchical multiple regressions were performed (one for each dependent variable). The

regression model consisted of two independent control variables; the Pre-D-IRAP scores and Mean scores on the AMMSA. The dependent variable consisted of the overall giving score in the first regression model and score on Question 4 in the second separate regression model. Multiple regression has additional assumptions to those specified for ANOVA's; namely, no perfect multicollinearity, non-zero variance, predictors are uncorrelated with external variables, and independent errors. Predictors do not need to be normally distributed (Field, 2009). Importantly, parametric approaches are deemed robust and perform well even if assumptions of normality are violated, unless the violations are multiple and severe (Kerlinger & Lee, 2000). On that basis, it was decided that a hierarchical multiple regression could be performed.

Multicollinearity is said to occur when there is a high correlation between independent variables ($r \geq .9$; Pallant, 2007). In the current study, the variance inflation factor (VIF), and the tolerance statistic was used to assess multicollinearity. The VIF helps to diagnose if one predictor has a strong linear relationship with another predictor (Field, 2009). A value of 10 is deemed to warrant concern (Myers, 1990). The VIF value in the current study was 1.02. With regards to the tolerance statistic, values of less than .2 are worthy of concern, with values of less than .1 indicating serious problems (Menard, 1995). The tolerance value in the current study was .83, which indicates the assumption of no perfect multicollinearity, was met.

With regards to the assumption of non-zero variance, this was met in the current study as the predictors both had some variation in value. Finally, to test whether the assumption of independent errors was met, the Durbin-Watson test (Durbin & Watson, 1951) was carried out, to test for serial correlations between errors, as recommended by Field (2009). Values less than 1 or greater than 3 are deemed to be problematic. The Durbin-Watson value in the current study was 2.25, indicating that this assumption was met.

The results of the two hierarchical multiple regression analyses will now be presented. With regards to predicting variance in overall giving score,

the AMMSA variable was entered at Step 1, explaining none of the variance in overall giving score. After entry of the Pre-D-IRAP variable at Step 2 the total variance explained by the model as a whole was 12.4%, $F(2, 34) = 2.41$, $p = .11$. The Pre-D-IRAP explained an additional 12.4% of the variance in overall giving score after controlling for AMMSA score, R^2 change = .12, F change $(2, 34) = 4.81$, $p = .04$. In the final model, only the Pre-D-IRAP measure was statistically significant, $\beta = 1.38$, $SE = .63$, $p = .04$. The direction of the relationship indicates that men who respond on the IRAP in a manner that indicates rape-supportive bias are more likely to produce lower overall giving scores (a greater proportion of tokens donated to The Blue Silence Foundation than Women's Refuge).

With regards to predicting variance in Question 4 scores on the RRS, the AMMSA variable was entered at Step 1, explaining none of the variance in overall giving score. After entry of the Pre-D-IRAP variable at Step 2 the total variance explained by the model as a whole was 11.5%, $F(2, 34) = 2.21$, $p = .13$. The Pre-D-IRAP explained an additional 11.5% of the variance of Question 4 scores after controlling for AMMSA score, R^2 change = .12, F change $(2, 34) = 4.41$, $p = .04$. In the final model, only the Pre-D-IRAP measure was statistically significant, $\beta = -.31$, $SE = .15$, $p = .04$. The direction of the relationship indicates that men who respond on the IRAP in a manner that indicates rape-supportive bias are more likely to rate the researcher as more approachable (as indicated by higher scores on Question 4 of the RRS).

3.12.2 Relationship between age and AMMSA. The relationship between age and reported acceptance of myths of sexual aggression was investigated in the large AMMSA sample ($N = 143$) as an additional analysis, using Pearson's r . Assumptions as previously noted (see section 3.2) were met including those for the variable of "Age". There was a significant medium negative correlation between the two variables, $r = -.23$, $N = 143$, $p = .01$, with increased age being associated with lower levels of adherence to myths of sexual aggression.

3.13 Type II Errors

As the research was exploratory and was investigating a novel methodology in the field of rape-supportive implicit cognition, when exploring the data, analyses were carried out to test a two-tailed hypothesis. This was in order to reduce the likelihood of Type II errors occurring, as power was limited due to having a relatively small sample size (Pallant, 2007).

4 Extended Discussion

4.0 Additional analyses, namely hierarchical multiple regression were conducted to explore the degree to which the Pre-*D*-IRAP scores and the AMMSA independently and combined, explained the variance in the overall giving score and Question 4 on the RRS. Previous research in the area of rape-supportive cognition has found that the implicit measures and explicit measures predicted self-reported sexual aggression more effectively when combined (Nunes, Hermann & Ratcliffe, 2013). The current findings were not in accordance with this previous research as only the implicit measure explained any variance for both behavioural outcome measures. This might be related to different types of behaviour being tapped into, for example in the present study behaviour was overt whereas in Nunes' et al. (2013) study, sexual aggression was measured using a likelihood to rape measure which might be measuring more elaborative responses than in the present study. This was not assessed because, including a measure of self-reported likelihood to rape or past sexual aggression was deemed to change participants' understanding of the true nature of the study and could have confounded the findings. Randomisation of participants was successful and so this will have accounted for any differences between groups in terms of previous engagement in sexual aggression. The finding that the AMMSA did not add anything to the model of predicting behaviour may indicate that it is

not as good at reducing social desirability as it claims to. For example, if it limited the effects of social desirability then it would be expected that results would be more convergent with the IRAP in the domain of socially-sensitive research such as this. These findings might also indicate that the behaviour being measured by the charity-box task was made up of responding which was towards the brief and immediate end of the spectrum. However, this is speculative, without a measure of time taken to respond.

A significant negative correlation was found with regard to age and AMMSA scores. This was in contradiction to previous research which suggests rape-supportive attitudes get stronger in older cohorts. This might be due to current increase in rape-supportive attitudes more generally with greater access to online pornography. Further research is needed to investigate this using a controlled design. Another explanation might be that the older men who took part in the study were not a representative sample of older men in the community. For example, men that took part were recruited from a university sample and so those that stay working within academic institutions are likely to be highly intelligent. High intelligence tends to be linked to a greater adherence to liberal attitudes and less adherence to right-wing/authoritarian attitudes, the latter of which are linked with adherence to rape myths (Suarez & Gadalla, 2010).

4.1 Previous research findings in the area can be explained by the REC model. For example, Blake and Gannon's (2010) conclusions that the lack of evidence of Polaschek and Ward's (2002) implicit theories model of rape-supportive cognition⁷ within a community male sample meant that their theory might need revising. From an REC perspective Blake and Gannon's (2010; 2012) findings can be explained by defining the type of behaviour captured within the rape proclivity measure, which appears to be based on extended and elaborate relational responding (elaborated explicit

⁷ Polaschek and Ward's (2002) implicit theories model proposes that implicit rape-supportive cognitive structures in memory serve to bias information processing and behaviour.

responses). For example, the rape proclivity measure (Bohner et al., 1998) requires participants to read date-rape scenarios, imagine themselves in that scenario as the perpetrator, before evaluating using a Likert scale, the degree to which they would have acted in the same way. It is not surprising that only the explicit measure captured this (Blake & Gannon, 2010) as the REC model would predict that the explicit measure captures elaborated explicit responses and so can predict elaborated behaviour. Whereas, from a REC perspective, the effects generated on the lexical decision task would be predicted to capture brief and immediate behaviour. In order to clarify their findings it would be useful for future studies to employ a research design which enables brief and immediate behaviour to be predicted, such as spontaneous behaviours.

For instance, one idea might be to simulate a court room scenario by getting participants to put themselves into the role of a judge and then to read fictional defence and prosecution case summaries of rape. Under time pressure they might then be asked to determine the offender's sentence length. Alternatively, also under conditions of time pressure they could be asked to rate how responsible they believe the perpetrator to be in the commission of the offence and how responsible they find the victims to be. Difficulties may exist with regards to determining what time constraint would be necessary for behaviour to be considered brief, rather than elaborate responding. Further research is currently being completed to address this within the IRAP field. The intellectual ability of each participant may also serve as a confounding factor, as individual differences in processing speed ability is hypothesised to affect the individual time constraints needed to ensure brief rather than elaborate responding.

Whilst Blake and Gannon's (2012) study found one attitude (implicit theory), namely, that *women are sex objects*, predicted scores on the rape proclivity measure, the REC model does not separate out brief implicit responses from elaborated explicit responses, instead it proposes they are on a continuum. So the finding that one result from a priming procedure predicts elaborated behaviour is not beyond the scope of explanation from

an REC perspective. To illustrate this point further, Widman and Olson's (2013) findings that a priming task predicted past sexually aggressive behaviour such as sexual assault and rape, in college men and community males, beyond the capacity of the direct/explicit measure, appears to contradict Blake and Gannon's (2010; 2012) findings. However, by applying the REC model to explain their results it can be seen that Widman and Olson (2013) measured behaviour to be predicted using a self-reported sexual experiences survey (SES; Abbey, Parkhill, & Koss, 2005) which might not require the same degree of elaborated explicit responses as in the rape proclivity measure. For example, it appears that simply asking if they had previously engaged in a behaviour would involve a less extended and elaborated relational response than asking them to imagine themselves in a scenario and evaluate their response to being in that situation (as in the rape proclivity measure), and so findings would be better predicted by a measure that captured conceptually related brief implicit responses than explicit/indirect measures, as found in Widman and Olson's (2013) study. The current study produced results in congruence with this explanation.

4.2 Limitations with regard to the current study relate to the lack of specificity in measuring time taken to produce behavioural responses. Arguably, if this had been measured then it would be more informative regarding whether elaborate or brief relational responding was being targeted.

Another limitation could possibly be the choice of control condition. Potentially the context of a job interview situation might elicit anxiety or relational frames relating to failure which might have impacted upon the context of the IRAP. Some research points to the links between negative emotional states and sexual aggression (Thornton, 2002). Therefore it may have served to confound any post treatment effects.

The IRAP stimulus set elicits further potential limitations. For example, following the data collection it emerged that some participants made

comments relating to the use of the stimulus *Women are subordinate* within the IRAP. In particular, participants discussed how there exists a dominant social discourse surrounding this view of women. Therefore, it is possible that responses to this stimulus might reflect an overlearned, societal truth, rather than any individually-held, rape-supportive beliefs.

The limitations of using the same target words for each individual, and arguably taking a more structuralist approach, are that individually-relevant rape-supportive attitudes may not have been assessed by the IRAP. For example, if a participant held particular beliefs such as, "Women that stay out late at night should realise they are asking for trouble (i.e. rape)," then the IRAP would not have captured this particular rape-supportive belief. Therefore, if such a belief had been the target of an intervention, then the pre and post testing contexts would not have been able to provide any information in terms of the malleability of that particular belief for that individual. The implications of this in the current study are that the results are less informative regarding the degree to which personally-relevant implicit beliefs are malleable within the current intervention context and regarding how well they predict behaviour. However, whilst this was a limitation of the study design, the benefits of adopting this design are that larger-scale analysis at the group-level could be conducted. It is recommended that future studies seek to employ idiographic IRAP stimuli and interventions, to explore this further.

Personal Construct Theory (PCT; Kelly, 1955) offers an interesting assessment approach to the idiographic selection of target words. The repertory grid (Kelly, 1969) is an instrument designed to capture the ways people make sense of their experiences in their own words. They can be useful in the study of the individual in idiographic detail, but also for the comparison of different respondents. The grids can be designed by the researcher to consist of elements that represent the area of investigation; in this case, this might be restricted to views of women and rape. The grids consist of a set of personal constructs that the participant uses to compare and contrast these areas, and a rating-system that evaluates the elements

in terms of how they are positioned in relation to the poles of each construct. This approach is particularly suitable for the elicitation of idiographic target words for individualised IRAPs, which require the identification of polarised beliefs, and can specify a person-centred focus for intervention. For example, individualised beliefs around women and rape can then be directly focussed upon in treatment. The systematic and thorough analysis within this approach lends itself well for use in research.

The influence of context in the eliciting of implicit responses is a key theoretical issue. The REC model proposes that contexts that include restrictions on time and accuracy will result in a greater likelihood that responses will be of low complexity and derivation. This is the theoretical assumption which underpins the IRAP. However, repeated administration of the IRAP provides increasing opportunity to derive particular relations, thus the speed and accuracy of the emitted response also increases (Barnes-Holmes et al., 2010). From this, it is hypothesised that maintaining the same time restrictions on responding over successive testing contexts, may result in responses being characterised by lower derivation. This may have been a limitation within the current study and highlights the need for further research to explore the nature of the effect of derivation reducing across repeated IRAP testing contexts, with a view to understanding how to adjust time constraints accordingly over successive tests.

Barnes-Holmes et al., (2010) investigated the impact of context on implicit racial attitudes. They found that manipulating the context of the IRAP, by administering it in public and private contexts, had a significant impact upon response latency; participants in the private context were found to respond more slowly than those in the public context. They concluded that this suggested that their unexpected finding of less racial stereotyping in the private context was explained by slower response latencies (meaning more elaborated, socially-desirable responses could unfold). These findings were consistent with the REC model. In the current study, this effect was minimised by employing modified IRAP software,

which includes feedback on response latency. For example, if participants respond too slowly the message "too slow" appears on screen.

Having the charity-box task at the end of the study may have left too much complex processing time between the pre-IRAP and the behavioural task, increasing the potential confounding factors. In hindsight, it might have been a better design to have the charity-box task right at the start, so as to reduce potential confounding factors, although this would have to be considered alongside the fact that it may appear to be a part of the study if it is conducted at the start.

4.3 Strengths of the study lie in its novelty of design and measurement procedure. For example, it is the first study in the area of sexual aggression to predict actual, in vivo, behaviour using the IRAP.

4.4 As this is a first study, further research is needed to explore the relationship between brief implicit responses and sexual behaviour as the findings are limited in terms of their generalisability, for example, from behaviour in the form of: providing less support to female victims of partner violence/sexual violence (*women's refuge charity-box*) in relation to male victims of sexual abuse; to acting out sexual aggression. If brief implicit responses are related to sexual aggression then it will be prudent to investigate this further in line with the Risk Needs Responsivity principles (Bonta & Andrews, 2007) to ascertain if actuarial risk has any mediating effect on behavioural prediction from the IRAP. Potentially if brief implicit responses indicate a rape-supportive bias to rape-supportive attitudes, and are linked to behaviour, then this should be the focus of future treatment. However, further research is needed to determine specifically what types of behaviour can be predicted, in whom, under what circumstances, in order to be more useful clinically. Further evaluation of the types of brief implicit responses that are pertinent in high risk sexual behaviour and how to modify them is warranted. Especially as the current treatment approaches

often base their effectiveness on self-reported measures, as it is hard to measure actual base rates of reoffending amongst treated individuals, upon release in the UK.

5 Critical Reflection

As I was carrying out this research I spent time reflecting on my reasons for conducting research in this area. I had previously completed my dissertation for the Masters in Applied Forensic Psychology qualification on implicit cognition in rapists but had approached the topic from a completely different perspective. For example, I had used a qualitative methodology, Interpretative Phenomenological Analysis (IPA) to investigate offenders' construction of how known dynamic risk factors manifested. From this research I had recommended that future studies should employ experimental paradigms to investigate this under-researched area. I was excited at the prospect of being able to achieve this whilst on the doctorate programme. I have previously worked in forensic settings such as prisons and secure hospitals facilitating the sex offender treatment programme and from this my interest was sparked in terms of researching offence-supportive attitudes. My work in these settings informed me of the limited knowledge available regarding offense-supportive attitudes and I was surprised that this was the case given the emphasis placed on restructuring cognition within treatment groups. This interested me further and I think formed the basis for my decision to investigate the malleability of implicit cognition using the IRAP.

My interest in using measures that are less sensitive to socially-desirable responding developed when completing risk assessments of sexual offenders, which I defended at oral hearings. For instance, I have sometimes found the information available to me, regarding how strongly

an offender holds offence-related attitudes following the completion of sex-offender treatment, to be limited by the potential for socially-desirable responding. The implications of over-estimating an offender's progress in treatment and subsequent changes in dynamic risk could mean that an offender is released when their risk is too great. Alternatively, an offender may have made progress in treatment but perhaps is less able to communicate subsequent changes in attitudes (as measured by explicit responding), due to poor social skills or intellectual difficulties. This could result in the over-estimation of current risk, potentially resulting in detaining the offender for longer periods than is necessary, impacting on the offender's human rights, and at extra cost to the public. The importance of developing research in the area of implicit cognition using relational measures that can ascertain the nature of responding in particular groups, the likelihood of that particular response being emitted, and the subsequent relevance for the prediction of behaviour appeared, to me, to be important.

I am acutely aware of the difficulties in using a measure such as the IRAP to make decisions about risk and do not believe that it will ever be used in this way, given the reliability issues. However, I can see how it might potentially add to information from other existing measures in the future if issues with reliability and validity are ironed out.

An additional reason I was attracted to a quantitative approach from a behavioural paradigm was perhaps due to my own epistemological position in relation to scientific knowledge. For example, I seek to generate and test hypotheses with the aim of finding an objective reality and accordingly employ a positivist approach in my research. The behavioural framework enables predictions about the world to be made and tested with the aim of providing further knowledge that feeds back into theories that can explain a wide range of phenomenon. This approach fits well with my personal approach to science.

I have developed my understanding of many aspects of research through the process of completing this study. For example, I have noticed an improvement in my ability to critically evaluate past research and attribute

this to the amount of reading I have engaged in. I have significantly developed my understanding of relational frame theory and the REC model and this new knowledge has enabled me to evaluate past research through the lens of the REC, and RFT models. During this process I learnt how difficult it is to generate empirical tests of theories such as the REC model, and associative models as, often findings can be explained from many different perspectives. However, my interest in acceptance and commitment therapy has been heightened as a result and so I am pleased to be able to direct my learning within my research into new areas within clinical practice.

Through the process of completing the research from the initial proposal to the final portfolio, I have encountered many points at which I have had to make decisions which shaped the rest of the research. For example, I have learnt of the importance of considering all possibilities when planning data collection. I conducted my data collection in March and whilst I had recognised that many students would be leaving the universities to return to their hometown over the summer, I had not anticipated how difficult recruitment would be around exam period shortly before they returned home. This meant that my research schedule was delayed. In future when carrying out research I will ensure I plan carefully when to collect data in given populations and will investigate thoroughly any potential events that may impact on recruitment.

Another decision I have reflected on related to planning the time and location of data collection. I learnt that at times towards the end of the day, other members of staff in the vicinity left work, leaving me on my own. On one occasion I had been locked in the building as they had locked up early. This raised my awareness of the potential risks associated with conducting research, particularly when it relates to a sexual nature. For example, I noticed that some participants appeared to be attracted to the study because it was about sex and this left me wondering about how to best manage the potential risks of conducting research of this nature. Following this, I ensured that I did not meet participants later than 3pm, and despite

the impact imposing additional time restrictions had on my data collection, I decided that it was important to implement this precaution for my own safety. It is important to learn from these experiences and to ensure that lone working procedures are developed when planning research. I reflected on the emphasis that is placed on ethical considerations for participants when applying for ethical approval, which is understandable given the inherent power imbalance when carrying out research. However, it made me consider the importance of ensuring personal protection is equally accounted for.

A difficult decision I had to make when designing the research was how to best measure conceptually relevant behaviour. Given the nature of the study I found it difficult to generate ways of capturing behaviour that were ethically acceptable. The most useful measure would be to track participants and observe the frequency with which they engage in sexually aggressive behaviour. However this approach is unethical and completely unfeasible. This left me with the task of choosing behaviours which could be deemed relevant. I overcame this difficulty by discussing options with other professionals in order to generate ideas, which I found particularly helpful. In hindsight, I would have liked to have specified more precisely the types of behaviours I was analysing, for example, the degree to which the behaviours were elicited from brief relational responses or elaborated relational responses. I could have achieved this by imposing a time restriction on the behavioural outcomes such as the charity-box task. I hypothesise that this would have enabled the IRAP to predict this type of behaviour more precisely as the behaviour would be more likely to be elicited from brief and immediate relational responses. This is an avenue worthy of exploration in future IRAP studies.

Extended References

- Abbey, A., Parkhill, M. R., & Koss, M. P. (2005). The effects of frame of reference on responses to questions about sexual assault victimization and perpetration. *Psychology of Women Quarterly*, *29*, 364–373.
- Ajzen, I., & Fishbein, M. (1977). Attitude-behaviour relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, *84*, 888-918.
- Barnes-Holmes, D., Murphy, A., Barnes-Holmes, Y., & Stewart, I. (2010). The implicit relational assessment procedure: Exploring the impact of private versus public contexts and the response latency criterion on pro-white and anti-black stereotyping among white Irish individuals. *The Psychological Record*, *60*, 57-80.
- Beck, A. T. (1976). *Cognitive therapy of the emotional disorders*. New York, NY: New American Library.
- Blair, I.V., Ma, J.E., & Lenton, A.P. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology*, *81*(5), 828-841. doi:10.1037//0022-3514.81.5.828
- Blake, E., & Gannon, T. A. (2010). The implicit theories of rape-prone men: An information-processing investigation. *International Journal of Offender Therapy and Comparative Criminology*, *54*(6), 895-914. doi:10.1177/0306624X09347732
- Blake, E., & Gannon, T. A. (2012). Investigating the implicit theories of rape-prone men using an interpretative bias task. *Journal of Legal & Criminological Psychology*. doi: 10.1111/j.2044-8333.2012.02056.x
- Bohner, G., Reinhard, M.-A., Rutz, S., Sturm, S., Kerschbaum, B., & Effler, D. (1998). Rape myths as neutralizing cognitions: Evidence for a causal

impact of anti-victim attitudes on men's self-reported likelihood of raping. *European Journal of Social Psychology*, 28, 257-268.

Bonta, J., & Andrews, D. A. (2007). *Risk-need-responsivity model for offender assessment and rehabilitation* (User Report 2007-06). Ottawa, Ontario: Public Safety Canada.

Bumby, K. (1996). Assessing the cognitive distortions of child molesters and rapists: Development and validation of the RAPE and MOLEST scales. *Sexual Abuse: A Journal of Research and Treatment*, 8, 37-54.

Burt, M. R. (1980). Cultural myths and supports for rape. *Journal of Personality and Social Psychology*, 38, 217-230.

Cortoni, F., Hanson, R. K., & Coache, M.E. (2010). The recidivism rates of female sex offenders are low: A meta-analysis. *Sexual Abuse: Journal of Research and Treatment*, 22(4), 387-401.

Clark, D. A. (2013). Cognitive restructuring. In S. Hofmann & D. Dozois (Eds.), *The Wiley handbook of cognitive behavioral therapy* (pp. 1-22). doi: 10.1002/9781118528563.wbcbt02

Cullen, C., Barnes-Holmes, D., Barnes-Holmes, Y., & Stewart, I. (2009). The implicit relational assessment procedure (IRAP) and the malleability of ageist attitudes. *Psychological Record*, 59(4), 591-620.

De Houwer, J. (2006). What are implicit measures and why are we using them? In R. W. Wiers & A. W. Stacy (Eds.). *The handbook of implicit cognition and addiction* (pp. 11-28). Thousand Oaks, CA: Sage.

Dovidio, J. F., Evans, N., & Tyler, R. B. (1986). Racial stereotypes: The contents of their cognitive representations. *Journal of Experimental Social Psychology*, 22, 22-37.

Drake, C. E., Kellum, K. K., Wilson, K. G., Luoma, J. B., Weintstein, J. H., & Adams, C. H. (2010). Examining the Implicit Relational Assessment

- Procedure: four preliminary studies. *The Psychological Record*, 60, 81-100.
- Durbin, J., & Watson, G S. (1951). Testing for serial correlation in least squares regression, II. *Biometrika*, 30, 159-178.
- Efron, B., & Tibshirani, R. J. (1998). *An introduction to the bootstrap/Bradley Efron, Robert J. Tibshirani*. Boca Raton, FL: CRC Press.
- Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (pp. 75-109). New York: Academic Press. doi:10.1016/S0065-2601(08)60318-4
- Fazio, R. H. (2007). Attitudes as object-evaluation associations of varying strength. *Social Cognition*, 25(5), 603-637.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behaviour Research Methods*, 41(4), 1149-1160. doi:10.3758/BRM.41.4.1149.
- Field, A. P. (2009). *Discovering statistics using SPSS: (and sex and drugs and rock and roll)*. London: SAGE.
- Field, H. S. (1978). Attitudes toward rape: A comparative analysis of police, rapists, crisis counselors, and citizens. *Journal of Personality and Social Psychology*, 36(2), 156-179.
- Field, A. P., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. London: SAGE.
- Flores, S. A. and M. G. Hartlaub (1998). Reducing rape-myth acceptance in male college students: A meta-analysis of intervention studies. *Journal of college student development*, 39,(5), 438-448.

- Foroni, F., & Mayr, U. (2005). The power of a story: New, automatic associations from a single reading of a short scenario. *Psychonomic Bulletin and Review*, *12*(1), 139-144.
- Friese, M., Hofmann, W., & Wänke, M. (2008). When impulses take over: Moderated predictive validity of explicit and implicit attitude measures in predicting food choice and consumption behaviour. *British Journal of Social Psychology*, *47*, 397-419.
- Gawronski, B., & Bodenhausen, G. V. (2007). Unraveling the processes underlying evaluation: Attitudes from the perspective of the APE model. *Social Cognition*, *25*, 687-717.
- Gawronski, B., Geschke, D., & Banse, R. (2003). Implicit bias in impression formation: Associations influence the construal of individuating information. *European Journal of Social Psychology*, *33*(5), 573-589.
- Gawronski, B., & LeBel, E. P. (2008). Understanding patterns of attitude change: When implicit measures show change, but explicit measures do not. *Journal of Experimental Social Psychology*, *44*(5), 1355-1361. doi:10.1016/j.jesp.2008.04.005
- 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. doi:10.1002/ab.20195
- Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the implicit association test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, *97*(1), 17-41. doi:10.1037/a0015575
- Han, H. A., Czellar, S., Olson, M. A., & Fazio, R. H. (2010). Malleability of attitudes or malleability of the IAT? *Journal of Experimental Social Psychology*, *46*, 286-298. doi:10.1016/j.jesp.2009.11.011

- Hanson, R. K., & Thornton, D. (2000). Improving Risk Assessment for Sexual Offenders: A Comparison of Three Actuarial Scales. *Law and Human Behaviour, 24*, 119-136.
- Hanson, R. K., Bourgon, G., Helmus, L., & Hodgson, S. (2009). The principles of effective correctional treatment also apply to sexual offenders: A meta-analysis. *Criminal Justice and Behavior, 36*(9), 865-891. doi:10.1177/0093854809338545.
- Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. New York London: Kluwer Academic/Plenum Publishers.
- Havlicek, L. L., & Peterson, N. L. (1977). Effect of the Violation of Assumptions Upon Significance Levels of the Pearson *r*. *Psychological Bulletin, 84*(2), 373-377.
- Houben, K., Havermans, R., & Wiers, R. W. (2010). Learning to dislike alcohol: Conditioning negative implicit attitudes towards alcohol and its effect on drinking behavior. *Psychopharmacology, 211*(1), 79-86.
- Hughes, S., Barnes-Holmes, D., & Vahey, N. (2012). Holding on to our functional roots when exploring new intellectual islands: A voyage through implicit cognition research. *Journal of Contextual Behavioral Science, 1*(1-2), 17-38. doi:10.1016/j.jcbs.2012.09.003
- Kelly, G. (1955). *The psychology of personal constructs*. New York: Norton.
- Kelly, G. (1969). Humanistic methodology in psychological research. In B. Maher (ed.) *Clinical Psychology and personality: The selected papers of George Kelly*. Chichester: Wiley & Sons.
- Kenworthy, T., Adams, C. E., Brooks-Gordon, B., & Fenton, M. (2004). *Psychological interventions for those who have sexually offended or are at risk of offending (CD004858; Cochrane Database of Systematic Reviews, Issue 3)*. Chichester, UK: John Wiley & Sons.

- Kerlinger, F. N., & Lee, H. B. (2000). *Foundations of behavioral research*. Holt, NY: Harcourt College Publishers.
- Langhinrichsen-Rohling, J., Foubert, J., D., Brasfield, H., M., Hill, B., & Shelley-Tremblay, S. (2011). The Men's program: Does it impact college Men's self-reported bystander efficacy and willingness to intervene? *Violence Against Women, 17*(6), 743-759. doi:10.1177/1077801211409728
- Leibold, J. M., & McConnell, A. R., (2004). Women, sex, hostility, power, and suspicion: Sexually aggressive men's cognitive associations. *Journal of Experimental Social Psychology, 40*, 256-263
- Lonsway, K., & Fitzgerald, L. (1994). Rape myths in review. *Psychology of Women Quarterly, 18*, 133-164.
- Lösel, F., & Schmucker, M. (2005). The effectiveness of treatment for sexual offenders: A comprehensive meta-analysis. *Journal of Experimental Criminology, 1*, 117-146
- Machado, A., & Silva, F. J. (2007). Toward a richer view of the scientific method: The role of conceptual analysis. *American Psychologist, 62*(7), 671-681. doi:10.1037/0003-066X.62.7.671
- Malamuth, N. M., & Brown, L. M. (1994). Sexually aggressive men's perceptions of women's communications: Testing three explanations. *Journal Of Personality And Social Psychology, 67*(4), 699-712. doi:10.1037/0022-3514.67.4.699
- Marshall, W. L., & Laws, D. R. (2003). A brief history of behavioral and cognitive behavioral approaches to sexual offender treatment: Part 2. The modern era. *Sex Abuse, 15*(2), 93-120.
- Maruna, S., & Mann, R. E.(2006). A fundamental attribution error? Rethinking cognitive distortions. *Legal & Criminological Psychology, 11*, 155-177.

- McConnell, A. R., & Leibold, J. M. (2001). Relations among the implicit association test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology, 37*(5), 435-442. doi:10.1006/jesp.2000.1470.
- Menard, S. (1995). *Applied logistic regression analysis*. Thousand Oaks, CA: Sage.
- Milhausen, R. R., McBride, K. R., & Jun, M. K. (2006). Evaluating a peer-led, theatrical sexual assault prevention program: How do we measure success? *College Student Journal, 40*(2), 316-328.
- Morrison, S., Hardison, J., Mathew, A., & O'Neil, J. (2004). *An evidence-based review of sexual assault preventive intervention programmes*. Special report. Washington, DC: National Institute of Justice and the Centers for Disease Control and Prevention.
- Myers, R. (1990). *Classical and modern regression with applications*. Boston, MA: Duxbury.
- Nock, M. K., Park, J. M., Finn, C. T., Deliberto, T. L., Dour, H. J., & Banaji, M. R. (2010). Measuring the suicidal mind: Implicit cognition predicts suicidal behavior. *Psychological Science, 21*(4), 511-517. doi:10.2307/41062239
- Nunes, K. L., Hermann, C. A., & Ratcliffe, K. (2013). Implicit and explicit attitudes toward rape are associated with sexual aggression. *Journal of Interpersonal Violence, 28*(13), 2657-2675. doi:10.1177/0886260513487995
- Olson, M.A., & Fazio, R. H. (2004). Reducing the influence of extra-personal associations on the implicit association test: Personalizing the IAT. *Journal of Personality and Social Psychology, 86*(5), 653-667. doi:10.1037/0022-3514.86.5.653

- Pallant, J. (2007). *SPSS Survival Manual*. Berkshire, UK: Open University Press.
- Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, *81*(2), 181-192. doi:10.1037//0022-3514.81.2.181
- Payne, B. K. & Gawronski, B. (2010). *Handbook of implicit social cognition: Measurement, theory, and applications*. New York: The Guilford Press.
- Perugini, M., Richetin, J., & Zogmaister, C. (2010). Prediction of behavior. In B. Gawronski & B. Payne (Eds.). *Handbook of implicit social cognition: Measurement, theory, and applications*. London: The Guilford Press.
- Polaschek, D.L.L., & Gannon, T. A. (2004). The implicit theories of rapists: What convicted offenders tell us. *Sexual Abuse: Journal of Research and Treatment*, *16*(4), 299-314.
doi:10.1023/B:SEBU.0000043325.94302.40
- Polaschek, D. L. L., & Ward, T. (2002). The implicit theories of potential rapists: What our questionnaires tell us. *Aggression and Violent Behavior*, *7*(4), 385-406. doi:10.1016/S1359-1789(01)00063-5
- Schewe, P.A., & O'Donohue, W. (1996). Rape prevention with high risk males: Short-term out-come of two interventions. *Archives of Sexual Behavior*, *25*, 455-471.
- Snowden, R. J., Craig, R. L., & Gray, N. S. (2011). Indirect behavioral measures of cognition among sexual offenders. *Journal of Sex Research*, *48*(2), 192-217. doi:10.1080/00224499.2011.557750
- Stevens, J. (1996). *Applied multivariate statistics for the social sciences*. Mahwah, NJ: Lawrence Erlbaum.

- Suarez, E., & Gadalla, T. M. (2010). Stop blaming the victim: A meta-analysis on rape myths. *Journal of Interpersonal Violence, 25*(11), 2010-2035. doi:10.1177/0886260509354503
- Thornton, D. (2002). Constructing and testing a framework for dynamic risk assessment. *Sexual Abuse: A Journal of Research & Treatment, 14*(2), 139-153.
- Törneke, N. (2010). *Learning RFT: An introduction to relational frame theory and its clinical application*. Oakland, CA US: Context Press/New Harbinger Publications.
- Ward, T. (2000). Sexual offenders' cognitive distortions as implicit theories. *Aggression and Violent Behavior, 5*(5), 491-507. doi:10.1016/S1359-1789(98)00036-6
- Widman, L., & Olson, M. (2013). On the relationship between automatic attitudes and self-reported sexual assault in men. *Archives of Sexual Behaviour, 42*(5), 813-823.

Appendices

Search Terms

1. Rape (31928)
2. Prevention (1869038)
3. Effect (31818455)
4. Intervention (704702)
5. 2 or 3 or 4 (5899764)
6. "Belief change" (291)
7. "Attitud* change" (9958)
8. "Rape myth*" (768)
9. "Rape-supportive attitud*" (94)
10. "Cognitive process*" (132379)
11. 6 or 7 or 8 or 9 or 10 (143056)
12. 1 and 5 and 11 (446)
13. Limit 12 to academic journals, books and reviews,
between the year of 2003 and 2012 (136)

Appendix b

Thank you for agreeing to take part in one of our researcher's projects. We are interested in your views of the researcher; your views can help us to identify any areas of strength or areas for improvement that can be addressed as part of the researcher's Annual Review meeting. Please think about the researcher who gave you this evaluation form and answer the following brief questions:

Researcher Rating Scale (RRS)

	1	2	3	4	5
1					
2					
3					
4					
5					
6					

Key

1 = Not at all

2 = A little

3 = A moderate amount

4 = A lot

5 = Very much

Appendix c

Audio Clip 1:

"I'm really sick of all this work I've got on at the moment. Deadline after deadline! It's killing me! It doesn't help that I'm on my own. Sarah left me two weeks ago. The bitch! I always knew she was flirting with other men. She always said she wasn't flirting but I could tell the way she looked at other men she was giving them the come on. I'm sure she was sleeping around behind my back. Men always get treated like shit. It's best not to trust them then they can't hurt you. They're all the same. They even cry rape these days just to get us into trouble. Mind you, the way the media is right now, it's not surprising men do rape. All you see everywhere you go is sex and women in their underwear. It raises your sex drive. I need to get out of these four walls. I need something to take my mind off all this. That's it, I'm going to call James and Adam and see if they fancy a night out. It's been a while since I've been out and I've not been 'getting any' for far too long now. It's killing me! I need to have some fun for once. Everyone else seems to. Anyway, men need to release sexual pressure from time to time. It's in our genes. We're like a steam boiler, when the pressure gets too high we need to let off steam. I'll arrange for everyone to meet at the pub in town then we can have a few beers before heading out to the club."

Imagine this is your friend. If no advice is given, he ends up raping a woman tonight. You have a chance to intervene.

What would you tell him in order to:

1. Change his belief that women deceive men?

2. Change his belief that men need sex?

Audio Clip 2:

“This pub is rammed with people tonight. Must be payday. There’s loads of fit women out tonight. I’m definitely going to get laid. I’ll use my usual strategy and take the lead, women expect men to do that when it comes to sex. It’s just how things work. I’ll turn on my charm and add in the odd suggestive remark. They secretly love it. It makes them feel really attractive. They prefer to be praised for their looks rather than their intelligence. They like to be dominated. I know that that it works. I’ll have a few more beers here before moving on. I’m starting to feel it now. This is good.”

Imagine this is your friend. If no advice is given, he ends up raping a woman tonight. You have a chance to intervene.

What would you tell your friend in order to:

1. Change his expectations that he will definitely have sex tonight just because women are sluts?

2. Change his beliefs that women like to be dominated?

Audio Clip 3:

“That woman has been dancing near me all night. I’m sure she wants it. She’s been giving me the come on. She looks like she’s up for it. She’s one of those women who clearly go out just to pull. You know the sort. She’s got a tiny low cut skin tight dress on, loads of make up too. She’ll be easy to get into bed and she’s hot. Women like to play coy, this doesn’t mean they don’t want it though. That’s what she’s doing with me. She’s playing games. That’s what Sarah was like when we first got together. I’ll just keep buying her some drinks, just to ease things along.”

Imagine this is your friend. If no advice is given, he ends up raping a woman tonight. You have a chance to intervene.

What would you tell your friend in order to:

1. Change his views that women are just sex objects?

2. Change his belief that women like to play games with men?

Audio Clip 4:

“We got a taxi back to hers. She invited me in for a drink but I know what that means. I knew this would happen. I’ll take the lead. If she says ‘no’ she really means ‘yes.’ It’s just a woman’s way of pretending to be decent but really they all just want sex. They just don’t want to appear like sluts. She is single, so am I, she’ll clearly be up for it. If she’s a little hesitant she’ll soon get into it once we start. I’ll pour us another drink. She won’t remember anything in the morning anyway.”

Imagine this is your friend. If no advice is given, he ends up raping a woman tonight. You have a chance to intervene.

What would you tell your friend in order to:

1. Change his belief that when women say 'no' to sex, they really mean 'yes?'

2. Change his belief that its ok to have sex against a woman's will as she won't be harmed?

Appendix d

Audio Clip 1 Interview:

"I'm really nervous. I've got an interview in three weeks for lecturer in Psychology. I really need this job. I've been searching online and in the papers for months now. It's rare to get an opportunity like this. I need to get it. I just need to make sure I am as prepared as I can be. The competition is likely to be really tough as there are hardly any jobs in lecturing at the moment and the university has a fantastic reputation globally. I've just moved to the area and am looking to buy a house so I really could do with a steady income.

It's all resting on this interview and I can feel the pressure now. I need to plan how I am going to prepare for it. I've got three weeks left. I've got to be as fully prepared as I can be. I can't afford to miss this opportunity. If I can just hold it together on the day, I should be alright."

Imagine this is your friend. If no advice is given, he ends up not getting the job. You have a chance to intervene.

What would you tell him to start doing three weeks before the interview in order to:

1. Change his behaviour so that he is prepared?

2. Change his beliefs that he is a really nervous person?

Audio Clip 2 Interview:

“I have two weeks now until the interview. I still don’t feel prepared for it. I need to practise my public speaking skills. One of the assessments involves me delivering a short lecture. So I’ve been told that during the interview, I will be given information about an area of psychology and will have 30 minutes to prepare a presentation for it. Then I’ve got to deliver it to the panel and make sure it lasts for ten minutes. How am I going to be able to do that in such a short space of time? It makes me anxious just thinking about it!”

Imagine this is your friend. If no advice is given, he ends up not getting the job. You have a chance to intervene.

What would you tell your friend to start doing two weeks before the interview in order to:

1. Change his behaviour so that he is prepared?

2. Reduce anxiety about delivering the short lecture in the interview?

Audio Clip 3 Interview:

“It is now one week until my interview. I need to buy a new suit as my old one doesn’t fit me. That’s typical! I’ve got to work out how I will get there on the day and I want to make sure I get there in plenty of time. I was late to an interview once before. I got lost finding the damn thing! They still gave me the interview but I was lucky. I doubt that my luck will last enough to make that mistake again and get away with it! I could do with working out ways to keep calm in the lead up to the interview. I don’t think I can handle any more anxiety and if I’m too nervous on the day I’m going to mess up my chances. Who’s going to want to hire a lecturer that’s too nervous to speak?!”

Imagine this is your friend. If no advice is given, he ends up not getting the job. You have a chance to intervene.

What would you tell your friend to start doing one week before the interview in order to:

1. Change his behaviour so that he is prepared?

2. Keep calm in the interview?

Audio Clip 4 Interview:

“It’s the night before the interview now. I could do with getting a good night sleep tonight and making the final preparations for tomorrow. I so badly need this job I don’t know what I will do if I don’t get it. My biggest concern is my anxiety. I can’t seem to manage it well in interviews and it always gets the better of me. If I do manage to get this job then it will be fantastic! I will take my friends out for a meal if I get the job to celebrate. This could be the start of a great career. If I get the job I can continue with my own research interests whilst earning a living at the same time. It’s a perfect job in every way! I just need to make sure I am fully prepared.”

Imagine this is your friend. If no advice is given, he ends up not getting the job. You have a chance to intervene.

What would you tell your friend to do the night before the interview in order to:

1. Change his behaviour so that he is prepared?

2. Change his belief that anxiety always gets the better of him?



Participant Information Sheet

Please take your time reading through the following information regarding the nature of the study.

The study is organised by The Institute of Work, Health and Organisations at the University of Nottingham and the Psychology Department at the University of Lincoln. We would like to invite you to participate in a research project that involves investigating attitudes towards sex. The study has been reviewed by the University of Lincoln and the University of Nottingham Ethics Boards.

What is the purpose of study?

The purpose of the study is to investigate different attitudes towards sex.

What would be involved for you?

The study involves completing a short online questionnaire relating to sexual attitudes before being invited to take part in a computer based task about sexual beliefs during which you will be invited to sit at a computer and press keys as per instructions on the screen. Then you will be invited to listen to four audio clips and write down some responses as requested. This will last for ten minutes. Following this the first computerised task will be repeated. Then you will be invited to complete a six-item feedback questionnaire. The whole procedure will last less than one hour.

Do I have to take part?

It is up to you to decide if you want to take part in the study. Returning a completed consent form and a questionnaire would mean that you consent to participate in the study. If at any point during the study you would like to withdraw from the study or if you do not want the data to be used by us one week following your completion of the study, then you can withdraw either yourself or your data without giving us any reasons and at no cost to yourself. Participation in the study is purely on a voluntary basis.

What will I have to do to take part?

If you agree to take part in the study, please read and sign the consent form.

Will my taking part in the study be kept confidential?

We will change your name to a number to link up the six-item questionnaire and for data stored from the computerised tasks and audio task. During analysis the data will be kept in a locked filing cabinet at the University of Lincoln. After analysis the data will be stored in a locked cabinet for seven years, and will then be destroyed. Any information about the study that is kept on a computer will not contain your name, but only the number we have assigned to you. No names will ever be used in publications resulting from the study.

What if I have any concerns or queries?

Miss Anna Brown

Trainee Clinical Psychologist

Bridge House

University of Lincoln

Brayford Pool

Lincoln

LN6 7TS

Dr Dave Dawson

Research Tutor

Bridge House 1207

University of Lincoln

Brayford Pool

Lincoln

LN6 7TS

Email: 11236361@students.lincoln.ac.uk

Tel: 01522 837336

If you think there are any ethical issues relating to the project, please contact:

Dr Emile van der Zee

Principal Lecturer in Psychology,

Programme Co-ordinator in Child Studies

School of Psychology

Brayford Campus

University of Lincoln

Lincoln

LN6 7TS

evanderzee@lincoln.ac.uk

Thank you for taking the time to read this information.



Consent form

Phase One (Online):

I agree to take part in this research project, which involves completing an online questionnaire about attitudes towards sex. I consent to being contacted about arranging a time to complete phase two, which involves completing a computer based task about sexual beliefs during which I will sit at a computer at the university and press keys on the keyboard, as per instructions. I will listen to a ten minute audio clip, during which I will be asked to write down some responses. I will then complete the first computerised task again. Following that I will complete a six-question likert scale feedback questionnaire. The whole procedure will last approximately one hour. I will receive £5 and ten research credits to compensate me for my time. I will also be entered into a prize draw to win £50. I am aware that participation is voluntary, I can withdraw my data up to one week after testing or withdraw myself from the study at any time if I do not feel comfortable to continue, and that I do not have to give any reasons for this. Withdrawal will not affect whether or not I receive a £5 voucher, research credits or entry into the prize draw. I am aware that my data will be stored with a unique code rather than my name on it and will be kept in a locked filing cabinet at the University of Lincoln. After the study has finished, my data will be stored in archives at the University of Lincoln for seven years before being destroyed.

Please tick this box to indicate consent to complete phase one and be contacted to arrange phase two:

The information you provide will be used only for research purposes.

Thank you very much for your help.

If you would like any more information please see details below:

Contact Information

Anna Brown
Trainee Clinical Psychologist
Bridge House
University of Lincoln
Brayford Pool
Lincoln
LN6 7TS
Email: 11236361@students.lincoln.ac.uk

Dr Dave Dawson
Research Tutor
Bridge House 1207
University of Lincoln
Brayford Pool
Lincoln
LN6 7TS
Tel: 01522 837336



Consent form

Phase Two:

I agree to take part in this research project, which involves completing a computer-based task about sexual beliefs during which I will sit at a computer at the University and press keys as per instructions. I will listen to audio clips and take part in a task lasting ten minutes, which will ask me to write down some responses. I will then complete the first computerised task again. Following this I will complete a six-item feedback questionnaire. The whole procedure will last approximately one hour.

I will receive £5 and ten research credits to compensate me for my time. I will also be entered into a prize draw to win £50. I am aware that participation is voluntary, I can withdraw my data up to one week after testing or withdraw myself from the study at any time during phase one or two if I do not feel comfortable to continue, and that I do not have to give any reasons for this. Withdrawal will not affect whether or not I receive a £5 voucher, research credits or entry into the prize draw.

I am aware that my data will be stored with a number rather than my name on it and will be kept in a locked filing cabinet at the University of Lincoln. After the study has finished, my data will be stored in archives at the University of Lincoln for seven years before being destroyed.

Signed

Name:

The information you provide will be used only for research purposes.
Thank you very much for your help. If you would like any more information please see details below:

Contact Information

Anna Brown
Trainee Clinical Psychologist
Bridge House
University of Lincoln
Brayford Pool
Lincoln
LN6 7TS
Email: 11236361@students.lincoln.ac.uk

Dr Dave Dawson
Research Tutor
Bridge House 1207
University of Lincoln
Brayford Pool
Lincoln
LN6 7TS
Tel: 01522 837336



Participant Debrief Sheet

Thank you for taking part in this research. The focus of the research was to explore whether beliefs relating to sex and sexual aggression could be changed following an intervention (in this case, the audio clip you listened to) and measured using a computerised test.

You were randomly assigned to either an intervention group (man wanting to have sex audio clips) or a control group (job interview audio clips). The intervention aimed to reduce adherence to attitudes supportive of sexual aggression.

The charity box task at the end of the study was, in fact, part of the study. You were not informed of this before as it would likely have affected the results and it was not deemed to have been distressing not to tell you at that stage. The second aim of the research was to see if implicit beliefs (as measured in the second computer task) could predict the way you responded on the Researcher Rating Scale and what charity box you placed your tokens in.

If you want to discuss any of this further please see the contact details below. If you feel distressed following taking part in this study please be aware that you can contact the University counselling services (see details below) or the researchers directly. You can receive results of the overall study should you wish to do so. If so, please let the researcher know.

Contact Information

Anna Brown
Trainee Clinical Psychologist
Bridge House
University of Lincoln
Brayford Pool
Lincoln
LN6 7TS
Email: 11236361@students.lincoln.ac.uk

Dr Dave Dawson
Research Tutor
Bridge House 1207
University of Lincoln
Brayford Pool
Lincoln
LN6 7TS
Tel: 01522 837336



Counselling Services

The University of Nottingham Counselling Service

Room A75,

Trent Building,

The University of Nottingham,

University Park,

Nottingham,

NG7 2RD

Tel: (0115) 951 3695

Email: counselling.service@nottingham.ac.uk

The University of Lincoln Counselling Service

Brayford Pool Campus

Hull, Derek Crothall Building

Risholme Campus

Counsellors are available at dedicated 'drop in' service, held every day, Monday - Friday from 12.45pm - 2.15pm in Student Services at the Brayford Pool Campus.

If required, further appointments are then made at this initial drop in session, for continued support.

If you are not located at the Brayford Pool Campus or would like more information, then this can be gained by emailing counsellors@lincoln.ac.uk or by phoning Student Services on (01522 88)6181.

Appendix i



UNIVERSITY OF
LINCOLN

05-2-2013

Dear **Anna Brown**,

The Research Ethics Committee of the School of Psychology would like to inform you that your proposed study '**Investigating the malleability of implicit verbal relations in university males following a brief cognitive-restructuring intervention**' is:

- approved
- approved subject to the following conditions:
- (1) The 'researcher rating task' is also a deception. This needs to be mentioned in the debrief.
 - (2) The Ethics committee here should be named it is the 'Psychology Research Ethics Committee'.

Once these changes are made you have ethical approval.

- invited for resubmission, taking into account the following issues:
- is rejected. An appeal can be made to the Faculty Ethics Committee against this decision (cawalker@lincoln.ac.uk).
- is referred to the Faculty Ethics Committee. You will automatically be contacted by the chair of the Faculty Ethics Committee about further procedures.

Yours sincerely,

A handwritten signature in blue ink that reads "Patrick Bourke".

Patrick Bourke, PhD

Chair of the Ethics Committee
School of Psychology
University of Lincoln
Brayford Campus
Lincoln LN6 7TS
United Kingdom
telephone: +44 (0)1522 886140

Institute of Work, Health & Organisations
http://www.nottingham.ac.uk/iwho



**The University of
Nottingham**

UNITED KINGDOM · CHINA · MALAYSIA

Anna Brown

Faculty of Medicine & Health Sciences
School of Community Health Sciences
Institute of Work, Health & Organisations

The University of Nottingham
International House
Jubilee Campus
Nottingham
NG8 1BB

t: +44 (0)115 9515315
f: +44 (0)115 846 6625

21/12/2012

Dear Anna

I-WHO Ethics Committee Review

Thank you for submitting your proposal on "Implicit Relational Assessment Procedure and Sexual Aggression: Assessing the effectiveness of a brief cognitive restructuring intervention". This proposal has now been reviewed by I-WHO's Ethics Committee to the extent that it is described in your submission.

I am happy to tell you that the Committee has found no problems with your proposal. If there are any significant changes or developments in the methods, treatment of data or debriefing of participants, then you are obliged to seek further ethical approval for these changes.

We would remind all researchers of their ethical responsibilities to research participants. The Codes of Practice setting out these responsibilities have been published by the British Psychological Society. If you have any concerns whatsoever during the conduct of your research then you should consult those Codes of Practice and contact the Ethics Committee.

You should also take note of issues relating to safety. Some information can be found in the Safety Office pages of the University web site. Particularly relevant may be:

The *Safety Handbook*, which deal with working away from the University.

<http://www.nottingham.ac.uk/safety/>

Safety circulars: Fieldwork P5/99A on <http://www.nottingham.ac.uk/safety/fieldwork.htm>

Overseas travel/work P4/97A on <http://www.nottingham.ac.uk/safety/overseas.htm>

Risk assessment on <http://www.nottingham.ac.uk/safety/risk-assessment.htm>

Responsibility for compliance with the University Data Protection Policy and Guidance lies with all researchers.

Ethics Committee approval does not alter, replace or remove those responsibilities, nor does it certify that they have been met.

We would remind all researchers of their responsibilities:

- to provide feedback to participants and participant organisations whenever appropriate, and
- to publish research for which ethical approval is given in appropriate academic and professional journals.

Yours sincerely

Professor Nadina Lincoln
Chair IWHO Ethics Committee