Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

James Hardeman Walker Date

The Psychometric Evaluation of Sexual Violence Prevention Instruments for Use with Male College Students

By

James Hardeman Walker
Master of Public Health

Epidemiology

_____

Julie A. Gazmararian, PhD, MPH
Committee Chair

_____

Laura F. Salazar, PhD
Committee Member

The Psychometric Evaluation of Sexual Violence Prevention Instruments for Use with Male College Students

By

James Walker

B.S. in Molecular and Cellular Biology, B.A in Earth and Planetary Science

Johns Hopkins University

2010

Faculty Thesis Advisor: Julie A. Gazmararian PhD, MPH

An abstract of

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology

2012

Abstract

The Psychometric Evaluation of Sexual Violence Prevention Instruments for Use with Male College Students

By: James Walker

The purpose of this study was to assess the psychometric properties of three novel scales and two novel indexes developed and used in the RealConsent study, a web-based sexual violence preventive intervention for male undergraduates implemented and evaluated at Georgia State University. These metrics assessed five important constructs designed for use in evaluations of sexual violence interventions: legal knowledge of informed consent, perceptions of informed consent, self-efficacy to intervene when a woman's consent is compromised, outcome expectancies for non-consensual sex, and outcome expectancies for intervening behaviors. Each scale was assessed for reliability and validity. Results showed that one of the three novel scales: Self-Efficacy to Intervene, and one of the two indexes: Perceptions of Informed Consent for Sex, showed adequate reliability and validity. These two metrics are therefore verified for future research efforts, while the other three may need further analysis or modification to enhance their utility in future studies.

The Psychometric Evaluation of Sexual Violence Prevention Instruments for Use with Male College Students

By

James Walker

B.S. in Molecular and Cellular Biology, B.A in Earth and Planetary Science

Johns Hopkins University

2010

Faculty Thesis Advisor: Julie A. Gazmararian PhD, MPH

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology

2012

Table of Contents

Background

Addressing the Threat of Violence Against Women

Physical and sexual violence is a severe and significant threat to female college students (1, 2, 3, 4, 5). Women aged 16-24 are particularly vulnerable, with some studies showing the prevalence of this victimization to be as high as 50% (1). Even higher rates are reported for psychological and emotional abuse (1, 2). Males in this age group show much lower rates of victimization. Surveys of male college students indicate that 31-35% self-report instigating at least one act of sexual coercion within a four year time frame (6).

Additionally, sexual assault incidence is underreported, and many incidents take place between people who know each other. A survey found that 84% of women raped knew their attacker and 57% of these rapes happened on dates (7). Fewer than 5% of completed and attempted rapes are reported to the police (8). Reasons for this underreporting include fear of reprisals, desire for privacy, fear of not being believed, and fear of victim blaming.

According to Department of Justice Statistics, on average nearly one in four women is victimized by either completed or attempted rape over the course of a five year college career (8). This number means that for every 1,000 women attending an institution of higher education, 35 completed rape incidents occur every year (8). Beyond physical trauma, there are severe and long lasting psychological and emotional consequences of this victimization. Therefore, there is a vital need for interventions to reduce the frequency of rape behavior and sexual victimization.

The RealConsent Study

The RealConsent study was a sexual violence preventive intervention designed for male undergraduate students. The goal of this intervention was to reduce the occurrence of sexual violence, as well as to increase the potential of men to intervene in situations where a woman's consent may be compromised. Only males were included, as prior studies have shown that effects of rape prevention programs are greater for gender-specific studies than for mixed-gender programs (9).

RealConsent was based upon two theoretical frameworks that attempt to explain the personal and social influences behind behaviors. Social Cognitive Theory (SCT) contends that behavior is affected by three primary influences: personal, behavioral, and environmental (10). Personal characteristics, such as self-efficacy to engage in a particular behavior, outcome expectancies associated with a behavior, and knowledge are key constructs related to behavior. In addition, SCT posits that an understanding of the socio-cultural context allows greater understanding of observed behavior and guides interventional approaches. In this instance, the context in which sexual violence occurs would be the presence of bystanders who may tacitly support violence against women; however, if encouraged to intervene, these same bystanders can change this context to prevent sexual violence. Social Norms Theory (SNT) is concerned with individual's perceptions of social norms (which may differ between individuals) and allows an understanding of how these perceptions shape behavior (11).

RealConsent was aimed at addressing the public health threat of sexual violence against women through its impact on several theoretical mediating variables. In any

public health intervention, outcome measurement is vital to understanding the effectiveness of the intervention. Accurate outcome measurement is dependent upon having reliable and valid measures of the study constructs. Reliability is an indicator of a scale's ability to perform consistently (12). Without adequate reliability, a scale may be inconsistent across different studies, different research subjects, and even across different time points for the same subject. Reliability is necessary because it ensures that the measure will function as intended in varying situations and populations. Validity is the ability of a scale to accurately reflect an underlying theoretical construct (12). Without adequate validity, a scale may not be properly measuring the underlying construct, and therefore may be difficult or impossible to interpret correctly. Scale validity is analogous to internal validity in an experimental study: it allows the researcher to recognize whether what appears to be measured is actually being measured. Adequate reliability and adequate validity must both be established before a scale can be widely and confidently used (12).

The RealConsent study used surveys to collect data about the many potential mediators behind male respondents' understanding of the meaning and reality of sexual and physical violence against women. These mediators include misperceptions in underlying social norms, rape myths, attitudes toward women, knowledge of legal definitions of rape, informed consent to have sex, traditional gender roles, empathy for rape victims, communication with a partner about sex, and self-efficacy to intervene when witnessing violence against women (13, 14, 15, 16, 17). For many of these mediators, existing scales and indexes were used, but for legal knowledge, perceptions of informed consent, self-efficacy to intervene, outcome expectancies for non-consensual

sex, and outcome expectancies for intervening behaviors, no currently verified metrics existed. To assess these mediators, RealConsent study staff created five new metrics. The purpose of the present psychometric analysis is to assess the validity and reliability of these new metrics in order to both better understand the impact of the RealConsent intervention, as well as to assess the potential utility of these measures for future studies.

Methods

Intervention

RealConsent was an interactive multi-media web-based program aimed at reducing violence against women through educational and emotional appeals. An evaluation of RealConsent was conducted with a randomized controlled trial of 750 male undergraduate students recruited from Georgia State University (GSU). The trial took place between March of 2010 and January of 2011 and was administered online.

The sexual violence preventive intervention was a comprehensive program consisting of six modules, each taking 30-45 minutes to complete. Each module was presented in an interactive multimedia format including video and audio narratives, interactive quizzes, and literature. The intervention focused on several specific areas: increasing knowledge of informed consent for sex and knowledge of the legal definitions of violence against women and rape; increasing favorable outcome expectancies for prosocial (intervening) behaviors and negative outcome expectancies for negative (enabling) behaviors; enhancing self-efficacy to avoid negative behaviors and to engage in intervening behaviors; modeling positive behaviors that do not support sexual

violence; male socialization, debunking rape myths, victim empathy, and the role of alcohol in date rape.

The study also included a control group who instead participated in a Web-based General Health Promotion (GHP) program. This program was developed by the Institute for Social Analysis (ISA) through a National Institutes of Health (NIH) grant and, similar to the intervention condition, is an interactive audio/video program. The GHP program approximates the VAW intervention on duration, format, and interactivity. The GHP, however, focuses on stress management, weight management, nutrition, and fitness.

Data Collection

To collect data on the efficacy of the intervention, online surveys were administered that included measures of respondents' attitudes, intentions, and beliefs about sexual victimization of women. Surveys were administered at three time points for both the intervention and control groups: 1) before the program start, to establish a baseline for comparison; 2) immediately following the conclusion of the program, to assess the impact of the intervention, and 3) six months after the previous survey, to assess the long-term effect.

Study Population

Inclusion criteria for the study included being between 18-24 years of age, male sex, self-reported heterosexual identity, and receipt of informed consent. The original participant sample was drawn from an email list of students from the registrar that met the study criteria for age. A random sample of 8,458 email addresses was drawn from the list, and email invitations were sent to potential participants. It is from this sample that the original group of 746 participants was drawn. The baseline survey was then

administered, and then the group was randomized into treatment (n=376) and control groups (n=367). Moderate attrition was observed at the end of the intervention period, as 268 participants from the treatment group and 183 participants from the comparison condition completed the second survey. Further attrition was seen at 6-month follow up: 123 participants from the treatment group and 92 participants from the comparison condition completed the final survey.

Measures

Many of the psychometric scales used in the RealConsent study have been used extensively in other studies and have been independently verified prior to study commencement (13, 14, 15, 16, 17). As the goal of RealConsent was to affect sexual violence perpetration and intervening behaviors, measures of these primary outcomes were included. All six primary outcome measures were assessed using existing and verified psychometric scales. Psychological abuse perpetration, physical abuse perpetration, and sexual abuse perpetration were assessed using the Revised Conflict Tactics Scale (RCTS) (18). Sexual coercion was assessed using the Sexual Experiences Survey (SES) (19). Date rape related risk behaviors were measured using the College Date Rape Attitude and Behavior Survey (CDRABS) (20). Intervening behaviors were assessed with the Reactions to Offensive Language and Behavior (ROLB) (21).

Development of New Instruments for RealConsent

The purpose of the present analysis was to assess five novel measures developed or adapted by the Principal Investigator for the RealConsent study and used in the evaluation. These five instruments measured knowledge, self-efficacy and outcome expectations regarding sexual violence against female college students and intervening.

The five instruments were: 1) Legal Knowledge Scale; 2) Perceptions of Informed Consent for Sex; 3) Self-efficacy to Intervene; 4) Outcome Expectancies for Engaging in Non-Consensual Sex; and 5) Outcome Expectancies for Intervening Behaviors.

Important to understanding the development and subsequent psychometric evaluation of these instruments is noting the difference between scales and indexes. If the items of a metric are effect indicators, then the metric is classified as a scale. Effect indicators are designed to measure the underlying construct by inferring an influence on an observable behavior (12). This means that the degree to which a respondent is affected by the underlying construct should affect the responses that the respondent gives to the items on the scale. On the other hand, if the items of the metric are causal indicators, then the metric is classified instead as an index (12). That is to say that the responses that an individual gives to these causal items define the respondent's level of the underlying construct. In short, the level of an underlying construct impacts the answers a respondent gives on a scale, while the responses on an index themselves define the level of a construct. Additionally, items on an index may not necessarily be correlated with each other, while the items on a scale would.

Of the five novel metrics created for the RealConsent study, two: the Legal Knowledge Index and Perceptions of Informed Consent for Sex were indexes. The other three: Self-Efficacy to Intervene, Outcome Expectancies for Engaging in Non-Consensual Sex, and Outcome Expectancies for Intervening Behaviors were deemed scales.

The Legal Knowledge Index was developed based upon Georgia law and Georgia State University (GSU) policies. This index consists of nine true/false questions that test

participants' knowledge of the law regarding rape and sexual assault both at the state level and at the local university level.

The Perceptions of Informed Consent for Sex index was developed following formative research with focus groups. It is a 14 item true/false index which assesses participants' perceptions about what actions constitute obtaining informed consent for sex. Self-Efficacy to Intervene was assessed with an 18 item scale adapted from both the ROLB Self-Behavior Subscale and items from Banyard et al. (22). These items ask the respondent how confident they are on a scale of 1 to 7 (with 1 being the least confident and 7 being the most confident) that they would intervene in a number of observed situations that are potentially harmful towards a woman.

Outcome expectancies for sexual violence against women were assessed using the 15-item Outcome Expectancies for Engaging in Non-Consensual Sex scale. This scale was developed by the Principal Investigator of RealConsent following focus group assessments. Respondents chose how strongly they agreed on a scale from 1 to 5 (with 1 being strongly disagree, 3 being neither agree nor disagree, and 5 being strongly agree) with different outcomes that could occur (both positive and negative) if the respondent did not get clear consent from a woman before having sex.

Finally, the Outcome Expectancies for Intervening Behaviors scale was 17 items, mostly adapted from Banyard et al. with additional novel items added by study staff (22). Respondents chose how strongly they agreed on a scale from 1 to 5 (with 1 being strongly disagree, 3 being neither agree nor disagree, and 5 being strongly agree) with potential outcomes that could occur if the subject intervenes in a situation where a woman is under threat of violence or consent may have been compromised.

Data Analysis

Reliability is assessed for each scale in three ways: 1) test-retest reliability; 2) Cronbach's alpha test for internal reliability; and 3) the split-half method test for internal reliability (12, 23). Each of these methods is described below.

Test-retest reliability.

Test-retest reliability is assessed by comparing survey responses across multiple time points (23). During the RealConsent study, surveys were administered first at the study start prior to the intervention, secondly immediately following the intervention, and finally at a six month follow-up. In order to avoid the interaction between study impact and changes in survey responses (the treatment effect), this analysis was performed with the control group only.

Test-retest reliability assumes that the underlying construct is stable and a reliable survey would return similar responses for each participant at each of the study time points (23). Test-retest reliability was assessed by comparing responses for each of the five scales at time 1 compared to time 2, time 2 compared to time 3, and time 1 compared to time 3. A test-retest Pearson correlation coefficient greater than or equal to 0.7 is considered an adequate demonstration of stability.

To assess the test-retest reliability of each of the five metrics, a composite score based on recoded ascending score responses for each scale was created by summing respondents' responses. Higher composite scores were associated with increased legal knowledge, higher cognitive understanding of informed consent, greater self-efficacy to intervene, higher severity and outcome expectancies for violence against women, and more positive outcome expectancies for intervening behaviors. This composite-score

methodology parallels the method developed to assess the reliability of the Rosenberg Self-Esteem Scale, a similarly multi-item Likert-type psychometric scale. (24)

Cronbach's alpha.

Internal reliability for each of the three scales was assessed using Cronbach's alpha (α). This was a measure of the inter-correlations between items of a single scale (23). Cronbach's alpha (α) was therefore taken independently for each scale and represents the ratio of the sum of the inter-item covariances to the overall variance for the scale. A Cronbach's alpha (α) greater than or equal to 0.7 is considered a sufficient demonstration of internal reliability for each scale.

For the two indexes (Legal Knowledge and Perceptions of Informed Consent) the Cronbach's alpha test of internal reliability was not performed. Items in an index may be heterogeneous, and therefore testing for internal reliability is not appropriate (25).

Split-half reliability.

The split-half method involved dividing each scale randomly in half and then taking Spearman-Brown (SB) correlations between each half-scale. The greater the correlation statistic, the more evidence this test demonstrates of the reliability of the scale (25). A Spearman-Brown reliability coefficient of greater than or equal to 0.7 is considered to be a sufficient demonstration of reliability for each scale by the split-half method. Similarly to the Cronbach's alpha test, the split-half test of internal reliability was not performed for the Legal Knowledge Index and the Perceptions of Informed Consent for Sex index due to the item heterogeneity inherent of indexes.

Validity assessment.

Validity was assessed for each of the five scales by using a factor analysis for construct validity. This method tests if each scale is valid by assessing the correlations between items and their underlying factors (25). A scale is considered to demonstrate construct validity if the items of each scale relating to the same theoretical dimension correlate with each other to a greater degree than they do with items of other dimensions (25). This was an exploratory factor analysis, as the theoretical subscales for these five novel scales have yet to be defined. However, indicator variables are expected to load onto a predicted number of factors as defined below.

To predict these expected factors, a theoretical framework must be selected. For this analysis, Social Cognitive Theory and Social Norms Theory will both be used to drive the prediction of component factors for each of the scales. In Social Cognitive Theory, behavior is affected by three primary influences: cognition, behavioral influences, and environmental events (10). Social Norms Theory focuses on the social norms that influence behavior and underscores that this influence is mediated by individuals' differing perceptions of these norms (11).

For the Legal Knowledge Index, the expected number of factors is two: based on Social Cognitive Theory, people monitor their own conduct in relation to moral standards (normative processes) and cognitive processes, and therefore these would be the two expected factors within the Legal Knowledge. For Perceptions of Informed Consent for Sex, the expected number of factors is four: normative influences, attitudinal influences, cognitive conceptualizations, and situational influences. For the Self-Efficacy to Intervene scale, the expected number of factors is four: Bandura identifies four

components to self-efficacy – experience, modeling, social persuasion, and physiological factors (26). For Outcome Expectancies for Engaging in Non Consensual Sex, the expected number of factors is three: cognitive, attitudinal, and normative. Likewise, for Outcome Expectancies for Intervening, the expected number of factors is three: cognitive, attitudinal, and normative. However, since this is an exploratory factor analysis, further unpredicted combinations are possible.

For this study an iterated principal axis factor methodology with varimax rotation was used, where factors were identified based on having eigenvalues >1. A varimax rotation delivers an orthogonal solution based on which factors are not highly correlated with one another. How strongly items load onto each identified scale is observed with a Pearson correlation coefficient (PCC). This statistic represents the ratio between the Pearson correlation coefficient that the items of the scale have with non-related factors and the Pearson correlation coefficient that the items of the scale have with related factors. This factor analysis correlation statistic has a range from 0 to 1, and a value of greater than or equal to 0.4 is considered to be adequate loading onto a scale. Additionally, what percent of variance is explained by each factor is assessed, with cumulative % variance accountability of >40% being considered to be sufficient evidence of construct validity.

IRB approval for this analysis was granted by the Institutional Review Board of Emory University. Because this analysis does not meet the definition of “research with human subjects” and the analysis uses data that contains no personal identifiers, no IRB review is required (eIRB #IRB00054060 – Title: *Psychometric Analysis of Novel Scale Measures of the Real Consent Intervention Study*)

Results

Test-retest Reliability

Test-retest reliability of the scales is presented in Table1. Though none of the five scales showed Pearson Correlation Coefficients (PCCs) greater than 0.7 across all three time intervals, the Perceptions of Informed Consent index, Self-Efficacy to Intervene scale, and Outcome Expectancies for Engaging in Non-Consensual Sex scale showed PCCs of greater than 0.7 for at least one time interval. The Legal Knowledge Index and the Outcome Expectancies for Intervening Behaviors scale did not show PCCs greater than 0.7 at any of the three time intervals.

Cronbach's Alpha

Internal reliability for the three scales (but not the two indexes) was assessed using Cronbach's alpha. The results are presented in Table 2a. Two scales: Self-efficacy to Intervene (α=0.958), and Outcome Expectancies for Intervening Behaviors (α=0.826), show adequate internal reliability by Cronbach's alpha (α>0.7). While one scale: Outcome Expectancies for Engaging in Non-Consensual Sex (α=0.685) showed modest internal reliability by Cronbach's alpha (α<0.7).

For the scale that did not show adequate reliability, Outcome Expectancies for Engaging in Non-Consensual Sex, an analysis was run to see what the internal reliability by Cronbach's alpha would become for the scale if each item was deleted in turn. Results are shown in Table 2b. Total comprehensive item-by-item deletion was assessed. For emphasis, as well as for simplicity's sake, only those items whose deletion demonstrated notable improvement are included in the table. It was found that removal of either item #1 or #12 from the Outcome Expectancies for Engaging in Non-Consensual Sex scale

resulted in an increase (above the $\alpha>0.7$ cutoff) in internal reliability. This minor modification of the scale allows Outcome Expectancies for Engaging in Non-Consensual Sex to demonstrate adequate reliability by the Cronbach's alpha test.

Split-Half Reliability

Internal reliability for the scales (but not the indexes) was also assessed using the split-half method. Spearman-Brown (SB) coefficients are show in Table 3. All Three scales: Self-efficacy to Intervene (SB = 0.936), Outcome Expectancies for Engaging in Non-Consensual Sex (SB = 0.732), and Outcome Expectancies for Intervening Behaviors (SB = 0.807) showed adequate evidence of internal reliability by the split-half method (SB > 0.7).

Validity Assessment

The results of the validity assessment are shown in Table 4. For the Legal Knowledge index two factors were expected. One factor relates to how people monitor their own conduct in relation to moral standards (normative processes). The second factor relates to the cognitive processes behind understanding of the concepts of justice and institutional law. As expected, two factors were observed. These two factors accounted for 45.27% of the total variance in the data, which is reasonable (>40%). Unexpectedly, however, these two factors may differ in content from the expected factors. It was expected that the two factors would assess normative and cognitive processes. Instead, both normative and cognitive influences were rolled into the first factor, and the second factor measured an unexpected concept: gender differences under the law. The two items that showed high loading (PCC >0.7) on the second factor identified the facts that: 1) under Georgia state law, only a man can be convicted of rape, and 2) under Georgia state

law, only a woman can be a victim of rape. These two items represented an unexpected factor that showed a high eigenvalue (>1) and was separate, based on an orthogonal, varimax solution that differentiates factors based on non-correlation, from the factor assessed by five out of a total nine items which is a cognitive understanding of the legal ramifications of non-consensual sex. The remaining two items did not load adequately (PCC <0.4) onto either factor (as seen by the 77.4% "Adequate" loading on Table 5).

For the Perceptions of Informed Consent for Sex scale, only 3 rather than the predicted 4 factors showed eigenvalues >1 based on the varimax rotation. Based upon which items loaded into which factors, the three factors identified by the analysis are: cognitive and attitudinal influences (e.g. "*Regardless of the circumstances, if a woman has had sex with a man, she has given consent.*"), situational influences (eg. "*If a woman undresses, or allows a man to undress her, she has given consent for sex.*"), and normative influences (e.g. "*If you are in a steady relationship, you still have to get consent*."). All fourteen items of the scale loaded adequately (PCC > 0.4) into the three factors.

For the Self Efficacy to Intervene scale, two factors were observed where four were expected. Based on which items loaded strongly into which factors, these two items were identified as intervening about attitudes (e.g. "*[I] Indicate my displeasure when I hear a sexist comment.*"), and intervening about behaviors (e.g. "*[I] Call 911 if I hear someone yelling help.*"). Fifteen out of eighteen items loaded well (PCC > 0.5) onto one or the other of these two factors, and three items loaded adequately (PCC > 0.4) onto both factors.

For the Outcome Expectancies for Engaging in Non-Consensual Sex scale, two factors were observed where three were expected. Based on which items loaded strongly onto which factors, these two factors could be described as the negative outcomes, or risks, (e.g. "*If I engage in sex without clear consent I could get charged with rape.*") versus the positive outcomes, or rewards, (e.g. "*If I engage in sex without clear consent it would make me more attractive to women.*") of non-consensual sex. Though these two "risk vs. reward" factors deviate from the three expected factors, all fifteen items of the scale load well (PCC > 0.5) on to one or the other of these two factors.

The final scale, Outcome Expectancies for Intervening Behaviors, showed three observed factors where three were expected. The three factors that were expected were cognitive, attitudinal, and normative influences. However, based on which items loaded strongly into which factors, cognitive and attitudinal influences were rolled into a single underlying factor, while normative influences were separated into two: influences of social desirability functions, and influences of the normative social influence of conformity. The first factor, cognitive and attitudinal, loaded strongly with such self-efficacy items as "*If I intervene, I can prevent someone from being hurt.*" The second factor, measuring social desirability, loaded strongly with items like "*If I intervene, I will feel like a leader in my community.*" The third factor, measuring the normative social influence of conformity, loaded strongly with items like "*Intervening might cost me friendships.*" Sixteen out of seventeen items loaded well (PCC > 0.5) onto one or another of the three factors, and one remaining item did not load adequately (PCC < 0.4) onto any factor.

Discussion

Reliability Assessments

Two scales: Self-Efficacy to Intervene and Outcome Expectancies for Engaging in Non-Consensual Sex, and one index: Perceptions of Informed Consent for Sex demonstrated adequate test-retest reliability. However, the final two metrics: Outcome Expectancies for Intervening Behaviors, and the Legal Knowledge Index may not be reliable measures as determined by this analysis of test-retest reliability.

Based upon Cronbach's alpha, both the Self-Efficacy to Intervene and the Outcome Expectancies for Intervening Behaviors scales showed adequate internal reliability. The third scale, Outcome Expectancies for Engaging in Non-Consensual Sex, did not. Therefore, an additional round of statistical testing was done to observe if the Cronbach's alpha would increase if any single item from each scale was dropped. It was found that if either item #1 *("If I engage in sex without clear consent It's all good because I still get laid")* or item #12 *("If I engage in sex without clear consent it would make me more attractive to women")* from the Outcome Expectancies for Engaging in Non-Consensual Sex scale was dropped, then Cronbach's alpha for this scale would increase to the point that the scale demonstrates adequate internal reliability. These two items may be inconsistent with the concept addressed by the scale and could be dropped to improve the scale's reliability.

Based on the split-half method, the three scales: Self Efficacy to Intervene, Outcome Expectancies for Engaging in Non-Consensual Sex, and Outcome Expectancies for Intervening Behaviors all showed adequate evidence of internal reliability. Because of the item heterogeneity of indexes, the Legal Knowledge Index and the Perceptions of

Informed Consent for Sex index were not assessed for internal reliability by either Cronbach's alpha, or the split-half method. Reliability for these indexes is therefore determined by test-retest consistency alone.

In summary, the Self-Efficacy to Intervene scale showed adequate reliability by all three assessments with no modification. The Outcome Expectancies for Engaging in Non-Consensual Sex scale showed adequate reliability assuming removal of at least one item (#1 or #12) as seen in the discussion of the Cronbach's alpha results. The Perceptions of Informed Consent for Sex index showed adequate reliability by test-retest. The final two instruments: the Legal Knowledge Index and the Outcome Expectancies for Intervening Behaviors scale did not demonstrate adequate reliability on one or more of the reliability assessments.

Validity Assessment

Because this was an exploratory factor analysis, conclusions about the validity of the scales are based upon factor loadings relative to established theoretical frameworks. For this analysis, it was seen that four of the metrics: Legal Knowledge, Perceptions of Informed Consent, Self-Efficacy to Intervene, and Outcome Expectancies for Intervening Behaviors, all consist of items that load strongly onto factors that are accounted for by either Social Cognitive Theory or Social Norms Theory. The last scale, Outcome Expectancies for Non-Consensual Sex, showed an unexpected result: all items loaded strongly onto one or the other of two factors, neither of which could be explained by either theoretical framework. These two factors were identified as either "Risks of Non-Consensual Sex" versus "Rewards of Non-Consensual Sex." This may indicate that the Social Cognitive Theory and Social Norms Theory frameworks may be insufficient, and

a theory that takes into account a cost-benefit analysis would be useful for explaining outcome expectancies regarding non-consensual sex.

Therefore, all five metrics showed adequate validity according to the established theoretical frameworks except for the Outcome Expectancies for Non-Consensual Sex scale, which may consist of items and responses governed by another theoretical framework other than those upon which the RealConsent study was based.

In any behavioral science intervention, outcome assessment is a necessary component of program evaluation and is essential to creating increasingly effective interventions in the future. To improve the efficacy of subsequent studies, researchers must understand how past studies have succeeded, as well as how they have failed. By assessing the usefulness of new metrics of intervention impact, the present analysis is a part of this goal to constantly improve the ability of public health to modify human behavior. Future studies that intend to address the threat of male physical and sexual violence against women can now do so with these two new tools: the Perceptions of Informed Consent for Sex index, and the Self-Efficacy to Intervene scale.

Future Directions

Though construct validity was assessed using the factor analysis method, an additional means to measure construct validity is through the *criterion method*. The criterion method assesses the construct validity of a psychometric scale by measuring the correlation between self-reported attitudes and observed change in behavior (25). A follow-up study which included an outcome measurement of observed behavior would

allow the construct validity of these metrics to be further assessed through the criterion method.

Additionally, *face validity* was not assessed in this analysis. To assess this type of validity, a panel of experts must be convened to decide by vote whether each of these novel scales and indexes appears to measure the intended underlying construct (25). A further analysis to assess the face validity may help either support or refute the conclusions made here about construct validity of these five metrics.

In addition to the successful verification of the Perceptions of Informed Consent for Sex index and the Self-Efficacy to Intervene scale, the Outcome Expectancies for Engaging in Non-Consensual Sex scale was shown to be reliable once at least one item (either #1 or #12) was dropped. With this modification, the Outcome Expectancies for Engaging in Non-Consensual Sex scale may prove to be an additional useful metric in the field of behavioral science and for women's health interventions.

The final two metrics designed by the RealConsent study team: the Outcome Expectancies for Intervening Behaviors scale and the Legal Knowledge index, while showing a lack of reliability by the present analysis, may still be a useful starting point for the creation of future scales and indexes, as they address particular behavioral mediators that have not been previously considered by existing metrics.

Table 1: Test-Retest Reliability Coefficients of Novel Metrics Developed for the RealConsent Study

Scales	Test-Retest Pearson Correlation Coefficient			
	From t=1 to t=2	From t=2 to t=3	From t=1 to t=3	Mean PCC (across time intervals)
1. Legal Knowledge Index	0.477* (n=179)	0.507* (n=88)	0.476* (n=89)	0.487
2. Perceptions of Informed Consent for Sex	0.687* (n=172)	0.701* (n=82)	0.653* (n=87)	0.680
3. Self-efficacy to Intervene	0.786 * (n=174)	0.711* (n=85)	0.538* (n=87)	0.678
4. Outcome Expectancies for Engaging in Non-Consensual Sex	0.718* (n=153)	0.628* (n=77)	0.621* (n=72)	0.656
5. Outcome Expectancies for Intervening Behaviors	0.678* (n=175)	0.645* (n=85)	0.505* (n=87)	0.609

*p<0.01

Table 2a: Cronbach's Alpha (α) Internal Consistency Assessment of Novel Scales Developed for the RealConsent Study

Scales	Cronbach's Alpha (α)			
	Time 1 (Pre-Intervention)	Time 2 (Post-Intervention)	Time 3 (6 Month Follow up)	Average
1. Legal Knowledge Index				
2. Perceptions of Informed Consent for Sex				
3. Self-efficacy to Intervene	0.952* (n=733)	0.967* (n=435)	0.955* (n=212)	0.958
4. Outcome Expectancies for Engaging in Non-Consensual Sex	0.651* (n=686)	0.686* (n=428)	0.719* (n=202)	0.685
5. Outcome Expectancies for Intervening Behaviors	0.821* (n=726)	0.860* (n=436)	0.799* (n=210)	0.826

*p<0.001

Table 2b: Change in Cronbach's Alpha (α) for Outcome Expectancies for Engaging in Non-Consensual Sex Scale if Items Deleted

	Time 1 (Pre-Intervention)	Time 2 (Post-Intervention)	Time 3 (6 Month Follow up)	Average
4. Outcome Expectancies for Engaging in Non-Consensual Sex (minus item # 1)	0.687* (n=686)	0.714* (n=428)	0.745* (n=202)	0.715
4. Outcome Expectancies for Engaging in Non-Consensual Sex (minus item # 12)	0.679* (n=686)	0.711* (n=428)	0.749* (n=202)	0.713

*p<0.001

Table 3: Split-Half Reliability Assessment of Novel Scales Developed for the RealConsent Study

Scales	Spearman-Brown Correlation Coefficient			
	Time 1 (Pre-Intervention)	Time 2 (Post-Intervention)	Time 3 (6 Month Follow up)	Average
1. Legal Knowledge Index				
2. Perceptions of Informed Consent for Sex				
3. Self-efficacy to Intervene	0.928* (n=733)	0.954* (n=435)	0.926* (n=212)	0.936
4. Outcome Expectancies for Engaging in Non-Consensual Sex	0.684* (n=686)	0.766* (n=428)	0.746* (n=202)	0.732
5. Outcome Expectancies for Intervening Behaviors	0.806* (n=726)	0.841* (n=436)	0.775* (n=210)	0.807

*$p<0.01$

Table 4: Exploratory Factor Analysis of Novel Metrics Developed for the RealConsent Study

Scales	Expected Factors	Observed Factors	Cumulative % Variance Accounted For	Factor Loading (% of Factors loading) by loadings (PCCs) greater than:		
				Adequate (>0.4)	Good (>0.5)	Strong (>0.7)
1. Legal Knowledge Index	2	2	45.27%	77.8%	66.7%	22.2%
2. Perceptions of Informed Consent for Sex	4	3	44.98%	100%	92.8%	35.7%
3. Self-efficacy to Intervene	4	2	62.73%	100%	94.4%	55.5%
4. Outcome Expectancies for Engaging in Non-Consensual Sex	3	2	64.20%	100%	100%	86.7%
5. Outcome Expectancies for Intervening Behaviors	3	3	51.288	94.1%	94.1%	50.0%

References

1.) Brener, ND, McMahon, PM, Warren, CW, Douglas, KA. Forced Sexual intercourse and associated health-risk behaviors among female college students in the United States. *Journal of Consulting and Clinical Psychology.* 1999; *67*(2), 252-259.

2.) Koss, MP, Woodruff, JW, Koss, PG. Relation of criminal victimization to health perceptions among women medical patients. *Journal of Consulting and Clinical Psychology.* 1990; 58, 147-152.

3.) Harned, M. Abused women or abused men? An examination of the context and outcomes of dating violence. *Violence & Victims.* 2001; 16(3), 269-285.

4.) Rouse, LP. Abuse in dating relationships: A comparison of Blacks, Whites, and Hispanics. *Journal of College Student Development.* 1988; *29*, 312-319.

5.) Rennison, CM. Intimate partner violence and age of victim, 1993-99 (Bureau of Justice Statistics Special Report NCJ 187635. Washington, DC: U. S. Department of Justice. 2001.

6.) White, J, Williams, LV, Cho, D. A social norms intervention to reduce coercive behaviors among deaf and hard-of-hearing college students. *The Report on Social Norms.* 2003; *2*(4).

7.) Warshaw, R. I Never Called It Rape: The Ms. Report on Recognizing, Fighting, and Surviving Date and Acquaintance Rape. New York, NY. Harper Perennial. 1994.

8.) Fisher, BS, Cullen, FT, Turner, MG. The Sexual Victimization of College Women. National Institute of Justice, Bureau of Justice Statistics. USDOJ: NCJ 182369. 2000.

9.) Earle, JP. Acquaintance rape workshops: Their effectiveness in changing the attitudes of first year men. *National Association of Student Personnel Administrators (NASPA) Journal.* 1996; *34*(1), 2-16.

10.) Bandura, A. Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall. 1986.

11.) Perkins, HW, Berkowitz, AD. Perceiving the community norms of alcohol use among students: some research implications for campus alcohol education programming. *International Journal of the Addictions.* 1986; 21, 961-976.

12.) Streiner, DL, Norman GR. Health Measurement Scales: A Practical Guide to Their Development and Use. New York, NY. Oxford University Press. 2008.

13 .) Loh, C, Gidycz, CA, Lobo, TR, Luthra, R. A prospective analysis of sexual assault perpetration: Risk factors related to perpetrator characteristics. Journal of Interpersonal Violence. 2005; 20(10). 1325-1348.

14.) Check, JV, Malamuth, NM, Elias, B, Barton, S. On hostile ground. *Psychology Today*,56-61. 1985.

15.) Price, E, Byers, E. The Attitudes Towards Dating Violence Scales: Development and initial validation. *Journal of Family Violence.* 1999;14(4), 351-375.

16.) Payne, DL, Lonsway, KA, Fitzgerald, LF. Rape myth acceptance: Exploration of its structure and its measurement using the Illinois Rape Myth Acceptance Scale. *Journal of Research in Personality*. 1999; 33 (1), 27-68.

17.) Dietz, S, Blackwell, K, Daley, P, Bentley, B. Measurement of empathy toward rape victims and rapists. *Journal of Personality and Social Psychology*. 1982; 43, p. 372-383.

18.) Straus, MA, Hamby, SL, Boney-McCoy, S, Sugarman, DB. The Revised Conflict Tactics Scales: Development and preliminary psychometric data. *Journal of Family Issues*. 1996; 17, 283-316.
19.) Koss, MP, Gidycz, CA. Sexual Experiences Survey: Reliability and validity. *Journal of Consulting and Clinical Psychology*. 1985; 53, 422-423.
20.) Lanier, CA, Elliot, MN. A new instrument for the evaluation of a date rape prevention program. *Journal of College Student Development.* 1997; 38(6), 673-676.
21.) Loh, C, Gidycz, CA, Lobo, TR, Luthra, R. A prospective analysis of sexual assault perpetration: Risk factors related to perpetrator characteristics. Journal of Interpersonal Violence. 2005; 20(10). 1325-1348.
22.) Banyard, VL, Plante, EG, Moynihan, MM. Rape Prevention through Bystander Education: Bringing a Broader Community Perspective to Sexual Violence Prevention. National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. 2005.
23.) Salazar, LF, DiClemente RJ, Crosby RA. Measurement and Design Related to Theoretically-Based Health Promotion Research and Practice. *Research Methods in Health Promotion*. San Francisco, CA. 2011; 253-286.
24.) Rosenberg, M. Society and the adolescent self-image. Princeton, NJ: Princeton University Press. 1965.
25.) DiClemente, RJ, Salazar, LF, Crosby, RA. Health Behavior Theory for Public Health: Principles, Foundations, and Applications. Jones and Bartlett Learning, LLC. 2013.
26.) Bandura, A. Self-Efficacy: The Exercise of Control. W.H. Freeman Press. 1997.