Spring 5-2019

VALIDATION OF THE CONFLICT IN ADOLESCENT DATING RELATIONSHIPS INVENTORY (CADRI) IN YOUNG ADOLESCENTS: AN ANALYSIS OF CONSTRUCT VALIDITY AND MEASUREMENT INVARIANCE

Stephanie J. Stroever
UTHealth School of Public Health

VALIDATION OF THE CONFLICT IN ADOLESCENT DATING RELATIONSHIPS INVENTORY (CADRI) IN YOUNG ADOLESCENTS: AN ANALYSIS OF CONSTRUCT VALIDITY AND MEASUREMENT INVARIANCE

by

STEPHANIE J STROEVER, BS, MPH

APPROVED:

MELISSA PESKIN, PHD

CASEY DURAND, PHD, MPH

ROBERT ADDY, PHD

CHRISTINE MARKHAM, PHD

DEAN, THE UNIVERSITY OF TEXAS SCHOOL OF PUBLIC HEALTH

VALIDATION OF THE CONFLICT IN ADOLESCENT DATING RELATIONSHIPS INVENTORY (CADRI) IN YOUNG ADOLESCENTS: AN ANALYSIS OF CONSTRUCT VALIDITY AND MEASUREMENT INVARIANCE

Stephanie J Stroever, BS, MPH

The University of Texas

School of Public Health, 2019

Dissertation Supervisor: Casey Durand, PhD, MPH

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

INTRODUCTION

Adolescent dating violence, defined by the Centers for Disease Control and Prevention as "physical, sexual, psychological, or emotional aggression within a dating relationship, including stalking," among adolescents in intimate relationships is a serious public health concern (Centers for Disease Control and Prevention, 2018). Per the 2017 national Youth Risk Behavior Surveillance (YRBS) survey of high school students in the United States, 10.7 percent of females and 2.8 percent of males that reported a romantic relationship in the prior 12 months experienced sexual dating violence (Kann et al., 2018). Additionally, 9.1 percent of females and 6.5 percent of males reported physical dating violence, which included but was not limited to, being hit, punched, slammed into a wall, or otherwise injured (Kann et al., 2018).

Dating violence occurs not only in high school-aged adolescents (i.e. 14-18 years old), but in younger adolescents as well. The baseline survey of the *Dating Matters* initiative found that among sixth, seventh, and eighth graders (ages 11-13) who had dated, 77 percent had perpetrated verbal or emotional abuse, 32 percent physical abuse, and 15 percent sexual abuse (Niolon et al., 2015).

Dating violence among adolescents can result in serious mental and physical health outcomes. Longitudinal studies have shown that compared to individuals who did not experience dating violence, victimized adolescents are at greater risk of tobacco and marijuana use, depressive symptoms and suicidal ideation, heavy episodic drinking and binge eating, and antisocial behavior (Ackard, Eisenberg, & Neumark-Sztainer, 2007; D. Exner-Cortens, Eckenrode, & Rothman, 2013). Adolescents who experience dating violence are also at greater

risk of future victimization in late adolescence and adulthood when compared to those who have not experienced dating violence (D. Exner-Cortens et al., 2013; Deinera Exner-Cortens, Eckenrode, Bunge, & Rothman, 2017).

PAPER 1

Background

Adolescent Dating Violence Measurement

The CDC defines adolescent dating violence as psychological, emotional, physical, and/or sexual aggression within a dating relationship (Centers for Disease Control and Prevention, 2018). Adolescent dating violence may also be conceptualized as harm or distress caused by a dating partner. Differences in individuals' perception of aggression and harm make adolescent dating violence a difficult concept to measure. Most measures are self-report and rely on an individual's interpretation of events as harmful or distressing. Since these perceptions cannot be clearly observed, adolescent dating violence is essentially a latent variable. Latent variables cannot be measured directly via observation. Rather, the definition and set of concepts that describe the variable must be operationalized from the abstract into observable indicators (Kline, 2016). Operationalization of adolescent dating violence into observable indicators should include questions about specific behaviors that reflect aggression, harm, or distress following interaction with a dating partner (i.e. hitting, slapping, shaming, threatening, or forcing unwanted sexual contact).

It is critical that the definitions and indicators used to measure adolescent dating violence be consistent across research and practice to accurately capture the magnitude of adolescent dating violence and evaluate prevention programs. However, as with other latent variables (i.e. depression), the definition and operationalization of adolescent dating violence often varies (Brown, 2015; Cascardi, Blank, & Dodani, 2016; Kline, 2016). Some measures only ask about

victimization (i.e. Juvenile Victimization Questionnaire, Date and Family Violence Abuse Scale), while others only ask about perpetration (i.e. Dating Violence Perpetration Acts Scale) (Smith et al., 2015). A systematic review of behavioral measures for adolescent dating violence (n = 130) found at least 48 different measures have been used to capture adolescent dating violence (Smith et al., 2015). The Conflict Tactics Scale – 2 (CTS-2) was most frequently used (24%), followed by the Safe Dates (22%) and Conflict in Adolescent Dating Relationships Inventory (CADRI) (15%) (Smith et al., 2015). These measures include indicators for physical, psychological, and sexual dating violence, although the method in which they are operationalized also varies. For example, the CTS-2 asks about more severe forms of sexual dating violence, while the CADRI includes milder forms (i.e. unwanted kissing) (Straus, Hamby, Boney-McCoy, & Sugarman, 1996; Wolfe et al., 2001). If the two measures were given to the same sample of adolescents, those who answered the CADRI confirming unwanted kissing, and therefore dating violence, would not be captured by the CTS-2 as having experienced dating violence. This discrepancy creates a problem with adolescent dating violence prevalence estimation and surveillance.

Additionally, adolescent dating violence measures do not always include all concepts of the definition. Prevalence estimates for adolescent dating violence will be different and potentially under-representative of the true scope of the problem if inconsistent measures are used that do not capture the full conceptualization of adolescent dating violence. For example, the YRBS only measures physical and sexual dating violence, while the Interpersonal Control Scale, Psychological Abuse Scale, and Dyadic Adjustment Scale only capture psychological

dating violence (Smith et al., 2015). Digital, or cyber, abuse is a relatively new concept in dating violence following the rapid expansion of technology and media use by adolescents over the last 20 years (Smith et al., 2015; Zweig, Dank, Yahner, & Lachman, 2013). Experts debate whether cyber dating violence serves as a unique construct or is another manifestation of psychological dating violence (Smith et al., 2015). Indeed, some items that measure cyber dating violence resemble psychological items (e.g. spreading rumors, calling names), while others are unique to electronics (e.g. posting private or embarrassing photos) (Zweig et al., 2013). A review of adolescent dating violence behavioral measures by Smith and colleagues in 2015 found that only 1 of 24 measures developed to capture dating violence included questions about electronic perpetration of dating violence (Smith et al., 2015). Not included in the Smith article, Cutbush et al. found acceptable goodness-of-fit using items from the Youth Internet Safety Survey and Cyber Dating Violence Scales to describe a first order cyber abuse factor (Cutbush & Williams, 2016). Though a promising lead, additional work is needed to validate cyber abuse measures in adolescent populations.

Lastly, adolescent dating violence measures vary in the extent to which they have been validated across study populations. A measure itself does not possess validity outside of the context in which it is applied (Streiner, Norman, & Cairney, 2015). Though some assumptions for validity can be made across samples similar in age, gender, and other grouping categories, it is important to verify construct validity in each unique sample studied (Streiner et al., 2015). For example, the CTS-2 was originally validated in college-aged adults (Smith et al., 2015; Straus et al., 1996; Wolfe et al., 2001). There is evidence that high school-aged adolescents are at different

developmental stages than those in college (Wolfe et al., 2001). Therefore, the operationalization of adolescent dating violence among high school students may need to be different to accurately measure the latent variable in this age group. The same is true when adolescent dating violence measures are used in middle school-aged adolescents. Once again, there are developmental differences in dating relationships that may result in the inaccurate or invalid capture and interpretation of adolescent dating violence.

Conflict in Adolescent Dating Relationships Inventory

Prior to the development of the CADRI, items used to measure adolescent dating violence were adapted from adult measures (i.e. CTS-2, Psychological Maltreatment of Women Inventory) and were questionably sensitive to developmental differences in dating relationships (Wolfe et al., 2001). The CADRI was developed and validated in a large sample of high school-aged adolescents to address this gap (Wolfe et al., 2001). Developers used focus groups to create an initial questionnaire that was then pilot tested among 14-16-year old students. They made changes to the questionnaire based on the factor structure from exploratory factor analysis and then tested the revised version in a large sample of 9th-11th graders (n = 1019) (Wolfe et al., 2001).

The CADRI measurement model included one second order factor (abuse) and five first order factors (threatening behavior, relational aggression, physical abuse, sexual abuse, and verbal emotional abuse). A second order factor is one that explains the shared correlation between first order factors, as seen with abuse and the defined dimensions of adolescent dating violence. The survey was written to capture both perpetrator and victim behaviors and was

developed to be gender specific (Wolfe et al., 2001). Threatening behavior in male perpetrators was operationalized as 4 items (e.g. “I deliberately tried to frighten her”), while relational aggression was operationalized as 3 items (e.g. “I spread rumors about her.”) Physical abuse included 4 items (e.g. “I pushed, shoved, or shook her”), as did sexual abuse (e.g. “I kissed her when she didn’t want me to”). Lastly, the verbal emotional abuse factor included 10 items (e.g. “I said things just to make her angry”). The measures were assessed on a 4-point Likert-type scale (0 – never, 1 – seldom, 2 – sometimes, and 3 – often) (Wolfe et al., 2001).

Validation studies. Confirmatory factor analysis revealed a modest fit to the second order model hypothesized by Wolfe et. al., with X^2 (272) = 605.41, goodness of fit index (GFI) = 0.90, comparative fit index (CFI) = 0.86, and root mean square of error approximation (RMSEA) = 0.06 (Wolfe et al., 2001). The second order factor structure was also verified in a study of a shortened version of the CADRI (CADRI-S) (Fernández-González, Wekerle, & Goldstein, 2012). However, when translated into Spanish, there was some inconsistency with items loading on the relational aggression and threatening behavior factors (Deinera Exner-Cortens, Gill, & Eckenrode, 2016). This difference may be the result of cultural or language differences and should be considered when converting the CADRI to a different language.

Of note, Shorey and colleagues found a lower-order factor structure with five independent, correlated first order factors to have the better fit to study data in a sample of high school students than the second order factor structure developed by Wolfe (Shorey et al., 2018). Chi-square difference testing between the nested and parent models resulted in a p-value less than 0.05, suggesting the lower-order model was the superior model (Shorey et al., 2018).

In practice. Historically, the CADRI has measured adolescent dating violence in studies with both middle school and high school-aged adolescents (Deinera Exner-Cortens et al., 2016; Smith et al., 2015) as well as young college aged adults (mean age = 20.8 years, SD = 2.05) (Markham et al., 2017). It has also been used to assess the stability of adolescent dating violence status over time (Choi & Temple, 2016) and to examine the age of onset for adolescent dating violence perpetration (Shorey et al., 2017). It was also used to measure the prevalence of adolescent dating violence and associated risk factors in high-risk middle school students (Niolon et al., 2015), to investigate latent classes of adolescent dating violence (Reidy et al., 2016), and to evaluate the impact of several adolescent dating violence prevention programs (Joppa, Rizzo, Nieves, & Brown, 2016; Rowe, Jouriles, & McDonald, 2015; Taylor, Mumford, & Stein, 2015). It was also used recently with a sample of sixth grade adolescents to evaluate the impact of an intervention (Me & You: Building Healthy Relationships*)* designed to encourage healthy relationships and reduce the odds of adolescent dating violence (Peskin et al., under review).

Cyber Dating Violence Measure

Little has been done to evaluate the psychometric properties of cyber adolescent dating violence measures despite use in prior research (Peskin et al., 2017; Picard, 2007; Temple et al., 2016; Zweig et al., 2013). One measure was developed by the Teen Research Unlimited (TRU) group in a study of tech abuse in teen relationships (Picard, 2007). Psychometric properties for the survey were not published with the research findings in the publicly available report. However, Zweig et al. adapted the measure used by TRU for a study of cyber adolescent dating

violence. The total measure included 16 items, with 6 from the original TRU study and 10 added specifically for the study of interest (Zweig et al., 2013). Internal consistency was reported via Cronbach's alpha metrics. The sexual cyber dating violence measures (4 items; i.e. partner sent naked photos of himself/herself when she/he knew they were unwanted) demonstrated good internal consistency with $\alpha = .810$ for victimization and $\alpha = .885$ for perpetration. Non-sexual cyber dating violence (12 items; i.e. sending threatening messages, using partners social media account without permission, etc.) also demonstrated high internal consistency with $\alpha = .891$ for victimization and $\alpha = .923$ for perpetration. Though internal consistency provides some information about the reliability of the measure, additional testing is needed to verify the construct validity of the measure in young adolescents.

Specific Aims

Despite its use, the CADRI has not been validated in middle school-aged adolescents, nor have measures for cyber dating violence. Thus, the purpose of this study was to determine if the factor structure of the CADRI, demonstrated to be valid in older adolescents, holds among younger adolescents. The primary aim was to use confirmatory factor analysis to establish construct validity of the CADRI in a diverse sample of sixth grade adolescents from a large, urban school district in Southeast Texas and characterize the relationship between the latent factors and their items. The second aim was to use confirmatory factor analysis to assess the discriminant validity of cyber dating violence as a distinct construct of adolescent dating violence and characterize the relationship between the items and latent factor. Adaptations to the models were also explored to strengthen the validity of the measure in this age group and allow

the adjusted CADRI to be used confidently in a young, middle school-aged population in future studies.

Methods

Sample and Data Collection

Data were collected during a randomized controlled trial of the Me & You: Building Healthy Relationships (Me & You) study in 2014 from 10 middle schools in a large, urban school district in Southeast Texas. The baseline data from sixth graders whose parents consented to their participation were used for this analysis. Of note, only participants who indicated they had ever had a boyfriend or girlfriend were included in the dating analytic sample. A total of 826 students completed the baseline survey, with 51.7% (n = 424) included in the dating analytic sample (Table 1).

Of the students who reported having had a relationship, the mean age was 12.4 (SD = 0.62) and 43.9% were female. The participants were predominately Hispanic (60.8%), with 30.0% African American. Data were collected via computer-assisted self-report surveys. Students completed the baseline survey at school on school-owned computers. Only participants that affirmed having ever had a boyfriend or girlfriend were asked dating violence questions.

Measure

The survey used items from the CADRI and sought to measure 6 facets of adolescent dating violence (Appendix B), including physical abuse (4 items; “I threw something at him/her”), verbal emotional abuse (10 items; “I did something to make him/her jealous”), relational aggression (3 items; “I spread rumors about him/her”), threatening behavior (4 items;

"I tried to frighten him/her on purpose"), sexual abuse (1 item; "I kissed him/her when he/she didn't want me to"), and cyber abuse (12 items; "I sent him/her sexual photos or naked photos of me that I knew he/she did not want").

The self-report survey asked participants to consider prior experience with a boyfriend or girlfriend and respond to the statements according to a dichotomous scale (0 = no, 1 = yes). Those who agreed with the statement were asked to mark "yes," and those who disagreed were asked to mark "no." Each set of questions was asked twice to capture both victim and perpetrator experiences. However, following the precedent set by prior research, only perpetration questions were included in the confirmatory factor analyses (Cutbush & Williams, 2016; Shorey et al., 2018; Wolfe et al., 2001).

CADRI adaptations in Me & You. Some language of the measure was simplified to ensure readability and understanding in sixth graders. These minor word changes were pre-tested by an advisory panel of teens and are noted in Appendix B. The cyber abuse items were not part of the original CADRI and were selected from prior studies of cyber dating violence (Zweig et al., 2013). The original CADRI also included three other sexual abuse items ("I touched him/her sexually when he/she didn't want me to;" "I forced him/her to have sex when he/she didn't want to;" and "I threatened him/her in an attempt to have sex with him/her") (Wolfe et al., 2001). However, the content of the intervention did not address this level of sexual intimacy. Therefore, those questions were excluded.

Additionally, the original CADRI used a Likert scale for response options, while this study used a dichotomous scale. Lastly, the introduction to each set of questions was altered

from the original CADRI. The original CADRI asked participants to consider whether they experienced the behavior during an argument with their partner (Wolfe et al., 2001). The Me & You study asked participants if the behavior had ever occurred. This change may increase the frequency of endorsement since it removes the context of conflict from the behavior. However, the developers of the Me & You study were most interested in any dating violence among sixth grade students, not just that which occurred during conflict. Therefore, this introduction to the survey is appropriate in this context.

Analysis Strategy

MPlus version 8.2 was used for factor analyses (Muthen & Muthen, 2018). We used a robust weighted least squares estimator (WLSMV) with a tetrachoric correlation matrix to accommodate dichotomous data that often violate key assumptions of the maximum likelihood estimator (Brown, 2015). We performed confirmatory factor analyses in two steps to test the primary aim of the study. First, we estimated a first order model to ensure items loaded appropriately on first order, correlated latent factors. Then, we tested the CADRI's second order factor structure (published originally by Wolfe et al.) to determine if it fit as well as the first order model (Wolfe et al., 2001).

Importantly, the addition of a second order factor cannot improve the fit of the model. Rather, the test for model fit determines if the addition of the second order factor decrements the overall fit of the model (Brown, 2015). The benefit of a second order factor structure is a reduction in the number of parameters estimated and a more parsimonious model. We performed chi-square difference testing to test the hypothesis that the first and second order factor structures

were not different. A p-value <0.05 indicated that the first order factor model was the better model for the data.

Overall goodness-of-fit was assessed with parsimony and comparative goodness-of-fit indices. Though the chi-square test of model fit has limitations with non-normal data (Brown, 2015), we included it in model fit assessment. We corroborated the results with additional indices including the root mean square error of approximation (RMSEA), comparative fit index (CFI), and Tucker-Lewis index (TLI). The RMSEA is a test of parsimony that is regularly reported in confirmatory factor analysis. Values below 0.06 suggest a decent model fit to the data. The CFI and TLI assess the fit of the hypothesized model to more restricted "null" models (Brown, 2015). Values closer to one (>0.95) suggest good model fit to the data. We also examined residuals and modification indices to identify any focal areas of ill fit. Residuals ≥ |2.58| and modification indices >4.0 were evaluated for theoretically sound modifications to the model.

Model Specification

Model 1. The first model tested in this analysis included four first order latent factors, with the latent factor for physical abuse (PHYS) informed by P1-P4. The threatening behavior factor (THREAT) was informed by TB1-TB4, while the relational aggression factor (RELATION) was informed by RA1-RA3. Lastly, the verbal emotional abuse factor (VERB) was informed by VE1-VE10. The Me & You study only included one question for sexual abuse, and the model was not identified with a single item factor. We also could not remedy this

identification issue by assigning a starting value, because the data were categorical (Muthen & Muthen, 2017). Therefore, the sexual abuse item, S1, was removed from the analysis.

Model 2. We included the same four first order factors from Model 1 in Model 2. We also added a single second order factor, ABUSE, to account for the hypothesized higher order dimension of adolescent dating violence. ABUSE was informed by the latent factors PHYS, THREAT, RELATION, and VERB.

Model 3. We assessed the secondary aim of this study, to determine the discriminant validity of cyber dating violence measures, with this model. Cyber dating violence items were added as a fifth first order latent factor, CYBER, to the lower order factor structure of Model 1. Of note, only items C1-C5 and C7-C12 were included in the model. The item, C6 ("I sent him/her text messages to check up on him/her"), was removed because it was ambiguous in its representation of dating violence. It is possible young adolescents performed that behavior outside of the scope of dating violence or aggression.

Results

Descriptive statistics. We assessed the missingness of the data and all items met a minimum coverage proportion of 0.10 (Muthen & Muthen, 2017). Missing data ranged from 0.5 to 3.1 percent, with TB4 ("I threatened to hit him/her or throw something at him/her") yielding the most missing data. Bivariate analysis showed the missing data for this item were among Hispanic (n =9) and African American (n = 4) participants and both boys (n = 7) and girls (n = 6). We also evaluated the frequency of endorsement for each indicator item. Twenty-one of thirty-four items (61.7%) met the acceptable range (5-95 percent) of endorsement for

psychometric testing (Streiner et al., 2015). The lowest endorsements were among cyber items, with a range of 1.9-6.0 percent positive endorsement. The highest endorsements were among verbal emotional dating violence items (8.6-23.8 percent positive endorsement). Appendix C further displays the results of preliminary data screening, including missing data and item endorsement frequency.

Confirmatory factor analyses. The results of all confirmatory factor analyses are presented in Table 1. The first order factor structure (Model 1) demonstrated a strong fit to the data with a RMSEA = 0.030 (90% CI = 0.021, 0.039), CFI = 0.979, and TLI = 0.976. Model 2 also fit the data well, with a RMSEA = 0.032 (90% CI = 0.022, 0.040), CFI = 0.977, and TLI = 0.974. However, the chi-square difference test had a p-value <0.05, which suggests the second order model (Model 2) did not fit as well as the first order model (Model 1). Model 1 with four first-order factors was the better of the two for these data.

Despite the superiority of Model 1 in overall goodness-of-fit, there were issues with the specification of the model. There was poor discriminant validity among the independent factors with inter-factor correlations between 0.77 (PHYS-RELATION) and 0.91 (PHYS-THREAT). Ideally, independent factor correlations should fall below 0.85. Correlations above that standard suggest too many factors were included in the model (Brown, 2015). Thus, we chose to run an exploratory factor (EFA) analysis to statistically assess the appropriate number of factors for these data.

EFA provides statistical insight into the simplest model (i.e. fewest factors, fewest non-zero paths) that best explains the data (Loehlin, 2004). We used the robust weighted least squares

estimator in MPlus with oblique rotation and allowed the model to be estimated up to four factors. Oblique rotation allows latent factors to be correlated (Loehlin, 2004). The two-factor solution resulted in a chi-square value of 200.84 (df = 41, p = 0.047), while the three-factor solution had a chi-square value of 168.94 (df = 60, p = 0.137). This is consistent with prior research that defines adolescent dating violence in terms of only two factors; one related to physical harm and one related to emotional and psychological harm (Cascardi et al., 2016; Centers for Disease Control and Prevention, 2018; Niolon et al., 2015; Wolfe et al., 2001). Therefore, we combined physical abuse (PHYS) with threatening behavior (THREAT) into one physical aggression factor (PHYSAGG) and, we combined verbal emotional abuse and relational aggression into one psychological abuse factor (PSYCH).

The two-factor first order model (Model 4) demonstrated a similarly good fit to the data, with a RMSEA = 0.032 (90% CI = 0.022, 0.040), CFI = 0.976, and TLI = 0.974. The chi-square difference test resulted in a p-value <0.05, which indicated the original four-factor model was a better fit to the data. However, there is some disagreement in the field as to whether a model with fewer factors is truly a nested model and appropriate for chi-square difference testing (Brown, 2015). Additionally, there was little change in fit in the RMSEA, CFI, and TLI between Model 1 and 4. Thus, Model 4 is more appropriate for these data given the poor discriminant validity of Model 1 and the greater parsimony of Model 4.

The factor structure and standardized parameter estimates of Model 4 are displayed in Figure 1. Additionally, unstandardized factor loadings, standard errors, and R-square values of parameter estimates are displayed in Table 2. Most of the CADRI items loaded well on their

respective latent factors. However, VE7 (“I kept track of who he/she was with and where he/she was”) loaded on the psychological factor much lower than other items (standardized $\lambda = 0.368$). It is difficult to identify the reason for the misfit given our current methods, though it is possible the item was mis-interpreted or ambiguous in this sample. Further qualitative work should be done to understand the appropriateness of this item in a sixth-grade population.

Modification indices suggested allowing the correlation to be estimated between the residuals of VE9 (“I accused him/her of flirting with another guy/girl”) and VE10 (“I threatened to end the relationship”). This resulted in a statistically significant improvement to the model ($\Delta X^2 = 21.224$, $\Delta df = 1$, $p < 0.001$). Additionally, the residual of VE9 (“I accused him/her of flirting with another girl/guy”) was allowed to correlate with VE7 (“I kept track of who he/she was with and where he/she was”), resulting in another statistically significant improvement to the model ($\Delta X^2 = 12.76$, $\Delta df = 1$, $p < 0.001$). Lastly, the residual of P2 (“I kicked, hit, or punched him/her”) was allowed to correlate with the residual of P3 (“I slapped him/her or pulled his/her hair”), with a change in chi-square of 8.401 ($\Delta df = 1$, $p = 0.004$). The overall change in chi-square fit with the three added residual correlations was 42.675 ($\Delta df = 3$, $p = <0.001$).

Cyber dating violence. The secondary aim of this study was to assess the discriminant validity of a cyber dating violence measure and to characterize the relationship between the items and latent factor (Model 3). However, assessment of endorsement frequency revealed very low positive endorsement for most of the items. Only two items, C3 (“I used his/her social networking account (Facebook, Twitter, Instagram, etc) without his/her permission.”) and C9 (“I made him/her afraid when he/she did not respond to my phone call, text, posting on social

networking page, IM, etc.") had positive endorsement over 5%. Since the prevalence of these behaviors was so low, which can detract from the accuracy of the scale's psychometric properties, we chose not to include them as a subscale in this analysis (Streiner et al., 2015). Future studies should assess the validity of cyber dating violence in this population.

Discussion

The purpose of this study was to determine the construct validity of the CADRI, originally validated in high school students, in a sample of sixth grade adolescents. Though similarities are likely, there may also be important developmental differences among younger adolescents that impact the validity of the CADRI in this age group. For example, they have much less dating experience than high school aged adolescents and likely have less refined conflict resolution skills (Wolfe et al., 2001). Their relationships are more likely to be shorter and with low levels of commitment. Additionally, milder forms of physicality (i.e. shoving) and verbal aggression (i.e. teasing) are often used by young adolescents to relate to one another (Wolfe et al., 2001). Therefore, it was important to validate this measure for this age group.

We used confirmatory factor analyses to test the hypothesis that the original structure of the CADRI was valid in younger adolescents (e.g. twelve years old). We found that most of the items on the CADRI were indeed valid in our sample. The overall model fit of the CADRI was good, with parsimony and comparative fit indices almost identical across models. However, there were differences in parameter estimates that are worth consideration when selecting the most appropriate model for these data.

Broad versus nuanced CADRI. It is imperative that confirmatory factor analyses have substantive justification in model specification (Brown, 2015). The conceptualization of adolescent dating violence is highly variable across the field, and our trio of models contributes to the growing body of evidence that adolescent dating violence can be both nuanced and broad in its conceptualization.

Adolescent dating violence is broadly defined as any psychological, physical, or sexual aggression between dating partners (Centers for Disease Control and Prevention, 2018). Physical aggression includes the use of force against a dating partner that causes harm, whereas psychological aggression includes both verbal (i.e. put-downs and ridicule) and non-verbal (i.e. stomping from room) behaviors. Psychological aggression in a dating relationship is considered abuse when it intentionally causes symbolic or emotional harm to one's partner (Straus et al., 1996). These two constructs appear to be distinct, and numerous studies have employed this or similar definitions to the measurement of adolescent dating violence (Foshee et al., 1996; Halpern, Oslak, Young, Martin, & Kupper, 2001; Haynie et al., 2013; Sabina, Cuevas, & Cotignola-Pickens, 2016; Smith et al., 2015). The Safe Dates Psychological and Physical Abuse Scale also measures these broad concepts among adolescents, as do measures of intimate partner violence among adults (Foshee et al., 1996; Foshee, McNaughton Reyes, & Ennett, 2010; Straus et al., 1996).

Our two-factor model corresponds with the distinction between psychological and physical adolescent dating violence, while also remaining broad enough to capture a less nuanced and more parsimonious model. Psychometric evaluation of the Conflict Tactics Scale

demonstrated an overlap between physical and psychological abuse in male high school students. One construct included serious physical aggression (i.e. hitting, punching), whereas the other included milder forms of physical aggression (i.e. threatening, shoving) and psychological aggression (Cascardi, Avery-Leaf, O'Leary, & Slep, 1999). Additionally, though the threatening behavior construct correlates strongly with physical dating violence, it does not actually involve physical contact (Wolfe et al., 2001). Cascardi et al. argue the underlying constructs of adolescent dating violence should be defined by the potential of a behavior to cause harm (whether physical or emotional) and not by presence or absence of physical contact (Cascardi et al., 2016). This overlap supports Wolfe's original hypothesis that adolescent dating violence is less differentiated and occurs along a continuum than intimate partner violence seen among adults (Wolfe et al., 2001).

Wolfe and colleagues hypothesized that adolescent dating violence is a single, higher order dimension of abuse, with subfactors designed to distinguish between abusive behaviors (Wolfe et al., 2001). Our study found that both the second order factor structure and the first order, two-factor structure fit well in our sample of sixth grade adolescents. While both adequately explain the data in our sample, we selected the ideal model for this sample based on the principle of parsimony (Brown, 2015; Streiner et al., 2015). The two-factor model is more parsimonious, because it appropriately explains the data while estimating fewer parameters than the higher order model. However, further studies are needed to replicate these findings and build a broader evidence base for measurement of adolescent dating violence in young adolescents.

Modifications to the model. We allowed several sets of residual correlations to be estimated that resulted in an improved model fit. First, the residual of VE9 ("I accused him/her of flirting with another girl/guy") was allowed to correlate with that of VE10 ("I threatened to end the relationship"). This residual correlation was also identified in the original validation studies of the CADRI among high school students (Wolfe et al., 2001). Though the source of this shared error is not clear, it is possible that these two items are correlated due the timing of the behavior or desire to establish control in the relationship. Accusing one's partner of flirting with someone else may be a precursor to threatening to end the relationship (Jackson, 1999; Wolfe et al., 2001). The residual of VE9 was also allowed to correlate with that of VE7 ("I kept track of who he/she was with and where he/she was"). Again, the correlation may represent an unmeasured effect of timing or control in the relationship. Adolescents who are concerned with infidelity may be more likely to track another's movement, as well as accuse him or her of being unfaithful (i.e. flirting with someone else).

Lastly, we allowed the residual of P2 ("I kicked, hit, or punched him/her") to be correlated with that of P3 ("I slapped him/her or pulled his/her hair"). It is possible these two items represent various methods for the same overall action of physical harm and did not need to be separated into distinct items. The residual correlations may also be the result of unmeasured method effects, such as those related to a self-report questionnaire or the order of the items in the survey (Brown, 2015). Social desirability bias may also play a role in the correlation of the physical abuse items (Wolfe et al., 2001). For example, it may be more acceptable to slap or pull hair, whereas punching, hitting, and kicking may be considered less acceptable.

Practical Implications. Based on our findings, researchers and clinicians can confidently use the CADRI to measure dating violence among young, racial/ethnic-minority adolescents. The two-factor model displayed in Figure 1 is the most parsimonious and includes the latent constructs of both physical and psychological aggression. Caution should be used when describing sub-types of physical and psychological aggression in this age group, as we found there was poor discriminant validity between them.

Since adolescent dating violence is a latent construct, it may be advisable for future analyses to be conducted in the context of structural equation modeling methods. These methods maintain the unique properties of the latent factor, including information about the amount of variance in the indicators explained by the latent factor, as well as the variance in the indicator that is not explained by the factor (i.e. the residual or error variance) (Kline, 2016). The measurement model also includes correlations that account for shared variance in the residuals (i.e. the correlation between VE9 and VE10). In traditional regression analyses, this extra information is lost when a latent variable is modeled as an observed variable. Additionally, in SEM, all indicators from the measure can be included in the analysis rather than reducing the data to a dichotomous variable. A latent variable modeled in SEM is a continuous variable that provides substantially more information about the participant than a dichotomized measure (Streiner et al., 2015). Dichotomization is often accomplished by acknowledging ADV as "present" if at least one indicator is endorsed and "not present" if none are endorsed.

As demonstrated by the factor loadings from confirmatory factor analyses, not all indicators explain ADV as fully as others.

For example, psychological aggression explains 68.8 percent ($r^2 = 0.688$) of the variance in VE5 ("I insulted him/her with put downs"), but only 13.6 percent ($r^2 = 0.136$) of VE7 ("I kept track of who she was with and where she was"). Similarly, the latent construct of physical aggression explains 80.3 percent ($r^2 = 0.803$) of the variance in P3 ("I slapped him/her or pulled his/her hair") but only 42.7 percent ($r^2 = 0.427$) of the variance in TB2 ("I deliberately tried to frighten her"). Therefore, allowing endorsement of any of the indicators to reflect ADV equally may have limitations.

If analyses are performed outside of structural equation modeling methods, it is advisable to carefully consider each item's representation of adolescent dating violence prior to setting a threshold for perpetration or victimization. As discussed above, not all items explain the underlying construct of adolescent dating violence as well as others. Further analyses should be conducted in this population to determine appropriate thresholds for inclusion when utilizing the data as dichotomous.

Limitations. There were several limitations in our study that are worthy of consideration. First, our sample was drawn from one geographical region of Texas and may not be generalizable to other regions of the state or United States. Additionally, our sample was predominately low income, urban Hispanic and African American adolescents. Our results may not be representative of all sixth-grade adolescents of different socioeconomic status or race and ethnicity. Our study relied on self-report measures of behavior and were subject to traditional recall and social desirability biases inherent in this method of measurement. There may have also been differences in interpretation of items on the survey among participants.

Another important limitation of this study was the low endorsement of cyber dating violence items. The low prevalence of cyber dating violence in this sample prevented us from testing the items for discriminant validity and characterizing the relationship with other dimensions of adolescent dating violence. Of note, our survey did not ask participants to identify whether they had access to digital devices (i.e. cell phones, computers, etc.) or participated in social media. A survey of thirteen to seventeen-year-old teenagers found that eight-five percent of African American participants had access to a smartphone, with seventy-one percent of white and Hispanic teens reporting access (Lenhart, 2015). However, the study did not include sixth-grade aged students. It is possible that the low frequency of affirmative endorsement was an artifact of limited access, and future studies of cyber dating violence should include a question about access to further characterize cyber dating violence.

Another limitation in our study was the use of a single item to measure sexual dating violence. The other sexual abuse items included on the CADRI were removed so as to be sensitive and developmentally appropriate to the young age of the sample. However, it is still possible that more severe sexual dating violence (i.e. unwanted sexual touching, forced intercourse) occurs at this age, and further research should be conducted to develop the most appropriate measure of sexual adolescent dating violence for this age group.

Lastly, we made minor adaptations to the CADRI to enhance readability and comprehension among our sixth-grade sample. However, it is possible changes to the items and the order in which they were presented led to unmeasured method effects. Additionally, the

psychometrics from this study only hold in the context of the given measure. Further adaptations to the measure may result in differences in parameter estimates in future studies.

Conclusions

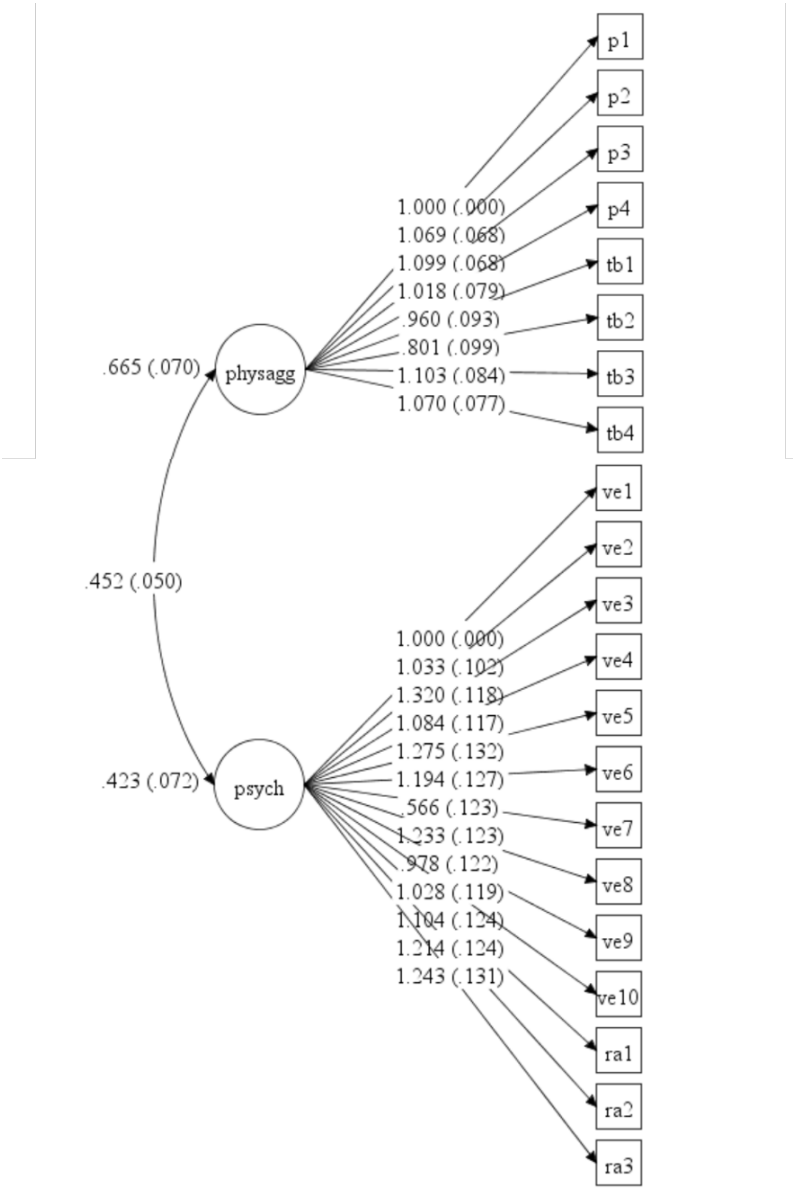
Studies have shown that adolescent dating violence may begin in adolescents as young as twelve years old. It is critical that experts who seek to prevent adolescent dating violence have the proper tools to assess the prevalence in this young population. This is the first study to assess the construct validity of the CADRI among sixth-grade adolescents. Researchers can go forward in the field with confidence that our adaptation of the CADRI is appropriate for the measurement of adolescent dating violence among low income, urban, racial/ethnic-minority adolescents.

Table 1. Overall goodness-of-fit statistics from model testing of the CADRI in the dating analytic sample of the Me & You Study (N = 423)

	X^2	df	p	ΔX^2	Δdf	p	RMSEA (90% CI)	CFI	TLI
Model 1	253.96	183	<0.001	-	-	-	0.030 (0.021, 0.039)	0.979	0.976
Model 2	263.26	185	<0.001	9.14*	2	0.010	0.032 (0.022, 0.040)	0.977	0.974
Model 3	Item endorsement too low for psychometric testing								
Model 4	267.66	188	<0.001	14.67*	5	0.012	0.032 (0.022, 0.040)	0.976	0.974
Model 4M	218.87	185	0.045	42.68**	3	0.000	0.021 (0.004, 0.031)	0.990	0.989

Model 1: Four 1st order factors (PHYS, THREAT, RELATION, VERB);
Model 2: One 2nd order factor (ABUSE) with four 1st order factors (see Model 1);
Model 3: Five 1st order factors (PHYS, THREAT, RELATION, VERB, CYBER);
Model 4: Two 1st order factors (PHYSAGG, PSYCH);
Model 4M: estimated correlation between P2-P3, VE9-VE10, and VE7-VE9
RMSEA = root mean square error of approximation; CFI = comparative fit index; TLI = Tucker-Lewis index; X^2 = chi-square test of model fit;
*Chi-square DIFFTEST versus Model 1
**Chi-square DIFFTEST versus Model 4

Figure 1. Unstandardized parameter estimates of a two-factor first order structure (Model 4) of the Conflict in Adolescent Dating Relationships Inventory (CADRI) in the dating analytic sample of the Me & You study.



Residual estimates found within parentheses. Correlations between residual variances fixed to zero; first indicator item for each factor fixed to 1.0 for scaling.

Table 2. Parameter estimates for Model 4 of the CADRI in the dating analytic sample of the Me & You Study

Factor	Indicator	Unstandardized Estimates		Standardized Estimates[b]	R-square
		Loading (SE)[a]	Threshold (SE)	Loading (SE)	
Physical Aggression (PHYSAGG)	P1	1.000 (0.000)	0.979 (0.073)	0.816 (0.043)	0.665
	P2	1.069 (0.068)	1.105 (0.077)	0.872 (0.034)	0.760
	P3	1.099 (0.068)	1.350 (0.087)	0.896 (0.034)	0.803
	P4	1.018 (0.079)	1.089 (0.076)	0.830 (0.043)	0.689
	TB1	0.960 (0.093)	1.602 (0.100)	0.783 (0.066)	0.613
	TB2	0.801 (0.099)	1.263 (0.083)	0.654 (0.069)	0.427
	TB3	1.103 (0.084)	1.717 (0.108)	0.899 (0.056)	0.809
	TB4	1.070 (0.077)	1.548 (0.098)	0.873 (0.047)	0.762
Psychological Abuse (PSYCH)	VE1	1.000 (0.000)	0.712 (0.067)	0.650 (0.055)	0.423
	VE2	1.033 (0.102)	0.826 (0.069)	0.672 (0.055)	0.451
	VE3	1.320 (0.118)	0.753 (0.068)	0.859 (0.034)	0.737
	VE4	1.084 (0.117)	0.906 (0.071)	0.705 (0.055)	0.497
	VE5	1.275 (0.132)	1.368 (0.087)	0.830 (0.048)	0.688
	VE6	1.194 (0.127)	1.054 (0.075)	0.777 (0.048)	0.603
	VE7	0.566 (0.123)	0.726 (0.067)	0.368 (0.073)	0.136
	VE8	1.233 (0.123)	0.903 (0.071)	0.802 (0.045)	0.644
	VE9	0.978 (0.122)	0.748 (0.068)	0.636 (0.061)	0.405
	VE10	1.028 (0.119)	1.049 (0.076)	0.669 (0.057)	0.447
	RA1	1.104 (0.124)	1.400 (0.089)	0.718 (0.063)	0.516
	RA2	1.214 (0.124)	1.516 (0.095)	0.790 (0.057)	0.624
	RA3	1.243 (0.131)	1.537 (0.096)	0.809 (0.056)	0.654

SE = standard error

[a]All factor loadings were statistically relevant with $p<0.05$ on Wald Test

[b]STDY standardization in MPlus version 8.2

PAPER 2

Background

Adolescent dating violence is a serious public health problem that affects young people across many subgroups of the population, including ages, genders, races, and ethnicities. Studies have shown that adolescent dating violence, or psychological, emotional, physical, and/or sexual harm caused by a dating partner (Centers for Disease Control and Prevention, 2018), occurs in adolescents as young as 10 years old and often continues into adulthood (Deinera Exner-Cortens et al., 2017; Halpern et al., 2001; Jennings et al., 2017; Niolon et al., 2015; Shorey et al., 2017). More recently, the definition of adolescent dating violence was expanded to include cyber abuse, or abuse that takes place via electronic media (Picard, 2007; Zweig et al., 2013).

Both males and females experience adolescent dating violence, often with reciprocal perpetration and victimization between dating partners (Fedina, Howard, Wang, & Murray, 2016; Foshee et al., 1996; Foshee, McNaughton Reyes, & Ennett, 2010; Halpern et al., 2001; Haynie et al., 2013; Jennings et al., 2017; Kann et al., 2018; Niolon et al., 2015). A review of adolescent dating violence victimization and perpetration found prevalence estimates ranged from 6 to 21.8 percent for boys, and 9 to 37.2 percent for girls (Jennings et al., 2017). The Safe Dates Project found that dating girls were more likely to report sexual dating violence victimization and non-sexual dating violence perpetration, and dating boys were more likely to report sexual dating violence perpetration (Foshee et al., 1996). Similarly, a longitudinal study of adolescent dating violence among Latino teens found boys reported higher levels of any dating violence victimization at baseline than girls, with girls reporting higher levels of sexual dating

violence victimization at follow up (Sabina et al., 2016). Additionally, Zweig et al. found females perpetrated more non-sexual cyber dating violence than males, while males perpetrated more sexual cyber dating violence (Zweig et al., 2013).

Adolescent dating violence has also been noted in most traditional racial and ethnic subgroups in the United States (i.e. white, African American, Hispanic, non-Hispanic, Asian, etc.) (Connolly, Friedlander, Pepler, Craig, & Laporte, 2010; East & Hokoda, 2015; Foshee et al., 2010; Lormand et al., 2013; Sabina et al., 2016). Data from the Dating Violence Among Latino Adolescents (DAVILA) study revealed approximately 20 percent of Latino teens had experienced dating violence victimization (Sabina et al., 2016), and the National Longitudinal Study of Adolescent Health (Add Health) revealed 32 percent of minority adolescents experienced dating violence victimization (Halpern et al., 2001). Additionally, a study of dating violence in predominately Hispanic (48.5%) and African American (36%) adolescents estimated as many as 52.6 percent of participants experienced either physical or non-physical dating violence victimization (Lormand et al., 2013). In a study of cyber dating violence, bivariate logistic regression found no significant differences between Hispanic and African American adolescents in the perpetration and victimization of cyber dating violence (Peskin et al., 2017).

Comparative analyses have found significantly higher odds of dating violence in minority adolescents than non-minority adolescents. For example, caucasian eighth and ninth-grade students that participated in the Safe Dates Project reported significantly less non-sexual dating violence victimization than African Americans and students in the “other” racial/ethnic group. Participants in the “other” category also reported significantly more sexual dating violence

victimization than those in the Caucasian group (Foshee et al., 1996). The Add Health study found black males have significantly higher odds (OR = 2.19; 95% CI = 1.48, 3.23) of experiencing physical and psychological dating violence victimization than white adolescent males (Halpern et al., 2001). Asian and Pacific Islander participants also reported a higher odds of experiencing physical and psychological dating violence victimization (OR = 2.19; 95% CI = 1.21, 3.94) than white male adolescents (Halpern et al., 2001). In another study, minority adolescents were more likely to report perpetrating moderate to severe physical dating violence than their non-minority counterparts (Foshee et al., 2008). Additionally, Foshee et al. found that African American adolescents were more likely to initiate perpetration of physical dating violence than white adolescents (Foshee et al., 2010).

Statistical Comparisons Across Groups

Comparative analyses have shown adolescent dating violence to be significantly different across subgroups of the population. However, the reason for this difference is not clear. It may reflect a true difference in prevalence of dating violence across groups. However, it could also be the result of measurement bias, including but not limited to differential item functioning or differential additive response styles (Kline, 2016; Streiner et al., 2015).

Differential item functioning occurs when a "person's score on the indicator, given his or her true score on the corresponding factor" depends on group membership (i.e. gender, race/ethnicity, cohort, etc.) (Kline, 2016). For example, Wolfe et al. compared responses in matched dating partners and found boys reported less psychological victimization than their female dating partners reported perpetrating. The authors hypothesized that boys may be less

likely to report threatening behavior from a female dating partner (i.e. “She deliberately tried to threaten me”) than a female because the “threat” was not perceived as such by boys (Wolfe et al., 2001). Differential additive response styles occurs when an unknown systematic difference across groups results in more or less endorsement of certain indicators (Kline, 2016). This bias may occur due to cultural differences across groups, cohort effects, or method effects (Kline, 2016).

The same biases discussed above may also be true across time. A behavior that may be perceived as severe or harmful in sixth grade students may not be perceived as harmful by twelfth grade students. For example, an unwanted kiss at the age of 10 may cause serious upset in a young adolescent in the early stages of dating, but not in a twelfth grade student with more experience. Given the threat of these potential biases, it is important that each measure used for multi-group comparisons be tested for measurement invariance.

Measurement Invariance

Definition. Measurement invariance holds when the observed scores of items on a measure equally capture the value of the underlying construct across different subgroups of the population, time of measurement, and/or administrative methods (Brown, 2015; Kline, 2016). A measure that is non-invariant may yield different observed scores of a construct across groups when the underlying value is actually the same, or the same observed scores when the latent variable values are different.

There are several consequences of using non-invariant measures for group comparisons. Most importantly, researchers may draw incorrect conclusions about the differences found in

statistical analyses. Statistical analyses are only as robust as the input data and may be unintentionally confounded by non-invariant measures. This is a particular problem when assessing health behaviors and outcomes for disparities across groups. Non-invariant measures may suggest a disparity where it does not exist and vice versa. Additionally, tests for program effectiveness might suggest null effects if a measure is used that is not invariant across time. Conversely, program effects may be found when they do not actually exist. It is important to verify measurement invariance across groups and time (and even methods) to ensure quality, accurate conclusions are drawn from the data analyses.

There are four different types of measurement invariance. *Configural* measurement invariance is considered the least restrictive and suggests the model structure is equivalent across groups. All other parameters are free to vary across groups, such as item coefficients (i.e. factor loadings) and intercepts, but the core structure is held constant (Kline, 2016). *Metric* (or weak) invariance is the next most restrictive. Not only is the model structure held consistent across groups, but the unstandardized factor loadings (i.e. item coefficients) are constrained to be equal across groups. Metric invariance is said to hold if the fit of the restricted model is not significantly worse than the baseline configural model (Kline, 2016). Traditionally, the chi-square difference test is used to compare model fit between the nested models (i.e. less constrained versus more constrained) (Brown, 2015). However, the change in the comparative fit index (CFI) can also be used to assess measurement invariance across models (Kline, 2016).

Scalar (or strong) invariance holds when the model structure, unstandardized factor loadings, and unstandardized intercepts are held constant across groups without a significant

decrement to the fit of the model. Chi-square difference testing compares the less constrained metric model to the scalar model. A non-significant test demonstrates scalar invariance (Kline, 2016). Scalar invariance is the minimal level of measurement invariance that should be achieved to confidently compare data across groups and draw meaningful conclusions (Kline, 2016). Lastly, *strict* measurement invariance is the most restrictive model and constrains all of the parameters listed above to be equal, as well as the error variances and covariances between groups. This is the least likely type of invariance to be achieved, and failure to achieve strict invariance does not rule out a measure's acceptability for group comparison (Kline, 2016).

It is also possible to test for partial measurement invariance if full invariance is not achieved. Full measurement invariance holds when the model meets invariance criteria with all constraints in place. When measures do not meet invariance criteria, researchers may allow some parameters to vary to see if partial invariance holds in the measure (Brown, 2015). This means that some parameters of the measure were invariant across groups and some were not. Often researchers will use this method to identify problem items for future studies.

Invariance in adolescent dating violence measures. As noted by Smith and colleagues (Smith et al., 2015), the most commonly used measures for adolescent dating violence include adaptations of the Conflict Tactics Scale (CTS), Families for Safe Dates (Safe Dates), and the CADRI. However, a review of the literature demonstrated few studies have been done to confirm measurement invariance across subgroups of the population or time. Additionally, no studies have been conducted that assess invariance of cyber adolescent dating violence across subgroups or time.

Nocentini et al. tested the CTS-2 physical perpetration subscale for invariance across countries (Italy and Canada) and gender. The scale demonstrated partial invariance across gender within each country and full invariance across countries. The authors suggested that aggression be assessed separately in males and females when used in Canada and Italy (Nocentini et al., 2011). Cascardi et al. found significant differences in model fit between U.S. high school males and females in a modified version of the CTS (M-CTS) (Cascardi et al., 1999). Though evidence is not exhaustive, results from these tests suggest further work is needed to ensure comparability between males and females prior to additional comparative analyses with the CTS.

The Safe Dates measure of physical and psychological dating violence perpetration demonstrated strict measurement invariance across gender in a sample of seventh grade students (Cutbush & Williams, 2016). Additionally, Goncy et al. tested the Safe Dates measure for invariance across gender and time in a large sample of middle school students (N = 3894) (Goncy, Farrell, Sullivan, & Taylor, 2016). They found configural invariance for gender in both their sixth and eighth grade samples. However, the chi-square difference tests of the scalar and configural models in both subsamples demonstrated a significant decrement to the fit of the model. The authors, however, noted very little difference in values of the comparative fit indices (CFI) and root mean square error of approximation (RMSEA). Thus, they concluded scalar measurement invariance held across gender. They found the same outcome in the test for longitudinal invariance, with a significant chi-square difference test. However, analyses demonstrated an improvement in the CFI and RMSEA scores of the constrained model. Therefore, the authors concluded the Safe Dates measure was invariant across time.

The CADRI has also been subjected to scrutiny for measurement invariance. Wolfe et al. tested for measurement invariance across gender and age groups with a sample of high school students (N = 335) upon establishment of the measure (Wolfe et al., 2001). Though the model fit decently in all groups, the authors found non-invariance in the sexual abuse and relational aggression subscales of the CADRI. The sexual abuse factor loaded significantly higher on the second order abuse factor in males (0.64) than females (0.47). Additionally, the relational aggression factor loaded on the second order factor much higher among ninth grade students (0.63) than tenth and eleventh grade students (0.21 and 0.13, respectively). They hypothesize that the differential item functioning between 9th and 11th graders, as well as boys and girls, may be the result of developmental differences in these groups. These findings are important in the context of research. Ideally, the measure would be adapted to eliminate differential item functioning. Shorey et al. also tested the CADRI for measurement invariance across sex, race/ethnicity, and time in a large sample of high school students (N = 1042). The CADRI demonstrated scalar invariance across time and sex in 5 of 6 waves of data. However, the CADRI only demonstrated partial invariance across race/ethnicity across waves of data (Shorey et al., 2018). They allowed the factor loadings and thresholds to vary in the sexual abuse items, "forced him/her to have sex" and "threatened him/her in an attempt to have sex" to achieve partial invariance in wave 1 of their data. They also freed constraints on additional items in wave 4 to achieve partial invariance, including "said things to his/her friends about him/her," "ridiculed him/her in front of others," and several others (Shorey et al., 2018).

As demonstrated above, there is limited evidence of measurement invariance across gender and time in adolescent dating violence measures. Multiple studies have assessed prevalence and risk differences for adolescent dating violence in racially and ethnically diverse groups of adolescents (East & Hokoda, 2015; Foshee et al., 2008; Halpern et al., 2001; Lormand et al., 2013; Sabina et al., 2016), yet measurement invariance has not been established for the CTS, Safe Dates, CADRI, or other measures for adolescent dating violence. This is an important gap in the development of valid adolescent dating violence measures.

Specific Aims

There is limited evidence of measurement invariance across gender and less for racial/ethnic subgroups of the population in measures of adolescent dating violence. However, there is research that shows a marked disparity in prevalence and risk of adolescent dating violence in minority adolescents. The primary aim of this study was to use multi-group confirmatory factor analysis (MG-CFA) to test the CADRI for measurement invariance across a sample of Hispanic and African American sixth grade boys and girls. We hypothesized that scalar invariance would hold for the CADRI across all subgroups of this population. The secondary aim for this study was to assess focal areas of ill fit in the configural and scalar models that may lead to non-invariance. The results of this study can provide researchers with added confidence that group differences found among young minority adolescents, when using the CADRI, is the result of a true disparity when measurement invariance is found across groups. However, if measurement invariance is not confirmed, further work should be done to eliminate

or adapt items that facilitate non-invariance and enhance the validity of the measure for use across groups.

Methods

Sample and Data Collection

The data for this study were collected as part of the baseline assessment of a study to evaluate the Me and You: Building Healthy Relationships (Me & You) adolescent dating violence prevention program. This program was provided to sixth grade students (mean age = 12.4 years, SD = .62) in 10 large, urban middle schools in Southeast Texas. All English-speaking students with parental consent were asked to complete a self-directed, computer survey prior to participation in the program. A total of 709 sixth grade students completed the baseline survey. Only students who reported ever having had a boyfriend or girlfriend were asked questions regarding dating violence and were included in the dating analytic sample (n = 424). Of these students, 43.9% were female, and the majority identified themselves as Hispanic (60.8%) or African American (30.0%). Students that identified themselves as "other" race and ethnicity were not included in this study due to low sample size (n = 39). Appendix A provides additional information regarding the sample characteristics.

Measure

The Conflict in Adolescent Dating Relationships Inventory (CADRI) was used to assess adolescent dating violence perpetration and victimization in this sample (Appendix B). The study also included measures for cyber dating violence, not originally included in the CADRI. The CADRI was divided into 6 subscales of adolescent dating violence, including physical abuse (4

items), threatening behavior (4 items), verbal and emotional abuse (10 items), sexual abuse (1 item), and relational aggression (3 items). Cyber dating violence (12 items) was the sixth subscale included in this survey. Each item was assessed on a dichotomous scale (0 = no, 1 = yes), with adolescents asked whether they experienced or perpetrated each behavior. Only perpetration data was used for these analyses, however, according to the precedent set by prior validation studies (Shorey et al., 2018; Wolfe et al., 2001).

Model Specification

Previously, we conducted confirmatory factor analyses of the Me & You dating analytic sample to determine the best model fit to the data (see Paper 1). We found the model with two first order factors which reflect the underlying constructs of physical and psychological aggression, to be the best fit. Review of item endorsement frequency revealed variables with less than five percent positive endorsement in several individual groups (Appendix D). Since these variables do not provide reliable estimates for psychometric testing in this sample, we removed TB1 ("I destroyed or threatened to destroy something he/she valued"), TB3 ("I threatened to hit him/her"), TB4 ("I threatened to hit him/her or throw something at him/her"), P3 ("I slapped him/her or pulled his/her hair"), RA3 ("I spread rumors about him/her"), and RA2 ("I said things to his/her friends to turn them against him/her") from the model for all analyses (Streiner et al., 2015).

Figure 2 displays the model we used to test invariance. The physical aggression factor, PHYSAGG, was informed by 3 physical abuse items (P1, P2, and P4) and 1 threatening behavior item (TB2). The psychological aggression factor, PSYCH, was informed by 10 verbal emotional

abuse items (VE1-VE10) and 1 relational aggression item (RA1). The single item for sexual dating abuse (S1) and cyber dating violence items (C1-C12) were excluded from both the gender and race/ethnicity models since they were not included in primary confirmatory factor analyses (see Paper 1). Sexual abuse was excluded from prior analyses because it was informed by only one item and resulted in an unidentified model. Cyber dating abuse items were excluded previously because positive endorsement of most items was below five percent. Low endorsement rates can negatively influence the validity of the confirmatory factor analysis.

Analysis Strategy

We used MPlus version 8.2 for categorical multi-group confirmatory factor analysis (MG-CFA) to test our hypotheses that the CADRI was invariant across racial and ethnic subgroups (i.e. African American and Hispanic) and gender (Muthen & Muthen, 2018). We used the free baseline approach for invariance testing, with the base model (i.e. configural or equal form) under the least constraints, with tighter restrictions placed on parameters in subsequent steps (Kline, 2016).

We conducted invariance testing in three steps, with each performed independently for race/ethnicity and gender subgroups (Brown, 2015). First, we completed CFA with a robust weighted least squares estimator (WLSMV) and delta parameterization to assess goodness-of-fit for each independent group. Secondly, we tested for configural invariance. We scaled each factor by fixing the first factor loading (P1 and VE1) to 1.0. The underlying latent factor must be assigned a scale when setting up confirmatory factor analyses since it inherently has no scale. One method is to set the scale with the first item of each factor (Brown, 2015). We also fixed the

factor means and residual correlations to zero and allowed all other parameters to be estimated. We used the chi-square value, Tucker-Lewis index (TLI), comparative fit index (CFI), and root mean square error of approximation (RMSEA) to determine goodness-of-fit of the equal form model. We used the following guidelines to identify models of good fit: TLI >.95, CFI >0.95, and RMSEA <0.06 (Brown, 2015).

We could not test metric invariance independently due to limitations with the estimator. The weighted least squares estimator with delta parameterization constrains the factor loadings and thresholds simultaneously (Muthen & Muthen, 2017). They cannot be tested separately. Thus, we tested for scalar invariance once configural invariance was confirmed. We fixed residual correlations to zero and held factor loadings and thresholds equal across groups. We also fixed the factor means to zero and scale factors to 1.0 in one group and allowed them to be freely estimated in the second (Muthen & Muthen, 2017). Within MPlus, we performed chi-square difference testing between the scalar and configural models. A p-value less than 0.05 indicated scalar invariance was achieved. The change in CFI was also assessed to verify measurement invariance in this analysis. A change of less than .01 suggested scalar invariance was achieved.

Results

Descriptive statistics. A total of 424 participants in the Me & You dating analytic sample completed the baseline survey. Missingness across the analytic sample was less than 10 percent for all items, and item endorsement frequency met psychometric testing standards for all remaining variables (Brown, 2015; Streiner et al., 2015). This information is detailed in Appendix B and D. Due to missing data patterns in the correlation matrix, the test for

measurement invariance across race/ethnicity and gender included a total sample size of 384 (n_H = 257, n_{AA} = 127) and 420 (n_{male} = 234, n_{female} = 186) respectively.

Multi-group confirmatory factor analysis across race/ethnicity. All fit statistics from the multi-group CFA are displayed in Table 3. The single group confirmatory factor analyses for Hispanic and African American participants resulted in decent fits to the data. The chi-square value in the Hispanic group (X^2 = 118.27, df = 89) was higher than in the African American group (X^2 = 93.32, df = 89), which may be a reflection of the unbalanced sample size. The RMSEA, CFI, and TLI were also acceptable across groups, with the Hispanic model fitting slightly better than the African American model (Table 3). Modification indices in the Hispanic group suggested allowing the correlation between VE10 ("I threatened to end the relationship") and VE9 ("I accused him/her of flirting with another guy/girl") to be estimated. This residual correlation was also allowed in the full model (see Paper 1). A chi-square difference test resulted in an improvement in model fit (ΔX^2 = 9.745, Δdf = 1, p = 0.002). Though not recommended as a relevant modification index, a chi-square difference test of the same residual correlation in the African American group also resulted in an improvement in model fit (ΔX^2 = 5.195, Δdf = 1, p = 0.022). Since both groups benefited from the change, the correlation was allowed to stand in subsequent testing for configural and scalar invariance.

Configural measurement invariance was achieved across race and ethnicity. The equal form model fit the data well, with RMSEA = 0.020 (90% CI = 0.000, 0.039), CFI = 0.994, and TLI = 0.993. The scalar model, which held factor loadings and thresholds equal across groups, also demonstrated a good fit to the data [RMSEA = 0.021 (90% CI = 0.000, 0.039), CFI = 0.992,

TLI = 0.991]. The chi-square difference test between the models resulted in a non-significant p-value (p = 0.226), which indicated the restrictions placed on the model did not decrement the fit of the model. The lower order, two-factor CADRI achieved scalar invariance among Hispanic and African American adolescents in this sample.

Multi-group confirmatory factor analysis across gender. Single group confirmatory factor analyses for male and female adolescents of the Me & You study resulted in adequate fits to the data (Table 3), though fit statistics were variable across groups. Males had better chi-square and RMSEA scores, while females had better CFI and TLI scores. There were no recommended modification indices in the male group. However, modification indices in the female group suggested a correlation between the residuals of VE10 ("I threatened to end the relationship") and VE9 ("I accused him/her of flirting with another guy/girl"). A chi-square difference test between the modified and base models showed a statistically relevant improvement to the female model ($\Delta X^2 = 17.43$, $\Delta df = 1$, $p = 0.000$). However, the added residual correlation did not improve model fit in the male group ($\Delta X^2 = 1.995$, $\Delta df = 1$, $p = 0.158$). We tested for invariance with the base model since both groups did not benefit from the added residual correlation.

The configural model resulted in a reasonable fit to the data, with RMSEA = 0.043 (90% CI = 0.030, 0.056), CFI = 0.967, TLI = 0.961. However, there were notable differences in parameter estimates (Table 4). The factor correlations between PHYSAGG and PSYCH were markedly different between male and female groups (standardized $\phi_{male} = 0.911$, $\phi_{female} = 0.799$). This comparison indicates that there is less construct discrimination between psychological and

physical aggression among males than females. Additionally, several factor loadings and thresholds were different in magnitude. For example, the unstandardized loading of TB2 ("I tried to frighten him/her on purpose) was lower in males than females ($\lambda_{male} = 0.610$, $\lambda_{female} = 0.976$). For males, every one unit increase in the latent dimension of physical aggression corresponded to a 0.610 unit increase in the y* (or the underlying continuous dimension) of the observed measure of purposely frightening his partner, while for females it was associated with a 0.976 increase in the y* of the observed measure. The standardized loadings for this item provide information about the change in standard deviations of the latent dimension of physical aggression as the response on the item changes from no to yes. In terms of the thresholds, the latent construct of physical aggression increases by 0.500 in males and 0.801 for females as responses change from no to yes. There are others that are similarly disparate, with most items reflecting higher factor loadings for females.

As expected with these large differences, the scalar model did not converge and scalar measurement invariance was not achieved for the CADRI across male and female participants of the Me & You study.

Discussion

The purpose of this study was to assess the invariance of the CADRI across different genders and race/ethnicities among a sample of sixth grade students. Our study found that an adapted version of the CADRI was not invariant across male and female sixth grade adolescents, but was invariant across Hispanic and African American groups. The adapted CADRI only achieved configural invariance across gender groups. We saw distinct differences in some of the

parameter estimates, and the test for scalar invariance was unsuccessful. We can conclude that individual item factor loadings and thresholds are relevantly different in male and female sixth grade students. This is not surprising given the differences in adolescent development and their approach to dating relationships (Furman & Wehner, 1997; Jackson, 1999; Wolfe et al., 2001). Investigators should use caution when assessing gender differences in prevalence and intervention effectiveness using this measure. These items, and their relationships to the latent constructs and each other, need to be further investigated to determine the extent of differential item functioning across genders. Of note, the residual correlation between the items "I accused him/her of flirting with another girl/guy" and "I threatened to end the relationship" was only relevant in the female group. This is an interesting finding that is not unique to our study. Wolfe and colleagues also found similar residual correlations with jealous accusations in the original validation studies of the CADRI (Wolfe et al., 2001). Qualitative studies with young adolescents may be important to tease apart the differences in this type of behavior among males and females.

Additionally, the correlation between physical and psychological aggression was much higher among males than females. This finding is supported by prior research that found underlying factors of adolescent dating violence were divided by the severity of aggression rather than the physicality of aggression among male adolescents (Cascardi et al., 1999). This difference is worth consideration in future studies, though it does not impact the configural invariance of the CADRI in this population.

We did find invariance among racial and ethnic sub-groups of our samples. This is one of the first studies to assess invariance of the CADRI across these groups. Though we found invariance, others have not been as successful. Shorey et al. found the CADRI was only invariant across race/ethnicity in two of six samples of high school students (Shorey et al., 2018). They loosened constraints to allow variation in factor loadings and thresholds in the scalar model to achieve partial invariance in the remaining samples. It is important to note that they included the White, non-Hispanic group in their tests for invariance. Our tests did not include this subgroup, as the sample size was too small (N = 39). This may account for the differences seen across studies, and future investigation is needed to ensure measurement invariance across all racial and ethnic subgroups of sixth grade adolescents.

Limitations. There are several limitations of this study. First, as discussed in paper 1, the sample included sixth graders from only one urban region of Texas. These findings may not be generalizable to other urban areas of the state or rural communities. Further studies should be done to verify the invariance of the CADRI in different populations. Secondly, only Hispanic and African American sixth grade students were included in the tests for invariance. Therefore, the findings may not hold in a sample that includes other races and ethnicities.

Another limitation is that our study did not include all items of the CADRI published by Wolfe et al. We did not include items of sexual abuse, because it was not previously included in the confirmatory factor analyses of this model in sixth grade students (see Paper 1). Additionally, we removed items from the model to ensure appropriate reliability of the parameter estimates given low endorsement on several items in this sample. Therefore, these results are only

applicable to the items tested in our adapted model. The excluded items should be interpreted cautiously in the context of multi-group comparison since invariance was not confirmed.

Interestingly, three of the threatening behavior items, including threats to destroy something valuable and threats to physically harm one's partner, were positively endorsed less than four percent of the time in male students. However, female students had higher endorsement of these items (range = 4.9-12.1%). These findings are not unique to our study. Prior studies of measurement validity also had lower than desirable rates of positive endorsement among these items (Shorey et al., 2018; Wolfe et al., 2001). This finding may reflect differences in the perception of threatening words or behaviors. Males may be less likely to perceive certain statements or actions as threatening, whereas females may be more sensitive to subtle threats (Wolfe et al., 2001). Wolfe included these items in the measure in order to capture more "prodromal" and subtle acts of dating aggression likely in adolescent relationships (Wolfe et al., 2001). Though they may have face validity and be relevant in this population, caution should be used when including these items in the scoring of the measure.

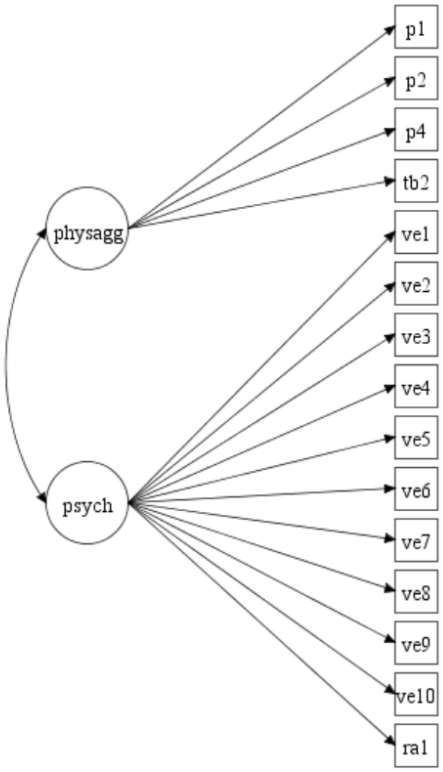
Lastly, this measure of adolescent dating violence was assessed with a self-administered survey. This method is subject to recall and social desirability biases, particularly when taken outside of the context of conflict (Wolfe et al., 2001).

Conclusions

Researchers and clinicians require valid measures to assess the differences in prevalence of adolescent dating violence across subgroups of the population. Our study provides evidence of invariance of the CADRI across Hispanic and African American groups of sixth grade

adolescents in low income, urban settings. Our study also contributes evidence to the field that the CADRI is not invariant across gender groups and should not be used to directly compare adolescent dating violence among males and females. Further studies are needed to determine the most valid methods for assessing gender differences in adolescent dating violence among sixth grade adolescents.

Figure 2. Two-factor model of the Conflict in Adolescent Dating Relationships Inventory (CADRI) adapted* for invariance testing across gender and racial/ethnic groups in the Me & You study among 6th grade adolescents



*Adaptation included removal of TB1, TB3, TB4, P3, RA2, and RA3 due to low frequency of positive endorsement

Table 3. Results of measurement invariance testing across race/ethnicity and gender of a modified CADRI in sixth grade adolescents.

	X^2	df	p	X^2_{diff}	Δdf	p	RMSEA (90% CI)	CFI	TLI
Single-group solutions									
Hispanic* (n = 257)	104.14	88	0.115	-	-	-	0.027 (0.000, 0.045)	0.988	0.985
African American* (n = 127)	86.38	88	0.529	-	-	-	0.000 (0.000, 0.046)	1.00	1.00
Male (n = 234)	119.15	89	0.018	-	-	-	0.038 (0.017, 0.055)	0.957	0.950
Female (n = 186)	130.80	89	0.003	-	-	-	0.050 (0.030, 0.068)	0.971	0.966
Measurement invariance – Race/Ethnicity (N = 384)									
Configural*	188.96	176	0.239	-	-	-	0.020 (0.000, 0.039)	0.994	0.993
Scalar*	202.78	187	0.204	13.96	11	0.235	0.021 (0.000, 0.039)	0.992	0.991
Measurement invariance – Gender (N = 420)									
Configural	248.40	178	0.000	-	-	-	0.043 (0.030, 0.056)	0.967	0.961
Scalar	Failed to converge								

N = sample size
df = degrees of freedom; RMSEA = root mean square error of approximation, CFI = comparative fit index; TLI = Tucker Lewis index
*Allowed VE9 and VE10 residuals to be correlated

Table 4. Unstandardized and standardized parameter estimates from multi-group confirmatory factory analysis of the configural (equal form) model across gender groups in the Me & You study

		Male [a,b]			Female [a,c]		
		Unstandardized		Standardized[d]	Unstandardized		Standardized[d]
Factor	Indicator	Loading (SE)	Threshold	Loading (SE)	Loading (SE)	Threshold	Loading (SE)
Physical Aggression (PHYSAGG)	P1	1.000 (0.000)	1.218	0.821 (0.078)	1.000 (0.000)	0.749	0.821 (0.075)
	P2	0.953 (0.134)	1.383	0.782 (0.085)	1.022 (0.117)	0.842	0.839 (0.061)
	P4	1.052 (0.121)	1.314	0.864 (0.071)	1.034 (0.126)	0.877	0.849 (0.059)
	TB2	0.610 (0.157)	1.335	0.501 (0.111)	0.976 (0.140)	1.171	0.801 (0.085)
Psychological Abuse (PSYCH)	VE1	1.000 (0.000)	0.895	0.545 (0.089)	1.000 (0.000)	0.501	0.693 (0.070)
	VE2	1.112 (0.218)	0.914	0.605 (0.087)	1.036 (0.121)	0.731	0.717 (0.069)
	VE3	1.463 (0.272)	0.965	0.797 (0.067)	1.330 (0.132)	0.517	0.921 (0.039)
	VE4	1.123 (0.236)	1.153	0.612 (0.089)	1.144 (0.144)	0.683	0.792 (0.062)
	VE5	1.453 (0.272)	1.454)	0.791 (0.073)	1.221 (0.154)	1.263	0.845 (0.059)
	VE6	1.511 (0.269)	1.109	0.823 (0.063)	1.068 (0.148)	1.001	0.740 (0.073)
	VE7	0.537 (0.192)	0.741	0.292 (0.102)	0.596 (0.156)	0.696	0.413 (0.099)
	VE8	1.460 (0.264)	1.193	0.795 (0.071)	1.191 (0.147)	0.612	0.825 (0.056)
	VE9	0.956 (0.165)	1.107)	0.521 (0.102)	1.068 (0.167)	0.391	0.739 (0.076)
	VE10	1.205 (0.245)	1.279	0.656 (0.079)	0.989 (0.142)	0.815	0.685 (0.073)
	RA1	1.091 (0.293)	1.552	0.594 (0.105)	1.052 (0.129)	1.240	0.728 (0.074)

SE = standard error

[a]All standardized parameter estimates had p-value < 0.05

PHYSAGG-PSYCH standardized correlation [b]male = 0.911; [c]female = 0.799

[d]STDY standardization in MPlus version 8.2

Table 5. Unstandardized and standardized parameter estimates from multi-group confirmatory factory analysis of the configural (equal form) model across racial/ethnic groups in the Me & You study

		Hispanic [a,b]			African American [a,c]		
		Unstandardized		Standardized[d]	Unstandardized		Standardized[d]
Factor	Indicator	Loading (SE)	Threshold	Loading (SE)	Loading (SE)	Threshold	Loading (SE)
Physical Aggression (PHYSAGG)	P1	1.000 (0.000)	1.095	0.819 (0.070)	1.000 (0.000)	0.797	0.869 (0.073)
	P2	0.956 (0.118)	1.177	0.782 (0.074)	1.070 (0.108)	0.906	0.930 (0.052)
	P4	1.013 (0.120)	1.150	0.829 (0.067)	1.023 (0.117)	0.876	0.889 (0.060)
	TB2	0.841 (0.137)	1.266	0.689 (0.093)	0.746 (0.164)	1.170	0.648 (0.125)
Psychological Abuse (PSYCH)	VE1	1.000 (0.000)	0.674	0.700 (0.062)	1.000 (0.000)	0.786	0.560 (0.110)
	VE2	0.931 (0.121)	0.932	0.652 (0.074)	1.203 (0.218)	0.668	0.673 (0.097)
	VE3	1.255 (0.113)	0.844	0.878 (0.044)	1.622 (0.326)	0.637	0.908 (0.056)
	VE4	1.136 (0.131)	0.947	0.795 (0.062)	1.247 (0.276)	0.882	0.698 (0.089)
	VE5	1.163 (0.130)	1.476	0.814 (0.060)	1.540 (0.325)	1.063	0.862 (0.072)
	VE6	1.137 (0.136)	1.070	0.796 (0.062)	1.236 (0.293)	0.967	0.962 (0.098)
	VE7	0.573 (0.130)	0.678	0.401 (0.091)	0.571 (0.282)	0.882	0.319 (0.126)
	VE8	1.200 (0.115)	1.090	0.839 (0.053)	1.416 (0.347)	0.668	0.793 (0.079)
	VE9	0.847 (0.124)	0.792	0.593 (0.083)	1.240 (0.311)	0.771	0.695 (0.092)
	VE10	0.909 (0.128)	1.065	0.636 (0.081)	0.881 (0.300)	1.136	0.493 (0.123)
	RA1	1.106 (0.138)	1.295	0.774 (0.070)	0.975 (0.282)	1.593	0.546 (0.145)

SE = standard error

[a]All standardized parameter estimates had p-value < 0.05

PHYSAGG-PSYCH standardized correlation [b]Hispanic = 0.856; [c]African American = 0.852

[d]STDY standardization in MPlus version 8.2

REFERENCES

Ackard, D. M., Eisenberg, M. E., & Neumark-Sztainer, D. (2007). Long-Term Impact of Adolescent Dating Violence on the Behavioral and Psychological Health of Male and Female Youth. *The Journal of Pediatrics*, *151*(5), 476–481. https://doi.org/10.1016/j.jpeds.2007.04.034

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research* (2nd ed.). In *Methodology in the Social Sciences* (2nd ed.). New York, New York: The Guilford Press.

Cascardi, M., Avery-Leaf, S., O'Leary, K. D., & Slep, A. M. S. (1999). Factor structure and convergent validity of the Conflict Tactics Scale in high school students. *Psychological Assessment*, *11*(4), 546–555. https://doi.org/10.1037/1040-3590.11.4.546

Cascardi, M., Blank, S., & Dodani, V. (2016). Comparison of the CADRI and CTS2 for Measuring Psychological and Physical Dating Violence Perpetration and Victimization. *Journal of Interpersonal Violence*, 088626051667018. https://doi.org/10.1177/0886260516670182

Centers for Disease Control and Prevention. (2018, June 11). Teen Dating Violence. Retrieved September 4, 2018, from Violence Prevention website: https://www.cdc.gov/violenceprevention/intimatepartnerviolence/teen_dating_violence.html

Choi, H. J., & Temple, J. R. (2016). Do Gender and Exposure to Interparental Violence Moderate the Stability of Teen Dating Violence?: Latent Transition Analysis. *Prevention Science*, *17*(3), 367–376. https://doi.org/10.1007/s11121-015-0621-4

Connolly, J., Friedlander, L., Pepler, D., Craig, W., & Laporte, L. (2010). The Ecology of Adolescent Dating Aggression: Attitudes, Relationships, Media Use, and Socio-Demographic Risk Factors. *Journal of Aggression, Maltreatment & Trauma*, *19*(5), 469–491. https://doi.org/10.1080/10926771.2010.495028

Cutbush, S., & Williams, J. (2016). Teen Dating Violence, Sexual Harassment, and Bullying Among Middle School Youth: Examining Measurement Invariance by Gender. *Journal of Research on Adolescence*, *26*(4), 918–926. https://doi.org/10.1111/jora.12244

East, P. L., & Hokoda, A. (2015). Risk and Protective Factors for Sexual and Dating Violence Victimization: A Longitudinal, Prospective Study of Latino and African American Adolescents. *Journal of Youth and Adolescence*, *44*(6), 1288–1300. https://doi.org/10.1007/s10964-015-0273-5

Exner-Cortens, D., Eckenrode, J., & Rothman, E. (2013). Longitudinal Associations Between Teen Dating Violence Victimization and Adverse Health Outcomes. *PEDIATRICS*, *131*(1), 71–78. https://doi.org/10.1542/peds.2012-1029

Exner-Cortens, Deinera, Eckenrode, J., Bunge, J., & Rothman, E. (2017). Revictimization After Adolescent Dating Violence in a Matched, National Sample of Youth. *Journal of Adolescent Health*, *60*(2), 176–183. https://doi.org/10.1016/j.jadohealth.2016.09.015

Exner-Cortens, Deinera, Gill, L., & Eckenrode, J. (2016). Measurement of adolescent dating violence: A comprehensive review (Part 1, behaviors). *Aggression and Violent Behavior*, *27*, 64–78. https://doi.org/10.1016/j.avb.2016.02.007

Fedina, L., Howard, D. E., Wang, M. Q., & Murray, K. (2016). Teen Dating Violence Victimization, Perpetration, and Sexual Health Correlates Among Urban, Low-Income, Ethnic, and Racial Minority Youth. *International Quarterly of Community Health Education*, *37*(1), 3–12. https://doi.org/10.1177/0272684X16685249

Fernández-González, L., Wekerle, C., & Goldstein, A. L. (2012). Measuring adolescent dating violence: Development of 'conflict in adolescent dating relationships inventory' short form. *Advances in Mental Health*, *11*(1), 35–54. https://doi.org/10.5172/jamh.2012.11.1.35

Foshee, V. A., Fletcher Linder, G., Bauman, K. E., Langwick, S. A., Arriaga, X. B., Heath, J. L., … Bangdiwala, S. (1996). The Safe Dates Project: Theoretical Basis, Evaluation Design, and Selected Baseline Findings. *American Journal of Preventive Medicine*, *12*(5), 39–47. https://doi.org/10.1016/S0749-3797(18)30235-6

Foshee, V. A., Karriker-Jaffe, K. J., Reyes, H. L. M., Ennett, S. T., Suchindran, C., Bauman, K. E., & Benefield, T. S. (2008). What Accounts for Demographic Differences in Trajectories of Adolescent Dating Violence? An Examination of Intrapersonal and Contextual Mediators. *Journal of Adolescent Health*, *42*(6), 596–604. https://doi.org/10.1016/j.jadohealth.2007.11.005

Foshee, V. A., McNaughton Reyes, H. L., & Ennett, S. T. (2010). Examination of Sex and Race Differences in Longitudinal Predictors of the Initiation of Adolescent Dating Violence Perpetration. *Journal of Aggression, Maltreatment & Trauma*, *19*(5), 492–516. https://doi.org/10.1080/10926771.2010.495032

Furman, W., & Wehner, E. A. (1997). Adolescent romantic relationships: a developmental perspective. *New Directions for Child Development*, (78), 21–36.

Goncy, E. A., Farrell, A. D., Sullivan, T. N., & Taylor, K. A. (2016). Measurement of Dating Aggression During Middle School: Structure, Measurement Invariance, and Distinction From General Aggression. *Journal of Research on Adolescence*, *26*(3), 509–523. https://doi.org/10.1111/jora.12208

Halpern, C. T., Oslak, S. G., Young, M. L., Martin, S. L., & Kupper, L. L. (2001). Partner Violence Among Adolescents in Opposite-Sex Romantic Relationships: Findings From the National Longitudinal Study of Adolescent Health. *American Journal of Public Health*, *91*(10), 1679–1685. https://doi.org/10.2105/AJPH.91.10.1679

Haynie, D. L., Farhat, T., Brooks-Russell, A., Wang, J., Barbieri, B., & Iannotti, R. J. (2013). Dating Violence Perpetration and Victimization Among U.S. Adolescents: Prevalence, Patterns, and Associations With Health Complaints and Substance Use. *Journal of Adolescent Health*, *53*(2), 194–201. https://doi.org/10.1016/j.jadohealth.2013.02.008

Jackson, S. M. (1999). Issues in the dating violence research. *Aggression and Violent Behavior*, *4*(2), 233–247. https://doi.org/10.1016/S1359-1789(97)00049-9

Jennings, W. G., Okeem, C., Piquero, A. R., Sellers, C. S., Theobald, D., & Farrington, D. P. (2017). Dating and intimate partner violence among young persons ages 15–30: Evidence from a systematic review. *Aggression and Violent Behavior*, *33*, 107–125. https://doi.org/10.1016/j.avb.2017.01.007

Joppa, M. C., Rizzo, C. J., Nieves, A. V., & Brown, L. K. (2016). Pilot Investigation of the Katie Brown Educational Program: A School-Community Partnership. *Journal of School Health*, *86*(4), 288–297. https://doi.org/10.1111/josh.12378

Kann, L., McManus, T., Shanklin, S., Flint, K., Queen, B., Lowry, R., … Ethier, K. (2018). *Youth Risk Behavior Surveillance: United States, 2017* (No. Vol. 67, No. 8). Centers for Disease Control and Prevention.

Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling* (4th ed.). In *Methodology in the Social Sciences* (4th ed.). New York, New York: The Guilford Press.

Lenhart, A. (2015). *Teens, Social Media & Technology Overview 2015*. Retrieved from Pew Research Center website: http://www.pewinternet.org/2015/04/09/teens-social-media-technology-2015/

Loehlin, J. C. (2004). *Latent variable models: an introduction to factor, path, and structural equation analysis* (4th ed). Mahwah, N.J: L. Erlbaum Associates.

Lormand, D. K., Markham, C. M., Peskin, M. F., Byrd, T. L., Addy, R. C., Baumler, E., & Tortolero, S. R. (2013). Dating Violence Among Urban, Minority, Middle School Youth and Associated Sexual Risk Behaviors and Substance Use. *Journal of School Health*, *83*(6), 415–421. https://doi.org/10.1111/josh.12045

Markham, C. M., Peskin, M., Addy, R. C., Wilkerson, J. M., Johnson Baker, K., Santa Maria, D. M., … Temple, J. R. (2017). Assessing the Need and Receptivity for an Integrated Healthy Sexual and Dating Relationships Intervention for Community College Students. *Journal of Applied Research on Children*, *8*(1), Article 1.

Muthen, L. K., & Muthen, B. O. (2017). *MPlus User's Guide* (Eighth). Los Angeles, CA: Muthen & Muthen.

Muthen, L. K., & Muthen, B. O. (2018). MPlus (Version 8.2). Los Angeles, CA: Muthen & Muthen.

Niolon, P. H., Vivolo-Kantor, A. M., Latzman, N. E., Valle, L. A., Kuoh, H., Burton, T., … Tharp, A. T. (2015). Prevalence of Teen Dating Violence and Co-occurring Risk Factors Among Middle School Youth in High-Risk Urban Communities. *Journal of Adolescent Health*, *56*(2), S5–S13. https://doi.org/10.1016/j.jadohealth.2014.07.019

Nocentini, A., Menesini, E., Pastorelli, C., Connolly, J., Pepler, D., & Craig, W. (2011). Physical Dating Aggression in Adolescence: Cultural and Gender Invariance. *European Psychologist*, *16*(4), 278–287. https://doi.org/10.1027/1016-9040/a000045

Peskin, M. F., Markham, C. M., Shegog, R., Temple, J. R., Baumler, E. R., Addy, R. C., … Emery, S. T. (2017). Prevalence and Correlates of the Perpetration of Cyber Dating Abuse among Early Adolescents. *Journal of Youth and Adolescence*, *46*(2), 358–375. https://doi.org/10.1007/s10964-016-0568-1

Picard, P. (2007). *Tech abuse in teen relationships study*. Northbrook, IL: Teen Research Unlimited.

Reidy, D. E., Ball, B., Houry, D., Holland, K. M., Valle, L. A., Kearns, M. C., … Rosenbluth, B. (2016). In Search of Teen Dating Violence Typologies. *Journal of Adolescent Health*, *58*(2), 202–207. https://doi.org/10.1016/j.jadohealth.2015.09.028

Rowe, L. S., Jouriles, E. N., & McDonald, R. (2015). Reducing Sexual Victimization Among Adolescent Girls: A Randomized Controlled Pilot Trial of My Voice, My Choice. *Behavior Therapy*, *46*(3), 315–327. https://doi.org/10.1016/j.beth.2014.11.003

Sabina, C., Cuevas, C. A., & Cotignola-Pickens, H. M. (2016). Longitudinal dating violence victimization among Latino teens: Rates, risk factors, and cultural influences. *Journal of Adolescence*, *47*, 5–15. https://doi.org/10.1016/j.adolescence.2015.11.003

Shorey, R. C., Allan, N. P., Cohen, J. R., Fite, P. J., Stuart, G. L., & Temple, J. R. (2018). Testing the factor structure and measurement invariance of the conflict in Adolescent Dating Relationship Inventory. *Psychological Assessment*. https://doi.org/10.1037/pas0000678

Shorey, R. C., Cohen, J. R., Lu, Y., Fite, P. J., Stuart, G. L., & Temple, J. R. (2017). Age of onset for physical and sexual teen dating violence perpetration: A longitudinal investigation. *Preventive Medicine*, *105*, 275–279. https://doi.org/10.1016/j.ypmed.2017.10.008

Smith, J., Mulford, C., Latzman, N. E., Tharp, A. T., Niolon, P. H., & Blachman-Demner, D. (2015). Taking Stock of Behavioral Measures of Adolescent Dating Violence. *Journal of Aggression, Maltreatment & Trauma*, *24*(6), 674–692. https://doi.org/10.1080/10926771.2015.1049767

Straus, M. A., Hamby, S. L., Boney-McCoy, S., & Sugarman, D. B. (1996). The Revised Conflict Tactics Scales (CTS2): Development and Preliminary Psychometric Data. *Journal of Family Issues*, *17*(3), 283–316. https://doi.org/10.1177/019251396017003001

Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health Measurement Scales: A Practical Guide to Their Development and Use* (5th ed.). New York, New York: Oxford University Press.

Taylor, B. G., Mumford, E. A., & Stein, N. D. (2015). Effectiveness of "Shifting Boundaries" Teen Dating Violence Prevention Program for Subgroups of Middle School Students. *Journal of Adolescent Health*, *56*(2), S20–S26. https://doi.org/10.1016/j.jadohealth.2014.07.004

Temple, J. R., Choi, H. J., Brem, M., Wolford-Clevenger, C., Stuart, G. L., Peskin, M. F., & Elmquist, J. (2016). The Temporal Association Between Traditional and Cyber Dating Abuse Among Adolescents. *Journal of Youth and Adolescence*, *45*(2), 340–349. https://doi.org/10.1007/s10964-015-0380-3

Wolfe, D. A., Scott, K., Reitzel-Jaffe, D., Wekerle, C., Grasley, C., & Straatman, A.-L. (2001). Development and validation of the Conflict in Adolescent Dating Relationships Inventory. *Psychological Assessment*, *13*(2), 277–293. https://doi.org/10.1037//1040-3590.13.2.277

Zweig, J. M., Dank, M., Yahner, J., & Lachman, P. (2013). The Rate of Cyber Dating Abuse Among Teens and How It Relates to Other Forms of Teen Dating Violence. *Journal of Youth and Adolescence*, *42*(7), 1063–1077. https://doi.org/10.1007/s10964-013-9922-8

Appendix A: Descriptive statistics for the dating analytic sample of the Me & You study used to determine the validity of the CADRI in low income, minority 6th grade adolescents (N = 424).

	Mean (SD)	N (%)
Age	12.4 (0.62)	-
Gender		
Female	-	186 (43.9)
Male	-	235 (55.4)
Missing	-	3 (0.7)
Race/Ethnicity		
Hispanic	-	258 (60.8)
Black/African American	-	127 (30.0)
Other	-	39 (9.2)
Missing	-	0 (0.0)

SD = standard deviation
N = sample size
Other category of race/ethnicity = White, Asian or Pacific Islander, and American Indian or Native American.

Appendix B: Survey based on items from the Conflict in Adolescent Dating Relationships Inventory (CADRI) used to assess perpetration of dating violence in a sample of sixth grade students.

Variable	Item	Scale
The following questions ask about things that you may have ever done toward a boyfriend or girlfriend (someone that you have dated, gone out with, gone steady with). Have you ever done any of the following to a boyfriend or girlfriend (someone that you have dated, gone out with, gone steady with)? Please mark YES or NO for each question. Remember, all of your answers will be kept private and no one will know your answers.[a]		
Physical abuse		
P1	I threw something at him/her	1 – yes 0 – no
P2	I kicked, hit, or punched him/her	1 – yes 0 – no
P3	I slapped him/her or pulled his/her hair	1 – yes 0 – no
P4	I pushed, shoved, or shook him/her	1 – yes 0 – no
Threatening behavior		
TB1	I destroyed or threatened to destroy something he/she valued	1 – yes 0 – no
TB2	I tried to frighten him/her on purpose[a]	1 – yes 0 – no
TB3	I threatened to hurt him/her	1 – yes 0 – no
TB4	I threatened to hit him/her or throw something at him/her	1 – yes 0 – no
Sexual abuse[b]		
S1	I kissed him/her when he/she didn't want me to	1 – yes 0 – no
Verbal emotional abuse		
VE1	I did something to make him/her feel jealous	1 – yes 0 – no
VE2	I brought up something bad that he/she had done in the past	1 – yes 0 – no
VE3	I said things just to make him/her angry	1 – yes 0 – no
VE4	I spoke to him/her in a hostile or mean tone of voice	1 – yes

		0 – no
VE5	I insulted him/her with put-downs	1 – yes 0 – no
VE6	I made fun of him/her in front of others[a]	1 – yes 0 – no
VE7	I kept track of who he/she was with and where he/she was	1 – yes 0 – no
VE8	During an argument, I blamed him/her for the problem.	1 – yes 0 – no
VE9	I accused him/her of flirting with another girl or guy.	1 – yes 0 – no
VE10	I threatened to end the relationship	1 – yes 0 – no
Relational aggression		
RA1	I tried to turn my friends against him/her	1 – yes 0 – no
RA2	I said things to his/her friends about him/her to turn them against him/her.	1 – yes 0 – no
RA3	I spread rumors about him/her	1 – yes 0 – no
Cyber abuse[c]		
C1	I posted embarrassing photos or other images of him/her online.	1 – yes 0 – no
C2	I sent threatening text messages to him/her.	1 – yes 0 – no
C3	I used his/her social networking account (Facebook, Twitter, Instagram, etc) without his/her permission.	1 – yes 0 – no
C4	I wrote nasty things about him/her on my profile page/timeline (on Facebook, Twitter, Instagram, etc).	1 – yes 0 – no
C5	I sent him/her so many messages (like texts, emails, chats) that it made him/her feel unsafe.	1 – yes 0 – no
C6	I sent him/her text messages to check up on him/her (where are you, what are you doing, who are you with).	1 – yes 0 – no
C7	I spread rumors about him/her using a cell phone, email, IM, web chat, or social networking site (on Facebook, Twitter, Instagram, etc).	1 – yes 0 – no
C8	I used information from his/her social networking profile/page to harass him/her or put him/her down.	1 – yes 0 – no

C9	I made him/her afraid when he/she did not respond to my phone call, text, posting on social networking page, IM, et.	1 – yes 0 – no
C10	I threatened to harm him/her physically through a cell phone, text message, social networking page, etc.	1 – yes 0 – no
C11	I sent him/her sexual photos or naked photos of me that I knew he/she did not want.	1 – yes 0 – no
C12	I pressured him/her to send a sexual or naked photo of him/herself.	1 – yes 0 – no

[a]These items had minor language adaptations from originally validated CADRI to enhance readability and understanding in sixth grade sample.
[b]The original CADRI included items with more severe forms of sexual dating violence. They were excluded for this sample.
[c]The original CADRI did not include cyber abuse items. These were adapted from prior studies on cyber dating violence (Peskin et al., 2017; Picard, 2007; Zweig et al., 2013).

Appendix C: Missing data and item endorsement frequency of the CADRI items in the full dating analytic sample of the Me & You Study, a prevention program designed to teach healthy relationship skills among 6th grade students in a low income, minority area of Southwest Texas (N = 424).

			Endorsement Frequency (%)	
Item		Missing (%)	No	Yes
P1	I threw something at him/her	0.7	83.6	16.4
P2	I kicked, hit, or punched him/her	1.9	86.5	13.5
P3	I slapped him/her or pulled his/her hair	1.4	91.1	8.9
P4	I pushed, shoved, or shook him/her	0.9	86.2	13.8
TB1	I destroyed or threatened to destroy something he/she valued	0.7	94.5	5.5
TB2	I tried to frighten him/her on purpose	1.9	89.7	10.3
TB3	I threatened to hurt him/her	1.2	95.7	4.3
TB4	I threatened to hit him/her or throw something at him/her	3.1	93.9	6.1
S1	I kissed him/her when he/she didn't want me to	1.4	91.4	8.6
VE1	I did something to try to make him/her jealous	0.9	76.2	23.8
VE2	I brought up something bad that he/she had done in the past	0.7	79.6	20.4
VE3	I said things just to make him/her angry	0.7	77.4	22.6
VE4	I spoke to her in a hostile or mean tone of voice	0.5	81.8	18.2
VE5	I insulted him/her with put-downs	0.9	91.4	8.6
VE6	I made fun of him/her in front of others	1.4	85.4	14.6
VE7	I kept track of who he/she was with and where he/she was	1.2	76.6	23.4

VE8	I blamed him/her for the problem	0.9	81.7	18.3
VE9	I accused him/her of flirting with another guy	1.4	77.3	22.7
VE10	I threatened to end the relationship	2.1	85.3	14.7
RA1	I tried to turn her friends against her	0.7	91.9	8.1
RA2	I said things to her friends about him/her to turn them against him/her	1.7	93.5	6.5
RA3	I spread rumors about him/her	1.4	93.8	6.2
C1	I posted embarrassing photos or other images of him/her online.	0.7	98.1	1.9
C2	I sent threatening text messages to him/her.	0.9	96.2	3.8
C3	I used his/her social networking account (Facebook, Twitter, Instagram, etc) without his/her permission.	0.7	93.3	6.7
C4	I wrote nasty things about him/her on my profile page/timeline (on Facebook, Twitter, Instagram, etc).	2.1	97.1	2.9
C5	I sent him/her so many messages (like texts, emails, chats) that it made him/her feel unsafe.	0.5	96.2	3.8
C7	I spread rumors about him/her using a cell phone, email, IM, web chat, or social networking site (on Facebook, Twitter, Instagram, etc).	1.9	98.1	1.9
C8	I used information from his/her social networking profile/page to harass him/her or put him/her down.	1.2	97.4	2.6
C9	I made him/her afraid when he/she did not respond to my phone call, text, posting on social networking page, IM, etc.	0.9	94.0	6.0
C10	I threatened to harm him/her physically through a cell phone, text message, social networking page, etc.	1.2	97.4	2.6
C11	I sent him/her sexual photos or naked photos of me that I knew he/she did not want.	0.9	97.6	2.4
C12	I pressured him/her to send a sexual or naked photo of him/herself.	2.1	98.3	1.7

% = percentage

Appendix D: Item endorsement frequency of the CADRI items in the dating analytic sample of the Me & You Study by gender and race/ethnicity sub-groups

	Endorsement Frequency N (%)							
	Male		Female (N = 186)		Hispanic (N = 257)		African American (N = 127)	
Item	No	Yes	No	Yes	No	Yes	No	Yes
P1	207 (88.8)	26 (11.2)	143 (77.3)	42 (22.7)	221 (86.3)	35 (13.7)	100 (78.7)	27 (21.3)
P2	209 (91.7)	19 (8.3)	148 (80.0)	37 (20.0)	221 (88.0)	30 (12.0)	103 (81.7)	23 (18.3)
P3*	224 (97.0)	7 (3.0)	154 (83.7)	30 (16.3)	230 (91.3)	22 (8.7)	114 (89.8)	13 (10.2)
P4	211 (90.6)	22 (9.4)	149 (81.0)	35 (19.0)	224 (87.5)	32 (12.50	102 (81.0)	24 (19.0)
TB1*	223 (96.1)	9 (3.9)	173 (93.0)	13 (7.0)	246 (96.1)	10 (3.9)	115 (91.3)	11 (8.7)
TB2	210 (90.9)	21 (9.1)	160 (87.9)	22 (12.1)	227 (89.7)	26 (10.3)	109 (87.9)	15 (12.1)
TB3*	223 (96.1)	9 (3.9)	175 (95.1)	9 (4.9)	243 (96.0)	10 (4.0)	119 (93.7)	8 (6.3)
TB4*	221 (96.9)	7 (3.1)	163 (90.6)	17 (9.4)	235 (94.4)	14 (5.6)	113 (91.9)	10 (8.1)
VE1	189 (81.5)	43 (18.5)	128 (69.2)	57 (30.8)	192 (75.0)	64 (25.0)	98 (78.4)	27 (21.6)
VE2	191 (82.0)	42 (18.0)	142 (76.8)	43 (23.2)	211 (82.4)	45 (17.6)	95 (74.8)	32 (25.2)
VE3	194 (83.3)	39 (16.7)	129 (69.7)	56 (30.3)	205 (80.1)	51 (19.9)	93 (73.8)	33 (26.2)
VE4	204 (87.6)	29 (12.4)	140 (75.3)	46 (24.7)	212 (82.8)	44 (17.2)	103 (81.1)	24 (18.9)
VE5	216 (92.7)	17 (7.3)	165 (89.7)	19 (10.3)	239 (93.0)	18 (7.0)	107 (85.6)	18 (14.4)
VE6	201 (86.6)	31 (13.4)	154 (84.2)	29 (15.8)	217 (85.8)	36 (14.2)	105 (83.3)	21 (16.7)
VE7	178 (77.1)	53 (22.9)	140 (75.7)	45 (24.3)	190 (75.1)	63 (24.9)	103 (81.1)	24 (18.9)
VE8	205 (88.4)	27 (11.6)	135 (73.0)	50 (27.0)	219 (86.2)	35 (13.8)	95 (74.8)	32 (25.2)

VE9	200 (86.6)	31 (13.4)	120 (65.2)	64 (34.8)	198 (78.6)	54 (21.4)	99 (78.0)	28 (22.0)
VE10	206 (90.0)	23 (10.0)	145 (79.2)	38 (20.8)	215 (85.7)	36 (14.3)	109 (87.2)	16 (12.8)
RA1	218 (94.0)	14 (6.0)	166 (89.2)	20 (10.8)	231 (90.2)	25 (9.8)	119 (94.4)	7 (5.6)
RA2*	221 (96.1)	9 (3.9)	166 (90.2)	18 (9.8)	239 (94.5)	14 (5.5)	116 (92.8)	9 (7.2)
RA3*	222 (95.7)	10 (4.3)	167 (91.3)	16 (8.7)	239 (93.7)	16 (6.3)	115 (92.7)	9 (7.3)

N = sample size
% = percentage
*Removed from analysis given frequency of positive endorsement <5%