**THE CAMPBELL COLLABORATION**

# Protocol: School-based Interventions to Reduce Dating and Sexual Violence: A Systematic Review

## Lisa De La Rue, Joshua R. Polanin, Dorothy L. Espelage, Terri D. Pigott

Submitted to the Coordinating Group of:

| | |
|---|---|
| ☐ | Crime and Justice |
| ☒ | Education |
| | ☐ Disability |
| ☐ | International Development |
| | ☐ Nutrition |
| ☐ | Social Welfare |
| ☐ | Other: |

Plans to co-register:

| | | | |
|---|---|---|---|
| ☒ | No | | |
| ☐ | Yes | ☐ Cochrane | ☐ Other |
| ☐ | Maybe | | |

Date Submitted: 21 January 2013
Date Revision Submitted: 21 June 2013
Approval Date:
Publication Date: 2 September 2013

## BACKGROUND

Violence and assaults experienced by adolescents are of great concern to researchers, parents, educators, and administrators who strive to help youth be healthy and happy. The impact of sexual assaults or sexual coercion and physical and psychological abuse in intimate dating relationships has a significant impact on young people. Consequences of dating violence include decreased mental and physical health and lower life satisfaction (Banyard & Cross, 2008). Teen dating violence impacts the psychological well-being of youth (Black, Tolma, Callahan, Saunders, & Weisz, 2008), with youth who are victims being more likely to experience depression and suicidal behaviors (Vézine & Hérbert, 2007). Additionally, longitudinal studies have identified long-term consequences of intimate partner violence to include depression, binge eating, substance abuse, and antisocial behavior (Foshee et al., 2012). While extensive research has not yet examined the impact of teen dating violence on academic outcomes, research in this area suggests that victims may have more negative views of school and that this may be the result of increased feelings of depression and substance abuse associated with victimization experiences.

Unfortunately, despite the alarming consequences noted above, researchers and educators struggle to prevent these problems. This is complicated by the fact that school-based prevention programs vary considerably, and often demonstrate small changes or no changes at all (Espelage, 2012; Espelage, Holt, & Isaia, 2007; Espelage & Low, *in press*). The present review will quantitatively synthesize prevention and intervention efforts implemented thus far in schools that sought to reduce or prevent the incidents of dating violence.

### *The Problem, Condition or Issue*

Adolescents spend a significant amount of time with their peers in school and in their neighborhoods. Although the majority of these relationships with peers provide positive social experiences, for some youth, some relationships may also involve victimization. Many assaults experienced by youth happen within interpersonal relationships, including in friendships and with romantic partners, with one in every four assaults committed by youth occurring in a domestic relationship (i.e., family members, intimate partners) (Snyder & McCurley, 2008). Abuse in dating relationships is an all too frequent occurrence, with 1 in 10 teenage relationships involving violence (Mulford & Giordano, 2008). This latter type of victimization is often described as teen dating violence (Mulford & Giordano, 2008; Offenhauer & Buchalter, 2011) and can include controlling behaviors, physical, verbal, psychological/emotional, and sexual abuse (Holt & Espelage, 2005; Offenhauer & Buchalter, 2011). The rates of teen dating violence in middle and high school are substantial. One school-based study of 9th -12th graders found a rate of 8.7% for physical dating violence, and another nationally representative study found a 1-year incidence rate of 3.6% for 13 to 17 year olds (Hamby, Finkelhor, & Turner, 2012). While both males and females experience teen dating violence, the exact nature of the violence tends to vary by gender. About one in four girls experience sexual and physical abuse (Latta & Goodman, 2011), while boys report

experiencing high amounts of psychological abuse (Molidor, 1995). Therefore, it is essential to provide information to adolescents about healthy dating relationships to reduce the incidence of teen dating violence, including sexual, physical, and mental abuse that young people may experience in unhealthy dating relationships.

The experience of dating violence has consequences for the overall well-being of students. These experiences also challenge a student's ability to be successful in school. Negative ramifications of teen dating violence include mental health problems, low academic achievement, and aggressive conflict-management (Offenhauer & Buchalter, 2011). Research has also found that girls who are victims of violence in relationships are at risk of increased discipline problems at school (Vézina & Hébert, 2007). These consequences extend beyond externalizing symptoms. Psychological symptoms of dating violence can include feelings of incompetence, anxiety, paranoia, severe depression, isolation from family and friends, and guilt and self-blame (Molidor, 1995). There can also be long-term consequences of dating violence, which can include isolation and emotional/interpersonal withholding of support as well as continued experiences with abuse such as constant harassment and degradation (Molidor, 1995). These adverse outcomes highlight the importance of implementing policies that address dating violence as a precursor to ensuring the well-being of students and supporting their educational opportunities.

Appropriate prevention efforts can reduce the risk of abuse and victimization for adolescents, and can contribute to their healthier development and overall well-being. However, policies and programs developed to meet the needs of young people who are victims of teen dating violence must be based on research, data, and best practices. Additionally, advocating for policy changes and funding to support school-implemented programs requires that there is solid empirical justification advocating the effectiveness of such prevention programs. Taken together, it is essential to gain a better understanding of the effectiveness of dating violence prevention programs implemented in schools, and this review aims to do this.

### *The Intervention*

Myriad programs exist within schools to prevent teen dating violence. Programs may exist at a universal level, where the school introduces various stimuli or psycho-educational directives. These can include lessons provided to all students and educational posters on walls and in hallways. Researchers and administrators may also implement programs directed solely at classrooms or even individuals. Many programs, especially those implemented with young students (i.e., elementary school students), focus on altering the school culture in an effort to decrease aggression and promote respect (Flannery et al., 2003; Haynes, 1998). These prevention efforts have the goal of shifting the culture of the school in positive directions, partly by encouraging bystander support, where students are supportive of victims of dating violence while also not accepting of teen dating violence behaviors perpetrated by their peers. Programs implemented for older youth, including during middle

school, spend more time trying to change dating attitudes and behaviors (Foshee et al., 1998; Macgowan, 1997). These programs focus on teaching the individual student skills that will foster healthy dating relationships.

The Safe Dates program is an example of a school-based prevention program for adolescents (Foshee & Langwick, 2004). This program includes a 45-minute theater production, a 10-session curriculum, and a poster contest. The intent of the program is to prevent the onset and promote a decrease of dating abuse perpetration and victimization by helping students recognize the difference between caring, supportive relationships and controlling, manipulative, or abusive dating relationships. This program has been evaluated with a randomized trial, and the long-term impacts have been explored with follow-up assessments up to 4 years after the intervention. Adolescents exposed to Safe Dates reported less psychological, physical, and sexual dating perpetration and less physical dating violence victimization (Foshee et al., 2005). The program effects were mediated primarily by changes in dating violence norms, gender role norms, and awareness of community services. The authors suggest that teen dating violence programs should target efforts to change dating violence and gender-role norms, and increase awareness of community services (Foshee et al., 2005). Additional studies have also noted the positive effects of teen dating violence prevention programs implemented in schools (Foshee & McNaughton-Reyes, 2009).

Therefore, current research supports the use of dating violence prevention programs in schools. However, there is still a gap in the knowledge about the overall effectiveness of these efforts. Specifically, there is a need to quantitatively synthesize empirical studies that have been conducted. Additionally, the different types of interventions, universal and psycho-educational, and those directed at individuals or classrooms, should be reviewed and synthesized.

### Why it is Important to do the Review

During the preteen and teen years, students are learning skills they need to form positive relationships with others (Center for Disease Control and Prevention, 2012), and dating violence prevention programs in schools serve as one way to help young people build the necessary skills to promote healthy relationships. Since this is an ideal time to promote the development of healthy relationships and to prevent patterns of dating violence that can last into adulthood, it is important to ensure programs for youth that target these domains are effective. In addition, given the cost of implementing such programs and the time allotted during busy school days to engage in the material, understanding what programs are effective will also support the appropriate use of time and resources in schools.

To date no previous authors have attempted to quantitatively synthesize empirical evaluations of school-based programs designed to prevent or reduce the incidence of dating violence among adolescents. As a result, there is a lack of knowledge surrounding the overall effectiveness of teen dating violence prevention programs in schools, and around what types of interventions are most effective. However, the information is available to integrate.

Research has consistently detailed the implementation of programs, and many have included empirical information about the effectiveness of said programs. Law (n.d.) and Calvillo (2010) detail several large-scale prevention programs using a narrative review technique, and note overarching benefits of prevention programs including altering the school climate and changing attitudes supportive of teen dating violence. Foshee and McNaughton-Reyes' (2009) narrative review of school-based dating violence interventions programs noted that changes in dating abuse norms and attitudes can be made with interventions. And when there are changes in dating abuse norms and attitudes, there are actual changes in dating abuse behaviors. The National Resource Center on Domestic Violence has also conducted a narrative literature review but failed to utilize a systematic search process or detail all programs completely.

These studies have demonstrated the effectiveness of school-based interventions in reducing violence within a narrative review framework (Foshee & McNaughton-Reyes, 2009; Topping & Barron, 2009) and set the foundation for a systematic review of the literature on dating violence prevention programs. While narrative reviews provide us with valuable insight, the subjective nature creates challenges when increasingly more studies are included. Therefore, the present review will be able to accommodate a large number of programs by quantitatively synthesizing multiple studies in order to evaluate the treatment effect of teen dating violence prevention programs, thus allowing us to better understand the effectiveness of such programs. Furthermore, Fellmeth and colleagues (2011) are conducting a review of educational skills-based interventions that aim to prevent initial or further relationship violence behaviors in individuals between the ages of 12 and 25. The present review will complement their work by examining the effectiveness of school based prevention programs that aim to change behaviors *and* attitudes supportive of teen dating violence, in addition to examining how these programs encourage peer support to reduce incidents of teen dating violence.

Dating violence can involve a multitude of behaviors ,including physical and emotional abuse, sexual violence and coercion, and/or behaviors that are intended to control or intimidate a dating partner. Students can be either perpetrators or victims of teen dating violence, with some youth being both a perpetrator and a victim. This review will include studies of programs implemented in schools (grades 4-12) that sought to address attitudes or beliefs supportive of teen dating violence, encourage bystander intervention or peer support, or sought to reduce the incidence of dating violence perpetration or sexual coercion in dating relationships. In order to best understand the effectiveness of the prevention programs, specific considerations must be made. First, there needs to be clear support that an intervention was indeed the cause of a change in attitudes and behaviors, therefore only those studies that implemented an experimental or quasi-experimental design with a control group should be included. Utilizing only studies that have included a control or comparison group will help clarify the specific effects of the prevention efforts. Pre-test-post-test and follow-up measures will help minimize the attribution of changes to experimenter, practice, attention, spontaneous maturation, or Hawthorne effects (Topping & Baron, 2009). Also,

the focus needs to remain on middle and high schools, as this population has specific needs and constraints. This includes a need for developmentally specific material that can be implemented effectively in schools and classrooms. These components will be included within this review.

This review will be the first to quantitatively synthesize empirical evaluations of school-based programs aimed to reduce behaviors or change attitudes related to dating violence. This review has the potential to impact school policy and prevent consequential victimization. If indeed programs are demonstrated to be effective at reducing behaviors or attitudes supportive of teen dating violence, the strong empirical support that is gained with this type of review will bolster efforts to advocate for administrative support and funding to implement prevention programs in schools. This review can also provide insight into the type of programming (i.e. universal, psycho-educational, individual or classroom level) that is most effective in preventing victimization in dating relationships.

## OBJECTIVES

1) Are dating violence prevention programs implemented in middle and high schools effective in changing attitudes or beliefs supportive of teen dating violence?

2) Are dating violence prevention programs implemented in middle and high schools effective in encouraging bystander intervention to stop the perpetration of dating violence and/or increase peer support for victims of dating violence?

3) Are dating violence prevention programs implemented in middle and high schools effective in reducing the incidence of dating violence perpetration, including reductions in mental and/or physical abuse, and/or sexual violence or coercion perpetrated in a dating relationship?

4) Are dating violence prevention programs implemented in middle and high schools effective in reducing incidences of dating violence victimization, including reductions in mental and/or physical abuse, and/or sexual violence or coercion experienced in a dating relationship?

5) Are there substantive or methodological variables that moderate the effect sizes?

    a. The variables that will be examined include: program characteristics, age, gender, location, racial composition, and SES.

## METHODOLOGY

The purpose of the review is to synthesize the best available evidence on the effects of prevention programs designed to reduce teen dating violence perpetration and victimization. The following criteria will be employed to select studies for the review.

**Criteria for including studies in the review**

*Types of research designs*

Only studies that have a well-defined control group will be included in the review. These may include wait-list controls, treatment-as-usual, and straw-man designs. Studies that compare a dating violence program to another dating violence program (i.e., treatment vs. treatment designs) will not be included. We will include studies where participants are assigned at the individual, group, school, district, or state level. Specifically, we plan to include the following designs:

> 1. Randomized controlled trials: Studies in which individual participants, classrooms, or schools were randomly assigned to control and treatment conditions.

> 2. Quasi-randomized controlled trials: Studies where assignment is through alternating last name (or some other quasi-random procedure) to intervention or control condition.

> 3. Quasi-experimental designs (i.e., where participants or groups are assigned to conditions non-randomly). Where applicable, we will calculate pre-test effect sizes to adjust the post-test effect size (Borenstein, Hedges, Higgins, & Rothstein, 2010). Presumably, most studies will use one of the following designs that we plan to include:

>> 1) Regression discontinuity designs: studies that used a cut-score to assign program eligibility and that examined program impact around the cut-off score.

>> 2) Propensity score matching that explicitly included baseline measures of enrolment; these designs ensure equivalence among the groups.

>> 3) Matching on individual or group pre-test demographics.

>> 4) Designs that included a variety of methods such as a combination of propensity score or covariate matching with difference-in-difference techniques.

We will also note that randomized experiments have been considered paramount to quasi-experimental designs due to a theoretical lack of selection bias. A burgeoning literature insists, however, that quasi-experimental designs, given appropriate pre-test observations, have the ability to produce similarly efficient and unbiased estimates of a treatment effect (Shadish & Cook, 2009). As it is often infeasible to randomly assign students to conditions, we have chosen not to exclude studies based on lack of random assignment. As a result, we believe that including such studies will increase the total study sample size and provide a more complete picture of the literature. To ensure that theoretically high risk of selection bias does not bias the average effect size, we will conduct sensitivity analyses that remove potentially biased studies and re-calculate the average effect size.

### Types of participants

The population of interest is 4th-12th grade students. Studies that include participants outside this range must provide summary statistics for the age groups of interest. For instance, a study might implement a program for students in 11th and 12th grade as well as freshman and sophomore college students. A study that provides summary statistics for the reviewed population, in this case 11th and 12th graders, will be included. However, studies that merely provide study-level statistics that fail to disaggregate information will be excluded.

### Types of interventions

The intervention of interest is a school-based prevention program designed to reduce teen dating violence. The intervention may also seek to change other outcomes (i.e., bullying perpetration, sexual harassment, etc.); however, *a clear goal*, as provided by the authors, must state that the program sought to explicitly reduce teen dating violence behaviors, change attitudes supportive of teen dating violence, increase bystander intervention to reduce perpetration, or increase peer support for victims of dating violence. Studies that merely measure these as secondary outcomes will be excluded.

### Types of outcomes

The primary outcomes of interest are:

1. Attitudes supportive of teen dating violence behaviors

2. Frequency of engagement in adolescent intimate partner violence behaviors, including perpetration of:

    a. Verbal aggression

    b. Relational aggression (controlling, jealously)

    c. Physical aggression/violence

    d. Sexual aggression/violence or coercion

3. Frequency of victimization in adolescent intimate partner violence behaviors, including being a victim of:

    a. Verbal aggression

    b. Relational aggression (controlling, jealously)

    c. Physical aggression/violence

    d. Sexual aggression/violence or coercion

4. Knowledge about teen dating violence and what behaviors constitute teen dating violence

5. Recognizing both safe and unhealthy behaviors in intimate partner disputes

6. Learning how to be a bystander who intervenes when dating violence is perpetrated or learning how to support a victim of dating violence

Outcome information may be reported via self-report questionnaires, teacher reports and observations, or researcher reports and observations. Examples of measures may include modified versions of the *Conflicts Tactics Scale* (Straus, Hamby, Boney-McCoy, & Sugarman, 1996), which assesses the frequency of use and type of behaviors used when dealing with conflict; dating violence knowledge, which can include "true" or "false" questions on definitions of abuse, resources for help, etc. (Taylor, Stein, Mumford, & Woods, 2013); and measures of victimization, which ask how many times incidents of sexual and nonsexual violence have occurred while on a date (Foshee et al., 2000). The outcome's metric is often determined by the primary author and is therefore difficult to hypothesize prior to coding.

### Types of time points

We plan to synthesize outcomes measured directly after intervention (i.e., post-test) and at follow-up. The first post-program measures will be considered outcomes that follow the intervention, and any subsequent measures will be considered as follow-up effects. In essence, two simultaneous syntheses will be presented. This technique allows for greater understanding of the programs' effects both short- and long-term while guarding against non-independence of effect sizes. Therefore, time will not be used as a moderator (as is sometimes the case). Instead program effects will be calculated as outcomes immediately following the completion of a program, and also as outcomes sustained for a period after a student has ended a program.

### Types of settings

The review will include studies conducted within middle and high schools. Studies that utilize community centers or other locations outside the brick-and-mortar schools will be excluded. We will include, however, all types of schools (i.e., publicly or privately funded). After-school programs conducted at the school will be included.

### Example of included study

Weisz and Black (2001) implemented a school-based intervention in an urban middle school to change students' knowledge about and attitudes toward dating violence. The authors recruited seventh-grade students by offering a $5 McDonald's gift card ($n = 44$). A control group was formed by randomly selecting students from the same school who chose not to participate in the program ($n = 20$). This study constituted a non-randomized design.

The authors assessed pre-test equality using an independent t-test; the results indicated no significant differences at pre-test. Results at post-test indicted significant increases in knowledge about and attitudes toward sexual violence for students who received the intervention ($gs$ = .20 & .17, respectively).

## Search Methods for Identification of Studies

### *Electronic searches*

Relevant studies will be identified using electronic database searches, government policy databanks, and Internet search engines. We will place neither language nor date restriction on the search. We plan to search from 1960 – February 2013.

The following "traditional" databases will be searched:

1. ERIC

2. PsycInfo

3. SocIndex

4. ASSIA

5. PubMed

6. Sociological Abstracts

7. GALE

8. Academic Search Premier

The traditional databases include international publications. However, we will also search CBCA Education, the British Education Index, and the Australian Education Index for citations outside the US.

The following "grey literature" databases will also be searched:

1. scientific.thomson.com databases

2. csa.com/factsheets databases

3. apa.org/psyextra database

4. Proquest (for dissertations and theses)

### *Search terms*

Search terms will be created using relevant key words that represent the studies of interest.

Table 1 lists each of the relevant search terms. These terms will then be arranged to produce a Boolean search phrase for each combination. An example of the search phrase that will be used in ERIC is presented in Appendix A.

### Searching other resources

Two further procedures will be conducted to ensure search breadth. First, the authors will scan and screen bibliographies of included studies. Hammerstrøm, Wade, and Jørgensen (2010) showed that this procedure has the ability to produce substantial additional studies even with the most robust searches of electronic databases. Second, the authors will contact high-profile researchers in this field. These contacts will ensure that relevant published studies are included and may engender the inclusion of "file-drawer" datasets yet unpublished (Rosenthal, 1991). We will identify high-profile researchers by identifying who has published extensively within the field of teen dating violence, and identifying who has received funding to explore teen dating violence prevention programs. Finally, we will search websites of foundations and organizations that aim to decrease the prevalence of sexual and dating violence among youth for studies that may not have been included from previous search efforts. Examples of foundations and organizations include Centers for Disease Control and Prevention, National Institute of Justice, Love is Respect, Robert Wood Johnson Foundation, and the American Association of University Women.

To identify potential grey literature outside of indexed databases, we will also search Google, Google Scholar, and Bing search engines. This will locate conference abstracts, government documents, and other online material. We plan to include any studies that meet the inclusion criteria listed above regardless of source.

### Targeting specific journals

One member of the review team will hand-search the *Journal of Counselling Psychology* and *Prevention Science* starting in 2000 to locate any additional studies or references. These journals were selected because they had high initial citation counts relative to the purpose of this review.

### Screening procedures

The screening process will occur in two distinct phases. First, LD and JRP will independently screen each title and abstract obtained from the search procedures described above for inclusion (see Appendix B). Each reviewer will code each citation according to pre-determined inclusion criteria. This information will be stored in an EXCEL database. Disagreements will be handled by discussion and consensus agreement with a third reviewer (DLE). The decisions available to the reviewer are: 1) Yes, include for full article scan, 2) Unclear so include for full scan, 3) Unclear but do not include (include reason), and 4) No, this article should be eliminated (include reason). Citations that meet the inclusion criteria will be retrieved for full review using the University of Illinois and Loyola University Chicago

library resources.

Second, each of the first and second authors will also independently screen the full articles for inclusion (see Appendix C). As with the previous procedure, we will screen the studies using the inclusion criteria, keeping track of that information in an EXCEL database. Should the citation be excluded at this stage, the reviewer will provide a brief description of the reason for dismissal. Again, a discussion and agreement will be reached for each disagreement using a third reviewer.

**Data Extraction**

***Extraction of study information***

At least two reviewers (LD, JRP) will independently code all of the included studies. The codebook (Appendix D) details the variety of study characteristics that will be coded. The codebook will be operationalized in a Microsoft Access database. Electronic coding is preferable to hand-coding because it reduces data entry errors (Cooper, 2010). Microsoft Access will be utilized because of the hierarchical nature of data extraction. For instance, a study may include multiple outcomes nested within multiple treatment groups. Access allows the reviewers to assign this multiplicity of information to one study, in turn limiting the amount of redundant coding.

All studies will be double-coded by an independent member of the review team (DLE or TDP) and inter-rater reliability (i.e., percentage match) will be calculated. Disagreements will be discussed and a consensus code used.

***Assessment of risk of bias in included studies***

The review team will assess the methodological quality of studies using the risk of bias tool developed by the Cochrane Methods group (Higgins, Altman, & Sterne, 2011). The risk of bias tool assesses study quality on nine indicators. The review team will not exclude studies based on the risk of bias assessment because this procedure has been shown to substantially bias meta-analytic results (Juni, Witschi, Bloch, & Egger, 1999). Rather, we will use the results of the risk of bias assessment as a categorical moderator (see moderator section below). In addition to the risk of bias assessment, we will also use attrition information as additional sources of information about study quality.

It should be noted that, because the risk of bias was developed for medical interventions, certain aspects of the tool may not be reported by the primary authors. It is uncommon for social science researchers to employ allocation concealment techniques and all but impossible to blind participants to psycho-educational interventions. Given that all studies are missing one of the nine indicators, however unlikely, we will not count that indicator toward a studies' overall level of bias. Instead we will use pre-test information to evaluate studies for risk of selection bias. We will calculate pre-test effect sizes to test for equivalence

of all outcomes reported, regardless of whether the outcome will be included in the review. Should more than half of the reported outcomes differ, the study will be labelled "high risk" and sensitivity analyses will be conducted by removing such studies. In addition, we will note if treatment and control groups differ on demographics variables or on any outcome measures of interest. Again, should demographic differences exist, we will conduct sensitivity analyses where these studies are removed from the synthesis.

### *Effect size calculation*

Effect sizes will be extracted from each study using relevant summary statistics. Traditional effect size calculation procedures will follow those laid out by Lipsey and Wilson (2001) and Hedges and Olkin (1985). David Wilson's online effect size calculator (2013) as well as the software program Comprehensive Meta-Analysis (Borenstein, Hedges, Higgins, & Rothstein, 2005) will be utilized to calculate each effect size.

#### *Discrete data*

Outcomes that measure incidence or a binary instance (yes/no) will be estimated via the odds ratio and 95% confidence interval as calculated by the 2 x 2 frequency table. It should be noted that all discrete effect sizes will be converted to the d-metric using Wilson's (2013) online effect size calculator (Lipsey & Wilson, 2001, pg. 187). The effect size measure (discrete vs. continuous) will be utilized as a categorical moderator to ensure that conversion does not bias the mean effect size.

#### *Continuous data*

The preferred summary statistics for continuous data are means and standard deviations (or standard errors). Cohen's d is easily calculated; the numerator is the mean post-test difference of treatment and control scores and the denominator is the pooled standard deviation. Lipsey and Wilson (2001) demonstrated, however, that Cohen's d effect sizes can be calculated from a variety of statistical information. For instance, t tests, F ratios, chi-squared values, correlation, and regression coefficients all represent statistical data that can be converted into a standard effect size.

To limit small-sample bias, all d effect sizes will be converted to Hedges' *g* effect size (Hedges & Olkin, 1985).

#### *Dependent effect sizes*

Dependent effect sizes can occur for myriad reasons. For instance, studies may include multiple treatment arms but only one control group; studies may include multiple time points; authors may publish the results of the same project in more than one report. To avoid biasing the average effect size, we will include only one effect size measure per study. We will pool the treatment group mean and standard deviation prior to calculating the effect size for studies that include multiple treatment arms with only one control group. Multiple time

points will be synthesized separately. As stated earlier, we will conduct two separate meta-analyses, one for the immediate post-test and one for the follow-up effects. For follow-up effects, it is likely studies will report follow-up effects at multiple time points after the intervention. In this case, we will take the final follow-up time point available and use that to calculate the effect size. Studies published in multiple reports will be handled by using the latest published iteration and therefore excluding all other versions.

It is important to note that primary studies can produce multiple effect sizes. Where clear evidence is given that separate studies are represented within one publication, the same publication can be represented twice. This is a rare occasion but paramount to consider given that the inclusion of one study effect size over another, without random selection, has the potential to bias meta-analytic results (Littell, Corcoran, & Pillai, 2008).

In a similar fashion, the review team will take pains to synthesize as many primary outcomes as possible. Again, we should note, however, that multiple effect sizes within the same construct category from the same study will not be synthesized. For instance, it is possible that one study will report multiple primary outcomes. Instead of synthesizing all outcomes as independent findings or simply averaging the effect sizes, multiple syntheses will be conducted for each construct. Separate analysis will be conducted for each primary outcome of interest. When multiple effects are provided on the same primary outcomes (e.g., if two effect sizes are provided for two different verbal aggression outcomes), these effect sizes will be averaged so as to have one effect size for this primary outcome.

*Unit of analysis*

A final important methodological consideration is the unit of analysis. Often, participants are randomly assigned into a treatment or control condition. In education research, however, entire classrooms or schools often are randomly assigned to condition. Calculating the variance assuming individual assignment, when group assignment is maintained, biases the study's effect and ultimately the average effect size. We will use statistical adjustments that correct for this phenomenon (Pigott, 2010).

**Meta-Analytic Procedures**

***Effect size synthesis***

Prior to synthesis, we will determine the appropriate effect size metric. Should the majority of studies provide a continuous measure of the effect, we will convert all discrete effect sizes to the continuous metric first. We hypothesize this will be the case. However, should most studies utilize discrete data, we will convert all effect sizes to discrete. The conversion procedures, as mentioned previously, will be performed via Wilson's (2013) online effect size calculator.

The review will utilize a random-effects model for effect size synthesis. This choice is

preferable given that the effect sizes derive from a theoretical population of effect sizes (Borenstein, Hedges, Higgins, & Rothstein, 2010). Thus, the random-effects model accounts for the distributional effects by including a between-groups variance component in addition to the within-group variance. The effect size and confidence interval therefore reflect a more conservative estimate relative to fixed-effect models. It should be mentioned that the fixed-effect estimate will not be reported.

The synthesis calculation is represented below:

$$\overline{ES} = \frac{\sum (w_i * ES_i)}{\sum \dots}$$

where $w_i$ represents the random-effect weight from study $i$ and $ES_i$ is the effect size from study $i$.

A z-test and confidence interval is calculated by estimating the standard error of the average weighted effect size. This calculation is represented below:

$$SE_{ES} = \sqrt{\frac{1}{\sum \dots}}$$

where $w_i$ again represents the random-effects weight from study $i$.

### Assessing heterogeneity

Following the average weighted effect size calculation, the authors' next procedure will be to assess for heterogeneity of effect sizes. Traditionally, heterogeneity is assessed using the Q statistic; this statistic is calculated using the following formula:

$$Q = \sum W_i * (ES_i - \overline{ES})^2$$

where $W_i$ is the random-effects weight for study $i$, $ES_i$ is the effect size for study $i$, and $\overline{ES}$ represents the overall mean effect size. The calculated Q statistic is distributed as a chi-square with $k - 1$ degrees of freedom and $k$ is the number of effect sizes (Hedges & Olkin, 1985). We also plan to estimate and report values for $\tau^2$ and $I^2$ to assess for heterogeneity in the effect sizes.

### Moderator analyses

Given a statistically significant $Q$ statistic, the authors will conduct moderator analyses to assess variation in treatment effects. Lipsey (2009) suggested utilizing three types of moderators: extrinsic, methodological, and substantive. Extrinsic variables are represented by the study's unchangeable characteristics, for instance date of publication, or published and not published. Methods variables can be represented by random vs. non-random assignment. Length of intervention or intervention location constitutes substantive moderators.

We plan to investigate the impact of 13 moderator variables. Following the suggestion of Lipsey (2009), they have been subdivided into 3 broad categories:

1) *Extrinsic*

a) Date of Publication: Trikalinos and Ioannidis (2005) showed that treatment effects may diminish over time as more sophisticated methodological approaches and designs are utilized. It is therefore important to test this assumption.

b) Publication Type: Rothstein, Sutton, and Borenstein (2005) wrote extensively on the impact of publication bias, especially with regard to publication type.

c) Funding: Although less common in social science, in medicine it has been shown that funded studies tend to produce more significant treatment effects (Dickersin, 1997).

2) *Methodological*

a) Design: We will be including quasi-experimental designs in addition to randomized experiments. Because quasi-experimental designs may inherently differ from randomized experiments, especially with regard to selection bias (Shadish, Cook, & Campbell, 2002), it is important to rule out treatment effect differences.

b) Risk of Bias: Higgins, Altman, and Sterne (2011) suggested conducted moderator analyses based on the potential level of bias. Per Table 8.7a, it is possible to group studies according to the "low risk of bias," "unclear risk of bias," and "high risk of bias." We plan to utilize these categories as level to test for differences among the effect sizes.

c) Metric: We plan to utilize a single effect size metric (i.e., only standardized-mean difference) based on the majority of effect sizes. As such, we will test whether the studies that utilized the different effect size metric biased by testing for differences.

3) *Substantive*

a) Program Type: If possible, we will group studies according to the type of program utilized.

b) Age: Given that we plan to include a wide range of students, it will be important to test of differences in treatment effect by age.

c) Gender

d) Location

e) Racial composition

f) SES

Borenstein, Hedges, Higgins, and Rothstein (2010) explicated several options for assessing conditional effect size differences. A distinction must first be made between categorical and continuous moderators. Categorical moderators will be assessed using ANOVA-like procedures that compare between- and within-level variance. Continuous moderators will be empirically assessed using meta-regression.

### Publication bias

Publication bias will be controlled for and assessed using various methodological and statistical techniques. First, the authors will take pains to locate and include studies not published via traditional methods. Sutton (2005) showed bias consequences associated with the exclusion of un-published studies; overall, studies that fail to find positive intervention effects tend to remain un-published. As a result, the meta-analytic estimate of the average weighted effect size is upwardly biased. Including un-published studies prevents this bias.

Second, the authors will evaluate empirically the presence of publication bias using funnel plot asymmetry in addition to Duval and Tweedie's (2000) Trim and Fill procedure. Funnel plot asymmetry is a simple graphical technique that plots the study's standard error against its effect size. Asymmetry detects potential bias of the outcomes. The trim and fill procedure's is also to locate potentially missing effect sizes, usually ones due to a non-positive treatment effect. The technique estimates an imputed effect size and recalculates the average weighted effect size. Although this procedure is not robust to criticism, it can be useful for sensitivity analysis purposes.

### Handling of missing data

Missing data is bound to occur in the meta-analytic context. Pigott (2009) suggested multiple methods for handling the inevitable situations. One practical solution often requires the meta-analyst to estimate an effect size given less than ideal summary statistics. These situations offer require little more than the correct software packages. We will use this technique when summary statistics are not available to calculate an effect size.

Another pragmatic solution will be to contact the primary author for further information. This will be especially useful for studies conducted and published recently. We will use this option when an effect size cannot be calculated or when moderator information is missing.

### Sensitivity Analyses

We will observe the distribution of effect sizes to check for potential outliers. Should an effect size rest more than 3 SDs from the mean, we will remove the effect size and recalculate the mean effect size.

We also plan to estimate the correlations among the moderators. Should substantial covariation exist (i.e., $r > .6$), we will conduct additional sensitivity analyses to address potential interactions among the moderators.

Finally, any quasi-experiments that fail to use one of the pre-determined acceptable methods will be removed from the total sample and the average effect size will be re-calculated.

### *Software*

Wilson's (2013) online effect size calculator will be utilized to calculate effect sizes. The stand-alone software Comprehensive Meta-Analysis (CMA; Borenstein, Hedges, Higgins, & Rothstein, 2005) will be used to calculate all meta-analytic estimates. CMA will also be used to produce data tables and graphs.

### Inclusion of Qualitative Research

We do not plan to include qualitative research.

## TABLES

*Table 1: Search Terms by Classification*

| Type | Outcomes | Design | Sample |
|---|---|---|---|
| Intervention | Sexual violence | Experiment* | 4-12 grade |
| Prevention | Sexual coercion | Quasi-Experiment* | High School |
| Program | Peer support | | Middle School |
| | Intimate Partner violence | | Middle Level |
| | Bystander | | |
| | Dating Violence | | |
| | Physical Violence | | |
| | Dating Aggression | | |
| | Dating Abuse | | |
| | Rape | | |

*Table 2: Moderator Variables*

| Variable | Moderator Type | Data Type | Description |
|---|---|---|---|
| Date of Publication | Extrinsic | Continuous | Year of publication |
| Publication type | Extrinsic | Categorical | Published vs. Unpublished |
| Funding | Extrinsic | Categorical | Funded vs. Unfunded |
| Design | Methodological | Categorical | Experimental vs. Quasi-experimental |
| Risk of Bias | Methodological | Categorical | ROB Ratings |
| Metric | Methodological | Categorical | d, OR, %, etc. |
| Program | Substantive | Categorical | Difference in program |

| characteristic | | | types |
|---|---|---|---|
| Age | Substantive | Continuous | Mean age |
| Gender | Substantive | Continuous | Mean percentage male |
| Location | Substantive | Categorical | Rural, urban, etc. |
| Racial composition | Substantive | Continuous | Mean percentage white |
| SES | Substantive | Continuous | Mean percentage SES |

## REFERENCES

Banyard, V. L., & Cross, C. (2008). Consequences of teen dating violence: understanding intervening variables in ecological context. *Violence Against Women, 14*, 998-1013. doi: 10.1177/1077801208322058

Black, B. M., Tolma, R. M., Callahan, M., Saunders, D. G., & Weisz, A. N. (2008). When will adolescents tell someone about dating violence victimization? *Violence Against Women, 14*, 741-758. doi: 10.1177/1077801208320248

Borenstein, M., Hedges, L.V., Higgins, J.P.T., & Rothstein, H.R. (2005). *Comprehensive meta- analysis, Version 2.* Englewood, NJ: Biostat.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods, 1*, 97-111. doi: 10.1002/jrsm.12

Calvillo, R.N. (2010). *Latino teens at risk: The effectiveness of dating violence prevention programs.* (Senior thesis). Claremont McKenna College, Claremont, USA. Retrieved from: http://scholarship.claremont.edu/cgi/viewcontent.cgi?article=1070&context= cmc _theses

Center for Disease Control and Prevention. (2012). *Understanding teen dating violence.* Retrieved from: http://www.cdc.gov/ViolencePrevention/pdf/TeenDating Violence2012-a.pdf

Cooper, H. (2010). *Research synthesis and meta-analysis: a step-by-step approach.* (4th ed). Thousand Oaks, CA: Sage Publications.

Dickersin, K. (1997). How important is publication bias? a synthesis of available data. *AIDS Education and Prevention: An Interdisciplinary Journal, 9*, 15-21.

Duval, S., & Tweedie, R. (2000). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95*, 89-98. doi: 10.1080/01621459.2000.10473905

Espelage, D. L. (2012). *Commissioned paper for National Institute of Justice Research Meeting on Longitudinal Data on Teen Dating Violence.* Retrieved from: http://www.nij.gov/nij/topics/crime/intimate-partner-violence/teen-dating-violence/longitudinal-data-meeting/tdv-longitudinal-data-research-meeting-paper.pdf

Espelage, D. L., Holt, M. K., & Isaia, A. (2007). Dating violence among adolescents— Understanding the roles of attachment, self-esteem, dominance, and need for

interpersonal control.  In K.A Kendall-Tackett & S.M. Giacomoni (Eds.), *Intimate partner violence* (pp. 14-15). Kingston, NJ: Civic Research Institute.

Espelage, D., & Low, S. M. (in press). Prevention of bullying, dating violence, & sexual violence among children and adolescents. In E. M. Vera, (Ed.) *Oxford Handbook of Prevention in Counseling Psychology.* New York:  Oxford University Press.

Fellmeth, G. L. T, Heffernan, C., Nurse, J., Habibula, S., & Sethi, D. (2011). Educational and skills-based interventions for preventing relationship and dating violence in adolescents and young adults (Protocol). *Cochrane Database of Systematic Reviews, Issue 6. Art. No.: CD004534.* doi: 10.1002/14651858.CD004534.pub3

Flannery, D. J., Vazsonyi, A. T., Liau, A. K., Guo, S., Powell, K. E., Atha, H., Vesterdal, W., & Embry, D. (2003). Initial behavior outcomes for Peacebuilders universal school-based violence prevention program. *Developmental Psychology, 39*, 292 – 308.  doi: 10.1037/0012-1649.39.2.292

Foshee, V. A., Bauman, K. E., Arriaga, X. B., Helms, R. W., Koch, G. G., & Linder, G. F. (1998). An evaluation of Safe Dates, an adolescent dating violence prevention program. *American Journal of Public Health, 88,* 45-50.

Foshee, V. A., Bauman, K. E., Ennett, S. T., Suchindran, C., Benefield, T., & Linder, G. F. (2005). Assessing the effects of the dating violence prevention program "Safe Dates" using random coefficient regression modelling. *Prevention Science, 6,* 245-258. doi: 10.1007/s11121-005-0007-0

Foshee, V. A., Bauman, K. E., Greene, W. F., Koch, G. G., Linder, G. F., & MacDougall, J. E. (2000). The safe dates program: 1-year follow up results. *American Journal of Public Health, 90,* 1619-1622.

Foshee, V. A., & Langwick, S. (2004). *Safe Dates: An adolescent dating abuse prevention curriculum* [Program Manual]. Center City, MH: Hazelden Publishing and Educational Services. Retrieved from: http://www.hazelden.org/web/public/document/safe_dates _scope_sequence.pdf

Foshee, V. A., & McNaughton-Reyes, H. L. (2009). Primary prevention of adolescent dating abuse perpetration: When to begin, whom to target, and how to do it. In D. J. Whitaker & J. R. Lutzker (Eds.) *Preventing partner violence: research and evidence-based intervention strategies* (pp. 141-168). Washington, DC: American Psychological Association.

Foshee, V. A., McNaughton-Reyes, H. L., Ennett, S. T., Cance, J. D., Bauman, K. E., & Bowling, J. M. (2012). Assessing the effects of Families for Safe Dates, a family-based teen dating abuse prevention program. *Journal of Adolescent Health, 51*, 349-356. doi: 10.1016/j.jadohealth.2011.12.029

Hamby, S., Finkelhor, D., & Turner, H. (2012). Teen dating violence: co-occurrence with other victimizations in the National Survey of Children's Exposure to Violence (NatSCEV). *Psychology of Violence, 2*, 111-124. doi: 10.1037/a0027191

Hammerstrøm, K., Wade, A., & Jørgensen, A-M. K. (2010) Searching for studies: a guide to information retrieval for Campbell Systematic Reviews. *Campbell Systematic Reviews, Supplement 1*, 1-75. doi: 10.4073/csrs.2010.1

Haynes, K. S. (1998). The one hundred-year debate: social reform versus individual treatment. *Social Work, 43*, 501-509.

Hedges, L. V. & Olkin, I. (1985). Statistical methods for meta-analysis. Orlando, FL: Academic Press.

Higgins, J. P. T., Altman, D. G., & Sterne, J. A. C. (2011). Chapter 8: assessing risk of bias in included studies. In J. P. T Higgins & S. Green (Eds.) *Cochrane Handbook for Systematic Reviews of Interventions* (V. 5.1.0). Retrieved from: www.cochrane-handbook.org

Holt, M. K., & Espelage, D. L. (2005). Social support as a moderator between dating violence victimization and depression/anxiety among African-American and Caucasian adolescents. *School Psychology Review, 34*, 309-328.

Jüni, P., Witschi, A., Bloch, R., & Egger, M. (1999). The hazards of scoring the quality of clinical trials for meta-analysis. *Journal of the American Medical Association, 282*, 1054-1060.

Latta, R. E., & Goodman, L. A. (2011). Intervening in partner violence against women: a grounded theory exploration of informal network members' experiences. *The Counseling Psychologist, 39*, 973 –1023. doi: 10.1177/0011000011398504

Law, K. M. (n.d.). *Youth sexual violence prevention programs: What we know and what we are doing.* Michigan Coalition Against Domestic and Sexual Violence: Okemos, MI. Retrieved from: http://www.mcadsv.org/resources/prevention/files/ Youth%20SV%20Prevention.pdf

Lipsey, M. W. (2009). Identifying interesting variables and analysis opportunities. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.) *The handbook of research synthesis and meta-analysis,* (pp. 147-158), New York: Sage Publication.

Lipsey, M. W. & Wilson, D. B. (2001). *Practical meta-analysis (2nd ed.).* Thousand Oaks, CA: Sage Publications.

Littell, J. H., Corcoran, J., & Pillai, V. (2008). *Systematic reviews and meta-analysis.* New York, NY: Oxford University Press.

Macgowan, M. J. (1997). An evaluation of a dating violence prevention program for middle school students. *Violence and Victims, 12*, 223 – 235.

Molidor, C. E. (1995). Gender differences of psychological abuse in high school dating relationships. *Child and Adolescent Social Work Journal, 12*, 119-134.

Mulford, C. & Giordano, P. C. (2008). Teen dating violence: a closer look at adolescent romantic relationships. *National Institute of Justice, 261*. Retrieved from: http://www.nij.gov/journals/261/teen-dating-violence.htm

Offenhauer, P., & Buchalter, A. (2011). Teen dating violence: a literature review and annotated bibliography. Retrieved from: https://www.ncjrs.gov/pdffiles1/nij/grants/235368.pdf

Pigott, T.D. (2009). Handling missing data. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.) *The handbook of research synthesis and meta-analysis,* (pp. 399-416), New York: Sage Publication.

Rosenthal, R. (1991). Meta-analysis: a review. *Psychosomatic Medicine, 53*, 247-271.

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2005). *Publication bias in meta-analysis: prevention, assessment and adjustments.* West Sussex, England: Wiley & Sons, Ltd.

Shadish, W. R., & Cook, T. D. (2009). The renaissance of field experimentation in evaluating interventions. *Annual Review of Psychology, 60*, 607-629. doi: 10.1146/annurev.psych.60.110707.163544

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston, MA: Houghton Mifflin.

Snyder, H.N., and McCurley, C. (2008). *Domestic assaults by juvenile offenders.* Washington, DC: U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention. Retrieved from: https://www.ncjrs.gov/pdffiles1/ojjdp/219180.pdf

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*, 283–316. doi: 10.1177/019251396017003001

Sutton, A. J. (2005). Evidence concerning the consequences of publication and related biases. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: prevention, assessment and adjustments* (pp. 175-192), West Sussex, England: Wiley & Sons, Ltd. doi: 10.1002/0470870168.ch10

Taylor, B. G., Stein, N. D., Mumford, E. A., & Woods, D. (2013). *Shifting Boundaries*: an

experimental evaluation of a dating violence prevention program in middle schools. *Prevention Science, 14*, 64-76. doi: 10.1007/s11121-012-0293-2

Topping, K. J., & Barron, I. G. (2009). School-based child sexual abuse prevention programs: A review of effectiveness. *Review of Educational Research, 79*, 431-463. doi: *10.3102/0034654308325582*

Trikalinos, T.A., & Ioannidis, J.P.A. (2005). Assessing the evolution of effect sizes over time. In Rothstein, H.R., Sutton, A.J., Borenstein, M. (Eds.), *Publication bias in meta-analysis: prevention, assessment and adjustments* (pp. 241-260), West Sussex, England: Wiley & Sons, Ltd.

Vézina, J., & Hébert, M. (2007). Risk factors for victimization in romantic relationships of young women: A review of empirical studies and implications for prevention. *Trauma, Violence, & Abuse, 8*, 33-66. doi: 10.1177/1524838006297029

Weisz, A.N., & Black, B.M. (2001). Evaluating a sexual assault and dating violence prevention program for urban youths. *Social Work Research, 25,* 89-100. doi:10.1093/swr/25.2.89

Wilson , D. B.(2013). *Practical meta-analysis effect size calculator*. Retrieved from: http://gemini.gmu.edu/cebcp/EffectSizeCalculator/index.html

## REVIEW AUTHORS

### Lead review author

| | |
|---|---|
| **Name:** | **Lisa De La Rue** |
| Title: | Doctoral Student |
| Affiliation: | University of Illinois Urbana-Champaign |
| Address: | 1101 S. Mattis Ave 2-6 |
| City, State, Province or County: | Champaign, IL |
| Postal Code: | 61821 |
| Country: | United States |
| Phone: | 408-772-5862 |
| Mobile: | 408-772-5862 |
| Email: | delarue2@illinois.edu |

### Co-authors

| | |
|---|---|
| **Name:** | **Joshua R Polanin** |
| Title: | Postdoctoral Fellow |
| Affiliation: | Vanderbilt University |
| Address: | 230 Appleton Place |
| City, State, Province or County: | Nashville, TN |
| Postal Code: | 37203 |
| Country: | USA |
| Phone: | 217.369.2046 |
| Mobile: | |
| Email: | joshua.r.polanin@vanderbilt.edu |

| | |
|---|---|
| **Name:** | **Dorothy L. Espelage** |
| Title: | Professor |
| Affiliation: | University of Illinois Urbana-Champaign |
| Address: | 1310 S. 6th St. 226A |
| City, State, Province or County: | Champaign, IL |
| Postal Code: | 61820 |
| Country: | United States |

| Phone: | 217-333-9139 |
|---|---|
| Mobile: | 217-766-6413 |
| Email: | espelage@illinois.edu |

| **Name:** | **Terri D. Pigott** |
|---|---|
| Title: | Professor |
| Affiliation: | Loyola University Chicago |
| Address: | 820 North Michigan Ave. |
| City, State, Province or County: | Chicago, IL |
| Postal Code: | 60611 |
| Country: | United States |
| Phone: | 312-915-6245 |
| Mobile: | |
| Email: | tpigott@luc.edu |

## ROLES AND RESPONSIBLIITIES

Please give brief description of content and methodological expertise within the review team. The recommended optimal review team composition includes at least one person on the review team who has content expertise, at least one person who has methodological expertise and at least one person who has statistical expertise. It is also recommended to have one person with information retrieval expertise.

Who is responsible for the below areas? Please list their names:

• Content: Lisa De La Rue and Dorothy Espelage

• Systematic review methods: Josh Polanin and Terri Pigott

• Statistical analysis: Josh Polanin and Terri Pigott

• Information retrieval: Lisa De La Rue and Josh Polanin

Lisa De La Rue will be responsible for collecting articles, coding, synthesizing the articles, and report writing. Joshua R. Polanin will be responsible for practical methodological aspects, study screening and coding, synthesis analysis, and report writing. Dorothy L. Espelage will serve as a consultant and will oversee the accuracy and presentation of writing on the topics of dating and sexual violence and sexual harassment. Terri D. Pigott will serve as a consultant and will oversee the methodological aspects of the review,

including study screening and coding.

## SOURCES OF SUPPORT

The Campbell Collaboration Education Coordinating Group provided a systematic review grant for this project.

## DECLARATIONS OF INTEREST

There are no conflicts of interest.

## PRELIMINARY TIMEFRAME

*December 2nd* (completed)
- Submit title registration to Campbell Collaboration

*January 21st*
- Submit protocol proposal to Campbell Collaboration

*February 7th*
- Contact potential authors for additional studies

*March 21st*
- Receive protocol reviews

*April 5th*
- Submitted revised protocol based on reviews

*April 5th*
- Coding begins
- LDR and JRP will code each study independently
- Members of Dorothy Espelage's (DLE) research team, trained by LDR and JRP, will double code each study

*August 6th*
- Coding ends; analysis and review writing begins
- JRP and Terri Pigott (TDP) will be responsible for the analysis
- Both have extensive expertise using meta-analytic datasets
- LDR and DLE will be responsible for producing the text of the report
- Each has experience writing reports and scientific literature

*September 15th*

- Submit review to Campbell Collaboration

## PLANS FOR UPDATING THE REVIEW

Lisa De La Rue will assume responsibility for maintaining the review in light of new evidence, comments and criticisms, and other developments, and updating the review at least once every three years. If Lisa De La Rue is no longer able to resume this responsibility, she will ensure an appropriate transfer of responsibility is made, and is agreed on by the Campbell Library.

## AUTHORS' RESPONSIBILITIES

By completing this form, you accept responsibility for preparing, maintaining and updating the review in accordance with Campbell Collaboration policy. The Campbell Collaboration will provide as much support as possible to assist with the preparation of the review.

A draft review must be submitted to the relevant Coordinating Group within two years of protocol publication. If drafts are not submitted before the agreed deadlines, or if we are unable to contact you for an extended period, the relevant Coordinating Group has the right to de-register the title or transfer the title to alternative authors. The Coordinating Group also has the right to de-register or transfer the title if it does not meet the standards of the Coordinating Group and/or the Campbell Collaboration.

You accept responsibility for maintaining the review in light of new evidence, comments and criticisms, and other developments, and updating the review at least once every five years, or, if requested, transferring responsibility for maintaining the review to others as agreed with the Coordinating Group.

### Publication in the Campbell Library

The support of the Coordinating Group in preparing your review is conditional upon your agreement to publish the protocol, finished review, and subsequent updates in the Campbell Library. The Campbell Collaboration places no restrictions on publication of the findings of a Campbell systematic review in a more abbreviated form as a journal article either before or after the publication of the monograph version in *Campbell Systematic Reviews*. Some journals, however, have restrictions that preclude publication of findings that have been, or will be, reported elsewhere and authors considering publication in such a journal should be aware of possible conflict with publication of the monograph version in *Campbell Systematic Reviews*. Publication in a journal after publication or in press status in *Campbell Systematic Reviews* should acknowledge the Campbell version and include a citation to it. Note that systematic reviews published in *Campbell Systematic Reviews* and co-registered with the Cochrane Collaboration may have additional requirements or restrictions for co-publication. Review authors accept responsibility for meeting any co-publication requirements.

I understand the commitment required to undertake a Campbell review, and agree to publish in the Campbell Library. Signed on behalf of the authors:

Form completed by: Lisa De La Rue

Date: 13 August 2013

## APPENDICES

### Appendix A. ERIC Search Example

(Prevention AND "Sexual violence" AND Experiment*) OR
(Prevention AND "Sexual violence" AND Quasi-Experiment*) OR
(Prevention AND "Sexual coercion" AND Experiment*) OR
(Prevention AND "Sexual coercion" AND Quasi-Experiment*) OR
(Prevention AND "Peer support" AND Quasi-Experiment*) OR
(Prevention AND "Intimate Partner violence" AND Experiment*) OR
(Prevention AND "Intimate Partner violence" AND Quasi-Experiment*) OR
(Prevention AND "Bystander" AND Experiment*) OR
(Prevention AND "Bystander" AND Quasi-Experiment*) OR
(Prevention AND "Dating Violence" AND Experiment*) OR
(Prevention AND "Dating Violence" AND Quasi-Experiment*) OR
(Prevention AND "Physical Violence" AND Experiment*) OR
(Prevention AND "Physical Violence" AND Quasi-Experiment*) OR
(Prevention AND "Dating Aggression" AND Experiment*) OR
(Prevention AND "Dating Aggression" AND Quasi-Experiment*) OR

(Intervention AND "Sexual violence" AND Experiment*) OR
(Intervention AND "Sexual violence" AND Quasi-Experiment*) OR
(Intervention AND "Sexual coercion" AND Experiment*) OR
(Intervention AND "Sexual coercion" AND Quasi-Experiment*) OR
(Intervention AND "Peer support" AND Quasi-Experiment*) OR
(Intervention AND "Intimate Partner violence" AND Experiment*) OR
(Intervention AND "Intimate Partner violence" AND Quasi-Experiment*) OR
(Intervention AND "Bystander" AND Experiment*) OR
(Intervention AND "Bystander" AND Quasi-Experiment*) OR
(Intervention AND "Dating Violence" AND Experiment*) OR
(Intervention AND "Dating Violence" AND Quasi-Experiment*) OR
(Intervention AND "Physical Violence" AND Experiment*) OR
(Intervention AND "Physical Violence" AND Quasi-Experiment*) OR
(Intervention AND "Dating Aggression" AND Experiment*) OR
(Intervention AND "Dating Aggression" AND Quasi-Experiment*) OR

(Program AND "Sexual violence" AND Experiment*) OR
(Program AND "Sexual violence" AND Quasi-Experiment*) OR
(Program AND "Sexual coercion" AND Experiment*) OR
(Program AND "Sexual coercion" AND Quasi-Experiment*) OR

(Program AND "Peer support" AND Quasi-Experiment*) OR
(Program AND "Intimate Partner violence" AND Experiment*) OR
(Program AND "Intimate Partner violence" AND Quasi-Experiment*) OR
(Program AND "Bystander" AND Experiment*) OR
(Program AND "Bystander" AND Quasi-Experiment*) OR
(Program AND "Dating Violence" AND Experiment*) OR
(Program AND "Dating Violence" AND Quasi-Experiment*) OR
(Program AND "Physical Violence" AND Experiment*) OR
(Program AND "Physical Violence" AND Quasi-Experiment*) OR
 (Program AND "Dating Aggression" AND Experiment*) OR
(Program AND "Dating Aggression" AND Quasi-Experiment*)

**Appendix B. Title and Abstract Screening Procedure**

1. Does the Title indicate that an intervention was conducted?
    a. Yes
    b. No

2. Does the Title indicate that sexual violence (or related outcome) was studied?
    a. Yes
    b. No

If Yes to BOTH, Continue.  If No to BOTH, STOP.  If No to only 1, Continue.

3. Does the Abstract indicate that a Two-Group design (experimental OR quasi) was utilized?
    a. Yes
    b. No

4. Does the Abstract indicate that this study utilized quantitative analysis?
    a. Yes
    b. No
    c. Unclear

5. Does the Abstract indicate that this study measured an included outcome?
    a. Yes
    b. No
    c. Unclear

6. Decision: Should this article be included?
    a. Yes, include for full article scan
    b. Unclear but Yes
    c. Unclear but NO (include reason)
    d. No, this article should be eliminated

**Appendix C. Full Article Screening**

1a. Does the article study and measure one of the related outcomes?
　　1. Yes
　　2. No
　　3. Unclear

1b. If yes, provide information (page numbers) supporting finding.

2. Does the article sample participants outside the 4-12 grade age range?
　　1. Yes
　　2. No (this is the answer we are hoping for)
　　3. Unclear

2b. If no, provide information (page numbers) supporting the finding.

3. Does the article utilize a Two-Group (at least) design?
　　a. Yes
　　b. No
　　c. Unclear

3b. If yes, provide information (page numbers) supporting finding.

4. Does the article present quantitative information?
　　a. Yes
　　b. No
　　c. Unclear

4b. What quantitative information does it provide?

5. Decision: Should the article be included for review?
　　a. Yes (all items answered Yes)
　　b. Yes but unclear--- send for further discussion (at least 1 of the items answered No)
　　c. No but may need further analysis (item 2 or 3 answered No)
　　d. No (item 1 answered No or more than 1 answered No)

**Appendix D. Codebook**

***Section A — Report Characteristics***

A01. What is the first author's name (APA format)?

A02. Number of authors.

A03. What was the year of publication?

A04. What type of report is this?
    1- Journal Article
    2- Book or book chapter
    3- Dissertation
    4- MA Thesis
    5- Private report
    6- Government report
    7- Conference paper
    8- Other (specify in answer)
    9- Unclear

A05.  Is this a peer-reviewed document?
    0- Not peer reviewed
    1- Peer reviewed
    9- Unclear

A06.  Was this research funded?
    0- No
    1- Yes
    9- Unclear

### *Section B — Setting Characteristics*

B01. What country was the program implemented?

    1- United States (specify state)

    2- Australia

    3- European Nation (specify)

    4- Other (specify)


B02. What type of school was used for the study?

    1- Public

    2- Private

    3- Mixture

    9- Unclear


*Treatment group*

B03.  What unit of analysis was defined as the treatment group?

    1- School

    2- Classroom

    3- Individual

    9- Unclear


B04.  What area was the unit of analysis located?

    1- Rural

    2- Urban

    3- Suburban

    4- Author stated "mixture"

    9- Unclear


B05.  What individual types were represented among the settings?

    (Place a 1 next to each applicable column, 0 to all others, 9 if not reported).

    1. Regular education

    2. Special education

    3. Free-reduced lunch

    4. Other (specify)

    5.  Author stated mixture of students or school-wide program

    6. No individual types given


*Control group*

B06.  What unit of analysis was defined as the control group?

    1- School

    2- Classroom

    3- Individual

9- Unclear

B07. What area was the unit of analysis located?

1- Rural

2- Urban

3- Suburban

4- Author stated "mixture

9- Unclear

B08. What individual types were represented among the settings?

(Place a 1 next to each applicable column, 0 to all others, 9 if not reported).

1. Regular education

2. Special education

3. Free-reduced lunch

4. Other (specify)

5. Author stated mixture of students or school-wide program

6. No individual types given

### Section C — Sample & Program Characteristics

*Treatment Sample & Program Characteristics*

C01. What was the SES of the students in the sample?

(Please indicate, with a 1, all that apply).

1. Low SES
2. Low-middle SES
3. Middle SES
4. Middle-upper SES
5. Upper SES
6. Only labeled as "mixed"
9. Can't tell

C02. What were the grade levels of the students in the sample?

(Please indicate, with a 1, all that apply?)

a. K
b. 1
c. 2
d. 3
e. 4
f. 5
g. 6
h. 7
i. 8
j. 9
k. 10
l. 11
m. 12
n. Labelled as "elementary school"
o. Labelled as "middle school"
p. Labelled as "junior high school"
q. Labelled as "high school"
r. No grade level information provided

C03. What percentage of males was represented in the treatment sample?

C04. What is the name of the program implemented (9 if cannot tell)?

C05. Who implemented the program?

(Indicate all that apply by putting a 1 if implemented, 0 if did not, 9 if can't tell).

a. Primary investigator
b. Research team

c. Graduate students

d. School administration

e. Teachers

f. External staff

g. Other (specify)

C06. What aspects of the program were involved?
(brief description; if can't tell indicate with 9)

C07. How long did the intervention last (in months)?

C08. Was there evidence that the program was implemented with reasonable fidelity?
0- No the program was not implemented with fidelity
1- Yes the program was implemented with fidelity
9- Can't tell either way

*Control Group Sample*

C09. What type of control group was utilized?
0- Wait list (complete through C13)
1- Treatment as usual (complete through C13)
2- Another type of program (continue)
3- No control group (stop coding)

C11. What was the SES of the students in the sample?
(Please indicate, with a 1, all that apply).
1. Low SES
2. Low-middle SES
3. Middle SES
4. Middle-upper SES
5. Upper SES
6. Only labelled as "mixed"
9. Can't tell

C12. What were the grade levels of the students in the sample?
(Please indicate, with a 1, all that apply?)
a. K
b. 1
c. 2
d. 3
e. 4
f. 5
g. 6

h. 7

i. 8

j. 9

k. 10

l. 11

m. 12

n. Labelled as "elementary school"

o. Labelled as "middle school"

p. Labelled as "junior high school"

q. Labelled as "high school"

r. No grade level information provided

**C13. What percentage of males was represented in the control sample?**

### Section D — Risk of Bias (Design Aspects)

D01. Random sequence generation?
    1- Low risk
    2- High risk
    3- Unclear risk

D02. Allocation concealment?
    1- Low risk
    2- High risk
    3- Unclear risk

D03. Blinding of participants and personnel?
    1- Low risk
    2- High risk
    3- Unclear risk

D04. Blinding of outcome assessment?
    1- Low risk
    2- High risk
    3- Unclear risk

D05. Incomplete outcome data?
    1- Low risk
    2- High risk
    3- Unclear risk

D06. Selective reporting?
    1- Low risk
    2- High risk
    3- Unclear risk

D07. Other sources of bias?
    1- Low risk
    2- High risk
    3- Unclear risk

D08. How are the units of analysis assigned to conditions?
    1- Random
    2- Quasi-random
    3- Non-random
    9- Unclear

**Section E — Outcome measures**

E01. Effect size type

    1- Trt/ctrl Post-test comparison

    2- Trt/ctrl Follow-up comparison


E02. Did the author conduct pre-test equivalence?

    0- No

    1- Yes

    9- NA or did not report


E03. If yes, were the groups equivalent?

    0- No

    1- Yes

    9- NA or did not report


*Dependent measure (separate for each measure)*

E04. What does the outcome measure?

    1- Teen dating violence knowledge

    2- Teen dating violence attitude

    3- Sexual violence knowledge

    4- Sexual violence attitude

    5- Rape construct (awareness, myths, etc.)

    6- Bystander awareness

    7- Other (specify)


E05. How was the outcome measure reported?

    1- Self-report

    2- Teacher report

    3- Administrator report

    4- Other (specify)


E06. Describe the measure.

*Effect size information*

E07. Metric

    1- d

    2- OR

    3- Other (specify)


E08. Author reported effect size

E09. Does the study report means and standard deviations for each condition?

    0- No

    1- Yes

EA. Treatment group N

EB. Treatment group mean

EC. Treatment group SD

ED. Control group N

EE. Control group mean

EF. Control group SD

EG. If means and SD not reported, list what quantitative information is provided.