

The Effectiveness of College Dating Violence Prevention Programs: A Meta-Analysis

TRAUMA, VIOLENCE, & ABUSE
2023, Vol. 24(2) 684–701
© The Author(s) 2021



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/15248380211036058
journals.sagepub.com/home/tva



Jennifer S. Wong¹ , Jessica Bouchard¹, and Chelsey Lee¹

Abstract

Due in part to their involvement with social activities on campus, college students experience an increased risk of dating violence. Recent legislation such as the Campus SaVE Act (which requires U.S. colleges to offer training on sexual assault, domestic violence, stalking, and sexual harassment to all incoming students) has contributed to the increase in prevention programming offered across postsecondary campuses, as well as subsequent research examining the effectiveness of these prevention efforts. The current study provides a systematic review and meta-analysis of college dating violence prevention programs. A systematic search of 28 databases and numerous gray literature sources identified an initial 14,540 articles of which 315 were deemed potentially eligible for inclusion. Studies were selected if they (1) evaluated a college dating prevention program/campaign, (2) reported one of five outcomes (knowledge, attitudes, or bystander efficacy, intentions, or behavior), (3) had a minimum sample size of 20 in the treatment group, (4) used a pre/post and/or comparison group design, and (5) were published in English or French between January 2000 and October 2020. We calculated 53 effect sizes from 31 studies and conducted separate meta-analyses on various categories of outcome measures. Findings suggest that college dating violence prevention programs are effective at increasing knowledge and attitudes toward dating violence, as well as bystander skills, but are not effective at increasing bystander behaviors. Findings from moderator analyses suggest that several program components influence the strength of treatment effects. Implications for improving the effectiveness of college dating violence prevention programs are discussed.

Keywords

dating violence, domestic violence, intervention/treatment, domestic violence, date rape, sexual assault

Dating violence is a serious issue impacting approximately 20%–30% of college students in the United States (Brewer et al., 2018; Wood et al., 2020). Dating violence refers to any physical, psychological, emotional, or sexual abuse that occurs within a nonmarried dating relationship, including stalking (Centers for Disease Control [CDC], 2021). The high rates of victimization are concerning given the deleterious outcomes that can result from dating violence experiences, which include physical injury (Amar & Gennaro, 2005), increased risk for depression, anxiety, and post-traumatic stress disorder (PTSD; Eshelman & Levendosky, 2012), and poor academic performance (Brewer & Thomas, 2019).

The introduction of the Campus SaVE Act in 2013 in the United States mandated that all Title IX schools¹ implement prevention education programming for addressing sexual assault, domestic violence, dating violence, and stalking (RAINN, n.d.). This legislation contributed to the increase in prevention programming offered across postsecondary campuses, as well as subsequent research examining the effectiveness of these prevention efforts. Several meta-analytic reviews investigating sexual violence prevention programs exist (e.g., Jouriles et al., 2018; Katz & Moore, 2013; Kettrey

& Marx, 2019); however, no such reviews have specifically focused on dating violence prevention for college populations. The current study gathers and synthesizes evidence on the effectiveness of college dating violence prevention programs.

Prevalence of Dating Violence Among College Students

Dating violence is prevalent among college/university students. Results from a nationally representative sample of undergraduate students ($N = 85,071$) in the United States suggests that 20% of students have experienced at least one form of intimate partner violence (IPV), including emotional violence, physical violence, sexual violence, or stalking (Brewer et al., 2018).

¹ School of Criminology, Simon Fraser University, Burnaby, British Columbia, Canada

Corresponding Author:

Jennifer S. Wong, School of Criminology, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia, Canada V5A 1S6.
Email: jenwong@sfu.ca

With respect to physical violence, Gover et al. (2008) found that 29% of participants across two U.S. universities had perpetrated physical violence on a romantic partner, while 20% reported being victimized. Comparable rates of emotional violence exist; Gidycz et al. (2007) studied 425 college males and found that approximately 25% of participants had perpetrated moderate verbal aggression and 60% had engaged in severe verbal aggression toward their dating partners. Similarly, Gover et al.'s (2008) study found that approximately 54% of students had perpetrated psychological abuse within the previous 12 months, and 52% had experienced psychological abuse victimization. With regards to sexual violence, Bhoohibhoya et al. (2019) found that 35% of a sample of 361 undergraduate students had at least one experience of sexual victimization by their dating partner. Gidycz et al. (2007) found a substantially lower estimate of perpetration, with 17% of participants reporting some form of sexual aggression against a dating partner. Given the stigma associated with sexual violence, it is likely that participants underreported their behavior.

Risk Factors for College Dating Violence Victimization

Notably, college students are vulnerable to sexual and dating violence due to their involvement with social activities on campus such as parties, the "hookup culture," and initiation into the Greek system (Jozkowski & Wiersma-Mosley, 2017). The Greek system and associated parties in particular, foster an environment in which both sexual and dating violence are common, including sexually competitive and aggressive behaviors (Jozkowski & Wiersma-Mosley, 2017), with fraternity members more likely to engage in sexually aggressive behaviors compared to nonfraternity members (Loh et al., 2005). Men in fraternities are also more likely to be accepting of rape myths and traditional gender norms, endorse hostility toward women, and be perceived as less culpable than those outside the Greek system, thus creating an environment conducive to sexual and dating violence, and increasing the risk of such behaviors (Canan et al., 2018; Seabrook & Ward, 2019; Seabrook et al., 2018).

Substance use among college students is one of the most common correlates of college dating violence experiences (e.g., Baker & Stith, 2008; Shorey et al., 2011); research by Shorey et al. (2015) found that alcohol use was significantly related to all forms of aggression measured (physical, psychological, and sexual). As well, Shorey et al. (2011) found that while research surrounding illegal drug use and dating violence victimization is limited, the existing literature suggests that illegal drug use is associated with increased risk of physical and psychological victimization.

Consequences of Dating Violence in College

Many short- and long-term negative consequences can result from the experience of dating violence. For example, victimization can result in physical injury (e.g., sprains, bruises,

black eyes; Amar & Gennaro, 2005), as well as long-term health consequences such as increased risk of developing chronic neck or back pain, frequent headaches, and sexually transmitted infections (Coker, 2000). Dating violence has also been linked to considerable psychological impacts, such as increased risk for depression, anxiety, and PTSD (Choi et al., 2017).

In addition, college dating violence can have negative academic implications. Brewer et al. (2018) found that students who had experienced either stalking or emotional, physical, or sexual abuse from an intimate partner had an increased risk of poor academic performance. This is consistent with findings from Brewer and Thomas (2019), who report that students perceive their abuse and subsequent physical or mental health symptoms as negatively impacting their academic success. Wood et al. (2020) also found that physical, psychological, sexual, and cyber abuse are significant predictors of lower academic performance.

College Sexual and Dating Violence Prevention Programs

Prevention and intervention strategies addressing sexual and dating violence on college campuses typically take one of two approaches: (1) general awareness/education and (2) bystander education. General awareness/education programs focus on increasing knowledge and changing attitudes by teaching students about the forms of IPV, healthy relationship strategies and skills, and harmful stereotypes regarding gender roles (DeGue et al., 2014). Another common approach is bystander education, which focuses on encouraging students to intervene when they witness signs of sexual or dating violence; bystander programs teach participants warning signs to look for as well as safe, appropriate ways to intervene in instances of violence or the perpetuation of harmful attitudes (Katz & Moore, 2013; Kettrey & Marx, 2019; 2020).

Though individual programs vary in delivery style and content, both approaches typically follow similar formats such as small groups and classroom workshops, or large group presentation formats. Programs usually involve interactive activities such as group discussions, role play scenarios, worksheets, or interactive videos (e.g., Amar et al., 2015; Fenton & Mott, 2018). These programs may occur as a single session or over multiple sessions, and are often delivered by specific program facilitators which include program staff members and student volunteers who receive instructional training (e.g., Amar et al., 2015). The large presentation style programs generally occur as a single session, and are more didactic in style with fewer opportunities for participant interaction (e.g., Borsky et al., 2018; Reid et al., 2013). Some programs operate in part or wholly online.

Overlap of Dating Violence and Sexual Violence Prevention Efforts for College Students

One noteworthy characteristic of prevention efforts at the college level is the overlap between dating violence and sexual

violence programs. Dating violence is a type of IPV that encompasses acts of physical, psychological, and sexual violence, as well as stalking, which occur within the context of a dating relationship (CDC, 2020). With respect to sexual violence, acts can be committed within the context of a dating or romantic relationship, but can also occur in other contexts (e.g., stranger, acquaintance, friend). As such, although sexual violence can occur within a dating relationship, dating violence and sexual violence can also occur separately. Given this overlap, it is reasonable for both dating violence and nonrelationship sexual violence to be addressed in a single program; while it is not necessary for the topics to be addressed by separate programs, the overlap does present a challenge when attempting to examine dating violence prevention program effectiveness. For example, although most dating violence prevention programs include some focus on sexual violence, not all sexual violence prevention programs address the broader concept of dating violence. It is not always clear in program descriptions if a program addresses dating violence *and* sexual violence, or just sexual violence. Similarly, as sexual violence can often be a component of dating violence, measures often include items addressing both constructs.

Existing Summative Research on Prevention Programs

While the effectiveness of sexual violence prevention programs has been frequently studied at the college level, dating violence has been substantially less studied. More specifically, we identified 10 systematic reviews or meta-analyses focused on college sexual violence and another 10 focused on bystander prevention approaches to both sexual violence and dating violence (e.g., recent meta-analyses include Jouriles et al., 2018; Katz & Moore, 2013; Kettrey & Marx, 2019; and Kettrey & Marx, 2020). Several recent systematic reviews of bystander prevention programs were also identified (e.g., Evans et al., 2019; Mujal et al., 2019; Storer et al., 2016). Notably, none of these studies examined dating violence prevention programs alone, but rather did so in combination with sexual violence prevention programs, and with an emphasis on the sexual violence related outcomes. See Appendix A for a summary of the existing reviews and meta-analyses in the field.

Three additional studies were identified that intentionally included dating violence prevention programs; however, these were not exclusively focused on college programs. More specifically, Carlos et al. (2017) systematically reviewed school based IPV prevention programs for teens and young adults. Twelve studies were included, which addressed various outcomes related to IPV including bystander intentions and behaviors, knowledge, attitudes, violence perpetration, and violence victimization. Results showed generally positive results immediately after the interventions, but these were not maintained in the long term. In a similar study, Crooks et al. (2019) reviewed IPV and sexual violence prevention programs for adolescents and young adults. The review found that existing programs have mixed effectiveness overall, though many programs had at least one positive effect. Last, Fellmeth et al. (2015) meta-

analyzed 33 adolescent and college dating violence prevention programs. Outcomes included knowledge, attitudes, behaviors, and skills, with no significant effects found for any of the categories.

Study Aim

Dating violence has been extensively studied among adolescent populations (e.g., de la Rue et al., 2017; Edwards & Hinsz, 2014; Lee & Wong, 2020), but has received much less focus among studies of college students. As well, numerous meta-analyses of college-based sexual violence prevention programs have been conducted; however, to date no meta-analysis has exclusively examined dating violence prevention programs targeting a college population. Given the pervasiveness of dating violence among college students and the vast negative consequences associated with these experiences, there is a pressing need to address, prevent, reduce, and ultimately eliminate such violence. The purpose of the current study is to examine the impacts of college dating violence prevention programs using systematic review and meta-analysis. Specifically, we investigate the effectiveness of dating violence prevention programs at improving college students' knowledge of and attitudes toward dating and sexual violence, and increasing bystander efficacy, intentions, and behaviors for intervening in dating and sexual violence situations.

Method

Systematic Literature Review

Four key constructs were used to develop a comprehensive search strategy (university/college, relationship violence, prevention program, and evaluation). To ensure that the list of key terms was exhaustive, the literature was consulted to find interchangeable terms (see Kugley et al., 2017). Boolean operators and wildcard markers were used to broaden the search, and the sets of search terms for each construct were developed over multiple trial-and-error iterations. The final search strategy is presented in Table 1.

The search protocol was applied to 28 databases (e.g., PsycINFO, Web of Science), and key terms were searched in the *abstract* and *title* fields. Additional search methods were also used, sources included Google and Google Scholar, the curriculum vitae of 19 key authors in the field, 17 journals that often publish articles on IPV (e.g., *Journal of Interpersonal Violence*), and 18 websites of governmental agencies and organizations that are associated with gender-based violence and/or dating/relationship violence research (e.g., *End Violence Against Women International*; *National Center on Domestic and Sexual Violence*). When an advanced search function was not available on a website, the first 100 hits for selected key words (e.g., "dating violence" AND college") were reviewed. We also searched the reference lists of included studies and relevant literature reviews and meta-analyses (i.e., backward searching) to identify any other relevant literature that was not uncovered by electronic database searches or the other search

Table 1. Systematic Search Terms for College Dating Violence Prevention Programs.

| Construct | Key Terms |
|-------------------------------|---|
| University/college | (university or college or “higher education” or campus or undergrad*) |
| Relationship violence | (“domestic violence” or “intimate partner violence” or “partner violence” or IPV or “partner abus*” or “dating violence” or “relationship violence” or “dating abuse” or “dating aggression” or “relationship aggression” or “couple violence” or “gender-based violence” or “gendered violence” or “violence against women” or “bystander*” or “bringing in the bystander” or BITB or “Green Dot” or “peer prevention” or “Step Up” or “Mentors in Violence Prevention” or MVP or “Know Your Power” or Hollaback or “Circle of 6” or “That’s Not Cool” or “Red Flag Campaign” or “Where Do You Stand” or “White Ribbon Campaign” or “Men Can Stop Rape” or “The Men’s Program” or “The Women’s Program” or “The Men’s Project” or “Coaching Boys into Men” or “Campus Violence Prevention Program” or “Real Men Respect” or “Speak Up Speak Out” or “sexual* violen*” or “sexual* aggressi*” or “sexual* abus*” or “sexual assault” or rape or “date rape” or “acquaintance rape”) |
| Prevention program Evaluation | (strateg* or interven* or prevent* or program* or treatment* or campaign* or training or education or online or poster) (eval* or impact* or outcome* or assess* or effect* or efficacy) |

methods. See Appendix B for the complete list of databases and other sources, including gray literature. The search was implemented by three reviewers (the study authors) between September 27, 2020, and October 5, 2020. All electronic databases were searched for articles that were published between January 1, 2000, and September 27, 2020.

Selection criteria. To be considered eligible for inclusion in the meta-analysis, studies were required to meet the following a priori inclusion criteria: (1) evaluates a program implemented in a college setting that targets the prevention of dating violence; (2) reports on one or more of the following outcomes: knowledge about dating violence, attitudes toward dating violence, bystander efficacy with respect to intervening, bystander intentions to intervene, or bystander behaviors in the context of relationship abuse or sexual violence (more details on this below); (3) provides sufficient data to allow for the calculation of an effect size; (4) has a minimum sample size of 20 subjects in the treatment group; (5) uses a pretest/posttest or a comparison group design; and (6) was published in English or French² between January 1, 2000, and September 27, 2020.³ Additionally, as studies that evaluate college-based interventions in which sexual violence is the primary focus have been well-reviewed in the academic literature (e.g., see Evans et al., 2019; Jouriles et al., 2018; Kettrey & Marx, 2020), studies that focused primarily on sexual violence were excluded when it was not clear that relationship violence was also included as a component in the program.

With respect to (2), upon close examination of the outcome measures and instruments used in the set of program evaluations in the current sample, it became clear that limiting outcomes to those pertaining only to dating violence would not be possible. There are two reasons for this; the first is conceptual and reflects that sexual violence falls within the general definition of IPV. In other words, sexual violence is one type of relationship violence, meaning that sexual violence occurring within college samples may in many cases be occurring between dating partners. The second reason is technical, in that the instruments used in many studies combined outcomes of dating violence and sexual violence into a single measurement

scale; outcomes were included only if dating violence was included as a portion of the scale items. See Appendix C-3 for more details on the measures used in each study.

Data collection. Following completion of the search, a merged list of potentially eligible studies was compiled. Three reviewers shared the task of independently screening the initially selected citations/abstracts to determine which articles should be retrieved for detailed review (with a yes/no decision); the database was split such that each citation was screened by two reviewers. Once the set of articles was retrieved, the reviewers independently applied the selection criteria to determine the set of studies for exhaustive coding and effect size calculation (again, each article was assessed by two reviewers), and a consensus on inclusion was reached.

Analytic Approach

Data were coded in Microsoft Excel for each study on a series of 77 variables.⁴ Two reviewers shared the task of independently extracting data from the selected studies; the third reviewer validated the coding of all studies for accuracy and any discrepancies in coding were discussed until a consensus was reached. Following data extraction, the effectiveness of college dating violence prevention programs was calculated using effect sizes.

We calculated 53 effect sizes based on five types of reported outcome data: (1) pretest and posttest means with standard deviations for a treatment and control group ($n = 29$); (2) posttest means and standard deviations for a treatment and control group ($n = 1$); (3) pretest and posttest percentage of participants who responded yes/no to the outcome for a treatment and control group ($n = 2$); (4) F test with unequal posttest sample sizes for a treatment and control group ($n = 4$); (5) pretest and posttest means, standard deviations, and either an F test or t test for a single treatment group ($n = 17$). Formulas are available upon request.

Notably, the current research included effects for studies that used two-group designs as well as single-group designs. Combining effect sizes from two-group and single-group

designs is a controversial issue in the field (e.g., see Carlson & Schmidt, 1999; Cuijpers et al., 2017). However, as evidenced by the results of the current systematic review, single-group research designs are common within the university setting, and, in many cases, are the most feasible approach and/or the only option through which to conduct research. By excluding these designs in meta-analyses, a considerable portion of the existing research would not be considered, providing only a partial summary of research and an incomplete overview of the treatment effect of interest (Card, 2011; Lee & Wong, 2020). Some authors argue that with certain statistical adjustments, effect sizes from these designs can be appropriately pooled (Borenstein & Hedges, 2019). In the current analysis, studies using a single-group design were included in the analysis after being transformed into a raw score metric (see Morris & DeShon, 2002)).

In addition, four of the study designs did not involve assigning individuals to the treatment and comparison groups; instead, students were assigned based on preexisting groupings such as fraternities/sororities, university classrooms, or entire universities. For these cases, we cluster-adjusted the effect sizes and standard errors to account for the nesting (Hedges, 2007), using intracluster correlation coefficients based on existing literature⁵ (empirically derived ICCs were not provided). Of the 53 effect sizes, 10 required cluster adjustments.

Further, to ensure independence of effect sizes both between and within studies (Card, 2011), we: (1) used only nonoverlapping samples (i.e., only one report of the same sample was used if it was included in multiple publications); (2) included multiple studies within single reports only if the samples were completely independent (i.e., if a single document reported on more than one experiment, using completely independent samples, each was coded as an independent study); and (3) did not double count control groups (i.e., if a study included two treatment groups and both were compared to the same control group, only one treatment group (the most commensurate to other programs across the set of included studies) was selected for effect size calculation).

Data synthesis. DerSimonian and Laird random effects models were chosen for the primary analyses; these models assume that between-study variability is due to factors other than random sampling error (Lipsey & Wilson, 2001). Heterogeneity was assessed using Q statistics and I^2 statistics. Potential publication bias (Sterne & Harbord, 2004) and small study effects (Egger et al., 2003) were assessed using funnel plots and Egger's test for funnel plot asymmetry. Influence analyses were also implemented for each of the models to assess whether any individual study had an extreme impact on the pooled effect (Tobias, 1999).

Heterogeneity. Two methods were used to examine heterogeneity. First, through an examination of the data collection instruments used in each study we considered whether the outcome measures being pooled were measuring conceptually similar concepts; previous research suggests that when meta-

analytic outcomes are disaggregated from larger pooled sets, results point to intervention effectiveness with respect to some outcomes but not others (e.g., Kettrey & Marx, 2019). Given this, we disaggregated our pooled outcome measures into five smaller, more commensurate sets. Second, we investigated between-study heterogeneity of effects through subgroup analysis in an approach which tests whether a categorical variable can explain some of the variability in effect sizes (Card, 2011; Lipsey & Wilson, 2001). Using this method, 12 dichotomous variables representing study, sample, and program characteristics were used to investigate heterogeneity in the pooled effect sizes for the knowledge/attitudes and the bystander efficacy/intentions outcomes. While we initially coded/attempted to code for a large number of variables across all studies that could potentially be useful descriptors of study, sample, and program characteristics, the selection of moderator variables for analytic purposes was more limited. This selection depended on whether the outcome had sufficient observations (i.e., that coding for a given program component was not missing large amounts of data) and whether there was sufficient variability between observations to conduct meaningful comparisons (we restricted observations to a minimum of four per subgroup). Further, if variables were strongly and significantly correlated, only one was selected.

Moderator variables included: (1) Research design (nonrandomized vs. randomized), (2) Treatment group sample size (participants vs. participants), (3) Sample gender (female vs. female), (4) Sample ethnicity (white vs. white), (5) Definitions of types of violence (yes/no), (6) Prevalence of violence (yes/no), (7), Impacts of violence (yes/no), (8) Definition of consent (yes/no), (9) Video component (yes/no), (10) Didactic delivery (yes/no), (11) Online exclusive delivery (yes/no), and (12) Role play (yes/no).

Results

Systematic Search Results

The search of the 28 databases resulted in an initial 11,532 studies for review. An additional 3,028 studies were identified through gray literature sources. Following data extraction, 29 studies (contributing 31 independent samples and 53 independent effect sizes) were deemed relevant for final inclusion based on the set of a priori inclusion criteria. Figure 1 presents the specific number of hits selected in each successive step of the search strategy.

Characteristics of included studies. Next, we provide an overview of the 31 independent samples included in the current study; Appendix C provides detailed descriptives of the 31 studies.⁶ As shown in Table 2, the large majority of the included studies are recent publications; 55% were published between 2017 and October 2020. Most of the studies were found in peer-reviewed journals (84%), almost all of the programs were implemented in the United States (94%), and all but one of the included studies employed a pretest posttest design (97%). Nearly 42% of the studies used a randomized control design ($n = 13$), 16%

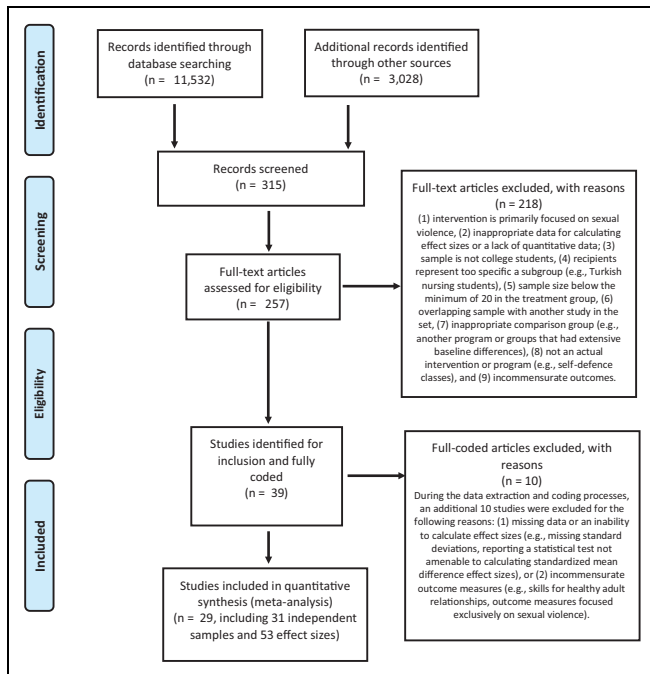


Figure 1. Systematic search results for dating violence prevention programs.

used a quasi-experimental design with a weakly matched comparison group (i.e., no use of propensity scores or other data-based techniques for matching; $n = 5$), and 42% used a single-group pre- and posttest design ($n = 13$). Correspondingly, 58% of the studies were rated as moderately rigorous on the Maryland Scale of Scientific Methods (SMS; Sherman et al., 1998)⁷ with scores of 2 or 3, and 42% were rated as strong on the SMS with a score of 4 or 5 ($n = 13$).

Posttest data were most often collected within two weeks of the end of the program (55%; $n = 17$). At pretest, treatment group sample size ranged from 28 to 2,444 participants ($M = 302.8$, $SD = 599.3$), and at posttest the sample size also ranged from 28 to 2,444 ($M = 281.9$, $SD = 582.0$). Females were more highly represented across study samples. In five studies the participants were at least 90% female (16%), and in 10 studies the proportion of females was 70%–89% (32%). With respect to ethnicity, 11 studies (36%) used a predominantly white sample (70%–89% white), nine studies (29%) had a mixed sample of 31%–69% white participants, and one sample (3%) was mostly composed of students identifying as ethnic minorities.

The programs were primarily delivered in a single session (77%) by program staff and/or university students (61%). While the duration of the interventions varied (0.3 to 12 hr; average duration 2.4 hr; $SD = 2.8$), they were typically less than 2 hr in length (68%). The most common program components focused on skill-building/development (81%) and bystander training (74%). Other common components included discussions of types of violent behaviors (42%), the prevalence of dating/sexual violence (39%), the definition of consent (39%), available campus resources (36%), and the consequences of dating/sexual violence (32%). With respect

Table 2. Study and Sample Characteristics.

| Study/Sample Characteristics | N (%) |
|---|------------------------------|
| Publication year | |
| January 2000 to December 2016 | 14 (45.2) |
| January 2017 to October 2020 | 17 (54.8) |
| Publication type | |
| Journal article | 26 (83.9) |
| Dissertation/thesis | 5 (16.1) |
| Geographic location | |
| United States | 29 (93.6) |
| Europe | 1 (3.2) |
| Mexico | 1 (3.2) |
| Research design | |
| Randomized control trial | 13 (41.9) |
| Quasi-experiment with weakly matched comparison group | 5 (16.1) |
| Single group pre- and posttest | 13 (41.9) |
| Number of groups | |
| Single-group design | 13 (41.9) |
| Two-group design | 18 (58.1) |
| Methodological rigor | |
| Moderate (Maryland Scale 2 or 3) | 18 (58.1) |
| Strong (Maryland Scale 4 or 5) | 13 (41.9) |
| Timing of posttest | |
| Immediately after program end (<2 weeks) | 17 (54.8) |
| 2–7 Weeks after program end | 7 (22.6) |
| 8–11 Weeks after program end | 3 (9.7) |
| 3+ Months after program end | 4 (12.9) |
| Treatment group sample size at pretest (range: 28–2,444) | $M = 302.8$ ($SD = 599.3$) |
| Treatment group sample size at posttest (range: 28–2,444) | $M = 281.9$ ($SD = 582.0$) |
| Attrition in analytic sample over 20% at posttest? | |
| No | 24 (77.4) |
| Yes | 7 (22.6) |
| Gender of participant sample | |
| All female (90%+) | 5 (16.1) |
| Mostly female (70%–89%) | 10 (32.3) |
| Mixed (31%–69% male/female) | 16 (51.6) |
| Race/ethnicity of participant sample | |
| Mostly white (70%–89%) | 11 (35.5) |
| Mixed (31%–69% white) | 9 (29.0) |
| Mostly minority (0%–30% white) | 1 (3.2) |
| Missing | 10 (32.3) |

Note. $N = 31$.

to program content delivery, interventions frequently used active participation (e.g., activities to practice skills; 77%) and/or group discussion activities (55%). Other common approaches included role play (48%), a didactic presentation style (39%), the use of video (39%), group activities (36%), and/or an online component (29%). See Table 3.

Meta-Analytic Results

Outcome Category 1: Knowledge/attitudes about dating violence. Figure 2 presents a forest plot of the effects of prevention

Table 3. Program Characteristics, Components, and Approach.

| Program Characteristics | N (%) |
|--|----------------|
| Type of program delivery | |
| In-person | 19 (61.3) |
| Online | 9 (29.0) |
| Campaign only (posters, ads, etc.) | 1 (3.2) |
| In-person + campaign | 2 (6.5) |
| Number of sessions (range: 1–8 sessions) | |
| 1 | 24 (77.4) |
| 2–8 | 4 (12.9) |
| Missing | 3 (9.7) |
| Number of hours (range: 0.3–12 hr) | |
| < 2 hr | 21 (67.7) |
| 2+ hr | 7 (22.6) |
| Missing | 3 (9.7) |
| M (SD) | 2.4 (SD = 2.8) |
| Type of program facilitator | |
| Program staff | 9 (29.0) |
| University students | 8 (25.8) |
| Staff + university students | 2 (6.5) |
| Online | 9 (29.0) |
| Staff + campaign | 1 (3.2) |
| Other | 2 (6.5) |
| Program components | Yes (%) |
| Incorporated focus on skill-building/development | 25 (80.7) |
| Incorporated bystander training | 23 (74.2) |
| Included definitions of violent behavior types | 13 (41.9) |
| Discussed prevalence of dating/sexual violence | 12 (38.7) |
| Discussed/defined consent | 12 (38.7) |
| Discussed available campus resources | 11 (35.5) |
| Discussed short- and/or long-term consequences of dating/sexual violence | 10 (32.3) |
| Addressed gender roles/inequities/stereotypes | 7 (22.6) |
| Included discussion of rape myths | 6 (19.4) |
| Discussed social norms in relation to IPV/sexual violence | 6 (19.4) |
| Discussed substance use in relation to IPV/sexual violence | 2 (6.5) |
| Program approach | Yes (%) |
| Incorporated any active participation component | 24 (77.4) |
| Incorporated discussion | 17 (54.8) |
| Incorporated role play | 15 (48.4) |
| Program included a film/video component | 12 (38.7) |
| Program used didactic presentation style | 12 (38.7) |
| Incorporated group activities | 11 (35.5) |
| Included an online component | 9 (29.0) |
| Program is mostly/exclusively delivered in video format | 7 (22.6) |
| Included written materials distributed to participants | 5 (16.1) |
| Included poster or campaign component | 3 (9.7) |

programs on the outcome of dating violence knowledge/attitudes. The pooled estimate of 0.418 ($n = 16$; 95% CI [0.263, 0.572]; $z = 5.29, p < .001$) is a statistically significant, positive result which suggests that prevention programs are effective at increasing participant knowledge about or attitudes against dating violence. The significant Q statistic of 527.09 and I^2 value of 97.2% suggest that there is heterogeneity within the sample

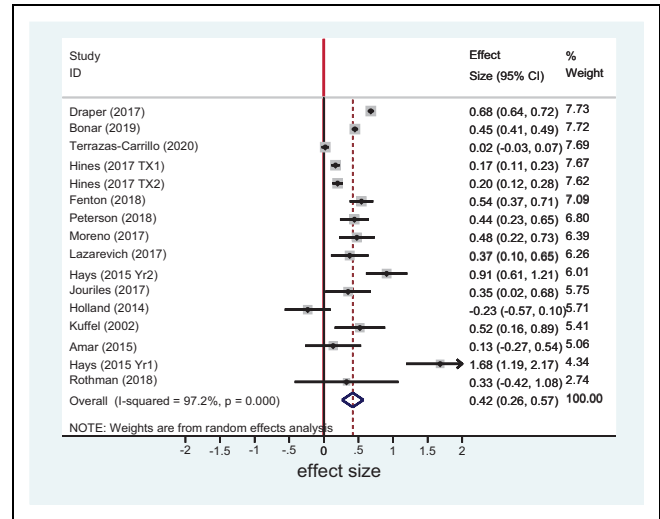


Figure 2. Forest plot for the meta-analysis on knowledge/attitudes about dating violence.

that can largely be attributed to factors beyond sampling error. Bias was assessed through a visual inspection of a funnel plot along with Egger’s test for funnel plot asymmetry (not shown here); Egger’s test of small study effects provides no indication of bias with a coefficient of -0.395 ($SE = 2.22, t = -0.18$). The results of the influence analysis (showing the resulting pooled estimates when each study is removed from the analysis one at a time) suggests that the overall effect is robust. Influence analysis tables are available in Appendix E.

Outcome Category 2: Bystander efficacy/intentions. The pooled effect of studies that measured outcomes of bystander efficacy and intentions was 0.484 ($n = 17$; 95% CI [0.276, 0.692], $z = 4.57, p < .001$). This result is positive and significant, indicating that prevention programs significantly and positively impact participants’ feelings of efficacy and their intentions to act as a bystander. The model had a statistically significant Q statistic of 170.15, with an I^2 statistic of 90.6%. The forest plot in Figure 3 shows that of the 17 included studies, eleven had statistically significant effects. Egger’s bias coefficient was not significant ($-2.003, SE = 1.37, t = -1.46$), and the influence analysis found a robust effect with no study so influential that its removal would change the overall pooled estimate to a nonsignificant finding.

Outcome Category 3: Bystander behaviors. Outcome measures of bystander behaviors yielded an effect size of 0.075 ($n = 11$; 95% CI [0.108, 0.258]; $z = 0.81, p = 0.42$). In contrast to the previous two findings, this result is not significant and suggests that, overall, prevention programs do not have an impact on bystander behaviors. The model resulted in a statistically significant Q statistic of 50.75, with an I^2 statistic of 80.3%. Egger’s bias coefficient was not significant ($1.385, SE = 1.75, t = 0.79$), and the influence analysis found a robust effect. The forest plot in Figure 4 shows that of the 11 studies, only three had statistically significant effects.

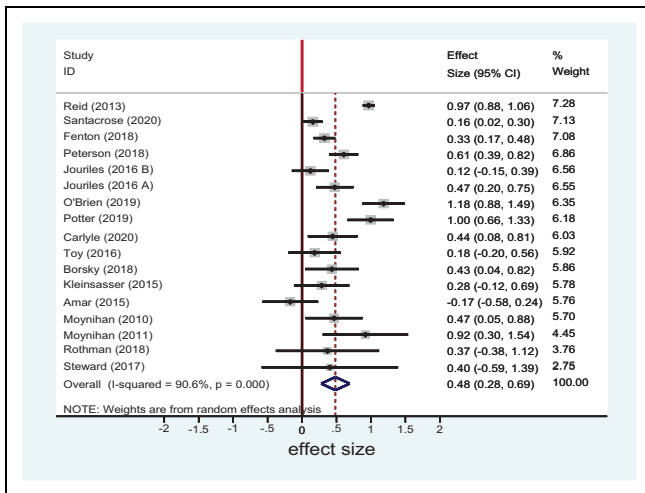


Figure 3. Forest plot for the meta-analysis on bystander intentions/efficacy.

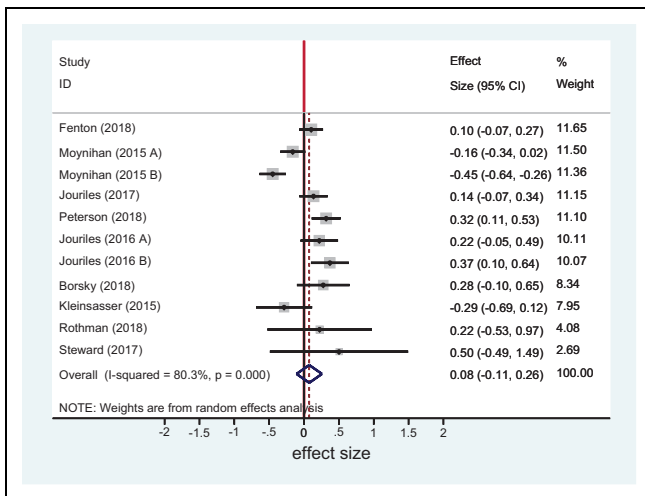


Figure 4. Forest plot for the meta-analysis on bystander behaviors.

Disaggregated Study Outcomes

Our first approach to addressing the heterogeneity present in the pooled effect sizes was to disaggregate the study outcomes into smaller sets with more commensurate outcomes combined. Forest plots are available in Appendix D; influence analysis tables are in Appendix E. An added benefit of disaggregating is that it enabled the inclusion of additional effect sizes overall; for example, if a study measured outcomes for both bystander efficacy and bystander intentions, it contributed one effect size to each grouping.

Knowledge about dating violence. The model of knowledge concerning dating violence yielded a positive, statistically significant pooled estimate of 0.513 ($n = 10$; 95% CI [0.339, 0.687]; $z = 5.79$, $p < .001$). All 10 studies are positive, and all but one effect size is statistically significant. This result suggests that

prevention programs are effective at increasing participants' knowledge about dating violence. The model resulted in a significant Q statistic of 284.08 and an I^2 value of 96.8%, suggesting there is still substantial heterogeneity within the sample.⁸ No publication bias, small study effects, or overly influential studies were noted.

Attitudes toward dating violence. Using a random effects model, we found a statistically significant, positive effect of 0.289 ($n = 8$; 95% CI [0.070, 0.509]; $z = 2.58$, $p < .05$) for attitudes toward dating violence. The Q statistic (65.28, $p < .001$) illustrates remaining heterogeneity across the set of eight studies, of which a majority can be attributed to factors other than sampling error ($I^2 = 89.3%$). Of the eight effect sizes, seven are positive and five have statistically significant results. Sensitivity analyses did not suggest any serious concerns of publication bias, small study effects, or influential studies.

Bystander efficacy. The overall effect for bystander efficacy was significant at 0.526 ($n = 13$; 95% CI [0.308, 0.744]; $z = 4.73$, $p < .01$), suggesting a positive impact of dating violence prevention programs on increasing participant feelings of efficacy to engage in bystander behaviors. The heterogeneity statistic was significant ($Q = 103.96$) and the I^2 value of 88.5% suggests that heterogeneity can be attributed to factors outside of sampling error. All of the effect sizes are positive, and nine of the 13 are statistically significant. No serious concerns of publication bias, small study effects, or influential studies were uncovered.

Bystander intentions. Last, the random effects model on bystander intentions to intervene resulted in a significant pooled estimate of 0.443 ($n = 11$; 95% CI [0.251, 0.634]; $z = 4.52$, $p < 0.001$). Substantial heterogeneity remained across the set of studies ($Q = 39.47$, $p < .001$; $I^2 = 74.7%$). Ten of the 11 effect sizes are positive, and eight of these effect sizes are statistically significant. The sensitivity analyses did not indicate any notable concerns.

Assessing Heterogeneity

Moderators of program impact. Our second approach to addressing the heterogeneity in the pooled effects for the knowledge/attitudes and the bystander efficacy/intentions outcomes was to examine moderators of program impact using 12 dichotomous variables representing study, sample, and program characteristics. Due to the small number of studies included in the “bystander behaviors” outcome (and subsequent small n s in variable subgroups), we did not conduct moderator analyses on the “bystander behaviors” outcome. For a summary of all moderator analyses, see Table 4.

Knowledge/attitudes about dating violence. Eleven variables were used to investigate heterogeneity in the set of 16 effect sizes representing program effects on knowledge/attitudes to dating violence (due to missing data across studies, not all moderator variables had 16 effect sizes). Eight variables were significant moderators of treatment impact. Specifically,

Table 4. Summary of Moderator Analyses.

| Type | Moderator | Variables | Knowledge/Attitudes | Bystander Efficacy/Intentions |
|---------------------------|-------------------------|---|---|---|
| Study | Research design | Non-RCT RCT | <i>Marginally significant</i> ES = .398, z = 35.46*** ES = .260, z = 3.27** | ES = .719, z = 20.75*** ES = .398, z = 8.24*** |
| | Tx group sample size | <100 100+ | ES = .094, z = 3.63*** ES = .464, z = 37.68*** | <i>Not significant</i> ES = .591, z = 8.15*** ES = .613, z = 20.07*** |
| Sample | Sample gender | 70%+ female <70% female | <i>Marginally significant</i> ES = .488, z = 8.93*** ES = .391, z = 34.47*** | ES = .745, z = 21.66*** ES = .335, z = 6.84*** |
| | Sample ethnicity | 70%+ white <70% white | ES = .328, z = 20.22*** ES = .071, z = 2.79** | ES = .844, z = 21.73*** ES = .371, z = 7.60*** |
| Program components | Definitions of violence | No | ES = .389, z = 20.69*** | ES = .259, z = 5.22*** |
| | | Yes | ES = .455, z = 29.6*** | ES = .770, z = 22.00*** |
| | Prevalence of violence | No | ES = .396, z = 21.29*** | ES = .371, z = 7.16*** |
| | | Yes | ES = .236, z = 8.64*** | ES = .712, z = 21.17*** |
| Impact of violence | No | ES = .381, z = 20.39*** | ES = .655, z = 20.58*** | |
| | Yes | ES = .279, z = 10.53*** | ES = .448, z = 7.37*** | |
| Definition of consent | No | ES = .191, z = 8.10*** | ES = .705, z = 21.05*** | |
| | Yes | ES = .570, z = 39.96*** | ES = .381, z = 7.32*** | |
| Program approach | Video component | No | ES = .217, z = 9.12*** | ES = .765, z = 21.33*** |
| | | Yes | ES = .444, z = 28.51*** | ES = .353, z = 7.25*** |
| | Online exclusive | No | N/A | ES = .723, z = 21.75*** |
| | | Yes | | ES = .323, z = 6.10*** |
| Didactic delivery | No | ES = .448, z = 28.32*** | ES = .325, z = 6.65*** | |
| | Yes | ES = .344, z = 22.03*** | ES = .771, z = 21.99*** | |
| Role play component | No | <i>Not significant</i> | ES = .677, z = 20.37*** | |
| | Yes | ES = .319, z = 4.68*** ES = .397, z = 35.28*** | ES = .439, z = 8.27*** | |

Note. N/A = not examined.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Bold font = ES is significantly larger

studies using larger sample sizes reported a significantly larger pooled effect than did studies using smaller samples ($ES = .464$ vs. $ES = .094$), and programs that served predominantly white participants (i.e., 70%+ White) had a larger pooled effect than did those serving samples with higher rates of ethnic minorities ($ES = .328$ vs. $ES = .071$). Additionally, all four program content variables were statistically significant moderators of treatment impact. Significantly larger pooled effect sizes were found for programs that included a focus on definitions of violence ($ES = .455$ vs. $ES = .389$), the definition of consent ($ES = .570$ vs. $ES = .191$), and did *not* focus on the prevalence of violence ($ES = .396$ vs. $ES = .236$) or the impact of violence ($ES = .381$ vs. $ES = .279$). With respect to program approach, programs with a video component ($ES = .444$ vs. $ES = .217$), and those *not* delivered using didactic methods ($ES = .448$ vs. $ES = .344$), were more impactful.

Bystander efficacy/intentions. Twelve⁹ moderator variables were examined for bystander efficacy/intentions; 11 were statistically significant moderators of treatment impact. Specifically, studies using a nonrandomized design showed a significantly larger pooled treatment effect ($ES = .719$ vs. $ES = .398$) than did studies using randomized designs, studies using samples of mostly white participants had larger effects

($ES = .844$ vs. $ES = .371$), and studies using predominantly female samples also had larger effects on bystander efficacy/intentions ($ES = .745$ vs. $ES = .335$). In addition, programs that included definitions of violence ($ES = .770$ vs. $ES = .259$) and discussed the prevalence of violence were more impactful ($ES = .712$ vs. $ES = .371$); conversely, pooled effects were larger for programs that did *not* include content on the impacts of violence on victims ($ES = .655$ vs. $ES = .448$), and did *not* focus on the definition of consent ($ES = .705$ vs. $ES = .381$). Last, programs that used a didactic teaching method were more effective at improving bystander efficacy/intentions to intervene ($ES = .771$ vs. $ES = .325$). On the contrary, programs that did not incorporate video ($ES = .765$ vs. $ES = .353$), did not use role play ($ES = .677$ vs. $ES = .439$), and did not use an exclusively online format ($ES = .723$ vs. $ES = .323$) had larger effects than programs that did.

Discussion

The present study examined the effectiveness of college dating violence prevention programs at increasing knowledge and improving attitudes toward dating violence, increasing bystander efficacy and intentions to intervene in dating violence

Table 5. Summary of Critical Findings.

| Aggregated Outcomes | | | | |
|-------------------------------|---------------------|-----------------|-----------------|---|
| Outcome category | No. of Effect Sizes | Pooled Estimate | Z (p Value) | Q Statistic (p Value) I ² Statistic |
| Knowledge/attitudes | 16 | .418 | 5.29 (p < .001) | Q = 527.09 (p < .001) I ² = 97.2% |
| Bystander efficacy/intentions | 17 | .484 | 4.57 (p < .001) | Q = 170.15 (p < .001) I ² = 90.6% |
| Disaggregated Outcomes | | | | |
| Outcome Category | No. of Effect Sizes | Pooled Estimate | Z (p Value) | Q Statistic (p Value) I ² Statistic |
| Knowledge | 10 | .513 | 5.79 (p < .001) | Q = 284.08 (p < .001) I ² = 96.8% |
| Attitudes | 8 | .289 | 2.58 (p < .05) | Q = 65.28 (p < .001) I ² = 89.3% |
| Bystander efficacy | 13 | .526 | 4.73 (p < .01) | Q = 103.96 (p < .001) I ² = 88.5% |
| Bystander intentions | 17 | .443 | 4.52 (p < .001) | Q = 39.47 (p < .001) I ² = 74.7% |
| Bystander behaviors | 11 | .075 | 0.81 (p = 0.42) | Q = 50.75 (p < .001) I ² = 80.3 |

situations, and increasing the prevalence of bystander behaviors. The systematic search identified 14,560 reports for review, resulting in the retrieval of 257 articles. After applying inclusion and exclusion criteria, 31 studies producing 53 independent effect sizes were included in the analysis. The interventions were most commonly short, single-session programs, with samples consisting primarily of white women, and outcomes were evaluated at immediate posttest.

Overall, our findings suggest that dating violence prevention programs are effective at improving elements of dating violence; Table 5 provides a summary of the critical findings. The positive and statistically significant effect for knowledge/attitudes is encouraging as it suggests that prevention programs are an effective approach for knowledge translation with respect to dating and sexual violence. Feminist theoretical perspectives suggest that dating and sexual violence behaviors can only be addressed if broader community norms and attitudes are also modified (Banyard et al., 2004); this is consistent with research, particularly within adolescent samples, that indicates acceptance and normalization of dating violence is a significant predictor of dating violence perpetration (Miller et al., 2020; Mumford et al., 2020). As such, changes in attitudes toward violence may have an indirect effect on changing abusive behaviors and may be an important step in eliminating this violence. With respect to bystander efficacy and intentions, the positive effect is also encouraging as bystanders play an important role in intervening in dating violence situations. That is, without having the skills and confidence to successfully intervene, bystanders may be unable to provide helpful assistance to victims in an emergency or learn how to prevent a potentially violent situation.

Our findings also suggest that college dating violence prevention programs are not effective at increasing bystander behaviors. This is consistent with research on adolescent dating violence, which indicates that existing prevention programs are effective at positively influencing knowledge and attitudes, but are less effective when it comes to modifying behaviors (e.g., Lee & Wong, 2020). This finding is not altogether surprising, as affecting changes in behaviors can be a more complex process than changing knowledge or attitudes (Gray et al., 2017). The empirical evidence discussed herein suggests that dating violence prevention programs may need to incorporate new techniques to effectively alter bystander behaviors. While the set of studies was too small to allow for moderator analysis ($n = 11$), based on the results for bystander efficacy/intentions it is likely that online program delivery is not the most effective approach to changing behaviors. Future dating violence prevention programs should consider incorporating evidence-based techniques for behavioral change, such as presenting information in a way that empowers students to want to problem solve (Katz & Moore, 2013). By framing the information on bystander behaviors as an effective and tangible solution to a problem, students may be more receptive and more likely to implement actions in their daily lives.

Heterogeneity

Disaggregated outcomes. The findings from the disaggregated outcomes provide further insight into the effectiveness of dating violence prevention programs at achieving a variety of participant outcomes. More specifically, the disaggregated

outcome groupings indicate a larger pooled effect of programs on knowledge than on attitudes toward dating violence ($ES = .513$ vs. $ES = .289$). The difference between bystander efficacy and intentions was less noteworthy ($ES = .526$ vs. $ES = .443$), but it is clear that the effects of programs on both of these outcomes were substantially larger than the treatment effect on bystander behaviors ($ES = .075$).

Subgroup analysis. Our findings point to several considerations in the design of prevention programs or revisions to existing programs, discussed in more detail below.

Study characteristics. It is well known that less rigorous research designs are more likely to overestimate intervention effects (e.g., Lipsey & Wilson, 2001). Consistent with the existing evidence, our findings suggest that studies using a nonrandomized design had a significantly larger impact on bystander efficacy/intentions and a marginally significantly larger impact on knowledge/attitudes. Further, given that strong impacts on knowledge/attitudes were observed for larger samples of participants, small study effects are unlikely to be driving the overall pooled estimate.

Sample characteristics. Our subgroup analyses suggest that the gender distribution of the sample was an important moderator of treatment effects. Specifically, studies that used samples of mostly female participants produced significantly larger pooled effects on knowledge/attitudes and bystander efficacy/intentions. This finding is worth additional consideration, as it may suggest that dating violence prevention programs are not equally effective for different gender groups. This finding is supported by some prior research which indicates that bystander programs produce differential (generally, larger) impacts for female participants than they do for male participants (e.g., Banyard et al., 2009; Borsky et al., 2018; Coker et al., 2017; Hines et al., 2019). While we are limited by the data in our ability to draw conclusions about the effects of programs for female versus male participants, future research should directly examine this issue. It may be that content could be altered to better target males, perhaps through using single gender treatment groups. Research is limited within the college population, but suggests that audience gender composition may have important impacts on effectiveness, with women showing greater effects in mixed-gender groups, while men show greater impacts within single-gender groups (though impacts overall are mixed; see Anderson & Whiston, 2005; Graham et al., 2019). Similar findings have been shown among adolescent samples with boys appearing to benefit more from single-gender groups, suggesting that audience composition for interventions should be considered (Black et al., 2012). Given the sensitive nature of the material, male participants may feel criticized or blamed and may react with hostility, impacting their receptiveness to the material (Malamuth et al., 2018; Spikes and Sternadori, 2018). Single-gender groups would allow for more targeted presentation of the material to cater to the needs of the specific population; participants may feel

more receptive to and comfortable engaging with content and less judged by others when surrounded by same-gender peers.

Our findings also suggest that the ethnic composition of the sample had a significant moderating effect for both knowledge/attitudes and bystander efficacy/intentions. Specifically, studies with predominantly white samples (i.e., 70%+) produced significantly larger effects than did studies with mixed ethnicity samples. This finding suggests that programs may benefit from considering cultural beliefs, values, and minority group experiences when targeting ethnic and racial minority groups. This is supported by intersectional perspectives which indicate that those from minority groups are at greater risk for dating and sexual violence and their experiences are marked by additional factors such as racism and discrimination (Burns et al., 2019; McMahon et al., 2020). Importantly, some research has found support for increased treatment effectiveness on ethnic minority IPV intervention participants with a greater incorporation of culturally sensitive material (e.g., Gondolf, 2007). When designing and/or selecting a program to deliver to college students, the ethnic composition of the population being targeted should be considered, and whether additional sensitivity to minority groups could be incorporated into the program.

Program components. Our findings suggest that several program components influence program effects. First, incorporating a discussion of definitions of types of violence was related to stronger treatment impacts for both knowledge/attitudes and bystander efficacy/intentions. This is important as several types of abuse are subsumed within the categories of "dating violence." Moreover, these behaviors are not simply restricted to physical violence alone; they also include more invisible types of abuse such as psychological, emotional, and economic violence that participants might not initially be aware constitute forms of abuse. This result implies that when participants are taught what behaviors are considered abuse, they are more likely to correctly identify abusive behaviors and are more likely to demonstrate confidence to intervene as bystanders.

Second, our results suggest that including program content that discusses the prevalence of violence is an important moderator of treatment effect. However, while interventions that included such content produced larger treatment effects for bystander efficacy/intentions, the opposite was true for knowledge/attitudes. These findings suggest that dating violence prevention programs that only target knowledge/attitudes may not need to include content on the prevalence of violence. With respect to bystander efficacy/intentions, it may be that learning the high rates of violence on campus encourages participants to be more aware of the severity of the problem and the need to reduce its prevalence. If so, participants may subsequently feel more willing to intervene when they observe abusive behaviors. Overall, including content on the prevalence of violence is likely important as long as the information is presented in a way that engages participants and to which they can directly relate (Michau et al., 2015; Radatz & Wright, 2016); perhaps with respect to rates of campus dating/sexual violence, or rates in the local community.

Third, interventions that incorporated a component in which the long-term impacts of violence on victims was discussed produced smaller effect sizes for both knowledge/attitudes and bystander efficacy/intentions than programs that did *not* discuss the impact of violence. This is a curious finding as it suggests that programs need not (or, should not) include a component that focuses on the impacts of violence. While we can only postulate as to a possible explanation, perhaps participants do not connect to this material, as they are already aware of the many negative impacts of violence, and are more likely to “zone out” when hearing this information.

Last, while interventions that include legal definitions of consent produced larger treatment effects for knowledge/attitudes, the opposite was true for bystander efficacy/intentions. The reason for this discrepant finding is not clear, however, it could perhaps be explained by the types of measures used to assess the two outcome categories, and whether legal definitions of consent were directly or indirectly related to the outcome category. For example, some measures of knowledge about dating violence directly included questions about definitions of consensual sexual activity, whereas measures of bystander efficacy/intentions rarely included direct assessments of definitions of consent (e.g., “*Think through the pros and cons of different ways I might help if I see an instance of sexual violence*” [from the Bystander Efficacy Scale; Banyard et al., 2007]). To explain, while the former *directly* measures participants’ understanding of definitions about consent, the latter is an *indirect* measure because although it insinuates general knowledge about consent, it does not directly measure how consent is legally defined. As such, because a program’s component about definitions of consent is assessed more directly with knowledge/attitudinal measures, the effect of such components on intervention outcomes is likely emphasized more in the knowledge/attitudes category than it is in the bystander efficacy/intentions category. An alternative explanation is that a detailed understanding of the legal definition of consent may not be necessary for respondents to understand bystander efficacy/intentions, such as the importance of being watchful of friends in risky situations like alcohol-fueled parties. Thus (perhaps unsurprisingly), learning how to recognize and identify (non)consensual behavior through situational examples (rather than definitions and legal jargon about consent) may be more important to teaching participants bystander skills (e.g., confidence and intentions to intervene). A further explanation could be that participants feel overwhelmed or unsure when presented with consent information and how this would translate into their actions as a bystander. In other words, determining whether behaviors between others are consensual versus nonconsensual may be challenging, and feelings of self-efficacy or intentions to intervene may be low. Ensuring that information on consent is presented and framed to participants with respect to how to intervene as a bystander may be important.

Program approach. Our findings suggest that virtual methods of delivery (i.e., the use of videos or an exclusively online format) were less effective for the development of bystander skills.

It may be that in-person interaction with a facilitator and/or other students in a group setting is particularly helpful for the transfer of bystander skills. We recommend that those programs seeking to increase bystander skills maximize the use of in-person delivery methods. However, for programs seeking primarily to change knowledge/attitudes toward dating violence, in-person interaction may not be as crucial. Last, we suggest caution with the use of role play in teaching bystander skills, as our findings indicate that this approach may not be an effective method for all participants. This is contrary to literature indicating that role-play can be an essential tool in situational skill development and behavior modification (e.g., Rao, 2011; Skoura-Kirk et al., 2020); it is possible that, due to the sensitive nature of dating/sexual violence, the participatory nature of role play was disconcerting to some participants. This possibility is increased because the large majority of programs used a single-session approach, in which role playing may not have felt like a comfortable activity for some (or many) participants. Multi-session programs, in which participants have time to develop a sense of trust and comfort with the facilitators and/or other participants, may be more amenable to role play activities. Additionally, norms and attitudes that normalize dating violence are pervasive within society and exposure to such perceptions are constant (Banyard et al., 2004); a single session of role-play may not be sufficient to translate into the confidence and skills needed for participants to act against these norms (and their peers) and actively engage in bystander behaviors. The short amount of time spent on these topics and skills may instead result in a backlash effect, wherein participants become less willing or feel less capable of intervening (Exner-Cortens & Cummings, 2021; Moynihan et al., 2010). For example, it is possible that participants may become more aware of potential situations and responses in which intervening is warranted, but are not provided enough time and practice to feel confident in their skills and abilities to engage in helpful bystander behaviors (Exner-Cortens & Cummings, 2021). Future research should investigate this program component more closely.

Following is a summary of our evidence-based recommendations and implications of the review for practice, policy, and research:

1. Program Outcomes
 - a. Bystander behaviors: Incorporate evidence-based techniques to help with behavioral change (e.g., in-person delivery, multi-session format, engaging and empowering content)
2. Program Components
 - a. Participant characteristics:
 - Consider the ethnic composition of the college population being targeted and whether additional sensitivity to minority groups should be incorporated
 - Consider the gender composition of the sample and whether content might be more impactful for men if delivered in gender-segregated groups

- b. Program components:
 - Include content that focuses on definitions of violence and, for programs primarily targeting bystander skills, include content on the prevalence of dating/sexual violence
 - Consider whether material focused on the impacts of violence and definitions of consent are being presented in an engaging manner
- c. Program approach:
 - Maximize the use of in-person delivery methods to teach bystander skills and limit the use of virtual components; knowledge/attitudes can be taught using online methods
 - Use interactive role play methods with caution when teaching bystander skills

Limitations

There are several limitations to the current study. First, as is always true with systematic literature reviews, it is possible that some studies were not identified. Further, publication bias is a risk in any meta-analysis; it is possible that some studies were conducted but not published due to null findings. Second, we did not restrict to two-group designs; instead, we extended to single-group pre- and posttest designs. As discussed previously, the literature includes a healthy debate on the appropriateness of pooling effect sizes from single-group and two-group designs in meta-analysis (see Borenstein & Hedges, 2019; Cuijpers et al., 2017). Given the numerous single-group designs identified through the systematic search, we believe the inclusion of these studies permits a more comprehensive examination of the literature with respect to the effectiveness of college dating violence prevention programs.

Third, as discussed previously, there is substantial overlap between the constructs of dating and sexual violence with respect to both program components and evaluation outcomes. While limiting studies to those that target *only* dating violence, and *only* outcome measures that assess dating violence would have presented a more concentrated set of results, we argue that these concepts are inextricably linked. In other words, because dating violence is a type of IPV that encompasses acts of physical, psychological, and sexual violence occurring within the context of a dating relationship (CDC, 2021), sexual violence is a component of dating violence. As such, by definition, dating violence programs will also seek to prevent sexual violence. Similarly, instruments such as the well-known *Bystander Intention to Help Scale* (Banyard, 2008), used by numerous studies in the current analysis, include items that measure both concepts, making their disaggregation impossible without item-level results.¹⁰ Fourth, as few studies measured dating violence perpetration or victimization directly, these outcomes could not be pooled across studies. In addition, the timing of posttest would have been an interesting moderator to include; this was not possible given that few studies assessed longitudinal outcomes.

Fifth, missing outcome data in some primary evaluation reports prevented the calculation of an effect size and resulted in their exclusion from the current study. Several study authors were contacted in the hopes of obtaining the necessary data; however, this effort resulted in usable information in only two cases. In addition, many studies suffered from missing data on program and sample characteristics. To prevent prohibitive rates of missing data for program components, we made assumptions about components when possible. In many cases these assumptions seemed reasonable, for example, a campaign primarily using poster materials is highly unlikely to have incorporated role play scenarios. Nevertheless, it is possible that in some cases our assumptions were faulty. Last, the moderator analyses dichotomized the samples into small subgroups; in some cases subgroups had samples of six and these results should be interpreted with caution. It is also possible that the effects may be confounded with other characteristics of these studies or programs.

Conclusion

Dating violence includes a number of unique behaviors and experiences that are distinct from other forms of sexual violence; a focus on education and prevention of such behaviors is necessary for a comprehensive approach to addressing violence on college campuses. Findings from the present study have important implications for policymakers and practitioners seeking to prevent dating violence among college students, suggesting that prevention programs are effective at increasing knowledge/attitudes toward dating/sexual violence as well as bystander skills. Results also suggest that more work is needed to improve effectiveness concerning bystander behaviors.

Authors' Note

Jouriles et al. (2016) evaluated two separate treatments; in analyses this was labeled Jouriles (2016a) and Jouriles (2016b). Moynihan et al. (2015) evaluated two separate treatments; these were labeled Moynihan (2015a) and Moynihan (2015b).


Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study received funding from BC Ministry of Public Safety, Office of Crime Reduction and Gang Outreach (Grant ID 24542).

ORCID iD

Jennifer S. Wong  <https://orcid.org/0000-0001-7523-7430>

Supplemental Material

The supplemental material for this article is available online.

Notes

1. Title IX refers to institutions that receive federal financial assistance from the U.S. Department of Education and includes approximately 7,000 postsecondary institutions (https://www2.ed.gov/about/offices/list/ocr/docs/tix_dis.html). Title IX is part of the Education Amendments of 1972, and prevents discrimination on the basis of sex to any person participating in an education program or activity that receives financial assistance from the federal government.
2. Restriction for language of publication were necessary, as our team did not have immediate access to resources that would have allowed us to confidently translate and code academic articles in languages other than English and French.
3. We limited the search period to the last 20 years as we wanted to ensure that we were summarizing the most current/relevant literature on the topic, while also maintaining an adequate sample size of studies for quantitative pooling.
4. Due to inconsistent reporting/missing data across studies, we were unable to code all 77 variables for all included studies. Specifically, we were often faced with missing data, precluding the calculation of an effect size. In these cases, study authors were contacted with a request for information (e.g., a pretest sample size or standard errors). Another inconsistent area of reporting across the primary evaluation studies was in respect to program components and approach. In many cases it was difficult to determine if a program truly did not include a component or if the study's description of the program simply failed to mention it. If coders felt at least 85% confident that the component was *not* included in the program (e.g., a discussion of gender stereotypes), the study was coded as a "no" for the component; if coders were less than 85% confident, the study was coded as "missing" for the component. Similarly, if coders were at least 85% confident that a given component such as group activities *was* included the study was coded as a "yes"; if not, it was coded as "missing". These assumptions were made to minimize the amount of missing data and allow for analyses of the moderating impact of program characteristics on program outcomes. All assumptions were agreed upon by two reviewers during the coding process, and were based on the content provided in the article concerning other program components and/or approach. For example, the evaluation by Hays et al. (2015) of the HEART program states "... five 60-minute psychoeducational group sessions were each offered three times. Sessions were held in community rooms of similar size within a university student center and residence halls" (p. 52). Despite the authors not stating that an online component was *not* included, we were 85% confident that online methods were not used in this study, and that a code of 0 was appropriate for this variable.
5. The choice of ICC for cluster adjustments differed based on the level of clustering used. As none of the four studies included empirically derived ICCs, we attempted to find other evaluations using similar samples to assist in deciding upon an appropriate ICC to use. The literature in this area was sparse, and we were unable to identify other study that reported ICC at the same levels of clustering for the same type of outcomes. Based on the available research, we decided on an ICC of .01 for the university-level effect sizes of Borsky et al. (2018; e.g., see Seo & Li, 2009). Holland (2014) assigned groups at the classroom level, with two university classrooms assigned to the treatment group and two classrooms assigned to the control group. Again, we turned to the literature to assist in the selection of an appropriate ICC; we did not identify any similar studies on which to base an estimate. However, given the fact that Holland's sample used students enrolled in undergraduate "Health Promotion and Behavior 1710 Health and Wellness classes", we expect the between-classroom variation to be very small (i.e., students in one class versus another class are unlikely to vary substantially). As such, we cluster-adjusted the Holland effect size with an ICC of .005. Last, both Rothman et al. (2018) and Steward (2017) assigned at the fraternity and/or sorority level; using research by Caudill et al. (2006) and LaBrie et al. (2008) we opted to use an ICC of .10 for the cluster adjustments on both the studies.
6. The included programs are described in more detail in Appendix C; the list is as follows: *Friends Helping Friends* (Amar et al., 2015), *Relationship Remix* (Bonar et al., 2019), New Student Orientation and *Red Flag Campaign* (Borsky et al., 2018; Carlyle et al., 2020), *Not Anymore* (Draper, 2017), *The Intervention Initiative* (Fenton & Mott, 2018), *Help End Abusive Relationships Today* (Hays et al., 2015), *Bringing in the Bystander* (Hines & Palm Reed, 2017; Moynihan et al., 2010; Moynihan et al., 2011; Moynihan et al., 2015), *F-iConsent* (Holland, 2014), *TakeCare* (Jouriles et al., 2016; 2017; Kleinsasser et al., 2015), *Choices* (Kuffel & Katz, 2002), *Dating Relationships in College Students workshop* (Lazarevich et al., 2017), *Love is Not Abuse* (Moreno, 2017), *Know Your Power* (Moynihan et al., 2015), *STOP Dating Violence* (O'Brien et al., 2019), "Traditional dating violence awareness education program" (no name provided; Peterson et al., 2018), *Mindflock* video game (Potter et al., 2019), *EKU-SAFE* (Reid et al., 2013), *Escalation* (Rothman et al., 2018), *Intervene* (Santacrose et al., 2020), *Sexpectations* (Steward, 2017), *Dating Relationships Involving Violence End Now* (Terrazas-Carrillo et al., 2020), and *Mentors in Violence Prevention* (Toy, 2016).
7. Developed by Sherman and colleagues (1998), the Maryland Scale for Scientific Methods (SMS) is a well-known scale used to assess the methodological rigor (i.e., scientific strength) of research designs with respect to internal validity. Research designs are scored on a 5-point scale, where Level 1 is the weakest design in terms of internal validity (e.g., posttest only design) and Level 5 is the strongest (e.g., randomized controlled trial).
8. For this reason, we argue that a random effects model remains more appropriate for pooling outcomes than a fixed effects model (for comparison, the fixed effect pooled estimate was 0.472, $z = 37.94$, $p < .001$). Random effects models are used accordingly for each of the disaggregated outcome measures.
9. There was insufficient variation in the online variable to use it in the analyses for knowledge/attitudes.
10. We also note that this issue is likewise present for the previously discussed existing meta-analyses of sexual violence prevention programs. Many of these studies use the same scales developed

by Banyard and colleagues, which include items tapping both sexual violence and dating violence constructs.

References

- Studies included in the meta-analysis are denoted by an asterisk
- Anderson, L. A., & Whiston, S. C. (2005). Sexual assault education programs: A meta-analytic examination of their effectiveness. *Psychology of Women Quarterly, 29*(4), 374–388.
- Amar, A. F., & Gennaro, S. (2005). Dating violence in college women: Associated physical injury, healthcare usage, and mental health symptoms. *Nursing Research, 54*(4), 235–242.
- *Amar, A. F., Tuccinardi, N., & Simpson, S. (2015). Friends helping friends: A nonrandomized control trial of a peer-based response to dating violence. *Nursing Outlook, 6*(3), 496–503.
- Banyard, V. L. (2008). Measurement and correlates of pro-social bystander behavior: The case of interpersonal violence. *Violence and Victims, 23*, 85–99.
- Banyard, V. L., Moynihan, M., & Crossman, M. (2009). Reducing sexual violence on campus: The role of student leaders as empowered bystanders. *Journal of College Student Development, 50*(4), 446–457.
- Banyard, V. L., Moynihan, M. M., & Plante, E. G. (2007). Sexual violence prevention through bystander education: An experimental evaluation. *Journal of Community Psychology, 35*, 463–481.
- Banyard, V. L., Plante, E. G., & Moynihan, M. M. (2004). Bystander education: Bringing a broader community perspective to sexual violence prevention. *Journal of Community Psychology, 32*(1), 61–79.
- Bhochhibhoya, S., Maness, S. B., Cheney, M., & Larson, D. (2019). Risk factors for sexual violence among college students in dating relationships: An ecological approach. *Journal of Interpersonal Violence, 36*, 1–25.
- Baker, C. R., & Stith, S. M. (2008). Factors predicting dating violence perpetration among male and female college students. *Journal of Aggression, Maltreatment & Trauma, 17*(2), 227–244.
- Black, B. M., Weisz, A. N., & Jayasundara, D. S. (2012). Dating violence and sexual assault prevention with African American middle schoolers: Does group gender composition impact dating violence attitudes? *Child & Youth Services, 33*(2), 158–173.
- *Bonar, E., Rider-Milkovich, H., Huhman, A., McAndrew, L., Goldstick, J., Cunningham, R., & Walton, M. (2019). Description and initial evaluation of a values-based campus sexual assault prevention programme for first-year college students. *Sex Education, 19*(1), 99–113.
- Borenstein, M., & Hedges, L. V. (2019). Statistically describing and combining study outcomes. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (3rd ed., pp. 207–243). Russell Sage Foundation.
- *Borsky, A. E., McDonnell, K., Turner, M. M., & Rimal, R. (2018). Raising a red flag on dating violence: Evaluation of a low-resource, college-based bystander behavior intervention program. *Journal of Interpersonal Violence, 33*(22), 3480–3501.
- Brewer, N., & Thomas, K. (2019). Intimate partner violence academic performance: The role of physical, mental, behavioral, financial health. *Social Work in Health Care, 58*(9), 854–869.
- Brewer, N., Thomas, K. A., & Higdon, J. (2018). Intimate partner violence, health, sexuality, and academic performance among a national sample of undergraduates. *Journal of American College Health, 66*(7), 683–692.
- Burns, V. L., Eaton, A. A., Long, H., & Zapp, D. (2019). Exploring the role of race and gender on perceived bystander ability and intent: Findings before and after exposure to an online training program to prevent sexual assault on campus. *Violence Against Women, 25*(8), 999–1017.
- Canan, S. N., Jozkowski, K. N., & Crawford, B. L. (2018). Sexual assault supportive attitudes: Rape myth acceptance and token resistance in Greek and non-Greek college students from two university samples in the United States. *Journal of Interpersonal Violence, 33*(22), 3502–3530.
- Card, N. A. (2011). *Applied meta-analysis for social science research*. Guilford Press.
- Carlos, D., Campeiz, A., Silva, J., Fernandes, M., Leitão, M., Silva, M., & Ferriani, M. (2017). School-based interventions for teen dating violence prevention: Integrative literature review. *Revista de Enfermagem Referência, 14*, 133–146.
- Carlson, K. D., & Schmidt, F. L. (1999). Impact of experimental design on effect size: Findings from the research literature on training. *Journal of Applied Psychology, 84*(6), 851–862.
- *Carlyle, K. E., Conley, A. H., & Guidry, J. P. D. (2020). Development and evaluation of the Red Flag campaign for the primary prevention of sexual and dating violence on college campuses. *Journal of American College Health, 9*, 1–5.
- Caudill, B. D., Crosse, S. B., Campbell, B., Howard, J., Luckey, B., & Blane, H. T. (2006). High-risk drinking among college fraternity members: A national perspective. *Journal of American College Health, 55*(3), 141–155.
- Centers for Disease Control. (2020). *Intimate partner violence*. <https://www.cdc.gov/violenceprevention/intimatepartnerviolenceindex.html>
- Centers for Disease Control and Prevention. (2021). Preventing teen dating violence. <https://www.cdc.gov/violenceprevention/intimatepartnerviolence/teendatingviolence/fastfact.html>
- Choi, H. J., Weston, R., & Temple, J. R. (2017). A three-step latent class analysis to identify how different patterns of teen dating violence and psychosocial factors influence mental health. *Journal of Youth and Adolescence, 46*(4), 854–866.
- Coker, A., Bush, H., Cook-Craig, P., DeGue, S., Clear, E., Brancato, C., Fisher, B., & Recktenwald, E. (2017). RCT testing bystander effectiveness to reduce violence. *American Journal of Preventative Medicine, 52*(5), 566–578.
- Coker, A., Hall Smith, P., Bethea, L., King, M., & McKeown, R. E. (2000). Physical health consequences of physical and psychological intimate partner violence. *Archives of Family Medicine, 9*(5), 451–457.
- Crooks, C. V., Jaffe, P., Dunlop, C., Kerry, A., & Exner-Cortens, D. (2019). Preventing gender-based violence among adolescents and young adults: Lessons from 25 years of program development and evaluation. *Violence Against Women, 25*(1), 29–55.
- Cuijpers, P., Weitz, E., Cristea, I. A., & Twisk, J. (2017). Pre-post effect sizes should be avoided in meta-analyses. *Epidemiological Psychiatric Science, 26*(4), 364–368.

- DeGue, S., Valle, L. A., Holt, M. K., Massetti, G. M., Matjasko, J. L., & Tharp, A. T. (2014). A systematic review of primary prevention strategies for sexual violence perpetration. *Aggression and Violent Behavior, 19*(4), 346–362.
- de la Rue, L., Polanin, J. R., Espelage, D. L., & Pigott, T. D. (2017). A meta-analysis of school-based interventions aimed to prevent or reduce violence in teen dating relationships. *Review of Educational Research, 87*(1), 7–34.
- *Draper, J. L. (2017). *Save our campus: Analyzing the effectiveness of an online sexual violence program* [Dissertation]. Rutgers, the State University of New Jersey.
- Edwards, S. R., & Hinsz, V. B. (2014). A meta-analysis of empirically tested school-based dating violence prevention programs. *SAGE Open, 4*(2), 1–8.
- Egger, M., Juni, P., Bartlett, C. J., Holenstein, F., & Sterne, J. A. C. (2003). How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. *Health Technology Assessment, 7*(1), 1–76.
- Eshelman, L., & Levendosky, A. A. (2012). Dating violence: Mental health consequences based on type of abuse. *Violence and Victims, 27*(2), 215–228.
- Evans, J. L., Burroughs, M. E., & Knowlden, A. P. (2019). Examining the efficacy of bystander sexual violence interventions for first-year college students: A systematic review. *Aggression and Violent Behavior, 48*, 72–82.
- Exner-Cortens, D., & Cummings, N. (2021). Bystander-based sexual violence prevention with college athletes: A pilot randomized trial. *Journal of Interpersonal Violence, 36*(1-2), NP188–211NP.
- Fellmeth, G., Heffernan, C., Nurse, J., Habibula, S., & Sethi, D. (2015). Educational and skills-based interventions to prevent relationship violence in young people. *Research on Social Work Practice, 25*(1), 90–102.
- *Fenton, R., & Mott, H. (2018). Evaluation of Intervention Initiative: A bystander intervention program to prevent violence against women universities. *Violence & Victims, 33*(4), 645–662.
- Gidycz, C., Warkentin, J., & Orchowski, L. (2007). Predictors of perpetration of verbal, physical, and sexual violence: A prospective analysis of college men. *Psychology of Men & Masculinity, 8*, 79–94.
- Gondolf, E. (2007). 'Culturally focused batterer counseling for African American men'. *Criminology and Public Policy, 6*(2), 341–366.
- Gover, A., Kaukinen, C., & Fox, K. A. (2008). The relationship between violence in the family of origin and dating violence among college students. *Journal of Interpersonal Violence, 23*(12), 1667–1693.
- Graham, L. M., Embry, V., Young, B.-R., Macy, R. J., Moracco, K. E., Reyes, H. L. M., & Martin, S. L. (2019). Evaluations of prevention programs for sexual, dating, and intimate partner violence for boys and men: A systematic review. *Trauma, Violence, & Abuse, 1–27*.
- Gray, M. J., Hassija, C. M., & Steinmetz, S. E. (2017). *Sexual assault prevention on college campuses*. Routledge, Taylor & Francis Group.
- *Hays, D., Michel, R. E., Bayne, H. B., & Myers, J. S. (2015). Counseling with HEART: A relationship violence prevention program for college students. *Journal of College Counseling, 18*, 49–65.
- Hedges, L. V. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics, 32*(4), 341–370.
- Hines, D., Bishop, L., & Palm Reed, K. (2019). Differential gender response to an empathy component of a sexual assault prevention program. *Victims and Violence, 34*(3), 397–413.
- *Hines, D., & Palm Reed, K. (2017). Bystander prevention of sexual and dating violence: An experimental evaluation of online and in-person bystander intervention programs. *Partner Abuse, 8*(4), 331–346.
- *Holland, K. M. (2014). *"iConsent": Evaluation of a universal dating and sexual violence prevention program for college students* [Dissertation]. University of Georgia.
- *Jouriles, E., McDonald, R., Rosenfield, D., Levy, N., Sargent, K., Caiozzo, C., & Grych, J. (2016). TakeCARE, a video bystander program to help prevent sexual violence on college campuses: Results of two randomized, controlled trials. *Psychology of Violence, 6*(3), 410–420.
- *Jouriles, E., Sargent, K., Salis, K., Caiozzo, C., Rosenfield, D., Grych, J., O'Leary, K., & McDonald, R. (2017). TakeCare, a video to promote bystander behavior on college campuses: Replication and Extension. *Journal of Interpersonal Violence, 35*(23-24), 5652–5675.
- Jouriles, E., Krauss, A., Vu, N. L., Banyard, V. L., & McDonald, R. (2018). Bystander programs addressing sexual violence on college campuses: A systematic review and meta-analysis of program outcomes and delivery methods. *Journal of American College Health, 66*(6), 457–466.
- Jozkowski, K. N., & Wiersma-Mosley, J. D. (2017). The Greek system: How gender inequality and class privilege perpetuate rape culture. *Family Relations, 66*(1), 89–103.
- Katz, J., & Moore, J. (2013). Bystander education training for campus sexual assault prevention: An initial meta-analysis. *Violence and Victims, 28*(6), 1054–1067.
- Kettrey, H. H., & Marx, R. A. (2019). The effects of bystander programs on the prevention of sexual assault across the college years: A systematic review and meta-analysis. *Journal of Youth and Adolescence, 48*(2), 212–227.
- Kettrey, H. H., & Marx, R. A. (2020 Online First). Effects of bystander sexual assault prevention programs on promoting intervention skills and combatting the bystander effect: A systematic review and meta-analysis. *Journal of Experimental Criminology, 1–25*.
- *Kleinsasser, A., Jouriles, E., McDonald, R., & Rosenfield, D. (2015). An online bystander intervention program for prevention sexual violence. *Psychology of Violence, 5*(3), 227–235.
- *Kuffel, S. W., & Katz, J. (2002). Preventing physical, psychological, and sexual aggression in college dating relationships. *Journal of Primary Prevention, 22*(4), 361–374.
- Kugley, S., Wade, A., Thomas, J., Mahood, Q., Jørgensen, A. M. K., Hammerstrøm, K., & Sathe, N. (2017). Searching for studies: A guide to information retrieval for Campbell systematic reviews. *Campbell Systematic Reviews, 13*(1), 1–73.
- LaBrie, J. W., Hummer, J. F., Neighbors, C., & Pedersen, E. R. (2008). Live interactive group-specific normative feedback reduces misperceptions and drinking in college students: A randomized cluster trial. *Psychology of Addictive Behaviors, 22*(1), 141–148.

- *Lazarevich, I., Irigoyen-Camacho, M. E., Velázquez-Alva, M. D. C., & Salinas-Avila, J. (2017). Dating violence in Mexican college students: Evaluation of an educational workshop. *Journal of Interpersonal Violence, 32*(2), 183–204.
- Lee, C., & Wong, J. S. (2020 Online First). Examining effects of teen dating violence prevention programs: A systematic review and meta-analysis. *Journal of Experimental Criminology, 1*–40.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis. Applied social research methods series: Volume 49*. Sage.
- Loh, C., Gidycz, C. A., Lobo, T. R., & Luthra, R. (2005). A prospective analysis of sexual assault perpetration: Risk factors related to perpetrator characteristics. *Journal of Interpersonal Violence, 20*(10), 1325–1348.
- Malamuth, N. M., Huppert, M., & Linz, D. (2018). Sexual assault interventions may be doing more harm than good with high-risk males. *Aggression and Violent Behavior, 41*, 20–24.
- McMahon, S., Burnham, J., & Banyard, V. L. (2020). Bystander intervention as a prevention strategy for campus sexual violence: Perceptions of historically minoritized college students. *Prevention Science, 21*(6), 795–806.
- Michau, L., Horn, J., Bank, A., Dutt, M., & Zimmerman, C. (2015). Prevention of violence against women and girls: Lessons from practice. *The Lancet, 385*(9978), 1672–1684.
- Miller, E., Jones, K. A., & Ripper, L. (2020). An athletic coach-delivered middle school violence prevention program: A cluster randomized trial. *JAMA Pediatrics, 174*(3), 241–242.
- *Moreno, A. A. (2017). *The impact of a dating abuse prevention program on college students' knowledge, attitudes, and conflict resolution skills* [Dissertation]. University of La Verne.
- Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods, 7*(1), 105–125.
- *Moynihan, M. M., Banyard, V. L., Arnold, J. S., Eckstein, R. P., & Stapleton, J. G. (2010). Engaging intercollegiate athletes in preventing and intervening in sexual and intimate partner violence. *Journal of American College Health, 59*(3), 197–204.
- *Moynihan, M. M., Banyard, V. L., Arnold, J., Eckstein, R. P., & Stapleton, J. G. (2011). Sisterhood may be powerful for reducing sexual and intimate partner violence: An evaluation of Bringing in the Bystander program with sorority members. *Violence Against Women, 17*(6), 703–719.
- *Moynihan, M. M., Banyard, V. L., Cares, A. C., Potter, S. J., Williams, L. M., & Stapleton, J. G. (2015). Encouraging responses in sexual and relationship violence prevention: What program effects remain 1 year later? *Journal of Interpersonal Violence, 30*(1), 110–132.
- Mujal, G., Taylor, Fry, Gochez-Kerr, T. H., & Weaver, N. (2019 Online First). A systematic review of bystander interventions for prevention of sexual violence. *Trauma, Violence, & Abuse, 22*(2), 381–396.
- Mumford, E. A., Taylor, B. G., & Giordano, P. C. (2020). Perpetration of adolescent dating relationship abuse: The role of conditional tolerance for violence and friendship factors. *Journal of Interpersonal Violence, 35*(5-6), 1206–1228.
- *O'Brien, K. M., Sauber, E. W., Kearney, M. S., Venaglia, R. B., & Lemay, E. P. Jr. (2019 Online First). Evaluating the effectiveness of an online intervention to educate college students about dating violence and bystander responses. *Journal of Interpersonal Violence, 36*(13-14), NP7516–NP7546.
- *Peterson, K., Sharps, P., Banyard, V., Powers, R. A., Kaukinen, C., Gross, D., Decker, M. R., Baatz, C., & Campbell, J. (2018). An evaluation of two dating violence prevention programs on a college campus. *Journal of Interpersonal Violence, 33*(23), 3630–3655.
- *Potter, S. J., Flanagan, M., Seidman, M., Hodges, H., & Stapleton, J. G. (2019). Developing and piloting videogames to increase college and university students' awareness and efficacy of the bystander role in incidents of sexual violence. *Games for Health, 8*(1), 24–34.
- Radatz, D. L., & Wright, E. M. (2016). Integrating the principles of effective intervention into batterer intervention programming: The case for moving toward more evidence-based programming. *Trauma, Violence, & Abuse, 17*(1), 72–87.
- Rao, D. (2011). Skills development using role-play in a first-year pharmacy practice course. *American Journal of Pharmaceutical Education, 75*(5), 1–10.
- Rape, Abuse, & Incest National Network. (n.d.). *Campus SaVE act*. <https://www.rainn.org/articles/campus-save-act>
- *Reid, C., Irvin, M., & Dye, H. (2013). Learning to prevent violence, EKV-Safe: A bystander intervention training. *Kentucky Journal of Excellence College Teaching & Learning, 10*, 9–19.
- *Rothman, E., Paruk, J., & Banyard, V. (2018). The escalation dating abuse workshop for college students: Results of an efficacy RCT. *Journal of American College Health, 66*(6), 519–528.
- *Santacrose, L. B., Laurita, A., & Marchell, T. (2020). Intervene: Modeling pro-social bystander behavior in college students through online video. *Health Communication, 35*(4), 397–409.
- Seabrook, R. C., & Ward, L. M. (2019). Bros will be bros? The effect of fraternity membership on perceived culpability for sexual assault. *Violence Against Women, 25*(12), 1471–1490.
- Seabrook, R. C., Ward, L. M., & Giaccardi, S. (2018). Why is fraternity membership associated with sexual assault? Exploring the roles of conformity to masculine norms, pressure to uphold masculinity, and objectification of women. *Psychology of Men & Masculinity, 19*(1), 3–13.
- Seo, D.-C., & Li, K. (2009). Effects of college climate on students' binge drinking: Hierarchical generalized linear model. *Annals of Behavioral Medicine, 38*(3), 262–268.
- Sherman, L. W., Gottfredson, D. C., MacKenzie, D. L., Eck, J., Reuter, P., & Bushway, S. D. (1998). *Preventing crime: What works, what doesn't, what's promising*. National Institute of Justice Research in Brief. U.S. Department of Justice.
- Shorey, R. C., Brasfield, H., Zapor, H. Z., Febres, J., & Stuart, G. L. (2015). The relation between alcohol use and psychological, physical, and sexual dating violence perpetration among male college students. *Violence Against Women, 21*(2), 151–164.
- Shorey, R. C., Stuart, G., & Cornelius, T. (2011). Dating violence and substance use in college students: A review of the literature. *Aggression and Violent Behavior, 16*(6), 541–550.
- Skoura-Kirk, E., Brown, S., & Mikelyte, R. (2020 Online First). Playing its part: An evaluation of professional skill development through service user-led roleplays for social work students. *Social Work Education, 1*–17.

- Spikes, C., & Sternadori, M. (2018). Boomerang effects of sexual-violence prevention messages on college men's attitudes. *Building Healthy Communities Journal*, 2(2), 45–57.
- Sterne, J., & Harbord, R. (2004). Funnel plots in meta-analysis. *The Stata Journal*, 4(2), 127–141.
- *Steward, J. M. (2017). *Building upon bystander intervention: A multi-component prevention programming approach for university sorority members* [Dissertation]. University of Tulsa.
- Storer, H., Casey, E., & Herrenkohl, T. (2016). Efficacy of bystander programs to prevent dating abuse among youth and young adults: A review of the literature. *Trauma, Violence, & Abuse*, 17, 256–269.
- *Terrazas-Carrillo, E., Garcia, E., Vásquez, D., Sabina, C., & Ponte, L. F. (2020). DRIVEN to stop violence: A group intervention to prevent dating violence among Latino college students. *Journal of Primary Prevention*, 41(4), 331–348.
- Tobias, A. (1999). Assessing the influence of a single study in the meta-analysis estimate. *Stata Technical Bulletin*, 8(47), 15–17.
- *Toy, J. (2016). *Mentors in violence prevention training and its effectiveness with resident assistants* [Unpublished master's thesis]. California State University, Long Beach.
- Wood, L., Voth Schrag, R., & Busch-Armendariz, N. (2020). Mental health and academic impacts of intimate partner violence among IHE-attending women. *Journal of American College Health*, 68(3), 286–293.

Author Biographies

Jennifer S. Wong is an associate Professor in the School of Criminology at Simon Fraser University. Her work employs applied, quantitative, and qualitative methods to study issues of delinquency/crime prevention and intervention, focusing primarily on evaluation and policy analysis in the areas of crime prevention and crime control

policy. Her recent work has examined evaluations of intimate partner violence intervention programs, the impact of prior deportation of illegal immigrants on recidivism, the relationship between media framing of homicide and the use of images in newspaper articles, and meta-analyses on gang prevention programs, restorative justice programs for at-risk youth, and intensive supervision programs.

Jessica Bouchard is a PhD candidate in the School of Criminology at Simon Fraser University. She has worked alongside Dr. Wong as a project coordinator on several large-scale community-based crime prevention evaluation research projects, which include designing an evaluation of an anti-gang messaging campaign, an evaluation of the metro Vancouver YWCA Youth Education Programs, and an evaluation of 12 intimate partner violence prevention programs in British Columbia. She has also worked with Dr. Wong on several projects that focused on systematically reviewing and meta-analyzing the effects of community-based interventions on recidivism, including intensive supervision probation and aftercare/re-entry programs for juvenile offenders, home confinement programs, Day Reporting Centers, halfway houses, Teen Court programs, and restorative justice diversion programs for at-risk youth.

Chelsey Lee is a doctoral student in the SFU School of Criminology. Much of her work has focused within the area of gender-based violence including topics such as university sexual assault policy, media representations of domestic violence and homicide, and adolescent dating violence prevention programs. She has also contributed to several meta-analyses focused on interventions to reduce criminal recidivism, including halfway houses and day reporting centers. Her recent work can be found in several scholarly journals such as *Crime and Delinquency*, *Journal of Experimental Criminology*, *Violence Against Women*, *Journal of Interpersonal Violence*, *Studies in Higher Education*, *Journal of Offender Rehabilitation*, and the *International Journal of Offender Therapy and Comparative Criminology*.