



A Randomized, Controlled Trial of the Impact of the Couple CARE for Parents of Newborns Program on the Prevention of Intimate Partner Violence and Relationship Problems

Richard E. Heyman¹ · Amy M. Smith Slep¹ · Michael F. Lorber¹ · Danielle M. Mitnick¹ · Shu Xu¹ · Katherine J. W. Baucom¹ · W. Kim Halford² · Phyllis Holditch Niolon³

Published online: 7 December 2018
© Society for Prevention Research 2018

Abstract

Effective, accessible prevention programs are needed for adults at heightened risk for intimate partner violence (IPV). This parallel group randomized controlled trial examines whether such couples receiving the American version of Couple CARE for Parents of Newborns (CCP; Halford et al. 2009) following the birth of a child, compared with controls, report fewer first occurrences of clinically significant IPV, less frequent physical and psychological IPV, and improved relationship functioning. Further, we test whether intervention effects are moderated by level of risk for IPV. Couples at elevated risk for IPV ($N = 368$) recruited from maternity units were randomized to CCP ($n = 188$) or a 24-month waitlist ($n = 180$) and completed measures of IPV and relationship functioning at baseline, post-program (when child was 8 months old), and two follow-ups (at 15 and 24 months). Intervention effects were tested using intent to treat (ITT) as well as complier average causal effect (CACE; Jo and Muthén 2001) structural equation models. CCP did not significantly prevent clinically significant IPV nor were there significant main effects of CCP on clinically significant IPV, frequency of IPV, or most relationship outcomes in the CACE or ITT analyses. Risk moderated the effect of CCP on male-to-female physical IPV at post-program, with couples with a planned pregnancy declining, but those with unplanned pregnancies increasing. This study adds to previous findings that prevention programs for at-risk couples are not often effective and may even be iatrogenic for some couples.

Keywords Prevention · Randomized controlled trial · Partner aggression · Intimate partner violence

Preventing intimate partner violence (IPV) is an urgent priority of the US public health leaders (e.g., U. S. Department of Health and Human Services 2014). This, in large part, is due to the endemic nature of IPV and its toll on public health. In the USA, nationally representative studies of couples physical IPV have a yearly prevalence of 15% (e.g., Schafer et al. 1998) and a lifetime prevalence of around 50% (see Heyman et al. 2015). Physical IPV victimization is associated with increased risk for

(a) physical injury and other adverse physical health outcomes (e.g., Smith et al. 2017) and (b) degraded mental health, including both concurrent (Beydoun et al. 2017) and subsequent psychiatric diagnoses (e.g., Dutton et al. 2006), as well as greater fear, concern for safety, and symptoms of PTSD (Smith et al. 2017). Furthermore, research on psychological IPV victimization indicates that it uniquely predicts degraded physical and mental health, even after accounting for effects of physical IPV victimization (Dutton et al. 2006).

For those focused on the primary prevention of IPV, the decision when to best deflect IPV trajectories is vexing, given that IPV can start as early as dating starts. Physical IPV is the most common source of violence in the lives of adolescents and emerging adults, with yearly prevalences of 9–23% of adolescents in US samples (reviewed by Betz 2007); 19–47% of emerging adults (Rennison and Addington 2014); and between 29 and 57% of engaged couples (e.g., Lawrence and Bradbury 2007). Prevention programs aimed at middle and high school-aged individuals have had some

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11121-018-0961-y>) contains supplementary material, which is available to authorized users.

✉ Richard E. Heyman
Richard.Heyman@NYU.edu

¹ New York University, New York, NY, USA

² University of Queensland, Brisbane, Australia

³ Centers for Disease Control and Prevention, Atlanta, GA, USA

traction (Sumner et al. 2015), but programs for those beyond adolescence are necessary, especially those at key emerging and early-adult crossroads.

The birth of a child is one such propitious prevention crossroad. First, new parents recognize the challenges facing them, providing a critical period for optimal openness to learning/improving relationship and parenting skills (Halford et al. 2003). Second, the rate of high-risk couples' (including those with low-level IPV) participation in prevention programs is higher prior to the birth of a child compared with before marriage (Petch et al. 2012a). Third, the perinatal period is an important one for prevention, as partners are at slightly elevated risk for IPV (Charles and Perreira 2007) and relationship satisfaction decline (Doss et al. 2009; Mitnick et al. 2009).

Several early RCTs supported the efficacy of skills-based prevention programs delivered to couples with new or young children (often their first child). Couples who received these programs reported better IPV-related relationship resilience—less decline in relationship satisfaction (e.g., Schulz et al. 2006; Shapiro and Gottman 2005) and less destructive observed communication (e.g., Feinberg et al. 2009; Shapiro and Gottman 2005)—compared with those who received no intervention. Pinquart and Teubert's (2010) meta-analysis revealed small effects on couple communication and very small, yet statistically significant, effects on relationship adjustment from pre- to post-program; effects on couple communication were larger when interventions included both prenatal and postnatal components (compared with only one of the two). Relatedly, by targeting relationship and/or parenting risk factors for IPV (though not IPV directly), some studies reported reductions in psychological IPV (Kan and Feinberg 2014) and physical IPV (Bair-Merritt et al. 2010; Feinberg et al. 2016).

Perhaps because of the small effect sizes and publication bias toward significant findings, these early RCT results may not be robustly generalizable and may even be illusory. For instance, one of the largest, best-powered studies in this area—the evaluation of the Building Strong Families (BSF) program with low-income, unmarried, predominantly racial/ethnic minority parents of newborns—pooled data across eight sites and found no differences between couples who received skills-based prevention and no-intervention control couples on relationship outcomes (e.g., satisfaction) or on IPV at the post-program assessment (Wood et al. 2010). Given the assumption that couples prevention interventions are innocuous at worst, of far greater concern was the finding at the 36-month follow-up assessment that women who participated in the prevention programs were significantly *more* likely to report experiencing more than one instance of severe IPV in the previous year compared with women in the control condition (Moore et al. 2012), despite two of the three BSF couples-oriented prevention programs including modules that explicitly targeted IPV (Dion et al. 2010). Thus, despite some promising results and the possible openness of parents to prevention during the perinatal period, further research is certainly

needed to test for both effectiveness and iatrogenic effects, especially for new parents with higher risk factors for IPV.

Another key issue is access to prevention services, for even if couples are especially open to prevention during this developmental window and even if interventions are effective, this matters little if couples cannot easily access the interventions. All of the programs discussed above were delivered in groups, which present logistical (e.g., time, distance; Sullivan et al. 2004) and social (e.g., discomfort with sharing personal information) barriers to participation. Alternatively, innovative applications of technology (e.g., video- and telephone-assisted interventions) have been used to create effective, low-cost, low-barrier couples prevention programs that combine skills training via prerecorded videos with personal coaching via telephone (e.g., Halford et al. 2004).

This paper describes a randomized controlled study (RCT) of one such program—the American version (Halford et al. 2009) of *Couple CARE for Parents of Newborns* (CCP). CCP, developed and tested in Australia (Halford et al. 2015), was designed for flexible delivery in a variety of settings (phone, face-to-face individual, group, Internet, combinations) and is a new-parents variant of the original *Couple CARE* (Halford et al. 2004). CCP targets behaviors and cognitions associated with IPV (Slep et al. 2011) and postnatal decline in relationship satisfaction (Shapiro et al. 2000)—such as hostile reciprocity, unrealistic expectations, poor conflict management, and stress—by assisting partners in the following: (a) assessing relationship strengths and weaknesses; (b) developing realistic parenting and compatible co-parenting expectations; (c) defining the relationship they want; (d) developing key relationship and parenting skills; and (e) identifying individual actions to strengthen their relationship and parenting skills (e.g., self-regulation; Halford et al. 2007). The full table of contents can be found in the online supplemental materials.

An Australian RCT compared *Couple CARE for Parents* with a mothers-only treatment control and found less decline in relationship satisfaction for women (but not men) who received CCP (Halford et al. 2010). A larger subsequent RCT found support for moderated effects on relationship satisfaction: high-risk women who received the couple-based program evidenced less decline in satisfaction compared with those in the control group, with a similar, but nonsignificant, trend for men (Petch et al. 2012b). Change in IPV following intervention was not examined (Petch et al. 2012b), although physical IPV at baseline was one of the factors included in the composite index of risk in this study (and was present in nearly one-third of the couples in the sample; Petch et al. 2012a). The Australian CCP trials focused on couples having their first child and delivered CCP in concert with birthing classes and home visitation for new parents (part of the Australian universal health system). Three-quarters of the program was delivered in weekend workshops during the final trimester, with two home visits after the child's birth.

Participants were almost entirely middle-class couples in long-term relationships.

The American version was developed initially as part of a project delivering services to low-income, unmarried parents, which necessitated several changes. First, our partners in the obstetrics departments of two regional hospitals from which we were recruiting indicated that prenatal recruitment would miss the majority of at-risk new parents; only recruitment on maternity units would provide access to the target population. Second, the change in recruitment timing would also change program delivery timing, necessitating vast rearrangement and reworking of the material to transform a largely prenatal program into a postnatal one. Third, the program content would have to be simplified and clarified to reduce complexity both in language and in the ways in which concepts were covered. Fourth, similarly, pilot testing indicated that the highly didactic style of the original Australian CCP relationship education videos was a poor fit for our couples. Instead, the American videos were professionally produced by a New York-based film production company in a documentary style using American couples (with diverse race/ethnicity backgrounds, socioeconomic statuses, living situations, and relationship statuses) who had participated in CCP pilot testing. Fifth, content explicitly addressing conflict escalation and IPV was added, including how to use a “pause, calm, and think” time-out-like strategy.

This RCT recruited couples who had not yet experienced, but were at elevated risk for, physical clinically significant IPV (CS-IPV; IPV acts resulting in injury, fear; or that have a high inherent potential for injury, as operationalized in psychiatric and health diagnostic systems; see Heyman et al. 2015). The following hypotheses were tested: couples receiving CCP, compared with control couples, will report (a) fewer first occurrences of physical CS-IPV; (b) less frequent instances of physical and psychological IPV; (c) improved functioning on a host of IPV-risk/protective factors (e.g., relationship satisfaction, dysfunctional relationship attributions, self-regulation, communication); and (d) decreased exposure of children to couple conflict.

Method

Participants

Trained research assistants recruited participants between September 2008 and October 2010 at maternity units in two large hospitals in the exurbs of New York City; follow-up continued until October 2012. Recruiters visited each maternity unit daily beginning mid-morning (after doctors completed rounds, breakfast trays have been collected, etc., but before visiting hours began). Recruiters knocked on the door of rooms that were not marked

requesting privacy. The recruiter introduced the program and asked if the mother would like to determine if she and her partner might be eligible. If interested, the mother was asked the screening questions. If the couple appeared to be eligible, the recruiter showed a professionally produced several minute promotional video describing the program and left an informational flier. Fathers were screened subsequently but before baseline assessment.

New parents in a committed relationship were invited to participate if (a) they could speak English, (b) at least one member was aged 30 years or younger, (c) at least one member had been verbally aggressive toward the other in the previous 6 months (based on self- or partner-report), and (d) they reported no male-to-female physical CS-IPV ever (see Fig. 1 for the CONSORT diagram, which shows the flow of participants through screening, assessment, and analysis). Couples who completed the baseline assessment were then randomized either to the eight-session CCP intervention ($n = 188$) or a 24-month waitlist control group ($n = 180$). Prior to the study, power analyses were conducted with the Optimal Design software (Raudenbush et al. 2011), suggesting good power (.80) given 300 dyads, four assessments, a modest intra-class correlation of .05, and a small standardized effect size (.25).

At least one partner completed one or more assessments on which the present analyses were based. On average, couples were established (living together $M = 5.40$ years; $SD = 3.42$; 59% married) and in their late twenties (men: $M = 29.31$, $SD = 5.23$; women: $M = 26.76$, $SD = 3.78$). Participants' racial/ethnic self-identification—for men and women, respectively—was as follows: non-Latino African-American (19% and 15.9%), Hispanic/Latino of any race (22.3% and 17.6%), non-Latino White (53.2% and 59.07%), and non-Latino multiracial/other (5.5% and 7.4%). About one-third of participants had undergraduate or advanced degrees (29.8% of men and 38.3% of women) and median annual family income was \$56,000 (interquartile range = \$30,600 to \$93,760), close to the national median but substantially lower than that for their county of residence (\$99,474 in the US Census 2007–2011 American Community Survey). About half of the pregnancies were reported as unplanned (48.8%). The study children were 50.4% girls and 49.6% boys; participants had between 1 and 6 ($M = 1.74$, $SD = 1.01$) children.

Procedure

All study procedures were approved by the university institutional review board.

Assessments and Randomization

Before recruitment began, Dr. Heyman created a password-protected spreadsheet with Microsoft Excel using the random

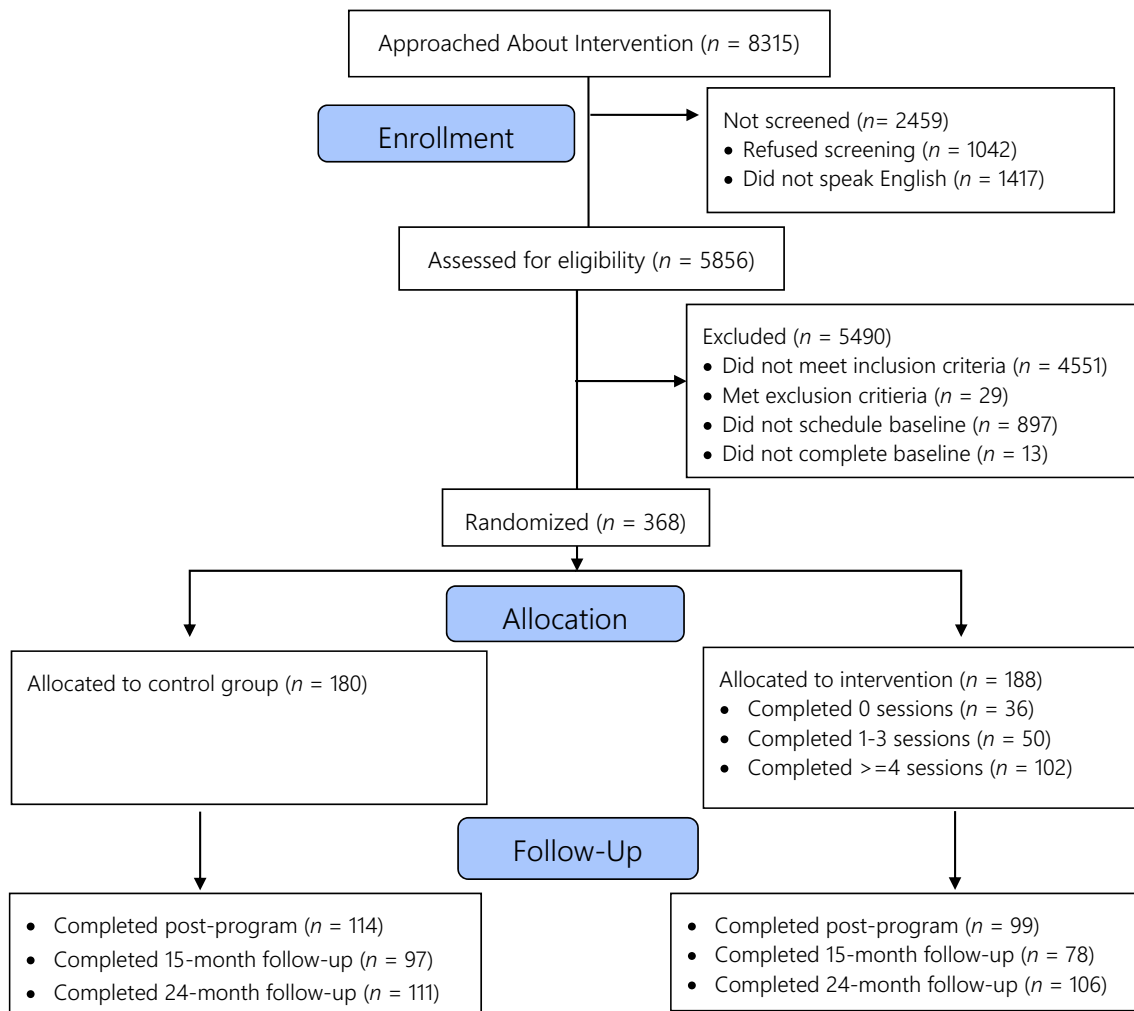


Fig. 1 CONSORT diagram of the CCP program

number function to create a list assigning 300 recruited couples to intervention or control. When recruitment exceeded the original target, the procedure was repeated for an additional 100 couples (i.e., *not* using block randomization, which led to slightly more couples being assigned to intervention over control [$n = 188$ vs. 180]). For both batches, random group assignments were placed in ordinally numbered, sealed envelopes prior to assessments began for that batch. All assessments were conducted by research assistants in the homes of participants, who were paid up to \$175 per person for completing four questionnaire assessments over the course of the study: baseline (when child was < 3 months old), when the child was 8 (post-program assessment), 15 (6-month follow-up), and 24 (16-month follow-up) months old. After consenting to participate and completing the baseline assessment, the research assistant opened a sealed envelope containing the assignment: CCP (intervention group) or CCP for toddlers (control group, wherein they would be offered a toddler program after the 24-month assessment period was completed).

CCP Intervention

The American version of CCP comprised eight sessions during the baby's first 8 months. Sessions were free and participants were not paid for attending. Sessions 1 and 4 were 1-h home visits; the others were typically conducted via 30–60 min telephone calls. Sessions began after the first assessment (when newborns were ≤ 3 months old) and were scheduled 1–3 weeks apart, with early sessions being more closely spaced.

Sessions 1–7 comprised 2–3 segments, with each segment including a video (typically viewed prior to the session on a pre-distributed DVD) that introduced key relationship or parenting skills. Video segments were typically 5–7-min long and included didactic content and demonstrations of the skills targeted in that session (e.g., playing with a young baby, asking for support). Couples watched the videos and completed activities from their workbooks prior to the session and discussed with the coach at the next session. The coach (a) clarified any concepts with which the couple may have been

struggling, (b) helped the couple identify and implement self-change objectives, and (c) was a source of support and knowledge during this challenging transition. Session 8 aimed to solidify prior gains and plan for maintaining them into the future.

Measures

In addition to the questionnaires completed by participants, CCP coaches, supervisors, and participants completed brief questionnaires about the intervention process after each session.

Demographic Factors

Demographic information included number of residents in the household, marital status, whether the pregnancy was planned, and maternal and paternal age and education (1 = some high school, 2 = high school graduate/GED, 3 = some college/vocational school, 4 = college graduate, 5 = some graduate school, and 6 = graduate degree received).

Outcomes: IPV and Relationship Functioning

Psychological and Physical IPV

The Revised Conflict Tactics Scale (CTS2; Straus et al. 1996) has been widely used in studies of IPV (e.g., Nixon et al. 2004; Suvak et al. 2013), and its scores consistently correlate with several factors in the nomological network of couple IPV (O'Leary et al. 2007). Two subscales, psychological IPV (e.g., sworn or cursed) and physical IPV (e.g., pushed), were used. Participants reported both perpetration and victimization frequencies in the past 6 months on a 7-point Likert-type scale (0 = *never*, 1 = *1 time*, 2 = *2 times*, 3 = *3–5 times*, 4 = *6–10 times*, 5 = *11–20 times*, and 6 = *more than 20 times*). Response categories 3 through 6 of physical IPV were combined to decrease the level of skewness in original responses; thus, the physical IPV variables are treated as ordinal in analyses. Item averages were calculated for the psychological IPV perpetration and victimization subscales. The psychological and physical IPV scores used for analysis were calculated based on the higher of one partner's perpetration responses and his/her victimization responses. CS-IPV was coded as present if either self report of perpetration or partner report of victimization of acts resulting in fear, injury, or potential for injury were endorsed (see Table 1 (available online) for complete list of items).

Relationship Satisfaction

The Couples Satisfaction Index (CSI; Funk and Rogge 2007) is a 32-item measure of intimate relationship satisfaction. One

global item uses a Likert-type scale (ranging from 0 = *extremely unhappy* to 6 = *perfect*) and the other items use a 6-point scale. It was developed using item response theory and correlates highly with several other validated measures of relationship satisfaction. CSI scores were the sum of participants' responses to 32 items, with higher CSI scores indicating greater satisfaction. Mean Cronbach's α s across time were .96 and .97 for males and females, respectively. Of partners who completed the baseline assessment, 14.2% of male partners and 25.1% of female partners reported clinically significant relationship distress (i.e., CSI scores less than 104.5; Funk and Rogge 2007).

Dysfunctional Relationship Attributions

The Relationship Attribution Measure (RAM; Fincham and Bradbury 1992) assesses responsibility and negative intent attributions for negative partner behavior. RAM scores are related to IPV (e.g., Holtzworth-Munroe and Hutchinson 1993) and have adequate 1-year stability (Fincham and Bradbury 1992). RAM scores in analyses were based on the item average of a shortened RAM measure (6 "partner criticizes" items and 6 "partner not paying attention" items). Participants responded using a Likert-type scale from 1 (*strongly disagree*) to 6 (*strongly agree*). Higher scores indicated higher levels of dysfunctional attributions. Mean α s across time were .91 for both males and females.

Self-Regulation in Relationships

The Behavioral Self-Regulation for Effective Relationship Scale (SRS; Wilson et al. 2005) is a 16-item self-report measure of capacity to assess one's own behavior in an intimate relationship and to set and implement relationship-enhancing goals. SRS self-report and partner-report scores prospectively predict relationship satisfaction (e.g., Halford et al. 2010). The item average (on 1 [*not at all true*] to 5 [*very true*] scale) was used for analysis, with higher scores indicating a more active approach to self-regulation. Mean α s across time were .86 for males and .85 for females.

Couple Communication and Conflict

The Conflicts and Problem-Solving Scales (CPS; Kerig 1996) assesses non-violent partner conflict strategies. The CPS has demonstrated convergent validity and 3-month test-retest reliability (Kerig 1996). Four subscales were used: Collaboration (8 items; mean α s across time were $\alpha = .77$ and $.79$ for males and females, respectively), Stalemate (7 items; $\alpha = .77$ and $.68$), Avoidance-Capitulation (8 items; $\alpha = .70$ and $.74$), and Child Involvement in Conflict (5 items; $\alpha = .76$ and $.69$). Each item was rated for self and partner on a scale from 0 (*never*) to 3 (*often*). Item averages, within and across reporter, were used

for analysis, with lower scores in Stalemate, Avoidance-Capitulation, and Child Involvement in Conflict, and higher scores in Collaboration indicating more positive behaviors.

Moderators

Cumulative Risk

A baseline cumulative risk index was based on five risks: high school education or less, family income equal or less than 150% of the federal poverty level, unplanned pregnancy, and scores in the highest quartiles of parent-child bonding problems and physical IPV at baseline. Each of the five risk factors was scored 1/0 (present/absent) and a cumulative risk factor score was calculated as the item average for each couple. This index was adapted from Petch et al. (2012b) to enhance comparability across evaluations of CCP. We did not include family of origin divorce or emotional distress during pregnancy as Petch et al. (2012b) did, and we added parent-child relationship functioning.

Intervention Process Variables

The intervention process measures are index measures that do not assume internal consistency; thus, we do not report alphas.

Fidelity

Intervention fidelity was assessed via coach ratings after each session and supervisor ratings of a random subsample (20%) of sessions. Nine items, including coach behaviors and session characteristics (e.g., “Coach reviewed reflections from previous unit”), were rated as “yes” (1) or “no” (0). All coach and supervisor ratings were averaged across sessions to form a composite measure, with higher scores indicating greater fidelity.

Alliance

Therapeutic alliance was assessed via coach, supervisor, and participant ratings on seven items (e.g., “The coach created an environment that encouraged the participants to open up”); ratings used a 5-point scale (ranging from 0 = *not at all* to 4 = *very much*). All coach and supervisor ratings were averaged across sessions to form a composite measure of alliance. Participant ratings were only made after sessions 2, 4, and 7 and were averaged.

Engagement

Engagement was assessed via supervisor and coach ratings of a random subsample (20%) of sessions. Three items (e.g., “Participants completed the assigned homework from the

previous session”) were rated as “yes” (1) or “no” (0). Coach and supervisor ratings were averaged across sessions, with higher scores indicating greater engagement.

Data Analysis

All hypothesis tests were conducted via structural equation modeling using Mplus version 7 software with full information robust maximum likelihood estimation to handle distributional nonnormality and missing values. All 368 couples who were randomized were included in analyses. Three statistical methods were used to model intervention effects: (a) intent to treat (ITT) assuming missing data were missing at random (MAR; Rubin 1976); (b) complier average causal effect (CACE) assuming MAR; and (c) CACE assuming latent ignorability (LI; Frangakis and Rubin 1999). Each method is described in more detail in the Data Analysis Supplement (available online). As shown in Fig. 1, the level of missing data was as follows: baseline: ($n = 20/188$ [11%] CCP and $n = 24/180$ [11%] control couples); post-program ($n = 89/188$ [47%] CCP and $n = 66/180$ [37%] control couples); 15-month follow-up ($n = 110/188$ [59%] CCP and $n = 83/180$ [46%] control couples); and 24-month follow-up ($n = 82/188$ [44%] CCP and $n = 69/180$ [38%] control couples).

CACE models intervention noncompliance (Jo and Muthén 2001). Compliance status is treated as a partially observable dichotomous variable. CACE divides the sample into partially latent “compliers” (i.e., couples who either completed four or more sessions [when in the intervention group] or would have if given the opportunity [when in the control group]) versus “non-compliers” (i.e., couples who either did not complete four sessions [when in the intervention group] or would not have if given the opportunity [when in the control group]). Four intervention sessions (54.0% of intervention group participants) were the threshold used because (a) it marked having attended more than half of the content-based sessions and (b) it included all the segments on couple communication and conflict management.

In CACE using the MAR assumption, response rates (i.e., provision of outcome data) were assumed to be equal for the compliers and non-compliers. Given that the mechanism of missing data in an intervention trial is unknown and response rates among compliers in the intervention group were much higher than those among non-compliers in the intervention group, we also used LI, assuming that the probabilities of observing missing values depend on the observed and the partially latent compliance class indicator and that the response rates of the compliers in the intervention and the control groups were equal.

Within each of the three analytic methods, we tested for intervention main effects on IPV (psychological and physical) and other relationship outcomes (relationship satisfaction, dysfunctional relationship attributions, self-regulation in

relationships, and couple conflict) for males and females. We then examined the moderation of these effects by cumulative risk, with follow-up analyses to probe significant interactions further. In addition, we tested for intervention effects on the presence of physical CS-IPV at post-program and follow-up [6-month and 16-month] assessments.

Figure 2 (available online) shows a path diagram of the CACE models. The intervention effect among compliers is indicated by the CCP effect on outcomes, controlling for the partially latent compliance status, and assuming that CCP had no effect among non-compliers in the intervention group. In our study, the probit link was used to estimate the association between an ordinal follow-up measure (e.g., physical IPV) and the intervention. In all other cases, the outcome measure was continuous (e.g., psychological IPV) and an identity link was used.

Covariates

Covariates were selected based on their ability to distinguish among complier, non-complier, and control couples (Jo and Muthén 2001). We screened 46 variables that could theoretically differentiate among these groups. Variables with a group difference at $p < .10$ or better in any pairwise comparison among compliers (couples who attended 4–8 sessions), non-compliers (attended 0–3 sessions), and control participants were selected as covariates in models estimating intervention effects. The selected covariates were (a) demographic: number of residents in the household, maternal age, maternal and paternal education, marital status, and whether the study child was the result of an unplanned pregnancy; (b) relationship: couple average dysfunctional relationship attributions; and (c) neonatal factors: couple average parent-infant bonding, infant distress to limitations and recovery from reactivity, child-related rigidity. (Descriptive statistics and references for these covariates at baseline are presented in Table 6 [available online].) All continuous covariates were grand mean centered. In some analyses, a given variable was removed from the baseline covariate set because it was either the moderator (e.g., unplanned pregnancy) or a component of the moderator (i.e., cumulative risk).

Effect Size

For ITT analyses, we reported the mean difference between the participants in the intervention versus control groups. For the CACE analyses, we reported the expected mean difference between the intervention and control groups among the compliers. For continuous outcomes, an effect size (d) and 95% confidence intervals were computed as the mean difference divided by the pooled standard deviation (SD) of the full sample at baseline (Jo and Muthén 2001).

To control the type-1 error rate, we applied the false discovery rate technique (FDR; Benjamini and Hochberg 1995). We defined a family of tests as tests for intervention effects on a given outcome at the last three waves of assessment (e.g., IPV at post-program and follow-ups). We report both the p -values before and after FDR correction (FDR p) in analyses where multiple follow-up measures were involved. Otherwise, only the unadjusted p value was reported.

Results

Baseline Differences in Intervention and Control Groups

As shown in Tables 3–5 (available online), the intervention and control groups were compared on 52 pre-intervention variables reflecting demographic factors, couple functioning, infant temperament, and parenting-related cognitions. After correcting for multiple testing, none of these baseline difference tests were significant.

Intervention Fidelity, Alliance, and Engagement

Coaches and supervisors reported that CCP was delivered as intended—session fidelity averaged 0.88 ($SD = 0.10$) from supervisors' ratings and 0.98 ($SD = 0.04$) from coaches' ratings. Participants, coaches, and supervisors all rated the participant-coach alliance as generally high: on a 0–4 scale, ratings for male participants were $M = 3.36$ ($SD = 0.62$), 3.26 ($SD = 0.51$), and 3.15 ($SD = 0.62$), respectively; for female participants, $M = 3.27$ ($SD = 0.61$), 3.18 ($SD = 0.54$), and 3.12 ($SD = 0.61$), respectively. Finally, participants were generally rated as engaged in the CCP content by coaches (males: $M = 0.94$, $SD = 0.11$; females: $M = 0.94$, $SD = 0.10$) and supervisors (males: $M = 0.94$, $SD = 0.14$; females: $M = 0.94$, $SD = 0.13$).

CCP as Primary Prevention of Clinically Significant Physical IPV

CCP had no significant effect on preventing the first occurrence of physical CS-IPV in ITT or CACE analyses (Table 2 [available online]). Across post-program and follow-ups, 18.4% of CCP and 18.6% of control couples developed physical CS-IPV (Table 1 [available online]).

Main Effects of CCP Intervention on IPV and Relationship Functioning

Table 7 (available online) presents descriptive statistics for outcomes.

ITT Main Effects

Using ITT, none of the hypothesized main effects on IPV (Table 8 [available online]) and other relationship functioning outcomes (Table 9 [available online]) were significant. Thus, assignment to the intervention group had no effect on target outcomes.

CACE Main Effects

Similarly, using CACE, no intervention main effects (under either MAR or LI) were significant for physical or psychological IPV (Table 8) nor for most of the relationship functioning outcomes (Table 9). However, the intervention may have worsened partners' collaboration with each other at the follow-ups (see Table 9). CCP had a significant (small- to medium-sized) iatrogenic impact on both men's and women's collaboration at 15 months assuming MAR (and $FDR\ p < .10$ under LI). A similar iatrogenic effect on collaboration was found for men at 24 months under MAR (and $FDR\ p < .10$ under LI).

Moderation Effects

With one exception, cumulative risk did not moderate effects of CCP on IPV (Table 10 [available online]) or other aspects of relationship functioning (lower panel of Table 9 [available online]). For male-to-female physical IPV at post-program, the intervention \times cumulative risk interaction was significant under LI (and $FDR\ p < .10$ under MAR). We examined the Johnson-Neyman (J-N) regions of significance and confidence bands (Preacher et al. 2006) to explore the nature of the interaction (see Figure 3 [available online]). Among the complier class, CCP reduced male-to-female physical IPV for men with lower cumulative risk but increased it in men with higher cumulative risk. Significant beneficial effects were found in men with cumulative risk at or below the 19th percentile whereas significant iatrogenic effects were found in men with cumulative risk at or above the 69th percentile. To better understand this finding, we unpacked the cumulative risk variable at post-program. Among the five components of cumulative risk, only unplanned pregnancy was a significant moderator (Table 11 [available online]). Follow-up analyses indicated that if the pregnancy was planned, CCP resulted in decreased male-to-female physical IPV (LI: simple slope = $-.67$, $p = .04$); if unplanned, CCP may have resulted in increased IPV (LI: simple slope = $.48$, $p = .09$).

Discussion

Basing public policy on science is an increasingly contentious issue in the USA (e.g., Gluckman 2016). Although policy is

messy—impacted by many factors other than evidence (e.g., ideology, lobbying, experience; Clark and Haby 2016)—science is also messy. Not only are study findings sometimes contradictory, but also scientific journals have a well-documented bias toward publishing studies with salient and statistically significant outcomes (e.g., Rosenthal 1979) and researchers are also less likely to write and submit studies with null findings (the “file drawer” problem; Franco et al. 2014). For behavioral health interventions, this leads to the overestimation of intervention effects (e.g., Driessen et al. 2015). Only when science “works” by reporting and publishing non-confirmatory reports on issues of high public policy importance can stakeholders expect that policy can “work” by making the correct cost/benefit decisions and revising those decisions when new evidence is presented that calls the original decision into question.

One such policy that may be informed by the analyses presented in this paper is the promotion of relationship education. Since 2000, US state and federal funders (see Johnson 2012) have spent close to one billion dollars on relationship programs for at-risk couples. The initial policy decisions to begin these initiatives were made before established science could soundly recommend these programs, and subsequent evaluations showing lack of impact in target populations (e.g., Lundquist et al. 2014; Wood et al. 2010) have failed to change the continued funding of the initiatives (e.g., Johnson 2014). Although recent meta-analytic studies have found positive treatment effects of prevention programs delivered to low-income couples, the overall between-group effect sizes are quite small (e.g., $d = .06$ on self-report measures of relationship outcomes; Hawkins and Erickson 2015). Importantly, Hawkins and Erickson (2015) did not find significant treatment effects in studies with samples similar to those targeted by some initiatives (i.e., unmarried couples, couples in a relationship of shorter duration, majority of couples below the federal poverty line). Thus, the implementation of traditional relationship education programs in less advantaged couples should be carefully considered.

This RCT tested a relationship education program for an at-risk population. We used an adapted version of a program for parents of newborns with prior demonstrated efficacy—Couple CARE for Parents (Halford et al. 2015). Our version of CCP was delivered with high intervention fidelity, good participant-coach alliance, and high participant engagement. Further, we employed state-of-the-science analytic tools to deal with variable session attendance (i.e., CACE analyses) in addition to the more traditional intent-to-treat analyses.

Participating couples were all at elevated risk for IPV and the overarching aim of this RCT was to test the primary preventive impact of CCP for clinically significant physical IPV (CS-IPV) at a high-risk developmental stage (i.e., parenting a newborn). However, contrary to hypotheses, CCP did not have a primary preventive main effect on physical CS-IPV,

nor did it have a secondary preventive main effect on physical or psychological IPV acts. Further, we did not find the hypothesized more proximal main effects on IPV risk and protective factors (i.e., relationship satisfaction, dysfunctional attributions, self-regulation, and communication) nor on exposure of children to non-aggressive couple conflict. Using CACE analyses, there may have even been some scattered iatrogenic effects, although given the number of tests and the effects in question being significant at only a single assessment wave, these may have been spurious even with efforts taken to limit the false discovery rate (Benjamini and Hochberg 1995).

In a large Australian RCT, CCP led to significantly less decline in relationship satisfaction in women, and greater reduction in clinically significant relationship distress in couples, classified as high- versus low-risk (Petch et al. 2012b). We examined whether risk for IPV moderated intervention effects in the current study using an adaptation of Petch et al.'s (2012b) index. Although we did find a moderated intervention effect on male-to-female IPV in the current study at post-intervention, our IPV result was in the opposite direction of Petch et al.'s (2012b) relationship satisfaction result. In our RCT, CCP participation reduced the commission of physical IPV acts by men with lower levels of cumulative risk (comprising low education, low income, unplanned pregnancy, poor bonding, and physical IPV), yet *increased* physical IPV by men with higher levels of cumulative risk. Thus, while participating in CCP, low-risk couples experienced some benefits, but the benefits were not enduring; those with high levels of cumulative risk may have suffered transient harm from participating. Further examination of the individual risk factors revealed that only unplanned pregnancy was a significant moderator of CCP's effect on male-to-female physical IPV—CCP decreased men's physical IPV from pre- to post-program in couples with planned pregnancies, but did not change physical IPV in those with unplanned pregnancies.

There are a number of methodological differences between the Australian trials and the current RCT that may explain the discrepancies in findings. First, there are major differences in the demographics of the samples recruited in the Australian trials compared with the current trial. The Australian trials included predominantly White middle-class couples who were satisfied and fairly established in their relationships (Petch et al. 2012b). In contrast, the current sample was racially diverse and young, and reported incomes near the threshold to be considered self-sufficient in the high-cost ex-urban county in which they lived. We did not exclude couples who were unhappy in their relationship; over one quarter of women in the sample reported clinically significant relationship distress in the initial assessment. Additionally, we did not require couples to be first-time parents. It is possible these demographic differences account for differences in outcome in the current investigation. Infrastructure for program delivery likely had an impact on recruitment and retention rates. In

Australia, CCP was incorporated into the universally delivered pre- and post-natal home visitation program for parents of newborns. Thus, the infrastructure in which to insert the program was already in place and well-established. In this US trial, young couples were asked to participate in the intervention at an extremely busy and stressful time, after the birth of a child, without such a surrounding program-delivery infrastructure, and it is reasonable to assume that this affected both recruitment and retention.

In addition to sample differences, there were also differences in the timing of assessments and CCP sessions in the current study compared with the Australian trials of CCP. Almost the entirety of the Australian version of CCP was delivered during pregnancy. This (a) ensures high dosage, (b) makes dropout nearly impossible, and (c) occurs at a considerably less stressful time (particularly for first-time parents, as in the Australian trials). The obstetricians at the university hospital with whom we partnered strongly dissuaded us from recruiting prenatally, since a large proportion of couples delivering there made sporadic, if any, contact with medical doctors during pregnancy. In large part due to targeting couples after the birth of a child, our timing of outcome assessments was also quite different from that in the Australian trials. Research on trajectories of change in relationship outcomes across the transition to parenthood has typically conducted assessments during pregnancy and again when the baby is 4–5 months old (Mitnick et al. 2009). The recent Australian trial examined the efficacy of CCP based on change in outcome across an equivalent time period (Petch et al. 2012b). In contrast, couples *began* CCP in the current study around when the post-intervention assessments occurred in this previous work. It is possible that we did not find an effect of CCP because the primary changes in outcomes had already occurred and dissipated. Consistent with this, there was little change in couples in the control condition.

US state and federal policy makers may have been using the common sense notion that relationship education would be benign at worst, making it appropriate for selective prevention efforts. CCP was designed and originally tested, for the most part, on (Australian) couples who had made lifetime commitments and were pregnant with, but had not yet delivered, their first child. In terms of men's physical IPV acts during the program period, such committed, planful couples appear to have benefited from the continued spotlight on their relationship, co-parenting, and outlook for the future. Yet, there are indications that with a diverse sample of American couples, the spotlight may have increased the likelihood of men's physical IPV for those at high levels of risk. Although our findings are too equivocal to influence intervention and policy decisions without replication, they should give all stakeholders pause to consider that the glare of a spotlight intended to be benevolent may, in fact, exert harmful pressure to vulnerable couples.

The strengths of this study were noted earlier—it tests an empirically supported intervention, delivered with fidelity and analyzed with a sophisticated approach. The limitations include transporting (with the partial collaboration of the developer) an evidence-based program into a similar-but-not-identical culture, with the attendant re-shooting and re-envisioning video content and re-ordering material (and adding material explicitly focused on reducing IPV). Although we attempted to retain the active ingredients of CCP, it is possible that we did not. Further, as described above, the shift of perinatal timing may be more impactful than expected. Other limitations include a modest sample size (368 randomized couples), modest retention in the program (of 188 couples randomized to CCP, 153 began it, and about one-third got less than half of the program), and challenges in enticing couples to participate in follow-up assessments (e.g., 106/188 couples assigned to CCP and 108/180 assigned to the control group completed the 24-month follow-up). And although variables such as relationship satisfaction and dysfunctional attributions can only be assessed via self-report and IPV can only be ethically and practically assessed via self-report, this study suffers from all the limitations of that mode of data collection (e.g., response biases, recall effects, self-presentation effects).

Finally, our principal findings emerged from advanced analytic models (e.g., CACE and LI models) that attempted to adjust for selection bias and different potential missing data mechanisms. Although ultimate establishment of these recent methodological tools is ongoing, validity, we argue that our use of techniques that attempt to limit bias in statistical inference was better than the alternative: failing to adjust for the ways in which some participants' behavior (noncompliance with treatment and data collection) worked against our research design.

In conclusion, this RCT adds to the growing literature indicating that relationship education may *not* be effective for at-risk new parents as a stand-alone intervention—neither as primary prevention for injurious IPV nor for bettering relationships. Policy makers might want to consider these and other empirical findings of non-effectiveness when considering the most appropriate investment to prevent violence and improve relationships among families at risk.

Funding This work was supported by a grant from the Centers for Disease Control and Prevention (5U49CE00124602).

Compliance with Ethical Standards

Conflict of Interest W. Kim Halford is the co-developer of Couple CARE and Couple CARE for Parents programs and receives royalties from the sale of these programs. He is also the Principal Investigator on grants which provide funding for intervention research on Couple CARE for Parents program. All the other authors declare that they have no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained for all individual participants included in the study.

Disclaimer The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Bair-Merritt, M. H., Jennings, J. M., Chen, R., Burrell, L., McFarlane, E., Fuddy, L., & Duggan, A. K. (2010). Reducing maternal intimate partner violence after the birth of a child: a randomized controlled trial of the Hawaii Healthy Start Home Visitation Program. *Archives of Pediatrics & Adolescent Medicine*, *164*, 16–23.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B*, *57*, 289–300.
- Betz, C. (2007). Teen dating violence: An unrecognized health care need [editorial]. *Journal of Pediatric Nursing*, *22*, 427–429. <https://doi.org/10.1016/j.pedn.2007.10.001>.
- Beydoun, H. A., Williams, M., Beydoun, M. A., Eid, S. M., & Zonderman, A. B. (2017). Relationship of physical intimate partner violence with mental health diagnoses in the nationwide emergency department sample. *Journal of Women's Health*, *26*, 141–151. <https://doi.org/10.1089/jwh.2016>.
- Charles, P., & Perreira, K. M. (2007). Intimate partner violence during pregnancy and 1-year post-partum. *Journal of Family Violence*, *22*, 609–619. <https://doi.org/10.1007/s10896-007-9112-0>.
- Clark, R., & Haby, M. (2016). Evidence-based policy: Why and how? In G. Carey, K. Landvogt, & J. Barraket (Eds.), *Creating and implementing public policy: Cross-sectoral debates* (pp. 98–112). New York: Routledge.
- Dion, M. R., Avellar, S. A., & Clary, E. (2010). *Implementation of eight programs to strengthen unmarried parent families*. Washington, DC: Mathematica Policy Research.
- Doss, B. D., Rhoades, G. K., Stanley, S. M., & Markman, H. J. (2009). The effect of the transition to parenthood on relationship quality: An eight-year prospective study. *Journal of Personality and Social Psychology*, *96*, 601–619. <https://doi.org/10.1037/a0013969>.
- Driessen, E., Hollon, S. D., Bockting, C. L. H., Cuijpers, P., & Turner, E. H. (2015). Does publication bias inflate the apparent efficacy of psychological treatment for major depressive disorder? A systematic review and meta-analysis of US National Institutes of Health-funded trials. *PLoS One*, *10*, e0137864–e0137823. <https://doi.org/10.1371/journal.pone.0137864>.
- Dutton, M. A., Green, B. L., Kaltman, S. I., Roesch, D. M., Zeffiro, T. A., & Krause, E. D. (2006). Intimate partner violence, PTSD, and adverse health outcomes. *Journal of Interpersonal Violence*, *21*, 955–968. <https://doi.org/10.1177/0886260506289178>.
- Feinberg, M. E., Kan, M. L., & Goslin, M. C. (2009). Enhancing coparenting, parenting, and child self-regulation: Effects of family foundations 1 year after birth. *Journal of Family Psychology*, *22*, 253–263. <https://doi.org/10.1007/s11212-009-0130-4>.
- Feinberg, M. E., Jones, D. E., Hostelter, M. L., Roettger, M. E., Paul, I. M., & Erenthal, D. B. (2016). Couple-focused prevention at the

- transition to parenthood, a randomized trial: Effects on coparenting, parenting, family violence, and parent and child adjustment. *Prevention Science*, 17, 751–764. <https://doi.org/10.1007/s11121-016-0674-z>.
- Fincham, F. D., & Bradbury, T. N. (1992). Assessing attributions in marriage: The relationship attribution measure. *Journal of Personality and Social Psychology*, 62, 457–468. <https://doi.org/10.1037/0022-3514.62.3.457>.
- Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science*, 345, 1502–1505. <https://doi.org/10.1126/science.1255484>.
- Frangakis, C. E., & Rubin, D. B. (1999). Addressing complications of intention-to-treat analysis in the combined presence of all-or-none treatment-noncompliance and subsequent missing outcomes. *Biometrika*, 86, 365–379. <https://doi.org/10.1093/biomet/86.2.365>.
- Funk, J. L., & Rogge, R. D. (2007). Testing the ruler with item response theory: Increasing precision of measurement for relationship satisfaction with the Couples Satisfaction Index. *Journal of Family Psychology*, 21, 572–583. <https://doi.org/10.1037/0893-3200.21.4.572>.
- Gluckman, P. (2016). The science–policy interface. *Science*, 353, 969. <https://doi.org/10.1126/science.aai8837>.
- Halford, W. K., Markman, H. J., Kline, G. H., & Stanley, S. M. (2003). Best practice in couple relationship education. *Journal of Marital and Family Therapy*, 29, 385–406. <https://doi.org/10.1111/j.1752-0606.2003.tb01214.x>.
- Halford, W. K., Moore, E. M., Wilson, K. L., Dyer, C., & Farrugia, C. (2004). Benefits of a flexible delivery relationship education: An evaluation of the Couple CARE program. *Family Relations*, 53, 469–476. <https://doi.org/10.1111/j.0197-6664.2004.00055.x>.
- Halford, W. K., Lizzio, A., Wilson, K. L., & Occhipinti, S. (2007). Does working at your marriage help? Couple relationship self-regulation and satisfaction in the first 4 years of marriage. *Journal of Family Psychology*, 21, 185–194. <https://doi.org/10.1037/0893-3200.21.2.185>.
- Halford, W. K., Heyman, R. E., Slep, A. M. S., Petch, J., & Creedy, D. K. (2009). *Couple CARE for Parents of Newborns*. Unpublished workbook. New York: New York University.
- Halford, W. K., Petch, J., & Creedy, D. K. (2010). Promoting a positive transition to parenthood: A randomized clinical trial of couple relationship education. *Prevention Science*, 11, 89–100. <https://doi.org/10.1007/s11121-009-0152-y>.
- Halford, W. K., Petch, J., & Creedy, D. K. (2015). *Clinical guide to helping new parents: The Couple CARE for Parents Program*. New York: Springer.
- Hawkins, A. J., & Erickson, S. E. (2015). Is couple and relationship education effective for low-income participants? A meta-analytic study. *Journal of Family Psychology*, 29, 59–68. <https://doi.org/10.1037/fam0000045>.
- Heyman, R. E., Slep, A. M. S., & Foran, H. M. (2015). Enhanced definitions of intimate partner violence for DSM-5 and ICD-11 may promote improved screening and treatment. *Family Process*, 54, 64–81. <https://doi.org/10.1111/famp.12121>.
- Holtzworth-Munroe, A., & Hutchinson, G. (1993). Attributing negative intent to wife behavior: The attributions of maritally violent versus nonviolent men. *Journal of Abnormal Psychology*, 102, 206–211. <https://doi.org/10.1037/0021-843x.102.2.206>.
- Jo, B., & Muthén, B. O. (2001). Modeling of intervention effects with noncompliance: A latent variable approach for randomized trials. In G. A. Marcoulides & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp. 57–87). London: Lawrence Erlbaum Associates.
- Johnson, M. D. (2012). Healthy marriage initiatives: On the need for empiricism in policy implementation. *American Psychologist*, 67, 296–308. <https://doi.org/10.1037/a0027743>.
- Johnson, M. D. (2014). Government-supported healthy marriage initiatives are not associated with changes in family demographics: A comment on Hawkins, Amato, and Kinghorn (2013). *Family Relations*, 63, 300–304. <https://doi.org/10.1111/fare.12060>.
- Kan, M. L., & Feinberg, M. E. (2014). Can a family-focused, transition-to-parenthood program prevent parent and partner aggression among couples with young children? *Violence and Victims*, 29, 967–980. <https://doi.org/10.1891/0886-6708.VV-D-12-00162>.
- Kerig, P. K. (1996). Assessing the links between interparental conflict and child adjustment: The conflicts and problem-solving scales. *Journal of Family Psychology*, 10, 454–473. <https://doi.org/10.1037/0893-3200.10.4.454>.
- Lawrence, E., & Bradbury, T. N. (2007). Trajectories of change in physical aggression and marital satisfaction. *Journal of Family Psychology*, 21, 236–247. <https://doi.org/10.1037/0893-3200.21.2.236>.
- Lundquist, E., Hsueh, J., Lowenstein, A. E., Faucetta, K., Gubits, D., Michalopoulos, C., & Knox, V. (2014). *A family-strengthening program for low-income families: Final impacts from the Supporting Healthy Marriage evaluation*. OPRE Report 2014-09A. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Mitnick, D. M., Heyman, R. E., & Slep, A. M. S. (2009). Changes in relationship satisfaction across the transition to parenthood: A meta-analysis. *Journal of Family Psychology*, 23, 848–852. <https://doi.org/10.1037/a0017004>.
- Moore, Q., Wood, R. G., Clarkwest, A., Killewald, A., & Monahan, S. (2012). *The long-term effects of Building Strong Families: A relationship skills education program for unmarried parents (technical supplement)*. Washington, DC: Mathematica.
- Nixon, R. D., Resick, P. A., & Nishith, P. (2004). An exploration of comorbid depression among female victims of intimate partner violence with posttraumatic stress disorder. *Journal of Affective Disorders*, 82, 315–320. <https://doi.org/10.1016/j.jad.2004.01.008>.
- O’Leary, K. D., Slep, A. M. S., & O’Leary, S. G. (2007). Multivariate models of men’s and women’s partner aggression. *Journal of Consulting and Clinical Psychology*, 75, 752–764. <https://doi.org/10.1037/0022-006X.75.5.752>.
- Petch, J., Halford, W. K., Creedy, D. K., & Gamble, J. (2012a). Couple relationship education at the transition to parenthood: A window of opportunity to reach high-risk couples. *Family Process*, 51, 498–511. <https://doi.org/10.1111/j.1545-5300.2012.01420.x>.
- Petch, J., Halford, W. K., Creedy, D. K., & Gamble, J. (2012b). A randomized controlled trial of a couple relationship and coparenting program (Couple CARE for Parents) for high- and low-risk new parents. *Journal of Consulting and Clinical Psychology*, 80, 662–673. <https://doi.org/10.1037/a0028781>.
- Pinquart, M., & Teubert, D. (2010). A meta-analytic study of couple interventions during the transition to parenthood. *Family Relations*, 59, 221–231. <https://doi.org/10.1111/j.1741-3729.2010.00597.x>.
- Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics*, 31, 437–448. <https://doi.org/10.3102/1076986031004437>.
- Raudenbush, S.W., Spybrook, J., Congdon, R., Liu, X., Martinez, A., Bloom, H., & Hill, C. (2011). *Optimal Design Plus Empirical Evidence (Version 3.0)*. [Software]. Retrieved on December 5, 2018 from <https://www.wtgrantfoundation.org>.
- Rennison, C. M., & Addington, L. A. (2014). Violence against college women: A review to identify limitations in defining the problem and inform future research. *Trauma, Violence, & Abuse*, 15, 159–169. <https://doi.org/10.1177/1524838014520724>.

- Rosenthal, R. (1979). The “file drawer problem” and tolerance for null results. *Psychological Bulletin*, *86*, 638–641. <https://doi.org/10.1037/0033-2909.86>.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, *63*, 581–592.
- Schafer, J., Caetano, R., & Clark, C. L. (1998). Rates of intimate partner violence in the United States. *American Journal of Public Health*, *88*, 1702–1704. <https://doi.org/10.2105/ajph.88.11.1702>.
- Schulz, M. S., Cowan, C. P., & Cowan, P. A. (2006). Promoting healthy beginnings: A randomized controlled trial of a preventive intervention to preserve marital quality during the transition to parenthood. *Journal of Consulting and Clinical Psychology*, *74*, 20–31. <https://doi.org/10.1037/0022-006X.74.1.20>.
- Shapiro, A. F., & Gottman, J. M. (2005). Effects on marriage of a psycho-communicative-educational intervention with couples undergoing the transition to parenthood, evaluation at 1-year post intervention. *The Journal of Family Communication*, *5*, 1–24. https://doi.org/10.1207/s15327698jfc0501_1.
- Shapiro, A. F., Gottman, J. M., & Carrère, S. (2000). The baby and the marriage: Identifying factors that buffer against decline in marital satisfaction after the first baby arrives. *Journal of Family Psychology*, *14*, 59–70. <https://doi.org/10.1037/0893-3200.14.1.59>.
- Slep, A. M. S., Foran, H. M., Heyman, R. E., Snarr, J. D., & USAF Family Advocacy Program. (2011). Risk factors for clinically significant intimate partner violence among active duty members. *Journal of Marriage and Family*, *73*, 486–501.
- Smith, S. G., Chen, J., Basile, K. C., Gilbert, L. K., Merrick, M. T., ... Jain, A. (2017). *The National Intimate Partner and Sexual Violence Survey (NISVS): 2010–2012 State Report*. Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention.
- Straus, M. A., Hamby, S. L., Boney-McCoy, S., & Sugarman, D. B. (1996). The revised Conflict Tactics Scales (CTS2): Development and preliminary psychometric data. *Journal of Family Issues*, *17*, 283–316. <https://doi.org/10.1177/019251396017003001>.
- Sullivan, K. T., Pasch, L. A., Cornelius, T., & Cirigliano, E. (2004). Predicting participation in premarital prevention programs: The health belief model and social norms. *Family Process*, *43*, 175–193. <https://doi.org/10.1111/j.1545-5300.2004.04302004.x>.
- Sumner, S. A., Mercy, J. A., Dahlberg, L. L., Hillis, S. D., Klevens, J., & Houry, D. (2015). Violence in the United States: Status, challenges, and opportunities. *JAMA*, *4*, 478–488. <https://doi.org/10.1001/jama.2015.8371>.
- Suvak, M. K., Taft, C. T., Goodman, L. A., & Dutton, M. A. (2013). Dimensions of functional social support and depressive symptoms: A longitudinal investigation of women seeking help for intimate partner violence. *Journal of Consulting and Clinical Psychology*, *81*, 455–466. <https://doi.org/10.1037/a0031787>.
- U. S. Department of Health & Human Services. (2014). *Healthy people 2020*. Retrieved from <http://www.healthypeople.gov/>. Accessed 5 Dec 2018.
- Wilson, K. L., Chalker, J., Lizzio, A., Halford, K., & Kimlin, S. (2005). Assessing how much couples work at their relationship: The Behavioral Self-Regulation for Effective Relationships Scale. *Journal of Family Psychology*, *19*, 385–393. <https://doi.org/10.1037/0893-3200.19.3.385>.
- Wood, R. G., Moore, Q., Clarkwest, A., Hsueh, J., & McConnell, S. (2010). *Strengthening unmarried parents' relationships: The early impacts of Building Strong Families (technical supplement)*. Washington, DC: Mathematica Policy Research.