Published in final edited form as:
Fam Relat. 2014 October ; 63(4): 484–495. doi:10.1111/fare.12083.

A Randomized Controlled Trial of Relationship Education in the U.S. Army: 2-Year Outcomes

Scott M. Stanley,
University of Denver

Galena K. Rhoades,
University of Denver

Benjamin A. Loew,
University of Denver

Elizabeth S. Allen,
University of Colorado Denver

Sarah Carter, **Laura J. Osborne**,
University of Denver

Donnella Prentice, and
University of Denver

Howard J. Markman
University of Denver

Abstract

This study examined the effectiveness of an evidence-based, community-delivered adaptation of couple relationship education (CRE; specifically, PREP, The Prevention and Relationship Enhancement Program) delivered at two Army installations. The study is a randomized controlled trial with two years of follow-up, examining marital quality and stability. Sample composition was 662 married couples with a spouse in the U.S. Army. Analyses yielded no evidence of overall enduring intervention effects on relationship quality but couples assigned to intervention at the higher risk site were significantly less likely than controls to be divorced at the two-year follow-up (8.1% vs. 14.9%, $p < .01$). This effect was moderated by ethnic minority status. Specifically, the impact of the intervention on divorce was strongest for minority couples. The findings add to the literature on who may benefit most from CRE.

Keywords

Relationship education; Marriage education; Divorce; Army couples; Randomized Controlled Trial

Corresponding Author: Scott M Stanley, Ph.D., Research Professor, University of Denver, Department of Psychology, 2155 S. Race St., Denver, CO 80208. 303-692-8932, scott@stanleyemail.com, sstanley@du.edu.

Marriages in the United States are at significant risk to end in divorce and/or to become distressed (e.g., Tejada-Vera & Sutton, 2010; Whisman, Beach, & Snyder, 2008). Couple relationship education (CRE) is a preventive approach for educating couples about strategies designed to improve their odds of maintaining a successful relationship. In this paper, we present findings from a randomized controlled trial of CRE provided to married couples in the U.S. Army. This study addresses three major weaknesses in the literature on the evaluation of CRE: the scarcity of tests of long-term impacts, the absence of evidence for relationship stability impacts within fully randomized trials, and a paucity of studies capable of testing if important socio-demographic variables such as ethnicity moderate the effects of such services.

Meta-analyses of CRE have found positive effects on dimensions of relationship quality (e.g., Blanchard, Hawkins, Baldwin, & Fawcett, 2009; Hawkins, Blanchard, Baldwin, & Fawcett, 2008). Further, some studies on CRE have demonstrated that couples at higher risk may benefit more from CRE than lower-risk couples (e.g., Allen, Rhoades, Stanley, Loew, & Markman, 2012; Halford, Sanders, & Behrens, 2001; Petch, Halford, Creedy, & Gamble, 2012). While the existing literature is encouraging, there are few studies that include long-term follow-up, especially among those that use randomized designs. Further, studies documenting impacts on relationship stability (e.g., divorce) are rare, and most that exist employed quasi-experimental designs (e.g., Hahlweg, Markman, Thurmaier, Engl, & Eckert, 1998; Markman et al., 1988). One published study using a randomized controlled trial showed a divorce reduction effect at a one-year follow-up. In that study, CRE was delivered by chaplains to couples in the U.S. Army (Stanley et al., 2010). The current report is from the same project; it includes divorce outcomes for a second site not available at the time of the earlier report, and presents analyses not only for divorce but also marital quality outcomes at two years following the intervention. The CRE approach tested was an adaptation of PREP (The Prevention and Relationship Enhancement Program; e.g., Markman, Stanley, & Blumberg, 2010; Stanley, Markman, Jenkins, & Blumberg, 2006).

One of the greatest current controversies in the CRE field concerns whether ethnicity and/or economic factors moderate CRE impacts. There is evidence that some ethnic minority groups have increased risks for marital instability (Bramlett & Mosher, 2002), and that couples with greater economic disadvantages have more marital difficulties (e.g., Conger & Elder, 1994). Karney and Bradbury (2005) suggested that economically vulnerable couples were less likely to benefit from CRE because their greater levels of contextual stress would interfere with such training. Johnson (2012) reiterated this concern, and further suggested that since CRE programs were initially evaluated with mostly white, middle class couples, such services may not be effective for ethnic minority couples and couples who are more economically vulnerable. Hawkins et al. (2013) countered Johnson on numerous points, but all sides of this argument have agreed upon the need for more research and analyses testing this type of question.

Bearing on this question, in one large, recent randomized federal (U.S.) study of CRE programs for low-income unmarried couples expecting a baby, there were generally no significant findings at 15 months when comparing the intervention and control groups, pooled across eight sites (the Building Strong Families study (BSF); Wood, McConnell,

Moore, Clarkwest, & Hsueh, 2010). However, intervention couples in which both partners were African-American (the largest ethnic minority group in the study) showed significant impacts at the 15 month follow-up. In another large federal study, programs that included CRE were delivered to married couples with low income levels in eight sites (the Supporting Healthy Marriages study; SHM). As with the BSF study, this study used a randomized design with intervention and control groups. At 12 months following training, the intervention group reported significantly better relationship quality than the control group, though the effects were modest (Hsueh et al., 2012). Echoing the BSF study's results, SHM couples in which partners were Hispanic (the largest ethnic minority group in the study) tended to report larger and more consistent positive impacts than non-Hispanic couples. Hence, the best represented ethnic minority groups in each of these two large randomized trials demonstrated the most gains at 15 and 12 months, respectively.

The Present Study

In this report, we present two-year outcome findings from a randomized trial of CRE delivered by chaplains in the U.S. Army to married couples with at least one spouse in the Army. This project utilizes a large sample for this type of study (N = 662 couples), allowing for tests of relatively rare outcomes such as divorce and tests for the impact of moderators such as ethnic minority status. As the federal studies above also tested comprehensive programs that included CRE, the present study represents a purer test of CRE in a long-term design.

The project also tests an important group: a sample of U.S. Army couples during a time of war (i.e., during Operation Iraqi Freedom (OIF) and Operation Enduring Freedom (OEF)). As such, this sample may be under higher stress than many samples tested in studies of CRE and therefore reflect a higher risk group than is typically served. Given that higher risk groups sometimes benefit more, not less, from interventions, this characteristic of this study may increase the ability to detect long-term impacts.

Prior reports from the current project document modest but positive pre-to-post intervention impacts for one of the two sites in the study on some relationship quality variables (Allen, Stanley, Rhoades, Markman, & Loew, 2011) and, as noted above, significant divorce reduction for intervention couples one year after intervention (Stanley et al., 2010). Here, we present findings on relationship quality variables from five waves of data (the post-intervention assessment and 4 follow-up assessments, controlling for pre-intervention baseline levels) and on divorce through two years after intervention. Our hypotheses were that intervention couples would be less likely to divorce (Hypothesis 1) and would maintain higher relationship quality (Hypothesis 2) than control group couples.

Given the importance of questions about CRE's use with groups historically not included in CRE's development and testing, we examined if ethnicity minority status moderated impacts. Based on the findings from the two federal trials noted earlier—suggesting that ethnic minority couples may benefit somewhat more from such interventions—we expected that there would be clearer intervention impacts for ethnic minority couples in the present analyses. Additionally, we examined moderators related to economic hardship. The contexts

of couples' lives are often linked to stressors that impact relationship quality, such as financial strain (e.g., Karney & Bradbury, 2005). Thus, we examined if self-reported income and economic strain moderated intervention impacts. However, we did not make a specific prediction about the direction of moderation. On one hand, a general risk perspective would suggest that couples with economic problems may gain the most from intervention because they have higher risks (e.g., Petch, Halford, Creedy, & Gamble, 2012). On the other hand, as noted earlier, it has been argued that some couples at lower income levels or under high economic strain may not be able to benefit as much because of their contextual stress (Johnson, 2012; Karney & Bradbury, 2005).

Method

Participants

This sample included 662 Army couples. At the baseline (pre) assessment, husbands averaged 28.5 years of age (*SD* = 5.9); wives 27.7 (*SD* = 6.2). Seventy-one percent of wives were white non-Hispanic, 11% were Hispanic, 10% African American, 1.8% Native American/Alaska Native, 0.9% Hawaiian or Pacific Islander, 0.8% Asian, and 4.6% endorsed mixed race/ethnicity. Sixty-nine percent of husbands were white non-Hispanic, 12% were Hispanic, 11% African American, 1.4% Native American/Alaska Native, 1.1% Asian, 0.6% Hawaiian or Pacific Islander, and 4.6% endorsed mixed race/ethnicity. Almost all husbands (97.4%) were Active Duty Army (2.3% were in the study due to their wives being Active Duty), whereas almost all wives (91.7%) were civilian spouses of Active Duty males (8.0% of wives were Active Duty Army or Reserves). Husbands' median income (endorsed by 31.2% of men) was between $30,000 and $39,999 a year, while wives' median income (endorsed by 67.6% of women) was under $10,000 a year. High school or an equivalency degree was the median highest degree (endorsed by 69% of husbands and 55% of wives). Couples had been married an average of 4.93 years, and 74% reported at least one child living with them at least part-time.

Procedures

We evaluated the intervention with Army couples at two installations (here called Sites 1 and 2). Site 1 included 478 couples and Site 2 included 184 couples. At Site 1, intervention couples received the training in classes comprised of other couples from their unit, with their unit chaplains leading the intervention. This is the typical fashion in which these services are offered in the U.S. Army, and is the approach reinforced by the U. S. Army Chief of Chaplains office. Additionally, Site 1 consisted of major war-fighting units which were preparing for deployment for the OIF surge at the time of intervention. Although we had completed the intervention phase with a large sample at Site 1, we were concerned about the high operational tempo at Site 1 and how it might impact the quality of the intervention. Our contacts in the Army Chief of Chaplains office provided the opportunity to run a smaller sample at a second site where couples were not as exposed to the stress of ramping up for deployment, and where couples would be trained by a small group of highly-trained family life chaplains. Hence, Site 2 couples were not trained by their own unit chaplains, nor trained in groups with couples from their units; and Site 2 procedures are not as typical of

the way these intervention services are most often delivered in the Army's Strong Bonds program.

There were no significant socio-demographic differences between the intervention and control groups. Further, couples at Sites 1 and 2 were not significantly different at pre-assessment on measures of relationship quality (as reported in Allen et al., 2011). However, there were a few significant socio-demographic differences between sites: couples at Site 2 were married longer (M = 6.08 years vs. M = 4.49), were older (M = 30.54 years vs. M = 27.24), had higher incomes (median total family income, Mdn = $55,000 vs. Mdn = $45,000) and rank (18% vs. 9% officers), and had lower rates of recent deployment at pre-assessment compared to Site 1 couples (47% vs. 72%). These differences may be markers for less sociodemographic and contextual stress for the Site 2 sample. However, couples at the two sites did not differ in self-reported economic strain. When possible, site was tested as a moderator in the analyses presented below. The intervention phase was completed at Site 1 in 2007 and at Site 2 in 2008.

A detailed review of procedures and the intervention can be found in (Stanley et al., 2010). A CONSORT (Consolidated Standards of Reporting Trials) flow chart is provided in Figure 1. To be eligible for the study, couples had to be married, with both partners aged 18 or over, fluent in English, and with at least one active-duty spouse in the Army. Couples could not have already participated in PREP and had to be willing to be randomly assigned to either intervention or an untreated control group, which can be considered treatment-as-usual. The study was presented as long-term research, wherein couples would be asked to complete multiple assessments (approximately 1 hour each) over the following years, with compensation increasing over time from $50 to $90 per person per assessment. Recruitment was conducted via brochures, media stories, posters, and referrals from chaplains; recruitment materials clearly specified eligibility requirements and compensation. Consent was obtained with printed forms by a research assistant. All procedures were approved by a university Institutional Review Board.

Prior to random assignment, each spouse separately completed baseline (pre-intervention) questionnaires under the supervision of study staff, on site. After couples completed this assessment, they were randomly assigned to intervention or the control group. 343 couples were assigned to intervention and 319 to the control group. There were 27 separate iterations of the intervention (23 at Site 1 and 4 at Site 2). At Site 1, chaplains allowed investigators a number of slots for study couples within the iteration of PREP they were providing to their units. At Site 2, investigators arranged for iterations consisting only of study couples. This difference resulted in varying numbers of study couples per iteration of the intervention. Two subgroups were associated with each iteration: couples randomly assigned to that iteration of PREP and couples randomly assigned to the control group. While this design conceptually resulted in 54 "cohorts" (27 iterations X 2 subgroups of couples), one iteration at Site 1 had only two study couples, both of which were assigned to PREP, thus the total number of cohorts is 53 (45 at Site 1 and 8 at Site 2). Approximately two weeks after the respective iteration's intervention was complete, each spouse separately completed another set of measures (post) under the supervision of study staff. Approximately every six months after post (through two years post-intervention), couples whose marriages remained intact

were given measures that could be completed online from anywhere in the world. Hence, the current analyses are based on data collected at pre and post intervention and four follow-up assessments; the fourth being two years after the post assessment.

Intervention

The intervention was a version of PREP (Markman et al., 2010; Stanley et al., 2006) adapted for use by Army chaplains with Army couples. The intervention as utilized was the regular version of PREP with some added content and adaptations focused on coping with deployments and reunions. The workshop consisted of two parts: a one-day on-post training followed, usually in the next week, by a weekend retreat at an off-post hotel, totaling approximately 14.4 hours of training. PREP is based on research regarding factors related to healthy marriage; intervention modules included communication problem solving skill training, negative affect management skills, insights into relationship dynamics, emotional support, stress and relaxation, principles of commitment, fun and friendship, forgiveness, sensuality and sexuality, expectations, core beliefs, and deployment/reintegration issues. From 20 of the 27 intervention iterations, we obtained adequate audio recording to code chaplain fidelity to the lesson material, from 1 (Very poor) to 5 (Excellent). The inter-rater correlation for coding fidelity was .88, and the average fidelity rating across iterations was 4.10.

Most couples assigned to the intervention attended some portion of it. At Site 1, 60% attended all, 23% attended part (e.g., either the duty day of training or the weekend retreat, but not both), and 17% attended none of the intervention. At Site 2, 54% attended all, 35% attended part, and 11% attended none of the intervention. Two couples assigned to the control group attended the intervention at Site 1.

Measures

Divorce—Divorce status was defined by either spouse reporting that they had obtained or filed for a divorce as of the two-year follow-up assessment. At Site 1, there were 54 divorces (11.3%) and 2 couples with missing data on this variable. One participant died (not combat-related) prior to the two-year time point, so divorce status is not applicable for that couple. For another couple we could not ascertain divorce status at two years post-intervention. At Site 2, there were 10 divorces (5.4%) and no missing data. Hence the number and rate of divorces at Site 1 is much larger than at Site 2.

Communication skills—Ten items from the larger Communication Skills Test (Saiz & Jenkins, 1995) were used to measure use of the communication skills taught in PREP while avoiding PREP-specific terminology (e.g., “speaker-listener skills”). Sample items include “When discussing issues, I allow my spouse to finish talking before I respond,” “When our discussions begin to get out of hand, we agree to stop them and talk later.” Prior research supports this measure’s reliability and validity (e.g., Stanley et al., 2005). Participants rated these items on a 1 to 7 scale, with greater endorsement corresponding to greater skill use, and responses were averaged to create a total score. At the pre-intervention assessment participants’ average total score was 4.08 (*SD* = 1.16); internal consistency was .85.

Marital satisfaction—The Kansas Marital Satisfaction Scale (KMS; Schumm et al., 1986) was used to measure marital satisfaction. This scale's three items assess satisfaction with one's marriage, with one's partner as a spouse, and with one's relationship with his or her spouse. The KMS has strong reliability and validity (Schumm et al., 1986), and provides an overall rating of marital satisfaction. Participants rated these items on a 1 to 7 scale, with greater endorsement corresponding to greater satisfaction. Responses to the three questions were averaged to create the scale score. At the pre-intervention assessment, participants' average total score was 5.67 (*SD* = 1.27); internal consistency was .94. This mean is equivalent to a suggested cutoff score for distressed versus non-distressed couples on this measure (Crane, Middleton, & Bean, 2000).

Positive bonding—From the larger Couple Activities Scale (Markman, 2000), positive bonding was measured with nine items that assess one's friendship, intimacy, fun, felt support, and sensual/sexual relationship with his or her partner. Various studies demonstrate evidence of reliability and validity for the measure (e.g., Allen, Rhoades, Stanley, & Markman, 2010). Sample items include "We regularly have conversations where we just talk as good friends," "We have a satisfying sensual or sexual relationship," "I feel emotionally supported by my partner," and "We regularly make time for fun activities together as a couple." Participants rated these items on a 1 to 7 scale, and responses were averaged to create a scale score. At the pre-intervention assessment participants' average scale score was 5.55 (*SD* = 1.14) with higher scores indicating higher levels of positive bonding; internal consistency was .88.

Divorce proneness—Three items were adapted from the short form of the Marital Instability Index (MII; Booth, Johnson, & Edwards, 1983) to measure divorce proneness. These items assess concern for one's marriage, consideration of separation or divorce, and couple-level discussion of divorce. Participants endorsed these items as yes (1) or no (0), and responses were averaged to create a scale score. At the pre-intervention assessment participants' average scale score was 0.36 (*SD* = 0.41); higher scores indicating higher levels of divorce proneness; internal consistency was .83.

Moderators—Tested moderators were ethnic minority status, income, and economic strain. For ethnic minority status, a dichotomous variable was created to differentiate couples in which at least one spouse self-identified as an ethnic/racial minority from those in which neither spouse did. Henceforth 'ethnic minority' and 'ethnic minority status' refer to this couple-level binary variable, with 60% of the couples in the sample being non-minority couples and 40% being ethnic minority couples. Income was measured in $10,000 increments. Economic strain was measured as the mean couple-level endorsement of two items assessed on a 1 to 7 scale of agreement: "Our income never seems to catch up with our expenses" and "I often worry about my poor financial situation." These items are from scales developed by Conger and Elder (e.g., 1994). The average score was 4.43 (*SD* = 1.56).

Results

All analyses presented in this report are based on the most conservative experimental design for such a study; using Intention to Treat Analyses (ITT). Hence, analyses are based on original random assignment to either intervention or control groups.

Tests Hypothesis One: Divorce Outcomes

To examine the impact of intervention assignment on divorce, we used a 2-level (multilevel) model in which couples were nested within cohorts and intervention assignment was modeled as a cohort characteristic. This design allowed us to appropriately account for the nested nature of the data, as couples received intervention in a group format. These analyses were conducted using HLM 7.0 (Raudenbush, Bryk, Fai, Congdon, & du Toit, 2011) and its Bernoulli routine for dichotomous outcomes. The equation was:

$$\begin{aligned} Level\ 1 \quad & (Couple): Probability\ (Divorce{=}1 \big| \beta){=}\phi \\ & Log\,[\phi/(1-\phi)]{=}\eta \\ & \eta{=}\beta_0 \\ Level\ 2 \quad & (Cohort): \beta_0{=}\gamma_{00}{+}\gamma_{01}\,(Intervention\ Group){+}\mathrm{u}_0 \end{aligned}$$

At Site 1, 8.1% (20/248) of the intervention group divorced compared to 14.9% (34/228) of the control group, B = -0.69, Odds Ratio = 0.50, $p < .01$. In contrast, at Site 2, 6.3% (6/94) of the intervention group couples divorced and 4.4% (4/90) of the control group couples divorced, B = 0.32, Odds Ratio = 1.37, $p = .60$. Findings from an overall test of divorce likelihood (pooled across both sites) are consistent with the prediction that the intervention group would be less likely to divorce than the control group; 7.6% (26/342) of the intervention group divorced compared to 11.9% (38/318) of the control group, B = -0.51, Odds Ratio = 0.60, $p = .06$. Whereas this result is significant using a one-tailed test (in line with the fact that the finding and the direction of it were predicted), it is obvious that this result is driven by Site 1 and that there is no significant divorce effect at Site 2. As noted in the methods section, while there were 54 divorces at Site 1, there were only 10 divorces at Site 2. Consistent with the earlier suggestion that Site 2 has characteristics of a lower risk sample, 11.3% (54/476) of couples at Site 1 divorced compared to only 5.4% (10/184) of the couples at Site 2, B = 0.80, Odds Ratio = 2.23, $p < .05$.

The very small number of divorces in the already smaller sample at Site 2 rendered further tests of divorce impacts involving Site 2 to be underpowered for tests of moderation by site or ethnic minority status. Indeed, power calculations for the test of moderation by site for divorce as an outcome suggest that no valid interpretation is possible from such an analysis. Hence, all further divorce analyses are based on the larger site (Site 1) that had both a substantially higher overall risk for divorce and where there was a substantial divorce reduction impact. While the lack of power for testing moderation of divorce outcomes by site is driven by there being so few divorces at Site 2, tests of moderation by site for relationship quality outcomes are not similarly affected; hence, those analyses include tests for moderation by site.

Tests of Moderators of Divorce Outcome (at Site 1)

We tested whether ethnic minority status, economic strain, and income moderated the impact of the intervention on divorce at Site 1. Income did not significantly moderate the impact of the intervention, but there was a significant effect for ethnic minority status (B = -1.29, Odds Ratio = 0.27, $p < .05$) to moderate divorce outcomes. For non-minority couples, 10.6% (15/141) of the intervention group divorced compared to 13.3% (19/143) of the control group. For ethnic minority couples, 3.8% (4/105) of the intervention group divorced compared to 15.7% (13/83) of the control group. The direction of this moderation effect held in the same direction no matter how we analyzed it; it held for couple-level ethnic minority status as presented, but it also held for individual-level ethnic minority status (by male and by female), for different ethnic minority groups (e.g., African American, Hispanic/Latino), and for any of the other possible couple-level designations (e.g., both partners having same or different ethnic minority status).

The test for moderation by economic strain on divorce as an outcome indicated a trend toward significance (B = -0.33, Odds Ratio = 0.72, $p = .06$), with couples who reported greater economic strain tending to show greater benefit from the intervention. Further analyses (not shown for space but available upon request) indicated that the economic strain and minority status moderation findings were independent. Finally, other analyses (also not presented for space) examined if pre-intervention marital quality accounted for or moderated the intervention and control group differences regarding divorce. No such effects were found.

Tests of the Hypothesis Two: Marital Quality Outcomes (Sites 1 & 2)

To examine change over time in marital quality following treatment, we used four-level models and HLM 7.0 (Raudenbush et al., 2011). Following suggestions by Atkins (2005) and prior publications from this data set (Allen et al., 2012), time-varying characteristics of individuals (e.g., marital satisfaction) were modeled at level 1, time-invariant individual characteristics (e.g., pre-intervention marital quality) were modeled at level 2, couple characteristics (i.e., for moderator analyses described below) were modeled at level 3, and cohort characteristics (e.g., intervention group) were modeled at level 4. Pre-intervention marital quality scores were controlled for by including them at level 2. Time was measured as months since post. Each measure of marital quality was examined in a separate model: communication skills, marital satisfaction, positive bonding, and divorce proneness. The basic equation for the initial tests of hypotheses follows.

$$
\begin{aligned}
Level-1\,(Time): &\quad Marital\ Quality_{tijk} = \pi_{0ijk} + \pi_{1ijk}(Time_{tijk}) + e_{tijk} \\
Level-2\,(Individual): &\quad \pi_{0ijk} = \beta_{00jk} + \beta_{01jk}(Pre\ Intervention\ Quality_{ijk}) + r_{0ijk} \\
&\quad \pi_{1ijk} = \beta_{10ijk} + \beta_{11ijk}(Pre\ Intervention\ Quality_{ijk}) \\
Level-3\,(Couple): &\quad \beta_{00jk} = \gamma_{000k} + u_{00j} \\
&\quad \beta_{01jk} = \gamma_{010k} \\
&\quad \beta_{10jk} = \gamma_{100k} \\
&\quad \beta_{11jk} = \gamma_{110k} \\
Level-4\,(Cohort): &\quad \gamma_{000k} = \delta_{0000} + \delta_{0001}(Intervention\ Group_k) + \nu_{000k} \\
&\quad \gamma_{010k} = \delta_{0100} \\
&\quad \gamma_{100k} = \delta_{1000} + \delta_{1001}(Intervention\ Group_k) + \nu_{100k} \\
&\quad \gamma_{110k} = \delta_{1100}
\end{aligned}
$$

In these models, Time was not centered, so the intercept could be interpreted as the estimated post score. Group was coded 0 = control, 1 = intervention (PREP). When the random effects for level 4 change over time were significant (ν_{100k}), they were retained. This was the case for communication skills, marital satisfaction, and divorce proneness, but not positive bonding. We also tested for differences by group in quadratic changes for all outcome variables. No such differences were evident, so the models below include only linear change. We also tested whether site moderated any findings for marital quality and found that it did not, so the presented models combine both sites. Lastly, we found that whether participants ultimately divorced or not did not significantly moderate any main findings, so we retained relationship quality trajectories for people who eventually divorced in the analyses presented below (Table 1).

Communication skills—Consistent with a prior report (Allen et al., 2011), we found that the intervention group reported significantly better communication skills than the control group at post, controlling for pre. Over time from post onward, both groups tended to decrease in communication skills, but the intervention group decreased significantly more such that the two groups were not significantly different at the two-year follow-up: intervention group (M = 4.23, SD = 1.22) and control group (M = 4.35, SD = 1.17), $F(1, 515) = .02$, $p > .50$).

Marital satisfaction, positive bonding, and divorce proneness—For marital satisfaction, positive bonding, and divorce proneness, we found no significant differences among the groups at post (as projected by the HLM trajectories) controlling for pre, no differences over time from post, and no differences at the two-year follow-up.

Moderators of Marital Quality Outcomes

By adding ethnic minority status, economic strain, or income to level 3 of the equation above, we tested if these variables moderated impacts of the intervention on communication skills, marital satisfaction, positive bonding, or divorce proneness. We found no evidence of moderation in any of these analyses.

Discussion

This study examined the effectiveness of CRE (PREP) in a randomized controlled trial at two years post-intervention. We found no evidence of long-lasting intervention effects for relationship quality but did find that couples in the intervention group were less likely to divorce than those in the no-treatment control group at the site with the larger, higher risk sample (Site 1). This effect was moderated by couple-level ethnic minority status with a trend for moderation by economic strain.

As found in a prior publication (Allen et al., 2011), there were significant pre-post differences for communication skills but not for marital satisfaction. In the prior work, we also found small but significant pre-post intervention effects on positive bonding for Site 1 but not for Site 2. In the current paper, we did not find pre-post differences on positive bonding collapsed across site. Analytic strategies are somewhat different between that report and this one; here, pre-post findings are not based on differences in means from pre to post but are estimated based on linear trajectories of change over time as a function of using multilevel modeling for the repeated follow-ups. Importantly, the analyses here did *not* yield support for our hypothesis that intervention couples would better maintain marital quality over time.

While the intervention group reported an initial gain in communication skills at post relative to the control group, the intervention group declined faster over time than the control group, resulting in a lack of significant difference in communication skills at two years post-intervention. This pattern could reflect decay in preventive effects, as has been observed in other CRE studies with longer-term follow-ups (e.g., Bodenmann, Pihet, Shantinath, Cina, & Widmer, 2006). Meta-analyses have also demonstrated some evidence of decay in relationship quality impacts; however, the matter is quite difficult to assess because few studies in this field have follow-ups past one year (Hawkins et al., 2008). It is possible that longer-term follow-ups with this sample could show the emergence of preventive impacts on marital quality; moreover, there may be additional moderators of effects that were not tested here.

In contrast to the lack of differences at two years post-intervention on relationship quality, there was significant divorce reduction associated with the intervention at Site 1, extending a similar finding from the one-year follow-up assessment (Stanley et al., 2010). Additionally, the overall divorce rate was very different between Sites 1 and 2, likely due to the fact that couples at Site 1 were younger, married fewer years, lower ranking, and more likely to be involved in war time deployments. These characteristics suggest higher risk, and Site 1 couples were, in fact, substantially more likely to divorce. As such, the divorce findings here fit a pattern of the higher risk group gaining the most from the intervention.

It is puzzling that there was an intervention effect for divorce but not 8marital quality. Few published studies of CRE have had large enough samples and long enough follow-ups to assess marital stability adequately, though long-term outcome studies on earlier versions of PREP, using quasi-experimental designs, found preventative impacts from earlier versions of PREP on the likelihood of couple dissolution (Hahlweg et al., 1998; Markman et al.,

1988). Interestingly, the current findings are similar to what was found for unmarried couples at the Oklahoma site of the Building Strong Families (BSF) study. Among the eight sites in the BSF study, Oklahoma had, by far, the highest participant involvement in the intervention, and it was also the only site to use a variant of PREP and to show a range of positive impacts at 15 months (Devaney & Dion, 2010). However, by 36-months, there was no longer any evidence of any relationship quality impacts but there was evidence of a stability impact, with intervention families being more likely to have lived together continuously from the birth of the focal child (Wood, Moore, Clarkwest, Killewald, & Monahan, 2012).

Why would intervention couples have lower odds of divorce but not higher levels of marital quality than control couples, over the two-year follow-up? Methodologically, prevention researchers (e.g., Halford, Sanders, & Behrens, 2001; Markman et al., 1988; Stanley, 2001) as well as professional evaluators (McConnell, Stuart, & Devaney, 2008) have discussed the possibility that strong stability impacts can inhibit the ability to detect relationship quality differences because of differential loss to follow-up for relationship quality outcomes; for example, more couples with lower relationship quality may leave the control group over time. We attempted to control for this potential differential selection in follow-ups by testing divorce at two years as a moderator of relationship quality outcomes, but found no such moderation effects. However, this issue is known to be extraordinarily difficult to resolve (McConnell et al.).

It is also possible that even modest short-term changes in relationship quality (as reported previously for this sample; Allen et al., 2011) reduce divorce risk. Given the external stressors such as deployment that military couples may experience, small initial impacts (measured by study procedures or not) during times of higher stress may function as protection, and may have helped some couples get through challenging times they may otherwise not have endured. Nevertheless, there is not much evidence in this sample for relationship quality impacts that would appear sufficiently powerful to explain the divorce reduction effect, raising the possibility that other mechanisms produced the impact on divorce, such as non-specific or unmeasured effects (see Hawkins, Stanley, Blanchard, & Albright, 2012). Such issues pertain to the general question of mechanisms of impact, which, as others have noted, are important to elucidate (e.g., Wadsworth & Markman, 2011) but difficult to demonstrate (e.g., Schilling, Baucom, Burnett, Allen, & Ragland, 2003; Stanley, Rhoades, Olmos-Gallo, & Markman, 2007).

Moderators of Impact

Neither income, economic strain, nor ethnic minority status moderated marital quality outcomes, but the divorce finding was moderated by ethnic minority status. Intervention group couples in which one or both partners were ethnic minorities experienced larger reductions in the likelihood of divorce than non-minority (White) couples. In fact, among ethnic minority couples, the divorce rate in the control group was 15.7% but only 3.8% in the intervention group, compared to 13.3% and 10.6%, respectively, for non-minority couples. Hence, the divorce rate for ethnic minority couples in the intervention group was a fraction of what was it was for ethnic minority couples in the control group. There was also

a trend for those in the intervention group who reported greater economic strain to show a stronger divorce reduction effect compared to those reporting lower economic strain. Further follow-ups with the sample will include more divorces, which will increase statistical power to analyze these and other potentially moderating effects in the future.

The fact that ethnic minority couples in the intervention group demonstrated greater marital stability appears to be consistent with the suggestion that those at higher risk benefit most from CRE (e.g., Allen et al., 2012; Halford, et al., 2001; Petch et al., 2012). At the same time, it is unclear that the moderation by ethnic minority status should be interpreted thusly in this sample. While ethnic minorities in the U.S. tend to be at higher average risk for divorce (especially African-Americans; Bramlett & Mosher, 2002), military service is associated with greater marital stability (compared to civilian life) among African Americans (Teachman & Tedrow, 2008). Further, ethnic minority couples in this sample were no more likely to divorce than other couples, so we cannot assume such couples in this sample are a higher-risk group overall.

Implications for Intervention

Johnson (2012) speculated that impacts in recent federally-commissioned CRE studies (BSF and SHM) were small because CRE interventions were mostly developed with non-minority, middle-class samples, rather than with lower-income and ethnic minority couples like the couples heavily recruited for those studies. Johnson (2012) also noted that mixed-race couples had worse outcomes than other couples in the BSF study (Wood et al., 2010). However, this conclusion appears to be a misinterpretation of the report, as no such specific analyses of mixed-race couples were conducted for that report (personal communication, Andrew Clarkwest, July 7, 2012). The finding in BSF was that African-American couples received, on average, the greatest benefit at 15 months out. However, analyses of longer term outcomes found that this moderation by ethnic minority status faded by the three-year follow-up point (Wood, Moore, Clarkwest, & Killewald, 2014). Related to arguments made by Johnson (2012), Amato (in press) has analyzed the 15 month outcomes from the BSF data set (the 3 year data is not available to others at the time of this writing). Using an index comprised of 11 variables, he tested if social and economic risk moderated program outcomes, finding that it did for relationship quality outcomes. Amato found that “the BSF study suggests that the treatment was especially effective for the most disadvantaged couples, at least with respect to the quality of their relationships.”

An underlying premise of Johnson’s paper was that CRE interventions may miss the mark for both ethnic minority couples and/or couples with higher social and economic risks. His call for further research on such issues is warranted. However, his suggestion that CRE may not work as intended for disadvantaged couples may be unwarranted given that the current study is the third large randomized trial of CRE (along with BSF and SHM) in recent years wherein either ethnic minority couples or couples with higher social and economic risks tended to gain as much or more than other couples from the services, at least over a couple of years’ time (see also Hawkins et al., 2013; Stanley et al., 2005; Stanley, Amato, Johnson, & Markman, 2006). While the present sample is not constituted by couples in poverty, and it has less ethnic minority representation than the BSF and SHM studies, Army couples do

represent a relatively diverse community that is probably at the lower end of the income scale compared to most couples in the history of CRE studies. On the other hand, married military couples tend to enjoy housing and other benefits which lower-income civilian couples are less likely to have. This means that we are unable, with this sample, to test if those experiencing poverty are more or less likely to benefit from CRE. Nevertheless, based on the present and prior findings, it seems promising that CRE interventions could be a reasonable relationship strengthening strategy for helping low-income couples, ethnic minority couples, and/or couples experiencing economic strain. Importantly, this point takes nothing away from the fact that there is room for improvement in overall effectiveness within the CRE field.

The fact that CRE lowered divorce rates for ethnic minority couples adds to an expanding set of findings concerning who benefits most from CRE, with most evidence suggesting that higher-risk couples benefit most (e.g., Allen et al., 2012; Halford et al., 2001; Petch et al., 2012) and some evidence suggesting that risk and type of CRE may interact in complex ways (Markman, Rhoades, Stanley, & Peterson, 2013). Though there are concerns that those who most need CRE are the least likely to seek it (e.g., Sullivan & Bradbury, 1997), a recent study suggests that when CRE is universally available (e.g., during pregnancy), higher-risk couples within the pool are more likely, not less, to attend (Engsheden, Fabian, & Sarkadi, 2013).

Limitations and Conclusion

Strengths of the current study include a relatively large sample for this field, a fully randomized design, low attrition, and the use of the most rigorous analytic approach (ITT analyses). Also, marital status was known for all but one couple two years post-intervention. In contrast, a weakness of this study was the dependence on self-report measures for the assessment of relationship quality. Future projects may benefit from including objectively-coded communication data, as observational methods have shown the largest impacts from CRE in meta-analyses (Blanchard et al., 2009). Further, while the ability to assess impacts in a community sample for services delivered by community-level providers is a strength, these same factors could have yielded underestimates of the effects that may be obtained in different or more controlled contexts. The sample has a lot of variability to assess differential intervention impacts for ethnic minority couples. However, the sample does not, generally, contain couples experiencing poverty, which limits the degree to which these data can speak to CRE impacts for couples with extreme economic vulnerabilities. Finally, while findings here are consistent with those emerging in non-military samples, the military context of this study could limit generalizability.

These limitations notwithstanding, this research makes a significant contribution to the CRE field, as it examined several important outcomes (divorce and four indices of relationship quality) in a large sample of couples in which one or both partners were in the U.S. Army. There were no lasting impacts on self-reported relationship quality but there were significantly lower divorce rates for couples assigned to PREP, especially among those in which one or both partners were racial or ethnic minorities. The context of CRE for this study is with couples who are generally under extraordinary stressors due to their military

involvement. Further, the significant divorce reduction effect was observed in the sample for the site that was at higher risk (as evidenced by the substantially higher risk for divorce at that site). Although the military sample may limit generalizability, these findings suggest that even in high-stress contexts, evidence-based CRE delivered in real-world settings can positively impact family stability.

Acknowledgments

The project described was supported by Award Number R01HD048780 from the Eunice Kennedy Shriver National Institute of Child Health & Human Development (NICHD). The content is solely the responsibility of the authors and does not necessarily represent the official views of NICHD or the National Institutes of Health. Scott Stanley and Howard Markman own a business that develops and sells relationship education curricula, including the intervention tested in this study. Galena Rhoades has been involved in developing relationship education curricula and receives royalties on the sales of some relationship education materials but not on the intervention tested here.

References

Allen ES, Rhoades GK, Stanley SM, Loew B, Markman HJ. The effects of marriage education for army couples with a history of infidelity. Journal of Family Psychology. 2012; 26(1):26–35.10.1037/a0026742 [PubMed: 22229880]

Allen ES, Rhoades GK, Stanley SM, Markman HJ. Hitting home: Relationships between recent deployment, post traumatic stress symptoms, and marital functioning for Army couples. Journal of Family Psychology. 2010; 24:280–288.10.1037/a0019405 [PubMed: 20545401]

Allen ES, Stanley SM, Rhoades GK, Markman HJ, Loew BA. Marriage education in the Army: Results of a randomized clinical trial. Journal of Couple and Relationship Therapy. 2011; 10(4): 309–326.10.1080/15332691.2011.613309 [PubMed: 22737042]

Amato, Paul R. Does social and economic disadvantage moderate the effects of relationship education on couples? An analysis of data from the 15-month Building Strong Families evaluation. Family Relations. in press.

Atkins DC. Using multilevel models to analyze couple and family treatment data: Basic and advanced issues. Journal of Family Psychology. 2005; 19(1):98–110.10.1037/0893-3200.19.1.98 [PubMed: 15796656]

Blanchard VL, Hawkins AJ, Baldwin SA, Fawcett EB. Investigating the effects of marriage and relationship education on couples' communication skills: A meta-analytic study. Journal of Family Psychology. 2009; 23(2):203–214.10.1037/a0015211 [PubMed: 19364214]

Bodenmann G, Pihet S, Shantinath SD, Cina A, Widmer K. Improving dyadic coping in couples with a stress-oriented approach: A 2-year longitudinal study. Behavior Modification. 2006; 30(5):571–597.10.1177/0145445504269902 [PubMed: 16894231]

Booth A, Johnson D, Edwards JN. Measuring marital instability. Journal of Marriage and the Family. 1983; 45(2):387–394.10.2307/351516

Bramlett, MD.; Mosher, WD. Vital and Health Statistics, Series 23. Washington, DC: U.S. Government Printing Office; 2002. Cohabitation, marriage, divorce, and remarriage in the United States.

Conger, R.; Elder, G. Families in troubled times: Adapting to change in rural America. New York: Aldine de Gruyter; 1994.

Crane DR, Middleton KC, Bean RA. Establishing criterion scores for the Kansas Marital Satisfaction Scale and the Revised Dyadic Adjustment Scale. American Journal of Family Therapy. 2000; 28:53–60.

Devaney, B.; Dion, R. 15-Month impacts of Oklahoma's Family Expectations Program. Washington DC: Mathematica Policy Research; 2010.

Engsheden N, Fabian H, Sarkadi A. Offering relationship education (PREP) for couples during pregnancy: Self-selection patterns. Family Relations. 2013; 62(4):676–685.10.1111/fare.12032

Hahlweg K, Markman HJ, Thurmaier F, Engl J, Eckert V. Prevention of marital distress: Results of a German prospective longitudinal study. Journal of Family Psychology. 1998; 12(4):543–556.10.1037/0893-3200.12.4.543

Halford W, Sanders MR, Behrens BC. Can skills training prevent relationship problems in at-risk couples? Four-year effects of a behavioral relationship education program. Journal of Family Psychology. 2001; 15(4):750–768.10.1037/0893-3200.15.4.750 [PubMed: 11770479]

Hawkins AJ, Blanchard VL, Baldwin SA, Fawcett EB. Does marriage and relationship education work? A meta-analytic study. Journal of Consulting and Clinical Psychology. 2008; 76(5):723–734.10.1037/a0012584 [PubMed: 18837590]

Hawkins AJ, Stanley SM, Blanchard VL, Albright M. Exploring programmatic moderators of the effectiveness of marriage and relationship education programs: A meta-analytic study. Behavior Therapy. 2012; 43(1):77–87.10.1016/j.beth.2010.12.006 [PubMed: 22304880]

Hawkins AJ, Stanley SM, Cowan PA, Fincham FD, Beach SRH, Cowan C, Rhoades GK, Markman HJ, Daire AP. A more optimistic perspective on government-supported marriage and relationship education programs for lower income couples: Response to Johnson (2012). American Psychologist. 2013; 68(2):110–111.10.1037/a0031792 [PubMed: 23421607]

Hsueh, J.; Alderson, DP.; Lundquist, E.; Michalopoulos, C.; Gubits, D.; Fein, D.; Knox, V. The Supporting Healthy Marriage Evaluation: Early Impacts on Low-Income Families. Administration for Children and Families, Office of Planning, Research, and Evaluation; Washington D. C: 2012.

Johnson MD. Healthy Marriage Initiatives: On the need for empiricism in policy implementation. American Psychologist. 2012; 67(4):296–308. [PubMed: 22468785]

Karney BR, Bradbury TN. Contextual Influences on Marriage: Implications for Policy and Intervention. Current Directions in Psychological Science. 2005; 14(4):171–174.10.1111/j.0963-7214.2005.00358.x

Markman, HJ. The Couple Activities Scale. University of Denver; 2000. Unpublished measure

Markman HJ, Floyd FJ, Stanley SM, Storaasli RD. Prevention of marital distress: A longitudinal investigation. Journal of Consulting and Clinical Psychology. 1988; 56(2):210–217.10.1037/0022-006X.56.2.210 [PubMed: 3372828]

Markman HJ, Rhoades GK, Stanley SM, Peterson KM. A randomized clinical trial of the effectiveness of premarital intervention: Moderators of divorce outcomes. Journal of Family Psychology. 2013; 27(1):165–172.10.1037/a0031134 [PubMed: 23421844]

Markman, HJ.; Stanley, SM.; Blumberg, SL. Fighting for your marriage. San Francisco: Jossey-Bass; 2010.

McConnell S, Stuart EA, Devaney B. The Truncation-by-Death Problem: What to do in an experimental evaluation when the outcome is not always defined. Evaluation Review. 2008; 37(1):157–186.10.1177/0193841X07309115 [PubMed: 18319423]

Petch JF, Halford W, Creedy DK, Gamble J. 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. 2012; 80(4):662–673.10.1037/a0028781 [PubMed: 22730950]

Raudenbush, SW.; Bryk, AS.; Fai, CY.; Congdon, R.; du Toit, M. HLM 7: Hierarchical linear and nonlinear modeling. Lincolnwood, IL: Scientific Software International, Inc.; 2011.

Saiz, CC.; Jenkins, N. The Communication Skills Test. University of Denver; Denver CO: 1995. Unpublished measure

Schilling EA, Baucom DH, Burnett CK, Allen E, Ragland L. Altering the course of marriage: The effect of PREP communication skills acquisition on couples' risk of becoming maritally distressed. Journal of Family Psychology. 2003; 17(1):41–53.10.1037/0893-3200.17.1.41 [PubMed: 12666462]

Schumm WR, Paff-Bergen LA, Hatch RC, Obiorah FC, Copeland JM, Meens LD, Bugaighis MA. Concurrent and discriminant validity of the Kansas Marital Satisfaction Scale. Journal of Marriage and the Family. 1986; 48(2):381–387.10.2307/352405

Stanley SM. Making a case for premarital education. Family Relations. 2001; 50(3):272–280.10.1111/j.1741-3729.2001.00272.x

Stanley SM, Allen ES, Markman HJ, Rhoades GK, Prentice D. Decreasing divorce in Army couples: Results from a randomized controlled trial of PREP for Strong Bonds. Journal of Couple and Relationship Therapy. 2010; 9:149–160.10.1080/15332691003694901 [PubMed: 20634994]

Stanley SM, Allen ES, Markman HJ, Saiz CC, Bloomstrom G, Thomas R, Schumm WR, Baily AE. Dissemination and evaluation of marriage education in the Army. Family Process. 2005; 44:187–201.10.1111/j.1545-5300.2005.00053.x [PubMed: 16013745]

Stanley SM, Amato PR, Johnson CA, Markman HJ. Premarital education, marital quality, and marital stability: Findings from a large, random household survey. Journal of Family Psychology. 2006; 20(1):117–126.10.1037/0893-3200.20.1.117 [PubMed: 16569096]

Stanley, SM.; Markman, HJ.; Jenkins, NH.; Blumberg, SL. PREP for Strong Bonds Leader's Manual. Denver: PREP Educational Products, Inc.; 2006.

Stanley SM, Rhoades GK, Olmos-Gallo PA, Markman HJ. Mechanisms of change in a cognitive behavioral couples prevention program: Does being naughty or nice matter? Prevention Science. 2007; 8:227–239. doi:1007/s11121-007-0071-8. [PubMed: 17578665]

Sullivan KT, Bradbury TN. Are premarital prevention programs reaching couples at risk for marital dysfunction? Journal of Consulting and Clinical Psychology. 1997; 65(1):24–30.10.1037/0022-006X.65.1.24 [PubMed: 9103731]

Teachman JD, Tedrow L. Divorce, race, and military service: More than equal pay and equal opportunity. Journal of Marriage & Family. 2008; 70(4):1030–1044.10.1111/j.1741-3737.2008.00544.x

Tejada-Vera, B.; Sutton, PD. National Vital Statistics Reports. Vol. 58. Hyattsville, MD: National Center for Health Statistics; 2010. Births, marriages, divorces, and deaths: Provisional data for 2009.

Wadsworth ME, Markman HJ. Where's the action? Understanding what works and why in relationship education. Behavior Therapy. 2011; 43(1):99–112.10.1016/j.beth.2011.01.006 [PubMed: 22304882]

Whisman MA, Beach SRH, Snyder DK. Is marital discord taxonic and can taxonic status be assessed reliably? Results from a national, representative sample of married couples. Journal of Consulting and Clinical Psychology. 2008; 76:745–755.10.1037/0022-006X76.5745 [PubMed: 18837592]

Wood, RG.; McConnell, S.; Moore, Q.; Clarkwest, A.; Hsueh, J. The Building Strong Families Project Strengthening unmarried parents' relationships: The early impacts of Building Strong Families. Administration for Children and Families, Office of Planning, Research, and Evaluation; Washington D. C: 2010.

Wood, RG.; Moore, Q.; Clarkwest, A.; Killewald, A.; Monahan, S. The long–term effects of Building Strong Families: A relationship skills education program for unmarried parents. Administration for Children and Families, Office of Planning, Research, and Evaluation; Washington D. C: 2012.

Wood RG, Moore Q, Clarkwest A, Killewald A. The long-term effects of Building Strong Families: A program for unmarried parents. Journal of Marriage and Family. 2014; 76:446–463.10.1111/jomf.12094

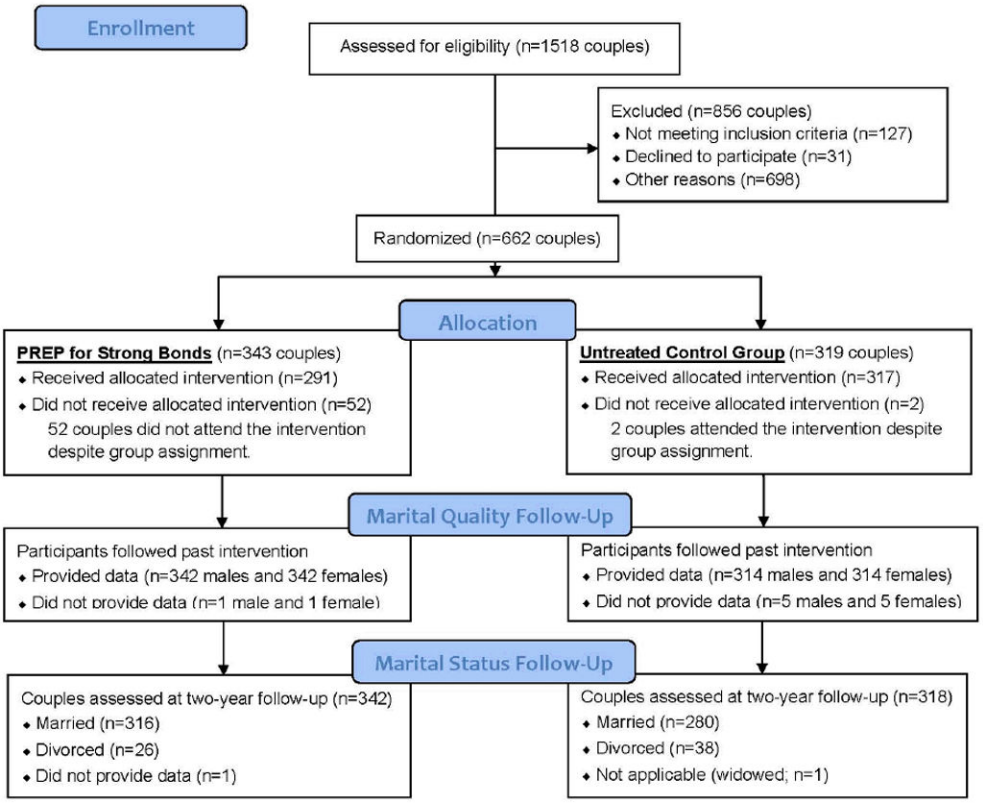

Figure 1.
Consolidated Standards of Reporting Trials

Table 1

Multilevel Modeling Analyses of Marital Quality Variables Controlling for Pre-Test Scores

Fixed effects	Communication Skills		Marital Satisfaction		Positive Bonding		Divorce Proneness	
	b	***SE***	**b**	***SE***	**b**	***SE***	**b**	***SE***
Post-Test								
Control	1.65***	.09	2.02***	.14	1.63***	.13	.06***	.01
PREP vs. Control	.15**	.06	.00	.06	-.00	.06	-.01	.02
Pre-Test	.65***	.02	.68***	.02	.71***	.02	.41***	.02
Linear Slope								
Control	.01**	.00	.02*	.01	.01	.01	.01***	.00
PREP vs. Control	-.01*	.00	-.01	.00	-.00	.00	.00	.00
Pre-Test	-.00***	.00	-.01***	.00	-.00**	.00	-.00**	.00

Note. *df* for *t*-tests of coefficients vary from 104 to 3612. Unstandardized coefficients are presented for the estimated post-test scores and the linear slope over time. For example, for communication skills, the intercept for the control group at Time = 0, controlling for pre-test communication skills, was 1.65 and it was significantly different from zero. In the same model, being assigned to PREP was associated with higher communication skills at post-test (.15; also significant). Further, pre-test scores were significantly associated with post-test scores on communication skills. Regarding linear slope (or change over time) on communication skills, the control group was estimated to have increased by .01 points each month following post. The PREP group increased significantly less per month (-.01) and higher pre-test scores on communication skills were also associated with less increase in communication skills over time.