



Published in final edited form as:

J Consult Clin Psychol. 2014 August ; 82(4): 592–604. doi:10.1037/a0037097.

The Marriage Checkup: A Randomized Controlled Trial of Annual Relationship Health Checkups

James V. Cordova, Ph.D.,

Clark University

C.J. Eubanks Fleming, M.A.,

Clark University

Melinda Ippolito Morrill, M.A.,

Clark University

Matt Hawrilenko, B.A.,

Clark University

Julia W. Sollenberger, M.A.,

Clark University

Amanda G. Harp, Ph.D.,

Harbor-UCLA Medical Center

Tatiana D. Gray, M.A.,

Clark University

Ellen V. Darling, M.A.,

Clark University

Jonathan M. Blair, Ph.D.,

Clark University

Amy E. Meade, Ph.D., and

McLean Hospital, Harvard Medical School Department of Psychiatry

Karen Wachs, Ph.D.

Psychological and Life Skills Association of Woodbridge

James V. Cordova: jcordova@clarku.edu; C.J. Eubanks Fleming: cjfleming@clarku.edu; Melinda Ippolito Morrill: mmorrill@clarku.edu; Matt Hawrilenko: mhawrilenko@clarku.edu; Julia W. Sollenberger: jsollenberger@clarku.edu; Amanda G. Harp: Amanda.harp@gmail.com; Tatiana D. Gray: tsettles@clarku.edu; Ellen V. Darling: edarling@clarku.edu; Jonathan M. Blair: jonathanmatthewblair@gmail.com; Amy E. Meade: aemeade@mclean.harvard.edu; Karen Wachs: Karen.wachs1@gmail.com

Abstract

Objective—This study assessed the efficacy of the Marriage Checkup for improving relationship health and intimacy.

Method—Cohabiting married couples ($N=215$, M_{age} women=44.5 years, men=47 years, 93.1% Caucasian) recruited from a northeastern U.S. metropolitan area through print and electronic media were randomly assigned to MC treatment or wait-list control. Treatment but not control couples participated in assessment and feedback visits, at the beginning of the study and again one year later. All couples completed nine sets of questionnaires over two years. Outcome measures included the Quality of Marriage Index, the Global Distress subscale of the Marital Satisfaction Inventory-Revised, the Intimate Safety Questionnaire, and the Relational Acceptance Questionnaire.

Results—A latent growth curve model indicated significant between group differences in intimacy at every measurement point after baseline (d ranged from .20 to .55, $M_d = .37$), significant between group differences in women’s felt acceptance for every measurement point after baseline (d ranged from .17 to .47, $M_d = .34$), significant between group differences in men’s felt acceptance through the 1-year-2 week follow-up (d across follow-up ranged from .11 to .40, $M_d = .25$), and significant between group differences in relationship distress through 1-year 6-month follow-up (d across follow-up ranged from .11 to .39, $M_d = .23$).

Conclusions—Longitudinal analysis of the MC supports the hypothesis that the MC significantly improves intimacy, acceptance, and satisfaction. Implications for dissemination are discussed.

Keywords

marriage checkup; marital satisfaction; brief intervention; intimacy; effectiveness research; multi-level modeling

Relationship health is a public health issue. Divorce and marital deterioration are ubiquitous, exacting substantial mental and physical health costs on individuals and society. By the time women and men reach the ages of 50–59, 39% and 41% respectively have been divorced at least once (Kreider, 2005). Furthermore, of those couples who remain together, a substantial number endure unhappy or abusive relationships.

Relationship distress has been associated with a higher incidence of both depression and substance abuse (e.g. Whisman, 2007), as well as diminished medical treatment adherence and even down-regulated immune system functioning (Whisman & Uebelacker, 2006). Forty percent of mental health patients named relationship difficulties as responsible, at least in part, for their psychiatric problems (Berger & Hannah, 1999). Given the additional negative effect of relationship deterioration on children’s health (e.g. Cummings, Goeke-Morey, & Papp, 2003), combined with the multiple negative health effects on individuals (Proulx et al., 2007), and the health effects associated with divorce and separation (Lucas, 2005), it is becoming increasingly clear that relationship health affects all other health systems.

Despite the epidemic of relationship health deterioration, there remain few avenues by which couples can effectively attend to their relationship health, short of tertiary couple therapy. Behavioral Couple Therapy (BCT; Epstein & Baucom, 2002; Jacobson & Christensen, 1998; Jacobson & Margolin, 1979) – is the current “gold standard,” having documented efficacy across multiple RCTs (e.g. Christensen, Atkins, Baucom, & Yi, 2010;

Shadish & Baldwin, 2005). Roughly two-thirds of couples improve with therapy, resulting in the average treated couple faring better than about 70–80% of untreated couples (Gurman, 2011). Clearly tertiary therapy is beneficial; however, the data also indicate that there are many couples who are not well served by traditional treatment. Tertiary treatment can only work when couples attend, and the percentage of distressed couples who attend therapy is alarmingly low.

Targeting prevention, educational programs have been developed to help couples sustain relationship health. These models usually focus on relationship education, increasing relationship awareness (e.g. PREPARE/ENRICH; Olson & Olson, 1999), or building relationship skills (e.g. PREP; Markman et al., 2001). Research has found that couples at low risk for marital discord are overrepresented in these programs, meaning that couples who stand to benefit the most do not participate as often as couples who may have remained well regardless (Halford, O’Donell, Lizzio, & Wilson, 2006; Sullivan & Bradbury, 1997).

A major barrier to help-seeking involves concerns about the time and cost of therapy, as well as the significant emotional challenge of admitting the need for marital therapy. The majority of people suffering from relationship distress do not seek professional help (Johnson et al., 2002). A recent statewide sample of adults indicated that only 37% of divorced persons had sought counseling before dissolving the marriage (Johnson et al., 2002). Of currently married couples, only 19% indicated they had ever participated in marital therapy (Johnson et al., 2002). Recent research indicates that couple help seeking is hindered by unique barriers as compared to individual help seeking. For example, seeking help as a couple requires the motivation and buyin of *both* partners, either of whom can refuse to participate (Fleming & Cordova, 2012). Additional barriers include partner’s lack of confidence in the outcome, preference to solve problems on their own, or logistical challenges such as cost, conflicting schedules, or lack of childcare (e.g. Uebelacker, Hecht, & Miller, 2006).

Other barriers to participating in prevention programs include healthy couples not feeling the need to participate, reluctance to share personal information about their relationship, and perceiving relationship educators as having lower professional status (Berger & Hannah, 1999; Bradbury & Karney, 2010). Also, couples who have never had therapy before are the hardest to attract, partly because they are more sensitive to stigmatization (Bringle & Byers, 1997).

Additionally, professional services are rarely used to maintain marital health (Wolcott, 1986). Couples who do not self-identify as distressed appear to be particularly reticent to participate in anything that resembles “therapy” (Morrill, Fleming, Harp, Sollenberger, Darling, & Cordova, 2011). Indeed, distressed couples wait an average of 6 years before seeking help, at which point their relationship likely has deteriorated dramatically (Notarius & Buongiorno, 1992, as cited in Gottman & Gottman, 1999). Studies also show that for people who divorced but never sought couple therapy, the top reason given was the belief that it was too late to make a difference (Wolcott, 1986). A major reason for this delay may be the process involved in the decision to attend couple therapy, which appears to involve three phases: 1) recognizing the problem, 2) considering therapy, and 3) taking steps to

engage in treatment (Doss, Atkins, & Christensen, 2003). Even the first stage only occurs after the relationship has been deteriorating for some time. Couples often do not self-identify as “distressed” until negative interactions have accumulated past a cognitive tipping point and serious relationship damage has occurred (e.g., Gottman, 1994).

Thus, there is a substantial need for early detection and preventative care for deteriorating couples before serious and irreversible relationship damage has occurred. There are currently no widely available means to fill this need. Mild-to-moderately distressed couples may view therapy as reserved for only the most severely distressed couples, and thus delay seeking treatment until its efficacy is seriously diminished by the chronicity and severity of the accumulated relationship dysfunction. Most distressed couples do not ever consider seeking help and those who do primarily consult medical doctors and clergy, rather than trained couple therapists (Veroff, 1981). Doss et al.’s series of studies (2003, Doss 2004, Doss 2009) concluded that nontraditional marital interventions need to become more widely available in order to attract a greater number of couples.

Given the above, it is likely that a fair number of couples at risk for significant relationship health deterioration exist who (1) may not yet be self-identifying as “needing help,” (2) are unlikely to seek tertiary couple therapy, and (3) are equally unlikely to seek preventive education workshops. By analogy, these couples are the relationship health equivalent of those for whom annual physical health checkups were designed – those who may benefit from early detection and preventative care despite their own perceptions of health.

The Marriage Checkup (MC) addresses this issue by providing a less-threatening option for couples to seek early preventative care before they have begun to identify as distressed. Intended to be the relationship health equivalent of the annual physical or dental checkup, the MC was designed to fill the gap in empirically supported relationship health care between tertiary care and preventative relationship education. The present study assesses the efficacy of the Marriage Checkup (Cordova, 2009; 2014; Cordova, Scott, Dorian, Mirgain, Yaeger, & Groot, 2005; Gee, Scott, Castellani, & Cordova, 2002; Morrill et al., 2011) as a brief, accessible maritalhealth intervention. The MC is a 2-session assessment and feedback intervention designed to be a safe and routine procedure for relationship health maintenance, early problem detection, and early intervention.

The Marriage Checkup: Previous Research

The current study is a longitudinal examination of both the short- and long-term relationship health effects of the MC. Utilizing both Motivational Interviewing (MI, Miller & Rollnick, 2002) and Integrative Behavioral Couples Therapy techniques (IBCT, Cordova, Jacobson & Christensen, 1998), the goal of the MC is to activate couples in the service of their marital health while simultaneously fostering a sense of greater acceptance and deeper intimacy.

Pilot research has demonstrated that the MC has high treatment tolerability (97% completion rate) and is safe for use with atrisk couples (Cordova, Warren, & Gee, 2001). Furthermore, longitudinal follow-up from the pilot studies demonstrated several positive outcomes of the MC intervention: (1) MC Couples significantly improved across a range of marital health variables compared to control, (2) MC couples maintained significant improvement two

years post MC; (3) MC wives' subsequent treatment seeking was affected by receiving a treatment recommendation; and (4) couples' affective tone after participating in the MC predicted later marital satisfaction (Gee et al., 2002). Further details of the MC format and results of previous studies are described in Cordova, 2014, Cordova et al., 2005, and Cordova, Warren, & Gee, 2001.

A principal goal of the MC is to attract a large enough sample of treatment-avoidant, at-risk couples to be useful as an indicated preventive intervention. Promoting the MC as an informational marital health checkup removes the "treatment" barrier, creating an atmosphere in which couples who are not openly distressed, or who are otherwise biased against "treatment," can address existing relationship concerns. Research has confirmed that the checkup format is effective at attracting couples considered at-risk for marital deterioration, but who are otherwise not seeking relationship treatment (Morrill et al., 2011). The current sample reported a broad range of relationship distress (as measured by the Marital Satisfaction Inventory – Revised, Global Distress Scale (MSI-R GDS; Snyder, 1997)) demonstrating the at-risk nature of a majority of the sample. About 44% had moderately distressed baseline scores, and 19% scored in the highly distressed range. Furthermore, 63% of MC participants had never sought couples therapy previously and over 32% of MC participants reported the MC as their first utilization of any mental health services (Morrill et al., 2011). These findings suggest that the MC's novel approach as an informational marital health checkup attracts a broad range of couples, many of whom fall into the "at-risk" category, and who might not otherwise seek marital health services.

Current Aims

The current study extends previous MC studies in important ways. First, the current study is the first to include two annual checkups with two years worth of follow-up assessments. Given that a checkup model, by definition, involves a regular schedule of repeated checkups, and that our previous studies were not able to study the effects of repeated checkups, the addition of a follow-up annual checkup gives us our first opportunity to study whether subsequent regular checkups are measurably beneficial. In addition, while previous MC studies have been relatively small pilot projects (e.g., $N=64$; Cordova et al., 2005), the current iteration of the MC was designed to recruit a much larger sample size ($N=215$) in order to augment the power and precision of our findings, to improve generalizability, and to allow for the use of more sophisticated analyses.

In the current study we tested the effects of the MC on two relationship satisfaction variables, as well as on intimacy and acceptance. These are the three primary areas of couples emotional health that we hypothesize to be influenced by the MC. Our hypotheses are that participation in the MC compared to a wait-list control condition will result in positive relationship health trajectories for intimacy, acceptance, and relationship satisfaction over the course of 2 years.

Methods

Participants

Participants included 215 couples recruited from a northeastern metropolitan area. Of the 430 individuals who participated in the study, there were 218 women and 212 men, for a total of 209 opposite-sex couples and 6 same-sex couples. Due to partner distinguishability on outcome variables, same-sex couples were excluded from the analysis. Participants ranged in age from 20 to 78 years, with an average age for women of 44.5 years ($SD = 10.8$), and average age for men of 46 years ($SD = 11.4$). The majority of participants were Caucasian (93.9%), followed by African-American (2.7%), Asian (2.7%), Hispanic (1.7%), and American Indian (.7%). Overall, couples had been married an average of 15.2 years (with a range of 22 days to 56 years), and had an average household income in the \$75,000 – \$99,000 range (with incomes ranging from under \$10,000 to over \$100,000). In the local metropolitan area, individuals have a median household income of \$66,389 (U.S. Census, 2009). About 82% of the individuals in the study had children, with an average of 2 children per couple. Eighty-eight percent of the sample had a high school degree, and 43.9% had a bachelor's degree or higher. About 84% of people in the metropolitan area graduated from high school, and about 30% had a bachelor's degree or higher. By comparison, our sample has a somewhat higher income and education level. Our sample has a similar racial make-up compared to the surrounding area.

To compare the initial distress level of our sample to community couples, we used the previously established cutoff score on the QMI of 30.5, with scores below 30.5 indicating significant distress (Funk & Rogge, 2007). About 20% of women and 15% of men fell into the distressed range; together, 26% of couples had at least one couple member who fell into the distressed range. Notably, an additional 7% of women and 7.5% of men fell within 2 points of the cutoff.

Procedures

Couples were recruited for the study through the use of flyers, paper and electronic advertisements, print and broadcast media, and word of mouth. Three hundred thirty four couple members initially contacted the study and were screened for eligibility. To be able to participate in the study, couples had to be both married and cohabitating, and could not currently be attending couples therapy. When couples were determined eligible to participate, the Research Coordinator (RC) randomly assigned an identification number from the master randomization list. The identification number was placed on a key and on the couple's pre-treatment questionnaires. Couples were enrolled in the study following return of the pre-treatment questionnaires and informed of their treatment condition.

Treatment couples completed additional questionnaires at the end of the feedback session, and at 2-week, 6-month, and 1-year follow-up, with control couples' questionnaires timed to coincide accordingly. Once the 1-year follow-up questionnaires were completed, treatment couples returned for a booster MC. All couples completed questionnaires at the 1-year feedback, 1-year 2-week, 18-month, and 2-year time points. Couples were paid for the completion of each packet, in escalating amounts ranging from \$25 to \$100. Participants

were paid a total of \$575 if the entire study was completed. The study was completed in compliance with the Institutional Review Board at the first author's institution.

By the 2-year time point, 22 of 101 control couples had dropped out of the study, whereas 35 of 108 treatment couples dropped out, leading to a total dropout rate of 27%. The participant flow is illustrated in Figure 1. Fisher's Exact Test indicated that the different dropout rate between treatment and control groups was not significant ($p = .09$). This dropout rate is similar to the 30% average drop out rate for longitudinal studies of this nature (Bradbury & Karney, 2010). We further probed between-group differences in dropouts by looking at three periods: before the intervention, between the intervention and booster, and after the booster. Only the preintervention difference was significant, with 1 of 101 control couples dropping out and 13 of 108 treatment couples dropping out ($p = .001$), indicating that some couples, having learned of their randomization, opted not to attend treatment. These couples were not significantly different from their counterparts who attended treatment on baseline distress, intimacy, or acceptance; however, they tended to have fewer years of education (mean difference = 2.00, $t(106) = 2.67$, $p = .009$ for women, and mean difference = 1.52, $t(106) = 2.06$, $p = .042$ for men), and lower household income with the mean value falling into the \$50,000–75,000 range as compared to the \$75,000–\$100,000 range for completers, $t(106) = 2.13$, $p = .035$. Therefore, this differential drop out rate appears to be more related to socioeconomic factors than relational ones.

Over the entire sample, dropouts tended to have lower income ($t(204) = 2.47$, $p = .014$), be more distressed ($t(78.02) = 2.43$, $p = .017$), have lower intimacy ($t(207) = 2.01$, $p = .046$), and have lower acceptance, $t(207) = 2.17$, $p = .031$. Treatment and control dropouts did not differ significantly on any variables that we examined.

Power Analysis

A power analysis was conducted using Cohen's (1988) recommendations. Analyses indicated that the study was sufficiently powered to detect the 0.40 effect size of the previous studies with the achieved sample size of $N = 215$.

Measures

The Quality Marriage Index (QMI; Norton, 1983)—The QMI is a 6-item measure that assesses a partner's evaluation of the quality of her or his marriage. The first five items in the measure are each ranked on a 7-point scale, ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Examples of these items include, “we have a good relationship,” and “my relationship with my partner makes me happy.” The final question asks participants to rate their overall level of happiness from 1 (*not at all happy*) to 10 (*extremely happy*). The sum of the items was used, with a possible range from 6 to 45. The measure has been extensively validated. In the current study, the internal reliability of the measure was very high (Cronbach's $\alpha = 0.97$).

The Intimate Safety Questionnaire (ISQ; Cordova, Gee, & Warren, 2005)—The ISQ is a 27-item self-report scale designed to measure the degree to which partners feel safe being vulnerable with each other across different domains of the relationship. Items include

“When I need to cry, I go to my partner,” and “I feel comfortable telling my partner things I would not tell anybody else.” Respondents rated each statement on a scale ranging from 0 (*never*) to 4 (*always*), with higher mean scores indicating higher levels of intimacy. In the current sample, internal reliability was high (Cronbach’s $\alpha = 0.91$). The ISQ has been found to be significantly correlated with the subscales of the Personal Assessment of Intimacy in Relationships Questionnaire (PAIR; Schaefer & Olson, 1981). In addition, the ISQ has been found to be significantly correlated with the Global Distress Scale of the Marital Satisfaction Inventory (Cordova, Gee, & Warren, 2005), providing support for its construct validity. We use the ISQ as our measure of intimacy in this study because it is a theory-driven questionnaire most consistent with our theory of change with regard to the MC. Additional details regarding the ISQ can be obtained from the first author.

The Relationship Acceptance Questionnaire (RAQ; Wachs & Cordova, 2007)—

The RAQ is a 26-item self-report measure of how accepting the participant feels of his/her partner and how accepted the participant feels by his/her partner. Participants responded on a scale of 1 (*strongly disagree*) to 5 (*strongly agree*) to such statements as “I feel like my partner accepts me as a person ‘warts and all’” and “I don’t dwell on my partner’s weaknesses.” Higher scores indicated higher levels of acceptance. This scale had high internal consistency for both partner felt acceptance (.94) and acceptance of partner (.91). In this study, we focused exclusively on how accepted the respondent felt by her or his partner, as this is believed to be more a more accurate measure of relational acceptance.

The Marital Satisfaction Inventory-Revised: Global Distress Scale (MSI-R-GDS; Snyder, 1997)—

The MSI-R GDS is a well-validated subscale measuring overall relationship satisfaction. In this sample, Cronbach’s alpha for the GDS was .93. Standardized mean T-scores are grouped by sex into categories such that scores between 39–49 indicate low distress, 50–59 indicate moderate distress, and 60 and above indicate severe distress.

The Marriage Checkup

The MC consists of Assessment and Feedback sessions, each lasting approximately 2 hours. During the Assessment session, couples were asked to discuss their reasons for seeking an MC, how they hoped to benefit, and about the history of their relationship (Buehlman, Gottman, & Katz, 1992). Partners engaged in two social support interactions and one problem solving interaction. The Assessment visit concluded with a therapeutic interview, which involved a discussion of both the partners’ strengths and areas of concern, using the techniques of IBCT (Jacobson & Christensen, 1998).

The feedback session was conducted approximately two weeks after the assessment session. Each couple’s feedback report was customized to their particular strengths and areas of concern. The session began with a review of the couple’s history, continued with a review of the couples’ strengths, as well as a summary of their questionnaire scores, and concluded by addressing couples’ concerns. Therapists presented a menu of options for how the couple might effectively address each area of concern and also asked couples to generate their own suggestions. Therapists used Motivational Interviewing techniques (Miller & Rollnick,

2002) to activate couples in the service of their marital health and IBCT techniques to promote increased acceptance, intimacy, and satisfaction.

Booster visits were conducted 1 year later in the same general format as described above. Couples were asked for an update about positive events in the past year, and also to follow-up on the concerns that they had discussed previously as well as new concerns. Both old and new concerns were reviewed during the feedback session.

Treatment Fidelity

Therapists included the first author and nine doctoral students, all of whom were trained and supervised weekly by the first author. For the assessment visits, 7 therapists saw a mean of 18.29 couples ($SD = 7.41$), with a range of 3–25. For the booster visits, 10 therapists (the original 7, plus an additional 3 who replaced those who finished the program) saw a mean of 12.80 couples ($SD = 9.91$), with a range of 2–29. An adherence scale was developed to assess therapist fidelity to the MC manual. Nineteen codes reflected therapist behavior during the assessment and feedback sessions of the MC. Four doctoral students served as coders, and two of the authors (JC and CJF) served as the coding trainers. Coding teams met weekly in pairs to discuss and reach consensus on final ratings. Out of the 101 treatment couple videos, 25 (24.8%) were selected via stratified sampling (per each therapist) and those tapes were rated for adherence. Each behavior was rated on a 5-point scale of therapist adherence ranging from 1 (*not at all*) to 5 (*extensively*). Due to low variability, traditional intraclass correlations were inappropriate. Consequently, exact percent agreement between codes was calculated, which is a value that represents the percentage of the time that raters agreed on a code. Exact percent agreement ranged from .76 to .85 across teams. Percent agreement within one level of the scale (which measures the percentage of the time that coders agree within one point) was also calculated across all raters. This level of percent agreement ranged from .96 to .99. The average adherence rating was 4.67 ($SD = .29$), indicating that therapists adhered to the MC manual.

Data Analysis

Analytical models were constructed using a latent growth curve framework. Couple was used as the unit of analysis and husband and wife growth curves were modeled in parallel. This approach allowed a flexible examination of partner distinguishability for variance components, slopes, intercepts, and treatment effects. It also allowed inclusion of multiple indicators of distress into a latent variable, enabling a separation of error due to trajectory disturbance and error specific to each measurement scale, resulting in increased power (Bollen & Curran, 2006).

Due to the study design, we expected to find brief periods of rapid change immediately after each intervention point, followed by longer periods of slower change, consistent with other intervention work (e.g. Keller et al., 2000). Thus, appropriate handling of nonlinearity became a central challenge to the analysis as we expected that higher order polynomial forms would not change shape sharply enough to capture accurate representations of change over time. Moreover, these forms could inappropriately smooth qualitatively distinct assessment points that occur close in temporal proximity, such as the 1-year follow-up (one

year after an intervention) and the booster feedback several weeks later. While some researchers have profitably employed piecewise linear models (2000) to capture such discontinuities, we opted instead to code treatment as a time-varying variable. Piecewise models in this analysis would have required four pieces and would still have imposed potentially inappropriate linearity assumptions over the follow-up period. The time-varying model parameterization is more parsimonious, computationally less intensive, allows treatment to have a different effect at each time point, and has the interpretive benefit of allowing between-group mean differences to be read directly from the model output.

A single dummy-coded variable, representing exposure to treatment, was coded as 0 for the control group and 1 for the treatment group. To maintain fidelity to an intent-to-treat analysis, couples that dropped out before the booster session were still coded as 1 after the booster session. All post-baseline measures were regressed directly on the treatment variable, channeling deviations in individual trajectories related to treatment through this path rather than through the slope, and generating a separate regression path for each time point. No path was included between treatment and baseline status, equivalent to coding all couples as zero at baseline and consistent with the time-varying approach.

Results

All growth curve analyses were conducted using Mplus Version 7.11 (Muthén & Muthén, 1998–2012). Other analyses were conducted using IBM SPSS Version 20. Overall significant effects for intervention were based on an intent-to-treat analysis including all enrolled participants. Descriptive statistics for all four primary outcome measures are shown in Table 1. Missing data were estimated using the full information maximum likelihood algorithm. Models were built following a paradigm of successive imposition of restrictions and tests of model fit. We note statistical decisions and describe the final models, but for brevity, leave out many of the intermediate deviance testing values. More information on the final models can be obtained from the first author. Model tests indicated that partners were not “exchangeable” on study variables, i.e. were distinguishable, so the study’s six same-sex couples were excluded from the analyses.

Marital Satisfaction

We used a multiple indicator model to examine change in marital satisfaction, which was measured as a latent variable combining the Quality of Marriage Index (QMI) and the global distress subscale of the Marital Status Inventory-Revised (GDS). The latent variable was scaled to QMI. Because GDS was not included in either of the feedback measures, in order to allow identification of the initial model, residuals of QMI at both feedback time points were constrained to equal the residuals of QMI at the following time point, just two weeks later. Due to small negative residual variances in these factors at the feedback time point, these residuals were fixed at zero.

Factor loadings of the slope were fixed to the number of weeks from baseline divided by 10 (to aid in convergence), modeling individual trajectories as calendar time. A linear trajectory accounted for significant variance, while higher order polynomials worsened the balance of fit and parsimony and did not account for significant variability; so individual trajectories

were modeled as linear. Note that because treatment is included as a time-varying variable, the model can still capture non-linear effects related to treatment. Partners' intercepts, slopes, and measurement residuals were allowed to correlate. The measurement models were found to be both time invariant and equivalent between partners, and therefore constrained to reflect these findings. After building individual curves and testing the measurement model, distinguishability between sexes was tested. Men and women's intercepts for QMI and GDS were significantly different (Wald (2) = 6.53, $p = .038$), whereas slopes were not significantly different, so men and women's slopes were constrained equal. Finally, the treatment effect was not significantly different between sexes (Wald (8) = 9.06, $p = .34$). Introducing an equality constraint produced a large increase in model fit; the Bayesian Information Criterion decreased by 37 points, where 10 points is generally considered a "very strong" indication of a better model (Raftery, 1995).

Due to a moderate level of kurtosis, we used a maximum likelihood estimator with robust standard errors. While the chi-square was significant ($\chi^2(542) = 911.81, p < .001$), indicating that the model was an imperfect fit to the data, chi-square tests are generally considered overly sensitive for a large dataset (Bollen & Curran, 2006). Otherwise, model fit was adequate: CFI = .94, TLI = .94, RMSEA = .06. Treatment effects from the final model, scaled to QMI points, are presented in Table 2, and graphically depicted in Figure 2. Due to known issues of variance shrinkage when using best linear unbiased prediction (Verbeke & Molenberghs, 2000), Cohen's d effect sizes were derived using standard deviations observed in the raw data, which produced a more conservative calculation than would have been found using model-based variance parameters. In a few cases, this caused the 95% confidence interval for Cohen's d to include zero when the p -value was at or below .05.

The control group's slope of marital satisfaction was not significantly different from zero (slope = .011, $p = .86$). Treatment effects can be interpreted as the degree to which treatment deflected the trajectory at a given time point. Effect sizes were largest immediately after treatment, stabilized at the two-week post measurement point, and were largely sustained throughout the first year of follow-up. The same pattern repeated over the follow-up period to the second intervention, although the between-group effect at the 2-year follow-up point dropped below statistical significance. Effect sizes ranged from $d = .11$ to $d = .39$, with a mean effect across follow-up of $d = .23$.

Intimacy

We followed the same modeling paradigm of imposing successive model restrictions and testing partner distinguishability to build the intimacy and acceptance models. In the final intimacy model, partner intercepts, slopes, and treatment effects were statistically indistinguishable, and therefore constrained equal.

Due to mild kurtosis, we used a maximum likelihood estimator with robust standard errors. While the chi-square indicates that model fit was not perfect ($\chi^2(177) = 271.14, p < .001$), it was otherwise adequate: CFI = .96, TLI = .96, RMSEA = .06. Model-based estimates and effect sizes are presented in Table 2. Trajectories are graphically depicted in Figure 3. The control group's slope of intimacy was not significantly different from zero (slope = $-.007, p = .17$). As can be seen in Table 2, we found statistically significant effects for intimacy

across the entire follow-up period (all p -values significant at $p < .001$ with the exception of 2-years, where $p = .002$). Cohen's d values were mostly in the small range through 6 months, and consistently in the small-to-medium range following that. Effect sizes for intimacy ranged from $d = .20$ to $.55$ (mean $d = .37$).

Acceptance

We found significantly different intercepts and treatment effects for men and women, so we allowed intercepts, slopes, and treatment effects to vary freely between partners. The chi-square was again significant ($\chi^2(167) = 305.44, p < .001$) but model fit was otherwise good: CFI = .95, TLI = .95, RMSEA = .06. Neither control men's nor control women's slopes were significantly different from zero (slope_m = .007, $p = .41$ and slope_w = -.003, $p = .64$). Treatment effects are presented in Table 2, and trajectories are shown graphically in Figure 4. For both men and women, treatment effects followed a distinct waxing-and-waning pattern, with sizeable bumps immediately after both intervention points, followed by a tapering off over the course of follow-up. For both sexes, effect sizes were in the small range, with women's ranging from $d = .17$ to $.47$ (mean $d = .34$), and men's ranging from $d = .11$ to $.38$ (mean $d = .25$). While treatment effects appeared similar between sexes over the first year of follow-up, they diverged over the second year, with women's effects largely sustained through 2 years (although at 2 years significance is borderline at $p = .050$), and the effect for men disappearing by 1-year 6-months.

Reliable Change Index

We used Jacobson and Truax's (1991) formula to calculate reliable change. To account for clustering in the chi-square analyses, for all variables except for acceptance, analyses were conducted at the couple level; partners' change scores were averaged and couples that scored above the threshold were coded as changed. We imputed missing data with subject-specific model-based estimates generated from empirical Bayes factors. To err on the conservative side, we removed the estimated treatment effect from all couples that dropped out before treatment, despite the fact that the model estimated the treatment effect over the entire sample. For parsimony, we present estimates at 6-month intervals. Reliable change estimates are provided in Table 3.

For QMI, reliable change required couples to average a 4.02-point gain from baseline. For GDS, reliable change required couples to average a 6.86-point gain from baseline. Reliable change did not reach statistical significance for either of the distress variables, although particularly for GDS, a clear pattern emerged of a higher proportion of treatment than control couples meeting criteria for change. GDS trended towards significance at 2-year follow-up, with 22 treatment couples meeting criteria, but just 11 control couples ($\chi^2(1, N = 209) = 3.53, p = .06$).

For intimacy, reliable change required couple members to average a 0.42-point increase. Between-group differences were significant at each follow-up point after 6-months, with 15 treatment couples and 4 control couples meeting criteria at 1-year ($\chi^2(1) = 6.23, p = .013$), and 17 treatment couples vs. 7 control couples meeting criteria at 2-years ($\chi^2(1) = 3.99, p = .046$).

For acceptance, women's and men's scores were computed separately. Reliable change required a 0.54 gain for women and a 0.60 gain for men. Women's differences were significant only at 6-months, whereas men's were significant through 1-year and trended at 1-year 6-months.

Sensitivity Analysis and Missing Data

We examined the data with several different models to determine the sensitivity of findings to the model parameterization. Most notably, because the full information maximum likelihood algorithm uses all data to estimate missing values, we wished to explore whether the brief spike immediately after treatment may have biased other estimates upwards. We removed both the feedback and 2 week post-feedback time points for both the initial assessment and the booster session, leaving just the longer-term follow-up measures taken 6 months apart. Results were consistent with the findings reported here.

We used pattern mixture models (Hedeker & Gibbons, 1997) to explore the sensitivity of the findings to patterns of missing data, with particular interest in whether attrition bias impacted the present findings. While we parameterized various model permutations, we found the best balance of theory and power by coding couples where both members dropped out before the final follow-up period as dropouts, and couples that completed data at the last time point as completers. Of primary interest was the interaction between dropout status and the treatment effect, which would suggest that asymmetric attrition biased parameter estimates of treatment effects. None of our models revealed any differences in trajectories or treatment-related deflections of trajectories, indicating that findings were robust to missing data patterns.

Discussion

The results of this randomized controlled trial of the Marriage Checkup support our hypothesis that the MC improves multiple domains of relationship health, and that yearly boosters are effective to augment and maintain those gains. In brief, relationship satisfaction, intimacy, and acceptance significantly improved compared to the control condition for up to two years following the MC when boosters were given at the 1-year time point. For satisfaction, MC couples were significantly more satisfied than control couples at all post-intervention time points but the final two-year follow-up. For intimacy, MC couples reported significantly more intimacy than control couples throughout 2 years of follow-up. Similarly for acceptance, MC couples reported significantly more acceptance over the 2 years.

Broadly, participation in the MC seemed to (1) reorient couples toward the most positive qualities of their relationships, (2) foster acceptance of common issues, differences, and patterns, (3) build intimacy bridges rooted in deeper compassionate understanding, and (4) generally improve relationship health by reactivating partners in the service of a more vibrant and engaged relationship. The effect of the MC appears to peak immediately following the intervention itself and then stabilizes into sustained improvement, with some waning over the course of the two-year follow-up. Thus, the results suggest a fairly robust

treatment effect and enough waning over the course of one or two years to support our presumption regarding the need for regular, annual, checkups.

Reliable change statistics were strongest for intimacy, with significantly more treatment than control couples meeting criteria throughout the entire follow-up period. As far as acceptance, while women's treatment effects were larger, more men met criteria for reliable change, suggesting that the men's effects may have been driven by a fewer number of individuals making larger gains, whereas women's effects may have been driven by more consistent gains across the sample. Treatment couples were not significantly more likely to meet reliable change criteria for either of the satisfaction variables, although they trended towards a significant difference at the 2-year point on the Marital Status Inventory Global Distress Scale.

Results of the current study compare favorably to previous findings (e.g. Cordova et al., 2005) with small to medium effect sizes for intimacy (mean d of .37) and felt acceptance (mean d of .30), and small effects for relationship satisfaction (mean d of .23). A meta-analytic review of marital education programs has found that while controlled studies of education programs have displayed a strong improvement in couple communications skills, relationship satisfaction gains have been more modest. In comparably designed studies that included both post-assessment and follow-up measures, the average effect size immediately post-intervention was .24, and .28 at follow-up (although the average duration of the follow-up period was shorter than the present study, typically lasting 3–6 months; Hawkins, Blanchard, Baldwin & Fawcett, 2008), versus .29 and .39 immediately post-intervention for the MC. On the other hand, studies of tertiary treatment (average of 23 sessions; Christensen et al., 2004) have found pre- to post-effect sizes on the order of .86. Considered in this context, the Marriage Checkup effects appear comparable to marital education programs.

Considering the obtained effect sizes, if the program were implemented on a larger scale, public health effects could be substantial. It has been argued (e.g., Kraemer, 2010) that even small effect sizes can have large public health consequences, depending on the risks and costs associated with treatment. The example often given is that of the effect size for a small dose of aspirin in preventing heart attacks (reported squared correlation = .0011; e.g., Rosenthal, 1994). Though the effect size in this example is comparably small, the public health effect at a population level is significant, especially considering the low cost and low risk of administration.

Although the format of the MC is substantially less intensive than tertiary treatment, the requirement for face-to-face visits does make the MC more intensive than wholly web-based interventions. Although we are in the process of developing a web-based MC, we believe that there will always be benefits of face-to-face contact including (1) better and more easily established therapeutic rapport, (2) greater motivational influence of the provider and greater ability to control the therapeutic environment, (3) greater ease with which the provider can detect relationship patterns and facilitate compassionate understanding and (4) more consistency with the health checkup model that participants are familiar with from physical and dental checkups – perhaps adding to the seriousness with which couples take the experience. At the same time, wholly web based interventions have the added benefit of

even further lowering the barriers to accessing care, essential in addressing the public health problem at hand.

Although promising, the results of this study speak to the efficacy, not the effectiveness, of the Marriage Checkup. Our intention has been to create the MC as an intervention that can be easily disseminated, with an eye toward dissemination through primary care facilities. Toward that end, supported by a grant from the Administration of Children and Families, we have begun dissemination trials through a Tennessee primary care system serving a mostly underserved population. In addition, we are piloting a similar primary care based version of the MC with our partners in the Air Force primary care system.

Similarly, a primary goal of the MC has been to reach at-risk couples who otherwise would not have sought help for marital concerns. Although the MC has been shown to be successful in attracting couples at risk for marital deterioration who might not have otherwise sought care, in this study we also found that couples with lower income and lower levels of education were more likely to drop out. Our supposition is that lower SES couples face more obstacles to effective treatment seeking and are therefore more likely to fall out of the study. These results suggest that effectively reaching lower SES couples continues to be an important challenge. We are currently attempting to address this in our work with Dr. Kristi Coop Gordon adapting the MC for use with lower income couples by more actively taking the MC to more accessible settings.

There were analytic limitations to consider. For the reliable change index, it was necessary to impute missing data. By choosing a model-based approach, we implicitly chose precision over unbiasedness (Singer & Willett, 2003). One drawback to this approach is variance shrinkage; by decreasing variability in the missing data, fewer couples will meet criteria for reliable change through chance alone. Because this study had more treatment than control dropouts, decreasing the variability of imputed data may have functioned to underestimate the number of treatment couples meeting change criteria. At the same time, using a trajectory-based approach is more likely to produce accurate estimates in long-term follow-up than other strategies such as multiple imputation, which tends to produce missing values that are “too healthy” (Engels & Diehr, 2003), a drawback in a sample where separation and divorce drive some dropout. Additionally, the interpretability of reliable change statistics is complicated because a prevention study like ours samples couples from across the satisfaction spectrum, including already satisfied couples at or near the ceiling of the satisfaction measures, compared to clinical samples in which all participants meet distress criteria prior to the intervention.

It is also important to consider the generalizability of the findings. With regard to diversity, our sample was drawn primarily from an educated Caucasian population, although this was not dramatically different from the surrounding community. This problem, which is common in both research and clinical settings, underscores the importance of actively recruiting diverse and traditionally underserved communities. A final important consideration for generalizability is transitioning from paid research participant to a paying community consumer. Although MC participants were compensated for their time, the amount of compensation was minimal in relation to the time requirement. The amounts of

payments and timing of disbursement (increasing over the course of two years) were carefully selected to be in line with the current standards of practice to assure that they were not coercive. Additionally, on a survey asking about perceived value, 72.8% of current MC couples said that they would be willing to pay \$50 or more for an MC, and 39.6% said that they would be willing to pay over \$100, indicating that they valued their participation in the MC above and beyond the influence of incentives.

Taken as a whole, the MC program of research thus far has demonstrated safety and acceptability, preliminary efficacy and longer-term positive results, attractiveness to a broader target population, and now both longitudinal efficacy and the benefits of a repeated annual checkup using the latest analytical techniques. A large scale, multi-year, dissemination trial of the MC in low-income populations began in 2011. We are also actively working on disseminating the intervention to make it as widely available as possible to those who could benefit from it, through publication of a clinician's treatment manual and the development of a web-based program that would make MC materials available to community clinicians and clients. In sum, regular checkups appear to improve and sustain relationship health across a range of variables and thus have the potential to become an important part of an overall comprehensive relationship health system.

Acknowledgments

Support for this project was provided by a grant (R01HD45281) from the Eunice Kennedy Shriver National Institute of Child Health and Human Development to the first author. Recruitment and interview data were published in a 2011 paper. The current project focuses on distinct data outcomes. Results of this study were presented at the 2012 conference for the Association of Behavioral and Cognitive Therapies.

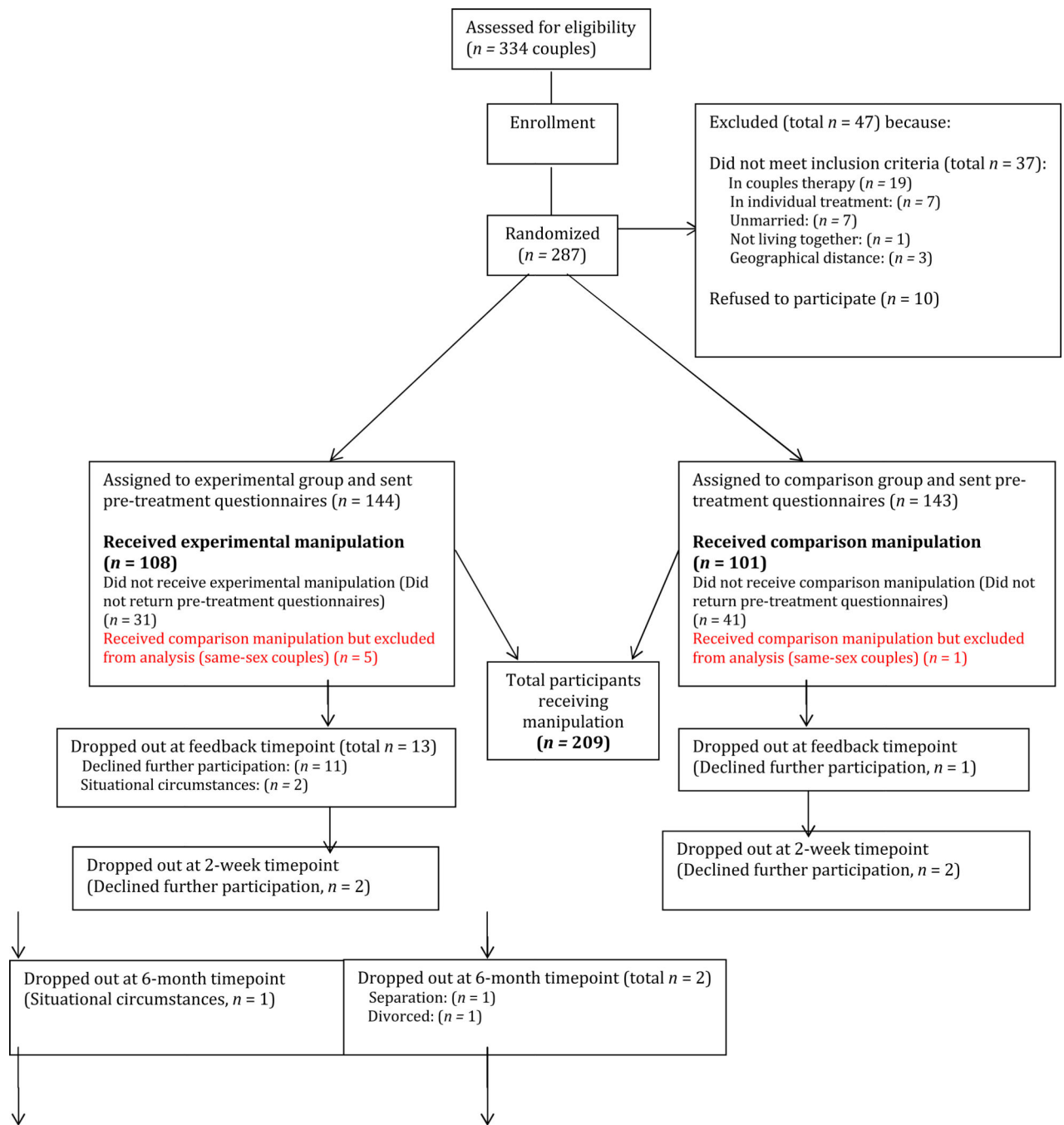
References

- Berger, R.; Hannah, MT. Preventive approaches in couples therapy. PA: Brunner/Mazel; 1999.
- Bollen, KA.; Curran, PJ. Latent Curve Models: A Structural Equation Perspective. Hoboken, New Jersey: John Wiley & Sons; 2006.
- Bradbury, TN.; Karney, BR. Intimate Relationships. New York: W. W. Norton & Company, Inc; 2010.
- Bringle RG, Byers D. Intentions to seek marriage counseling. *Family Relations*. 1997; 46:299–304.
- Buehlman K, Gottman JM, Katz L. How a couple views their past predicts their future: Predicting divorce from an oral history interview. *Journal of Family Psychology*. 1992; 5(3–4):295–318.
- Christensen A, Atkins DC, Baucom B, Yi J. Marital status and satisfaction five years following a randomized clinical trial comparing traditional versus integrative behavioral couple therapy. *Journal of Consulting and Clinical Psychology*. 2010; 78(2):225–235. [PubMed: 20350033]
- Christensen A, Baucom B, Atkins DC, Berns S, Wheeler J, Simpson LE. Traditional versus Integrative Behavioral Couple Therapy for Significantly and Chronically Distressed Couples. *Journal of Consulting and Clinical Psychology*. 2004; 72:176–191. [PubMed: 15065953]
- Cohen, J. Statistical power analysis for the behavioral sciences. Second Edition. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers; 1988.
- Cordova, JV. The Marriage Checkup. Maryland: Jason Aronson; 2009.
- Cordova, JV. The Marriage Checkup Practitioner's Guide: Promoting Lifelong Relationship Health. Washington, DC: American Psychological Association; 2014.
- Cordova JV, Gee CG, Warren LZ. Emotional Skillfulness in Marriage: Intimacy as a Mediator of the Relationship Between Emotional Skillfulness and Marital Satisfaction. *Journal of Social and Clinical Psychology*. 2005; 24:218–235.

- Cordova JV, Scott RL, Dorian M, Mirgain S, Yaeger D, Groot A. The marriage checkup: A motivational interviewing approach to the promotion of marital health with couples at-risk for relationship deterioration. *Behavior Therapy*. 2005; 36:301–310.
- Cordova JV, Warren LZ, Gee CB. Motivational interviewing with couples: An intervention for at-risk couples. *Journal of Marital and Family Therapy*. 2001; 27:315–326. [PubMed: 11436424]
- Cordova JV, Jacobson NS, Christensen A. Acceptance vs. change in behavioral couples therapy: Impact on client communication processes in the therapy session. *Journal of Marital and Family Therapy*. 1998; 24:437–455. [PubMed: 9802004]
- Cummings ME, Goeke-Morey MC, Papp LM. Children's responses to everyday marital conflict tactics in the home. *Child Development*. 2003; 74(6):1918–1929. [PubMed: 14669904]
- Doss BD, Atkins DC, Christensen A. Who's dragging their feet? Husbands and wives seeking marital therapy. *Journal of Marital and Family Therapy*. 2003; 29(2):165–177. [PubMed: 12728776]
- Doss BD, Rhoades GK, Stanley SM, Markman HJ. Marital therapy, retreats and books: The who, what, when and why of relationship help seeking. *Journal of Marital and Family Therapy*. 2009; 35(1):18–29. [PubMed: 19161581]
- Doss BD, Simpson LE, Christensen A. Why do couples seek marital therapy? *Professional Psychology: Research and Practice*. 2004; 35(6):608–614.
- Engels JM, Diehr P. Imputation of missing longitudinal data: a comparison of methods. *Journal of clinical epidemiology*. 2003; 56(10):968–976. [PubMed: 14568628]
- Epstein, NB.; Baucom, DH. Enhanced cognitive-behavioral therapy for couples: A contextual approach. Washington, DC, US: American Psychological Association; 2002.
- Fleming CE, Cordova JV. Predicting relationship help seeking prior to a Marriage Checkup. *Family Relations*. 2012; 61:90–100. [PubMed: 22577242]
- Funk JL, Rogge RD. Testing the ruler with item response theory: Increasing precision of measurement for relationship satisfaction with the couples satisfaction index. *Journal of Family Psychology*. 2007; 21(4):572–583. [PubMed: 18179329]
- Gee CB, Scott RL, Castellani AM, Cordova JV. Predicting 2-Year marital satisfaction from partners' reaction to a marriage checkup. *Journal of Marital and Family Therapy*. 2002; 28:399–408. [PubMed: 12382549]
- Gottman, JM. What predicts divorce? The relationship between marital processes and marital outcomes. Hillsdale, NJ: Lawrence Erlbaum Associates; 1994.
- Gottman, JM.; Gottman, JS. The marriage survival kit. In: Berger, R.; Hannah, MT., editors. Preventative approaches in couples therapy. Philadelphia, PA: Brunner/Mazel; 1999. p. 304-330.
- Gurman AS. Couple therapy research and the practice of couple therapy: Can we talk? *Family Process*. 2011; 50(3):280–292. [PubMed: 21884071]
- Halford WK, O'Donnell C, Lizzio A, Wilson K. Do couples at high risk of relationship problems attend premarriage education? *Journal of Family Psychology*. 2006; 20(1):160–163. [PubMed: 16569101]
- 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. [PubMed: 18837590]
- Hedeker D, Gibbons RD. Application of random-effects pattern-mixture models for missing data in longitudinal studies. *Psychological Methods*. 1997; 2:64–78.
- Jacobson, NS.; Christensen, A. Acceptance and change in couple therapy: A therapist's guide to transforming relationships. New York: W. W. Norton & Co, Inc; 1998.
- Jacobson, NS.; Margolin, G. Marital therapy: Strategies based on social learning and behavior exchange principles. New York: Brunner/Mazel; 1979.
- Jacobson NS, Truax P. Clinical Significance: A Statistical Approach to Defining Meaningful Change in Psychotherapy Research. *Journal of Consulting and Clinical Psychology*. 1991; 59(1):12–19. [PubMed: 2002127]
- Johnson, CA.; Stanley, SM.; Glenn, ND.; Amato, PR.; Nock, SL.; Markman, HJ., et al. Marriage in Oklahoma: 2001 baseline statewide survey on marriage and divorce (S02096 OKDHS). Oklahoma City: Oklahoma Department of Human Services; 2002.

- Keller MB, McCullough JP, Klein DN, Arnow B, Dunner DL, Gelenberg AJ. A comparison of nefazodone, the cognitive behavioral-analysis system of psychotherapy, and their combination for the treatment of chronic depression. *The New England Journal of Medicine*. 2000; 342:1462–1470. [PubMed: 10816183]
- Kraemer HC. Epidemiological Methods: About Time. *International Journal of Environmental Research and Public Health*. 2010; 7(1):29–45. [PubMed: 20195431]
- Kreider, RM. Current Population Reports. Washington, DC: U.S. Census Bureau; 2005. Number, Timing, and Duration of Marriages and Divorces: 2001; p. 70-97.
- Lucas RE. Time does not heal all wounds: A longitudinal study of reaction and adaptation to divorce. *Psychological Science*. 2005; 16(12):945–950. [PubMed: 16313658]
- Markman, HJ.; Stanley, SM.; Blumberg, SL. *Fighting for your marriage*. San Francisco: Jossey-Bass, Inc; 2001.
- Miller, WR.; Rollnick, S. *Motivational interviewing: Preparing people for change*. 2nd ed.. New York: Guilford Press; 2002.
- Morrill MI, Fleming CE, Harp AG, Sollenberger JW, Darling EV, Cordova JV. The Marriage Checkup: Increasing access to marital health care. *Family Process*. 2011; 50(4):471–485. [PubMed: 22145720]
- Muthén, LK.; Muthén, BO. *Mplus User's Guide*. Seventh Edition. Los Angeles, CA: Muthén & Muthén;
- Norton R. Measuring Marital Quality: A critical look at the dependent variable. *Journal of Marriage and the Family*. 1983; 45:141–151.
- Notarius, C.; Buongiorno, J. Wait time until professional treatment in marital therapy. Washington D.C: Unpublished paper, Catholic University of America; 1992.
- Olson, DH.; Olson, AK. PREPARE/ENRICH Program: Version 2000. In: Berger, R.; Hannah, M., editors. *Preventative approaches in couples therapy*. Philadelphia: Brunner/Mazel; 1999. p. 196-216.
- Proulx C, Helms HM, Buehler C. Marital quality and personal well-being: A meta-analysis. *Journal of Marriage and Family Therapy*. 2007; 69:576–593.
- Raftery AE. Bayesian model selection in social research. *Sociological methodology*. 1995; 25:111–164.
- Rosenthal, R. Parametric measures of effect size. In: Cooper, H.; Hedges, LV., editors. *The handbook of research synthesis*. New York, NY: Russell Sage Foundation; 1994. p. 231-244.
- Schaefer MT, Olson DH. Assessing intimacy: The PAIR inventory. *Journal of Marital and Family Therapy*. 1981; 7:47–60.
- Shadish WR, Baldwin SA. Effects of behavioral marital therapy: A meta-analysis of randomized controlled trials. *Journal of consulting and clinical psychology*. 2005; 73(1):6–14. [PubMed: 15709827]
- Singer, JD.; Willett, JB. *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press; 2003.
- Snyder, DK. *Marital Satisfaction Inventory, Revised (MSI-R): Manual*. Los Angeles: Western Psychological Services; 1997.
- 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. [PubMed: 9103731]
- Uebelacker LA, Hecht J, Miller IW. The Family Check-Up: A pilot study of a brief intervention to improve family functioning in adults. *Family Process*. 2006; 45(2):223–236. [PubMed: 16768020]
- U.S. Census. 2009. <http://www.census.gov/population/socdemo/computer/2007/tab01.xls>
- Verbeke, G.; Molenberghs, G. *Linear Mixed models for Longitudinal Data*. New York: Springer-Verlag; 2000. p. 64-74.
- Veroff JB. The dynamics of help-seeking in men and women: A national survey study. *Psychiatry: Journal For The Study Of Interpersonal Processes*. 1981; 44(3):189–200.
- Wachs, K.; Cordova, JV. *The Relational Acceptance Questionnaire (RAQ)*. Worcester, Massachusetts: Unpublished measure, Department of Psychology, Clark University; 2007.

- Whisman MA. Marital Distress and *DSM-IV* Psychiatric Disorders in a Population-Based National Survey. *Journal of Abnormal Psychology*. 2007; 116(3):638–643. [PubMed: 17696721]
- Whisman MA, Uebelacker LA. Distress and impairment associated with relationship discord in a national sample of married or cohabitating adults. *Journal of Family Psychology*. 2006; 20:369–377. [PubMed: 16937993]
- Wolcott IH. Seeking help for marital problems before separation. *Australian Journal of Sex, Marriage & Family*. 1986; 7(3):154–164.



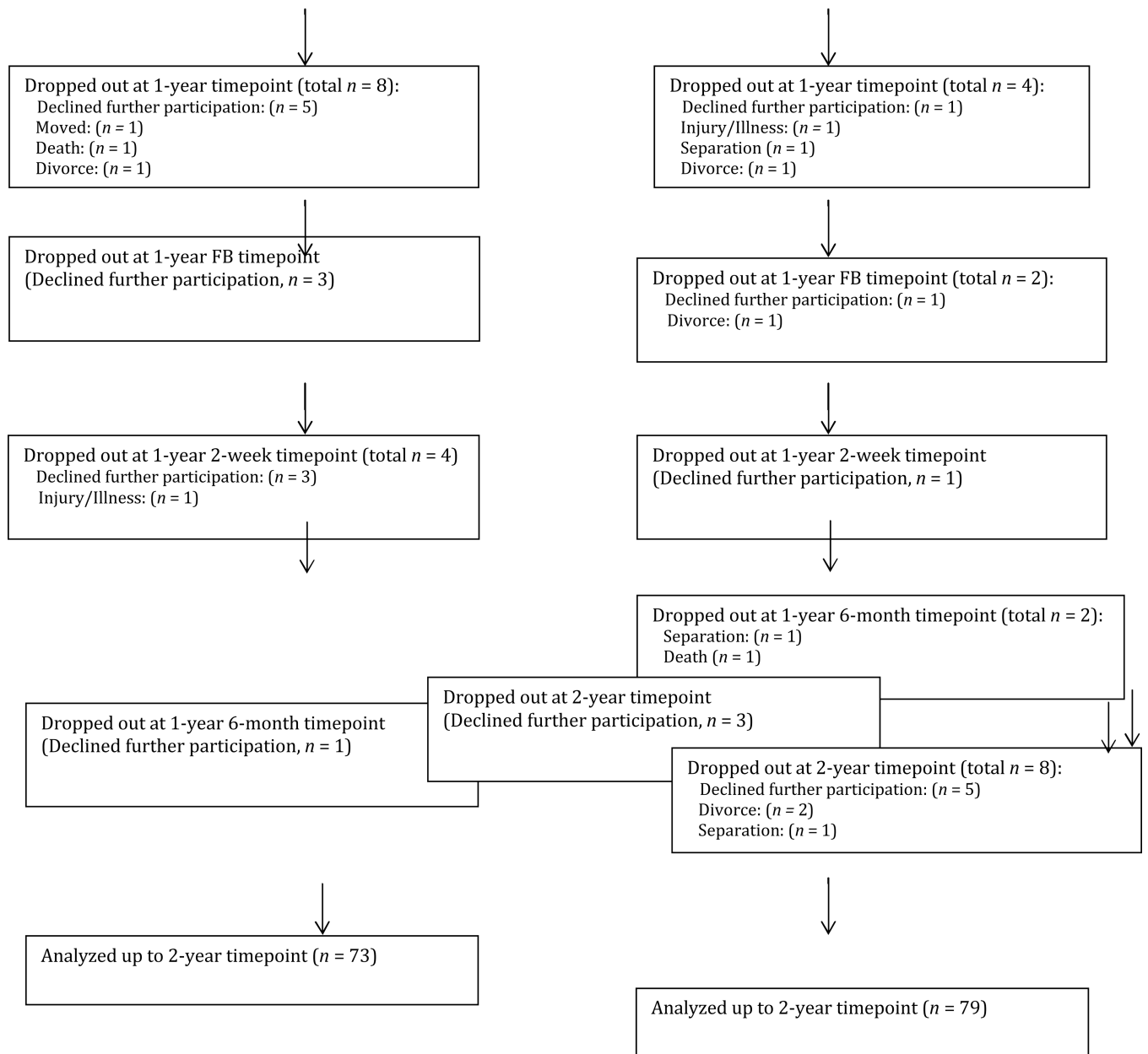


Figure 1.
Participant Flow.

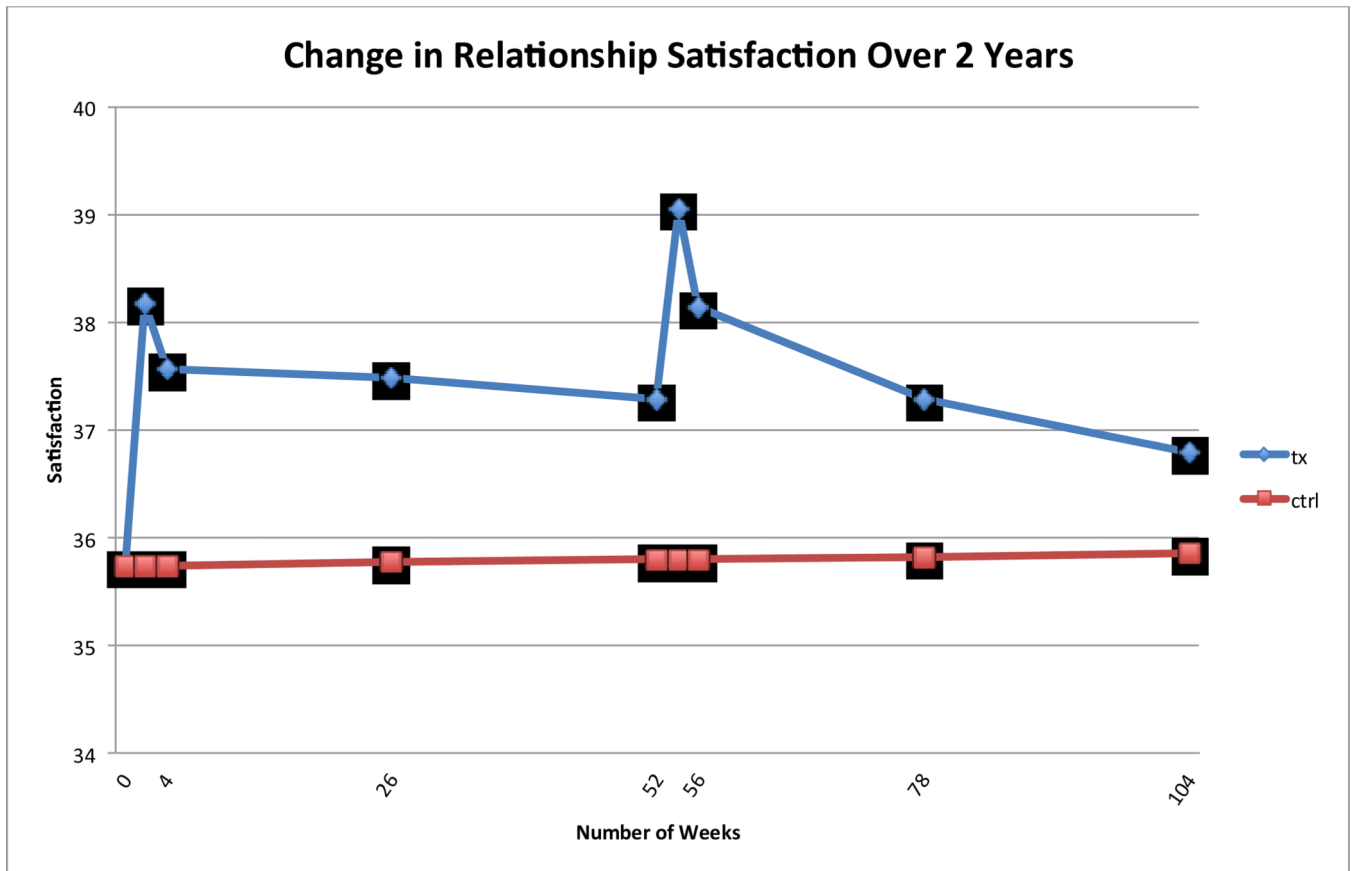


Figure 2.

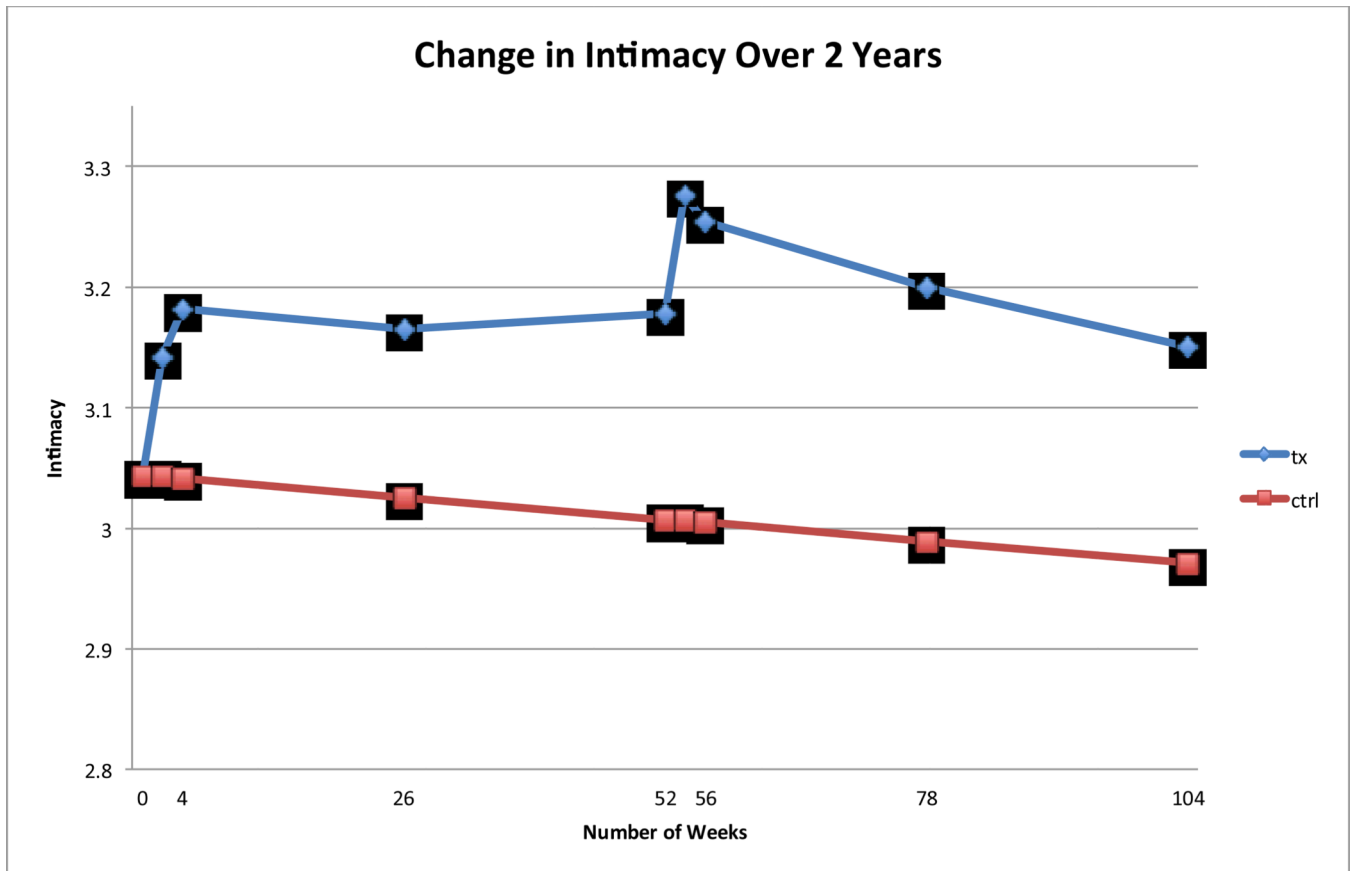


Figure 3.

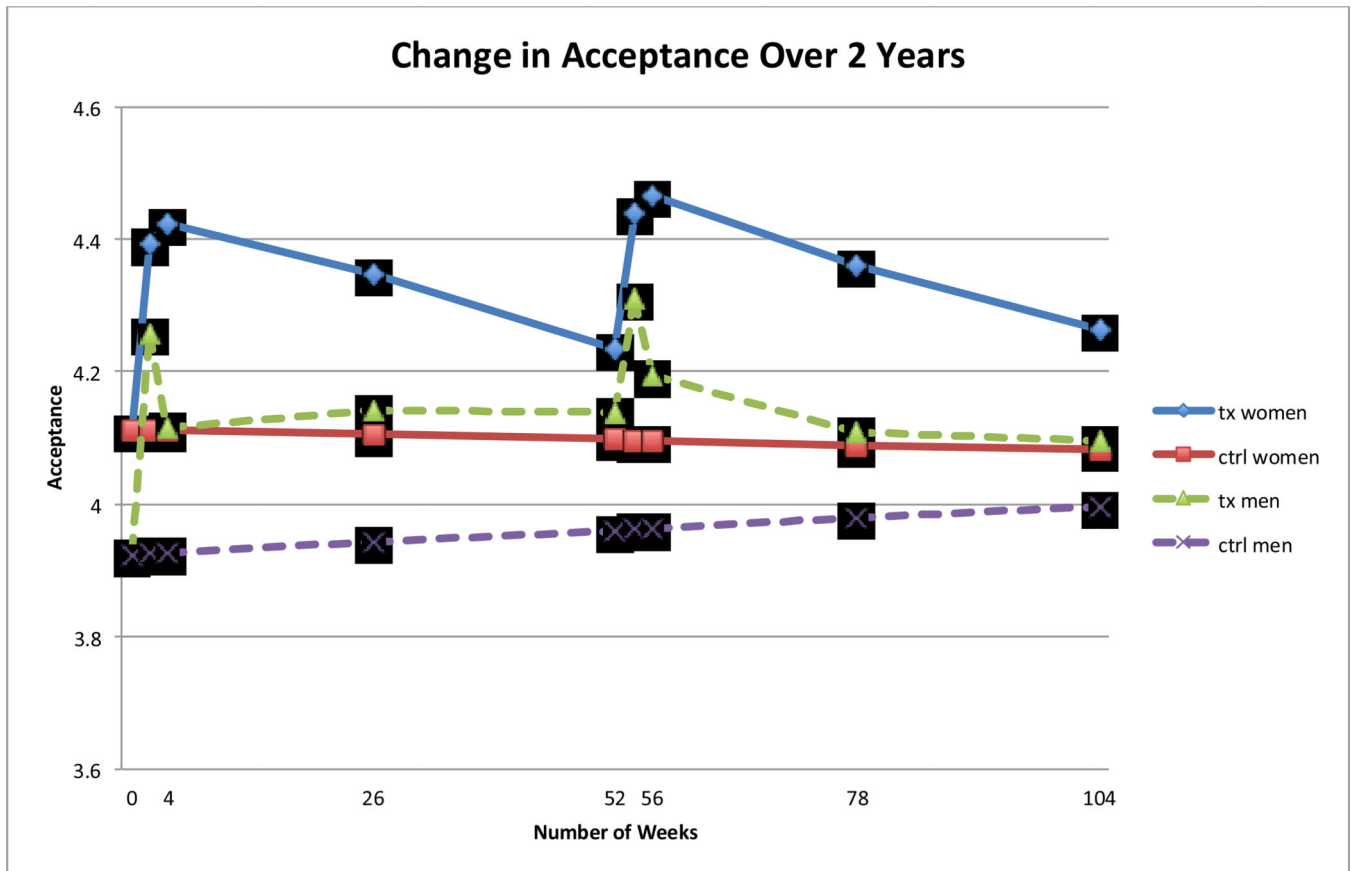


Figure 4.

Table 1
Descriptive Statistics for Marital Satisfaction, Global Distress, Intimacy, and Acceptance

Measure	Base	FB	2wk	6mo	1yr	1yr FB	1yr 2wk	1yr 6mo	2yr
QMI									
Control	36.37 (8.39)	36.80 (7.67)	36.48 (7.91)	37.46 (7.33)	37.32 (7.43)	37.79 (7.82)	37.14 (8.13)	37.37 (8.24)	37.66 (8.29)
MC	35.97 (8.35)	38.41 (7.07)	38.30 (6.81)	37.42 (8.04)	37.79 (7.64)	40.23 (5.12)	38.84 (6.69)	38.24 (6.86)	37.73 (7.34)
GDS									
Control	52.71 (9.48)	-	50.96 (9.40)	50.47 (9.14)	50.66 (9.12)	-	50.79 (9.70)	51.21 (9.14)	51.34 (8.50)
MC	52.50 (9.24)	-	50.72 (8.61)	50.12 (8.86)	49.14 (9.25)	-	48.38 (8.81)	49.42 (9.09)	49.95 (9.08)
Intimacy									
Control	3.08 (0.46)	3.08 (0.46)	3.12 (0.48)	3.16 (0.49)	3.12 (0.51)	3.15 (0.51)	3.09 (0.62)	3.08 (0.58)	3.12 (0.55)
MC	3.01 (0.53)	3.12 (0.48)	3.15 (0.47)	3.16 (0.50)	3.19 (0.48)	3.31 (0.44)	3.26 (0.50)	3.21 (0.51)	3.18 (0.52)
Acceptance									
Control	4.05 (0.77)	4.06 (0.77)	4.13 (0.80)	4.14 (0.78)	4.14 (0.81)	4.19 (0.71)	4.08 (0.85)	4.13 (0.83)	4.19 (0.82)
MC	3.97 (0.89)	4.30 (0.74)	4.24 (0.81)	4.23 (0.77)	4.22 (0.75)	4.42 (0.62)	4.38 (0.70)	4.30 (0.76)	4.25 (0.76)

Note. Standard deviations are in parentheses. QMI = Quality of Marriage Index. GDS = Global Distress. Descriptives ignore the clustering due to couples and are based only on available participants per assessment point.

Table 2

Unstandardized Treatment Effects for Satisfaction, Intimacy, and Acceptance

Time	Treatment	SE	p	Cohen's d	95% CI for Cohen's d
Satisfaction					
Feedback	2.45	0.34	<.001	0.29	[0.21, 0.37]
2-weeks	1.83	0.42	<.001	0.22	[0.12, 0.32]
6-months	1.72	0.46	<.001	0.21	[0.10, 0.31]
1-year	1.48	0.59	.012	0.18	[0.04, 0.32]
Booster feedback	3.26	0.54	<.001	0.39	[0.26, 0.52]
1-year 2-weeks	2.33	0.55	<.001	0.28	[0.15, 0.41]
1-year 6-months	1.46	0.70	.034	0.17	[0.01, 0.34]
2-years	0.95	0.81	.276	0.11	[-0.08, 0.30]
Intimacy					
Feedback	0.10	0.02	<.001	0.20	[0.12, 0.28]
2-weeks	0.14	0.03	<.001	0.28	[0.18, 0.38]
6-months	0.14	0.03	<.001	0.28	[0.16, 0.39]
1-year	0.17	0.04	<.001	0.34	[0.19, 0.49]
Booster feedback	0.27	0.04	<.001	0.55	[0.40, 0.71]
1-year 2-weeks	0.25	0.04	<.001	0.51	[0.36, 0.67]
1-year 6-months	0.21	0.05	<.001	0.42	[0.23, 0.62]
2-years	0.18	0.06	.002	0.36	[0.13, 0.60]
Acceptance - women					
Feedback	0.28	0.05	<.001	0.36	[0.23, 0.48]
2-weeks	0.31	0.06	<.001	0.40	[0.25, 0.54]
6-months	0.24	0.06	<.001	0.31	[0.16, 0.45]
1-year	0.14	0.07	.033	0.17	[-0.01, 0.35] [†]
Booster feedback	0.34	0.07	<.001	0.44	[0.27, 0.60]
1-year 2-weeks	0.37	0.08	<.001	0.47	[0.28, 0.66]
1-year 6-months	0.27	0.09	.001	0.34	[0.12, 0.57]
2-years	0.18	0.10	.050	0.23	[-0.03, 0.49] [†]
Acceptance - men					

Time	Treatment	SE	p	Cohen's <i>d</i>	95% CI for Cohen's <i>d</i>
Feedback	0.33	0.05	<.001	0.38	[0.27, 0.49]
2-weeks	0.19	0.06	.001	0.22	[0.09, 0.34]
6-months	0.20	0.06	.001	0.23	[0.10, 0.36]
1-year	0.18	0.07	.012	0.21	[0.04, 0.37]
Booster feedback	0.35	0.07	<.001	0.40	[0.25, 0.54]
1-year 2-weeks	0.23	0.08	.002	0.26	[0.10, 0.43]
1-year 6-months	0.13	0.09	.160	0.15	[-0.06, 0.35]
2-years	0.10	0.10	.340	0.11	[-0.12, 0.35]

Note. † indicates that these confidence intervals include zero due to the use of variances derived from the observed data, which are larger than those in the modeled data.

Table 3

Percentage of Participants Meeting Criteria for Reliable Change

	Marriage Checklist				Control				χ^2	<i>p</i>
	Declined (<i>n</i>)	Stable (<i>n</i>)	Improved (<i>n</i>)	Declined (<i>n</i>)	Stable (<i>n</i>)	Improved (<i>n</i>)	Declined (<i>n</i>)	Stable (<i>n</i>)		
QMI										
6-months	3.7 (4)	76.9 (83)	19.4 (21)	12.9 (13)	74.2 (75)	12.9 (13)	1.66	198		
1-year	13.9 (15)	67.6 (73)	18.5 (20)	16.8 (17)	67.4 (68)	15.8 (16)	0.26	.609		
1-year 6-months	7.4 (8)	71.3 (77)	21.3 (23)	15.8 (16)	65.4 (66)	18.8 (19)	0.2	.650		
2-years	14.8 (16)	63.9 (69)	21.3 (23)	11.9 (12)	67.3 (68)	20.8 (21)	0.01	.930		
GDS										
6-months	3.7 (4)	78.7 (85)	17.6 (19)	3 (3)	87.1 (88)	9.9 (10)	2.58	.108		
1-year	5.6 (6)	73.1 (79)	21.3 (23)	9.9 (10)	75.2 (76)	14.9 (15)	1.46	.230		
1-year 6-months	9.3 (10)	72.2 (78)	18.5 (20)	11.9 (12)	74.2 (75)	13.9 (14)	0.83	.362		
2-years	6.5 (7)	73.1 (79)	20.4 (22)	9.9 (10)	79.2 (80)	10.9 (11)	3.53	.060		
Intimacy										
6-months	2.8 (3)	87 (94)	10.2 (11)	4 (4)	91 (92)	5 (5)	2.02	.155		
1-year	0.9 (1)	85.2 (92)	13.9 (15)	8.9 (9)	87.1 (88)	4 (4)	6.23	.013		
1-year 6-months	0.9 (1)	82.4 (89)	16.7 (18)	12.9 (13)	81.2 (82)	5.9 (6)	5.91	.015		
2-years	4.6 (5)	79.7 (86)	15.7 (17)	11.9 (12)	81.2 (82)	6.9 (7)	3.99	.046		
Acceptance - Women										
6-months	3.7 (4)	76.9 (83)	19.4 (21)	12.9 (13)	78.2 (79)	8.9 (9)	4.71	.030		
1-year	9.3 (10)	74 (80)	16.7 (18)	9.9 (10)	79.2 (80)	10.9 (11)	1.46	.227		
1-year 6-months	6.5 (7)	72.2 (78)	21.3 (23)	12.9 (13)	71.3 (72)	15.8 (16)	0.72	.396		
2-years	9.3 (10)	72.2 (78)	18.5 (20)	9.9 (10)	71.3 (79)	15.8 (16)	1.45	.228		
Acceptance - Men										
6-months	3.7 (4)	69.4 (75)	26.9 (29)	8.9 (9)	80.2 (81)	10.9 (11)	8.59	.003		
1-year	3.7 (4)	72.2 (78)	24.1 (26)	10.9 (11)	78.2 (79)	10.9 (11)	6.23	.013		
1-year 6-months	5.6 (6)	69.4 (75)	25 (27)	13.9 (14)	72.3 (73)	13.9 (14)	3.35	.067		
2-years	9.3 (10)	63 (68)	27.8 (30)	11.9 (12)	70.3 (71)	17.8 (18)	2.34	.126		

Note. QMI = Quality of Marriage Index. GDS = Global Distress. Reliable change required couple members to average a 4.02 point increase from baseline for QMI, 6.86 decrease for GDS, 0.42 increase for intimacy, 0.54 increase for women's acceptance, and .60 increase for men's acceptance. Chi-square tests are based on number improved vs. number stable/declined.