

RANDOMIZED CONTROLLED TRIAL OF THE MARRIAGE CHECKUP: DEPRESSION OUTCOMES

Tatiana D. Gray 
Clark University

Matt Hawrilenko
VA Puget Sound Healthcare System

James V. Cordova
Clark University

The association between relationship functioning and depressive symptoms is well established. This study examined the effects of the Marriage Checkup, a brief two-session Assessment and Feedback relationship intervention, on depressive symptoms. Two hundred and nine married couples participated in the Marriage Checkup and were randomized into Treatment (N = 108) and Waitlist-Control Conditions (N = 101). Compared to the control condition, intervention participants reported significant improvements in depressive symptoms ($d = 0.55$), with an even greater effect for those who were reporting more severe baseline depression symptoms ($d = 0.67$). These outcomes are comparable to those within long-term individual psychotherapy, couple therapy, and pharmacology trials, making this the briefest intervention to date to demonstrate significant improvements in depressive symptoms. Clinical implications are discussed.

Intimate relationship quality is strongly associated with mental health outcomes. The link between depressive symptoms and marital discord is widely recognized (Beach, Katz, Kim, & Brody, 2003; Christian, & O'Leary, & Vivian, 1994; Whisman, 2007) and the bidirectional nature of the association is well established (Whisman, 2001). Individuals suffering depressive symptoms are more likely to have distressed relationships, and individuals in distressed relationships are over three times more likely to experience Major Depression and over five times more likely to experience Dysthymia (Whisman, 1999). The consistency of these findings highlights the important role interpersonal relationships play in overall mental health, and the promising potential for addressing depression at a relational level.

Couple Interventions for Depression

Given the association between marital distress and depression, several researchers have developed couple-based interventions in an effort to address individual mood disorders. Relationship scientists have gathered compelling evidence for the effectiveness of couple therapy for treating the symptoms of depression (Barbato & D'Avanzo, 2008; Beach & O'Leary, 1992; Bodenmann et al., 2008; Cohen, O'Leary, & Foran, 2010). Barbato and D'Avanzo (2008) conducted a meta-analysis of 567 subjects and found that couple therapy for depression as well as individual therapy for depression worked with the added benefit of also reducing relationship distress. In addition,

Tatiana D. Gray, PhD, Clark University, Worcester, MA; **Matt Hawrilenko**, PhD, VA Puget Sound Healthcare System, Seattle, WA; **James V. Cordova**, PhD, Clark University, Worcester, MA.

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 last author.

Recruitment and interview data were published in a 2011 article and relationship outcomes were published in a 2014 article. The current project focuses on distinct data outcomes.

Address correspondence to Tatiana D. Gray, Department of Psychology, Clark University, 950 Main Street, Worcester, Massachusetts 01610; E-mail: tsettles@clarku.edu

several intervention studies specifically targeting relationship satisfaction have found a concurrent reduction in depression symptoms (Whisman & Beach, 2012). Furthermore, research indicates that the reduction of depression symptoms is, at least partially, mediated by improvements in relationship satisfaction (Beach & O'Leary, 1992).

More recently, researchers examined the efficacy of an online relationship program on relationship functioning and individual outcomes such as depression (Doss et al., 2016). Partners completed online activities and participated in four 15-min calls with a study staff member, totaling 8 hrs of participation time. Individuals in the treatment condition experienced significant reductions in depressive symptoms ($d = 0.50$) compared to those in the waitlist condition, with an even greater effect for those who reported more significant depressive symptoms at baseline ($d = 0.71$). Taken together, the literature in this area suggests that couple therapy is an effective and promising treatment for individuals suffering from depression, whether the primary aim of the intervention is the depression or relationship discord.

Individual Pharmacological and Psychotherapy Interventions for Depression

Despite the emergence of relational treatments for depression, the majority of existing research examines individual-level treatments. A significant volume of research in this area involves pharmacological treatments. This is significant for couples because many individuals treated for depression pharmacologically are in long-term relationships and are motivated to seek treatment either by their partner or because of the effects of their depressive symptoms on their relationship. Hundreds of randomized controlled trials have claimed that antidepressants are more effective in the treatment of depression than placebo (Cipriani et al., 2018; Mulrow et al., 1998). However, there are discrepancies in reported effect sizes within these trials and while antidepressants have demonstrated superiority over placebo, only 25–30% of people achieve full remission and 40–50% of people do not respond at all (Kocsis, 2003). Notably, when unpublished data are taken into consideration (controlling for the publication bias), Kirsch and colleagues discovered that pharmacological treatments fail to demonstrate a significant difference from placebo (2008). One thing that has been demonstrated is that the baseline depressive symptom severity is a significant predictor of treatment outcomes with antidepressant medications (Fournier et al., 2010; Khan et al., 2002; Kirsch et al., 2008), such that the effect size of the drug–placebo difference tends to range from small for those experiencing mild to moderate depressive symptoms (e.g., $d = 0.20$) to large for those experiencing more severe symptoms (e.g., $d = 0.81$; Fournier et al., 2010). These findings indicate that the effectiveness of pharmacological treatments for depression varies significantly based on baseline symptom severity. Taken together, the existing literature paints a complicated, unresolved, and potentially biased picture of pharmacological treatments of depression.

Psychotherapy treatments for depression have existed nearly as long as pharmacological solutions, and their efficacy has been clearly demonstrated. A meta-analysis of 12 studies of empirically supported treatments (ESTs) for depression found that the median effect-size for reducing depressive symptoms was small ($d = 0.30$), with 54% of participants (who completed treatment) reporting improvement (Westen & Morrison, 2001). In terms of long-term effects, the one study, of the 12, that followed clients for a significant period of time, found no difference between the treatment and control condition at 18 months, suggesting that the majority of people did not sustain improvement in depressive symptoms following therapy (Westen & Morrison, 2001). Additional studies have shown similar results, reporting a 48% response rate after completing 16 sessions (e.g., Keller et al., 2000). A common theme in the existing literature is the assumption that depression, and particularly chronic depression, requires more frequent therapy sessions in order to yield positive results (e.g., Thase et al., 1994).

When pharmacological and psychotherapy treatments for depressions are directly compared, researchers have found that participants who received cognitive therapy were less likely to relapse after treatment than those who discontinued antidepressant medication, and no more likely to relapse than those who continued their medication (Evans et al., 1992; Hollon et al., 2005). Overall, the research indicates that antidepressant medication and psychotherapy are about equally effective in treating depression. Psychotherapy is the preferred treatment for mild to moderate depressions (National Institute for Health & Care Excellence, 2009), the combination of some forms of psychotherapy with pharmacological treatments may be superior to either approach alone when

the patient is experiencing severe symptoms (Arnow & Constantino, 2003; Keller et al., 2000), and psychotherapy tends to have more enduring effects than pharmacological approaches (Hollon et al., 2005).

The Marriage Checkup. The present study investigated the effect a relationship-oriented annual checkup had on depressive symptoms. The Marriage Checkup (MC) is a brief, preventative intervention designed to be the relationship health equivalent of physical and dental health checkups. Given the supposition that relationships are a comprehensive health system (Cordova, 2013), the MC was developed to attend to the health and well-being of our relationships in the same way that we attend to the health and well-being of our teeth and bodies. The MC, comprised of an Assessment and Feedback session, draws on therapeutic techniques from Motivational Interviewing (MI; Miller & Rollnick, 2002) and Integrative Behavioral Couple Therapy (IBCT; Jacobson & Christensen, 1996). More specifically, couples are guided through a discussion about their relationship history and what drew them together, invited to identify and celebrate their primary strengths as a couple, and finally work with the clinician to create a deeper and more compassionate understanding of their chief concerns (for more information on the Checkup protocol, please see Cigrang et al., 2016). Targeted toward all couples at any stage in their relationship, the MC is designed to reduce barriers to help seeking present in traditional Couple Therapy contexts (Morrill et al., 2011).

The goals of the Marriage Checkup, in the service of building intimacy between partners, are threefold. First, the MC aims to reorient the couple toward the strengths and positive qualities of their relationship. This should theoretically diminish depressive symptoms through a similar mechanism demonstrated within the gratitude literature, in that the depressive symptoms are often characterized by a hyper focus on negative thoughts, feelings, and experiences. Within this literature there is evidence to suggest that shifting the focus of attention to more positive qualities of experience results in improved mood outcomes (Lambert, Fincham, & Stillman, 2012). In addition, within the treatment theory of the MC, increasing couple focus on the positive qualities of the relationship enhances intimacy by creating a more positive emotional context within which vulnerable experiences can be met. Second, the MC works to foster acceptance of common relationship concerns and patterns the couple may be experiencing. This should work to diminish depressive symptoms via an increase in the type of psychological flexibility (in the relationship context) targeted by third-wave acceptance approaches (Hayes, Luoma, Bond, Masuda, & Lillis, 2006). Within the MC, meeting common concerns with acceptance facilitates intimate events, which in turn increase the experience of felt intimacy (Cordova & Scott, 2001). Finally, the MC intends to reactivate the partners to wholeheartedly attend to the health and well-being of their relationship. This theoretically should diminish depressive symptoms through a similar mechanism to that demonstrated in the behavioral activation literature (Jacobson et al., 1996). The Marriage Checkup is intended to address a range of relationship issues, including those related to partners' mental health. While the MC does not specifically treat psychopathology, it does address any mental health issues raised by the couple.

In a randomized controlled trial, the MC demonstrated significant gains in relationship satisfaction, intimacy, and acceptance (Cordova et al., 2014). Couples showed sharp gains in all outcomes in the weeks immediately following both their first and second checkup. Those gains mostly held steady in the year following the first checkup, and slowly eroded over the year following the second checkup. Based on the existing literature, we had three primary hypotheses for the current study. First, couples in the treatment group would show reductions in depression symptoms after participating in their Marriage Checkup, compared to couples in the control group. Second, depressed individuals at baseline would show larger treatment effects than nondepressed individuals. And finally, the treatment effect would be mediated by relationship satisfaction.

Study Overview

The present study tested the efficacy of a randomized control trial of the Marriage Checkup on depressive symptoms and was built on previous work in several ways. First, in line with the checkup model, the Marriage Checkup is substantially briefer than previous approaches, lasting a total of three therapy hours: 90 min for the Assessment Session and 90 min for the Feedback Session. The majority of couple-based interventions targeting mood disorders and relationship

distress involve 16–20 hrs of therapy (Bodenmann et al., 2008; Jacobson, Dobson, Fruzzetti, Schmalig, & Salusky, 1991; O’Leary & Beach, 1990) and the briefest known therapy study to date required 10 therapy hours (Cohen, O’Leary & Foran, 2010). Cutting the time requirement even further, Doss et al. (2016) offered an 8-hour IBCT-based online program for couples (though this intervention did not include any time with an actual therapist and was predominantly self-driven). The MC approach used in the current study required only a three hour time commitment, less than half of the time of even the briefest existing options for couples.

Second, the MC uses an acceptance and motivational enhancement-based approach to help couples. Nearly all of the studies examining couple therapy for depression emphasize changing behaviors and cognitions (Beach & O’Leary, 1992; Emanuels-Zuurveen, & Emmelkamp, 1996; Foley, Rounsaville, Weissman, Sholomskas, & Chevron, 1989; Jacobson et al., 1991; Teichman, Bar-El, Shor, Sirota, & Elizur, 1995; Whisman & Beach, 2012). Only a handful of studies employed a noncognitive or behavioral-based approach. Of the few, one looked at the efficacy of systemic couple therapy for depression (Leff et al., 2003), one examined Emotionally Focused Therapy (EFT), an attachment-based approach, compared to treatment as usual for depression, (Wittenborn, Liu, Ridenour, Lachmar, Mitchell, & Seedall, 2018), and a third tested an online IBCT program for relationship distress and individual outcomes (Doss et al., 2016). The potential benefit of an acceptance and motivational-based treatment is that these approaches lend themselves to briefer and yet equally effective interventions (Miller & Rollnick, 2002). The current study is innovative in that it is the first study to examine the effects of an in-person, dyadic, acceptance-based approach to the treatment of relationship health, examining depression as a major outcome variable. Finally, the current study evaluates the use of an annual Booster checkup and utilizes cutting edge data analytic techniques to model the phenomenon in a more accurate and nuanced way than previously possible.

METHODS

Participants

Participants were 215 couples recruited from a New England city; 209 were opposite-sex couples, and six were same-sex couples. Same-sex couples were excluded from the analysis for two reasons: we did not have enough same-sex couples to make meaningful inferences about them separately from opposite-sex couples and we planned to test sex-specific effects. The average age of the participants was 44.5 years for women ($SD = 10.8$) and 46 years for men ($SD = 11.4$) with a range of 20–78 years. The majority of participants were Caucasian (93.3%), followed by Black (2.7%), Asian (2.7%), Hispanic (1.7%), and American Indian (0.7%). The average length of the marriage was 15.2 years ($SD = 12.0$, ranging from 22 days to 56 years). On an average, couples had two children, and a household income between \$75,000 and 99,000. The majority of the sample (88%) graduated high school, and 43.9% had a bachelor’s degree or higher.

Procedures

When couples were determined eligible to participate, the Research Coordinator randomly assigned an identification number from the master randomization list. The identification number was placed on a key and on the couple’s pretreatment questionnaires. Couples were enrolled in the study following return of the pretreatment questionnaires and informed of their treatment condition. Couples in the treatment condition completed baseline questionnaires and then attended their MC, which consisted of an Assessment and Feedback Session. One year after their initial checkup, treatment couples participated in a Booster MC, where they followed the same procedures as the first year. Couples in the wait-list control condition were offered an MC after 2 years. All participants completed questionnaires throughout 2 years of follow-up. The current analysis includes five time points: baseline, and measures collected every 6 months through 2-year follow-up. More detailed information on procedures and participant flow can be found in Cordova et al. (2014).

Measures

Depression. We measured depression with the Center for Epidemiologic Studies Depressions Scale (CES-D; Radlof, 1977). The CES-D is a 20-item self-report scale designed to measure

depressive symptom frequency. Scores range from 0 to 60, with higher scores indicating greater symptoms. Typical items include “I was bothered by things that usually don’t bother me” and “I felt that everything I did was an effort.” Internal reliability for this scale was high for women (Cronbach’s α from .90–.92) and men (α from .87–.92).

Relationship satisfaction. We measured relationship satisfaction with the Quality of Marriage Index (QMI; Norton, 1983) and the global distress subscale of the Marital Satisfaction Inventory-Revised (GDS; Snyder, 1997).

Quality of marriage index—The QMI is a six-item self-report scale designed to measure general satisfaction within married relationships. Scores can range from 6 to 45 with higher scores indicating greater levels of satisfaction, and lower scores indicating greater levels of distress. Items include “my relationship with my partner is strong,” “my partner and I are really a team,” and “my relationship with my partner makes me happy.” Response options on the first five questions range from 1 (*strongly disagree*) to 7 (*strongly agree*). The sixth item assesses the general degree of happiness within the relationship and is answered on a 10-point scale ranging from 1 (*extremely low*) to 10 (*extremely high*). Internal reliability for this measure was extremely for both women (Cronbach’s α from .94–.97) and men (α from .92–.97).

Marital satisfaction inventory-global distress subscale—The GDS measures overall marital satisfaction with 22 true or false questions describing the relationship. Example items include, “There are many things about our relationship that please me,” “Our relationship has been very satisfying,” and “The good things in our relationship far outweigh the bad.” Items were summed and converted to t-scores (Snyder, 1997). Internal consistency in this sample was good for women (α from .92–.94) and men (α from .91–.94).

Analytic Strategy

We used latent change score models in a structural equation modeling framework to examine changes in depression symptoms across time, and multivariate latent change models to examine the dynamic relationships between depression and relationship satisfaction across couple members (McArdle, 2009). Latent change score models are an extension of latent growth models (Meredith & Tisak, 1990). Similar to growth curves, they include a slope term that captures the average trajectory of change. In addition to modeling trajectories, latent change models include change scores between each wave that capture deviations from the average amount of change. For the univariate models, this was a convenient way to model nonlinear within-person and between-person trajectories. For the multivariate models, latent change scores were preferable to other common approaches (e.g., growth curves) because they allowed more flexibility to examine the dynamics of change from wave-to-wave, allowing us to disentangle contemporaneous from time-lagged associations between variables, and dynamic processes that occurred across partners and over time. Multigroup models were employed to explicitly model differences in means and variances between treatment arms.

Models were built in a stepwise fashion. Preliminary models indicated moderate correlations between partners’ scores for satisfaction ($r = .54-.61$) and smaller correlations for depression ($r = .18-.40$), indicating that it would be more appropriate to model processes at the individual-level rather than to assume processes occurring at the couple level (i.e., a common fate model). First, separate models were fit to men and women to ensure good local fit of means and variances. The best fitting model for each sex was then combined and covariances were included between partners’ scores at each time wave. After finding the best fitting combined model for men and women, we examined differences between the treatment and control groups with a multigroup invariance testing strategy. The baseline model allowed all parameters to vary across intervention arms. To increase parsimony, we used the free baseline approach (Stark, Chernyshenko, & Drasgow, 2006) beginning with an unconstrained model and testing simpler models by imposing successive equality constraints between treatment and control groups. Differences between nested models were tested stepwise using the Wald Test (Wald, 1945). The Wald Test compares the fit of a model where parameters are constrained equal to one and where the same parameters are freely estimated. It is chi-square distributed and asymptotically equivalent to a chi-square test. Global fit for all models was evaluated with the chi-square test, where a nonsignificant result indicates a good fit of the model to the data, and the Comparative Fit Index (CFI), Tucker Lewis Index (TLI), and

root mean square error of approximation (RMSEA), where values $\geq .05$ indicated acceptable fit for the CFI and TLI, and below .06 indicated acceptable fit for the RMSEA (Hu & Bentler, 1999).

All structural equation modeling analyses were conducted in Mplus version 7.3. Because our depression outcome was nonnormally distributed (skew range: 0.70–1.57, kurtosis range: –0.48 to 2.72), we used a maximum likelihood estimator with robust standard errors (i.e., a sandwich estimator called with the MLR estimator in Mplus) (Muthén & Muthén, 1998–2011). We calculated standard errors for indirect effects using bias-corrected bootstrapping, which is preferable to model-based standard errors because it accounts for nonnormality of indirect effects (Williams & MacKinnon, 2008).

Missingness increased over time to a maximum of 28% at 2 years for women, and 31% at 1 year 6 months for men (27% at the final wave). Little's Missing Completely At Random test indicated there was patterning to missingness ($\chi^2(152) = 208.73, p = .002$), which is common in longitudinal studies with dropout. Maximum likelihood estimators are unbiased under conditions up to missing at random (Enders & Bandalos, 2001), meaning that if missingness/dropout was dependent on the level of a variable already included in the model (e.g., depression) then estimates would remain unbiased. Previous analyses using pattern mixture modeling in this sample have indicated that missing at random was a reasonable assumption for these data (Cordova et al., 2014). To calculate the treatment effect (Hypothesis 1), we fit a multigroup model and parameterized the latent change component of the model to reflect additive change. The trajectory component of the model (e.g., the linear slope) was constrained equal between the treatment and control groups. The average change score for the control group was centered at zero and change scores were freely estimated in the treatment group. This parameterization is known as an additive-change model (McArdle, 2017), meaning that the amount of change in the control group was entirely captured by the linear slope, but the treatment group had an additional, nonlinear component of change captured by the change scores at each wave. A significant time-varying change score meant the amount of change for that period was significantly different between groups. The total treatment effect at each wave was calculated by summing all change scores up to that point. Cohen's *d* effect sizes were calculated by dividing cumulative treatment effects by the pooled baseline standard deviation (Cohen, 1988). Because preliminary analyses indicated that participants in the treatment group had higher baseline levels of depression, we took two steps to adjust these differences. First, we adjusted for baseline differences by calculating treatment effects as the amount of change that occurred over treatment rather than as differences in endpoints only (i.e., difference-in-differences analysis; Dimick & Ryan, 2014). Second, because change may be related to baseline scores, we controlled this by regressing change scores on the baseline intercepts.

To examine moderation (Hypothesis 2), we included a dummy variable that characterized high baseline severity using the recommended CES-D cut score of 16. We tested whether high baseline severity was associated with larger changes at each time point and larger changes in the treatment group than the control group. To calculate the moderation effect, we subtracted the difference between groups at each time point using the delta method via the MODEL CONSTRAINT command in Mplus (Raykov & Marcoulides, 2004). To calculate the cumulative moderation effect, we summed these differences across time. For effect size estimates, we calculated the total effect for the group with clinically significant symptoms by adding the main effect to the moderation effect.

To examine mediation (Hypothesis 3), we used a multivariate latent change model (McArdle, 2009). Relationship satisfaction was included as a latent variable measured by the QMI and GDS. Factor loadings were fixed to 1 at all time points for the QMI, and loadings for the GDS were held equal across time and across men and women to ensure that marital satisfaction was measuring the same construct in both partners.

In the multivariate latent change model, we examined how husbands' and wives' changes in relationship satisfaction related to their own and each others' changes in depressive symptoms (i.e., an actor-partner model; Cook & Kenny, 2005) both contemporaneously and with a 6-month time lag (See Figure 1). We hypothesized that we would see stronger contemporaneous effects (changes occurring across the same wave) and weaker time-lagged changes, as we expected depressive symptoms and relationship satisfaction to be tightly intertwined. We tested a series of nested models using the free baseline approach (Stark, Chernyshenko, & Drasgow, 2006) to determine

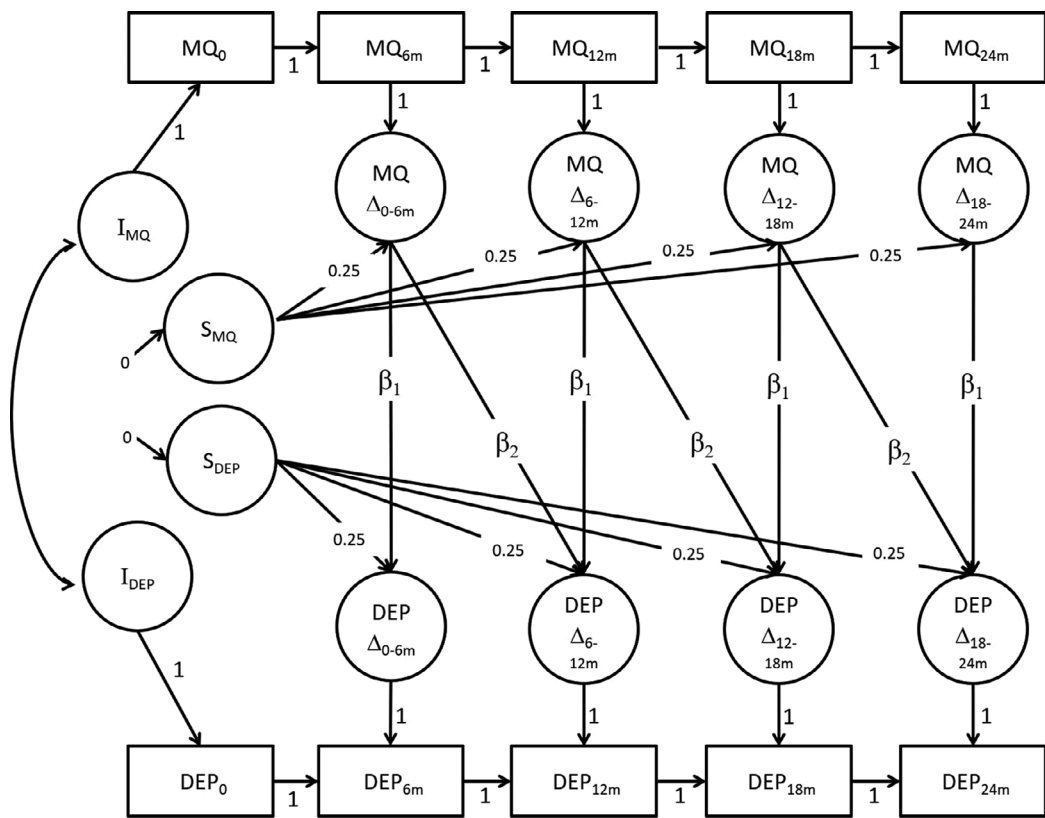


Figure 1. Actor Partner Model of changes in relationship satisfaction related to changes in depressive symptoms.

whether the relationship between depression and marital satisfaction was the same across time, and whether actor and partner effects differed by sex. To calculate Cohen's *d* for indirect effects, we multiplied the treatment effects on marital quality by the respective actor and partner effects of marital quality on depression, then divided by the baseline standard deviation of depression. We calculated effects for each specific wave, as well as the total cumulative effect across follow-up.

RESULTS

H1: treatment effects

The final model provided an excellent fit to the data, $\chi^2(28) = 31.16$, $p = .31$, CFI = 1.00, TLI = 0.98, RMSEA = 0.03. We examined numerous intermediate models, and their fit statistics are provided in a supplementary online document, Table S1. Effect sizes are presented in Figure 2 (Table 1).

A linear slope provided a good fit to the control group's trajectory of depression symptoms. The slope was not statistically significantly different from zero, indicating that the control group maintained a similar level of depression across the entire study. The treatment effect did not vary across men and women ($\chi^2(10) = 5.11$, $p = .88$), so it was constrained equal across sex to provide an average effect. The treatment group experienced a sharp, nonlinear gain in the time interval following the first checkup ($d = 0.22$) and did not significantly change compared to the control group thereafter. Although the size of the effect did not change, the treatment effect became nonsignificant at 18 and 24 months (p values .067 and .10, respectively).

Change score variances indicated that individuals' symptom levels varied markedly from wave to wave, underscoring the episodic nature of symptoms. The average fluctuation of women's depression scores was 22% lower in the treatment group than the control group, suggesting that

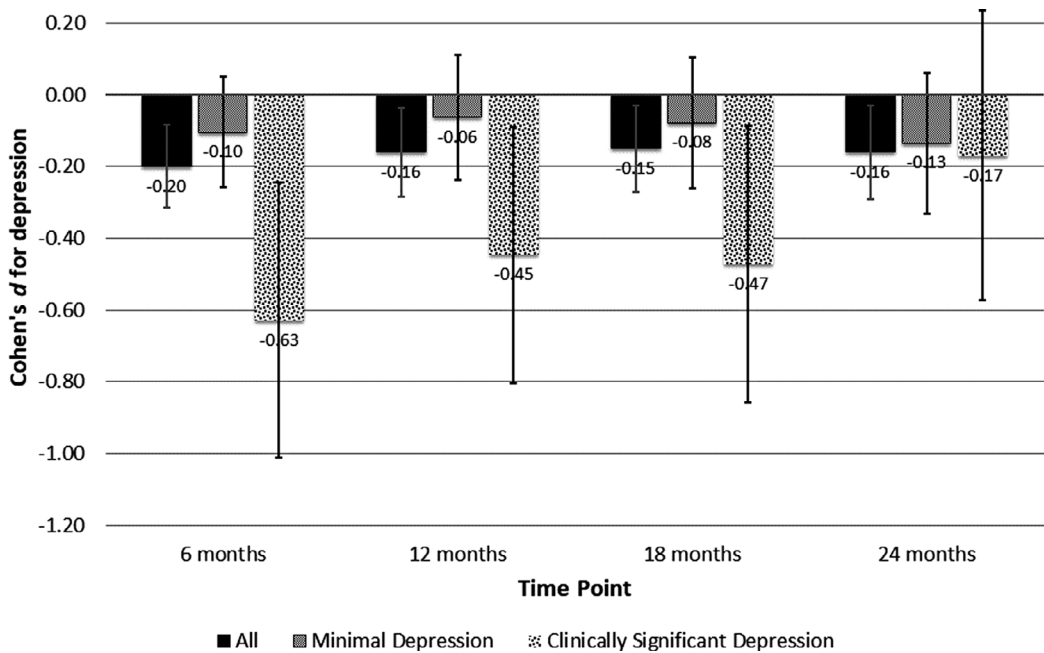


Figure 2. Treatment effect sizes across baseline depression symptom severity.

treatment not only reduced the mean level of depression symptoms, but it also stabilized their fluctuation. This was not the case for men.

H2: moderation effects

Twenty percent of men and 32% of women had depression scores above the cutoff for clinical significance. The final model constrained men's and women's slopes, treatment effects, and moderation effects equal. Fit of the final model was adequate, $\chi^2(57) = 91.54$, $p = .003$, CFI = 0.97, TLI = 0.93 RMSEA = 0.08. Because the chi-square was significant, we examined local fit by looking at the size of the raw versus model-estimated residuals. Residuals for all paths of interest were small, and no patterning to the residuals suggested that major changes to the model would significantly improve fit. We took this as evidence that the model was valid for the areas of primary substantive interest.

We found a statistically significant moderation effect for baseline severity that indicated the main effects for treatment on depression were driven primarily by larger effects for people with greater severity (see Table 2 for model parameters and Figure 2 for effect sizes). The between-group difference in the effect of baseline severity on change in depression scores was medium in size ($d = -0.53$) and statistically significant. The total effect for more severe individuals (main effect + moderation effect) was also in the medium range ($d = 0.63$). This difference eroded over time and was nonsignificant at 18 months ($p = .054$), and near zero at 24 months.

H3: mediation effects

Actor and partner effects of changes in relationship satisfaction on changes in depression were invariant across sex and time, so they were constrained equal to form a single actor effect and single partner effect (Table 3). The model with lagged changes did not fit better than the model with only contemporaneous effects, meaning that changes in depression due to relationship satisfaction occurred in close proximity to one another. For parsimony, the lagged paths were trimmed from the final model. Despite a significant chi-square, $\chi^2(354) = 482.67$, $p < .001$, the final mediation model otherwise provided a good fit to the data, CFI = 0.97, TLI = 0.96, RMSEA = 0.04. Here again, despite the significant chi-square, inspection of local fit via model residuals showed residuals that were small or zero for the paths of primary interest, indicating that inferences made from the

Table 1
Multigroup Model for Treatment Effects

Model parameter	Control group		Treatment group		Significance tests for between-group contrasts
	Estimate	<i>p</i>	Estimate	<i>p</i>	
Women					
Intercept	11.72	<.001	12.83	<.001	.32
Average change score	−0.29 ^{a1}	.16 ^{a1}	−0.29 ^{a1}	.17 ^{a1}	—
Treatment effect					
Δ 0–6 m	—	—	−1.97 ^{b2}	.001 ^{b2}	.001
Δ 6–12 m	—	—	0.36 ^{b3}	.57 ^{b3}	.57
Δ 12–18 m	—	—	0.09 ^{b4}	.88 ^{b4}	.88
Δ 18–24 m	—	—	−0.83 ^{b5}	.90 ^{b5}	.91
Cumulative treatment effect					
Δ 0–6 m	—	—	−1.97 ^{b6}	.001 ^{b6}	.001
Δ 0–12 m	—	—	−1.61 ^{b7}	.031 ^{b7}	.034
Δ 0–18 m	—	—	−1.52 ^{b8}	.071 ^{b8}	.067
Δ 0–24 m	—	—	−1.6 ^{b9}	.10 ^{b9}	.10
Change score <i>SD</i>					
<i>SD</i> Δ 0–6 m	8.85	<.001	7.23	<.001	$\chi^2(4) = 10.68,$ <i>p</i> = .030
<i>SD</i> Δ 6–12 m	9.27	<.001	8.84	<.001	
<i>SD</i> Δ 12–18 m	9.04	<.001	5.85	<.001	
<i>SD</i> Δ 18–24 m	8.74	<.001	6.08	<.001	
Men					
Intercept	7.01	<.001	10.18	<.001	.003
Average change score	0.21 ^{a2}	.23 ^{a2}	0.21 ^{a2}	.23 ^{a2}	—
Treatment effect					
Δ 0–6 m	—	—	−1.97 ^{b2}	.001 ^{b2}	.001
Δ 6–12 m	—	—	0.36 ^{b3}	.57 ^{b3}	.57
Δ 12–18 m	—	—	0.09 ^{b4}	.88 ^{b4}	.88
Δ 18–24 m	—	—	−0.83 ^{b5}	.90 ^{b5}	.90
Cumulative treatment effect					
Δ 0–6 m	—	—	−1.97 ^{b6}	.001 ^{b6}	.001
Δ 0–12 m	—	—	−1.61 ^{b7}	.031 ^{b7}	.031
Δ 0–18 m	—	—	−1.52 ^{b8}	.071 ^{b8}	.071
Δ 0–24 m	—	—	−1.60 ^{b9}	.10 ^{b9}	.10
Change score <i>SD</i>					
<i>SD</i> Δ 0–6 m	4.73	<.001	6.75	<.001	$\chi^2(4) = 7.19,$ <i>p</i> = .13
<i>SD</i> Δ 6–12 m	6.71	<.001	6.49	<.001	
<i>SD</i> Δ 12–18 m	6.63	<.001	6.26	<.001	
<i>SD</i> Δ 18–24 m	7.27	<.001	8.10	<.001	

Note. Linear slope constrained equal across groups so treatment effect would be captured in change scores. a# = The letter *a* reflects parameters constrained equal across treatment and control groups. The number following it indicates which parameters have been constrained equal to each other. b = parameters constrained equal across sexes. Δ = change score. m, months; *SD*, standard deviation.

model were valid. Given the large number of model parameters, the significant chi-square was likely driven by an accumulation of small deviations between the observed and model-estimated data.

Table 2
Multigroup Model for Moderation by Baseline Depression Status

	Control group		Treatment group		Between-group contrasts: treatment–control	
	Estimate	<i>p</i>	Estimate	<i>p</i>	difference	<i>p</i>
Cumulative effect of baseline severity						
Δ 0–6 m	1.33	0.45	–3.92	0.065	–5.25	0.012
Δ 0–12 m	–1.22	.43	–5.06	.004	–3.84	.041
Δ 0–18 m	–1.04	.56	–4.96	.010	–3.92	.054
Δ 0–24 m	–4.12	.019	–4.56	.017	–0.37	.86
Cumulative treatment effect—low depression group						
Δ 0–6 m	—	—	—	—	–1.05	0.19
Δ 0–12 m	—	—	—	—	–0.62	0.49
Δ 0–18 m	—	—	—	—	–0.78	0.40
Δ 0–24 m	—	—	—	—	–1.34	0.18
Cumulative Treatment effect—high depression group						
Δ 0–6 m	—	—	—	—	–6.29	0.001
Δ 0–12 m	—	—	—	—	–4.46	0.014
Δ 0–18 m	—	—	—	—	–4.71	0.017
Δ 0–24 m	—	—	—	—	–1.70	0.41

Note. Δ = change score. m, months.

Actors' changes in relationship satisfaction were strongly related to changes in depression across the same time period. Partner effects were small and nonsignificant. The indirect effect of treatment action on depression through relationship satisfaction accounted for about three-quarters of the treatment effect. Notably, the indirect effect was statistically significant at 2 years despite the main effect for the model testing Hypothesis 1 dipping below significance, likely because including relationship satisfaction in the model increased statistical power.

DISCUSSION

Overall, results of the current study suggest that a brief, acceptance- and motivation-based relationship health checkup can significantly improve participant's depressive symptoms, particularly for those individuals with the most severe symptoms. Examination of the overall treatment effect indicated that depressive symptoms significantly declined over the first 6 months in the treatment group. Furthermore, though there was not an *additional* decline in symptoms at the time of the Booster Session, treatment couples maintained their initial improvement over the course of the study. Though the statistical significance of the difference between groups was marginal at 18 and 24 months (despite no diminishment in effect size), these between-group differences remained significant at follow-up when statistical power was increased by considering relationship satisfaction as a mediating variable. It should be noted that treatment couples received an additional checkup at 1 year, likely contributing to the sustained effects. It is also notable that control couples' depressive symptoms remained stable for the assessed 2 years, suggesting that the average levels of depressive symptoms were not spontaneously improving in the absence of intervention. Although episodes of depression meeting criteria for MDD tend to remain chronic for only 20% of people

Table 3
Direct and Indirect Mediation Effects

Parameter	<i>b</i>	<i>p</i>	95% CI	Cohen's <i>d</i>
Δ Marital quality \rightarrow Δ Depression				
Actor effect	-0.81	.014	—	—
Partner effect	0.13	.68	—	—
Actor's indirect effect (Tx \rightarrow Δ Actor's marital quality \rightarrow Δ Actor's depression)				
6 months	-1.31	—	[-2.66, -0.54]	-0.13
1 year	-0.23	—	[-0.94, 0.40]	-0.02
1 year 6 months	0.02	—	[-0.65, 0.77]	0.02
2 years	0.13	—	[-0.76, 1.05]	0.01
Partner's indirect effect (Tx \rightarrow Δ Partner's marital quality \rightarrow Δ Actor's depression)				
6 months	0.20	—	[-0.24, 1.02]	0.02
1 year	0.04	—	[-0.07, 0.45]	0.00
1 year 6 months	0.00	—	[-0.28, 0.18]	0.00
2 years	-0.02	—	[-0.54, 0.15]	0.00
Total effects				
Actor total through 2 years	-1.39	—	[-2.99, -0.42]	-0.14
Partner total through 2 years	0.21	—	[-0.24, 1.11]	0.02

Note. Men's and women's effects were constrained equal and time invariant in the final model. CI = confidence interval. Cohen's *d* was calculated as the size of the effect in depression that was explained by changes in marital satisfaction.

(Kocsis, 2003), our data suggest that in the absence of treatment, some measurable level of depressive symptoms remain consistent over the 2 years that our participants were followed.

Furthermore, we believe innovative to this study, results indicated that, for women, treatment stabilized the amplitude of fluctuations in depressive symptoms over the 2-year study period, suggesting that their overall mood was not simply more positive, but less susceptible to dips toward greater depression. In contrast, control couples' symptom fluctuations remained comparably large over the course of the study such that they demonstrated lower lows from baseline. It is our contention that a significant marker of improvement in depression involves just this type of long-term diminishment in the fluctuation of depressive symptoms. In other words, given the consensus that depression is an episodic condition (Arnou & Constantino, 2003; Keller et al., 2000; Klein & Santiago, 2003), effective treatment might best be marked by shortening the duration of any particular acute episode, lengthening the asymptomatic time between episodes, and diminishing the negative amplitude of episodic symptoms.

We next wanted to consider whether the MC would have a differential effect on depression scores depending on the partners' initial level of depressive symptom severity. First, we found a relatively large percentage of partners scored above the cutoff for clinical significance, indicating that the MC attracted people across the full continuum of distress, both individually and relationally. Furthermore, while it was conceivable that partners with more severe depressive symptoms might benefit less from a brief relationship health intervention, our results suggest exactly the opposite. Examination of depressive symptom severity as a moderator revealed that more severely depressed partners improved significantly more than less depressed partners in the treatment group compared to the control group, with a fairly significant effect size. The size of the effect did, however, erode by the final 24-month assessment point. These results suggest that targeting the relationship health of more severely depressed individuals through a brief relationship health checkup can rapidly improve depressive symptoms on par with lengthier couple-based treatments, (Barbato & D'Avanzo, 2008; Beach & O'Leary, 1992; Bodenmann et al., 2008; Cohen, O'Leary, & Foran,

2010; Doss et al., 2016), individual treatments directly targeting depression (Westen & Morrison, 2001), and standard pharmacological treatments (Fournier et al., 2010; Kocsis, 2003). Although it is unclear from the existing literature whether lengthier treatments for depression might show beneficial effects within a shorter time span, the possibility exists. Future studies of longer treatments should assess the possibility that significant change might occur within the very early part of treatment. That being said, it is worth noting that the MC is the briefest relationship intervention to show this type of effect size on depressive symptoms.

Finally, though the difference between more and less severely depressed individuals eroded over time, the overall intervention effect remained significant, suggesting that the slope of improvement differed by group, such that people who were most depressed and accrued the most benefits in the first year of the study were not the same people who benefitted most in the second year. One possibility is that, given the episodic nature of depression, different sets of couples were distressed at the beginning of the second year and consequently received more benefit.

We next examined the mediation effect of relationship satisfaction on the association between treatment and depressive symptoms. Initial examination of the mediation effect shows that changes in depression occurred contemporaneously with changes in relationship satisfaction, and that the effect was not lagged—or at least, not lagged in a time window captured by our 6-month intervals. Three quarters of the treatment effect on depression was mediated by changes in relationship satisfaction and the indirect effect remained significant at 2 years. Thus, our results are consistent with the previous literature suggesting that relationship therapies have the benefit of both improving relationship health (Whisman & Beach, 2012) and alleviating depressive symptoms simultaneously (Barbato & D'Avanzo, 2008), and that alleviation of depressive symptoms is a dependable therapeutic effect of improving relationship health. Furthermore, these results are consistent with the notion that relationship health is a health domain on par with others like physical health and mental health and is inextricably intertwined with physical and mental health outcomes. It is conceivable that mental health, particularly depression, is best conceptualized as inextricably relationally contextualized, such that relationship health quality is potentially causally determinative of mental health outcomes, as well as being causally determined by mental health symptoms. In the current case, one could argue that when people are more depressed, they are more likely to evaluate their relationship more negatively. And, while this may be the case, our analyses suggest the causal arrow also points in the opposite direction, opening up questions about what drives that effect. For example, it is possible that, given the lower change score variances for women, the MC results in a greater acceptance of, and more compassion for, the normal day-to-day differences and challenges in a long-term intimate relationship, resulting in overall mood being less affected by the presence of the relationships' perpetual issues. The MC uses couple's perpetual issues to build intimacy bridges through compassion elicitation, and it may be that the resulting sense of being deeply understood and accepted within an ongoing intimate relationship addresses some part of the fundamental cause of depressive symptoms. As noted, we suspect that the mechanisms of change are (a) shifts in the focus of attention to neglected positive qualities of experience (Lambert, Fincham, & Stillman, 2012), thereby benefitting the intimacy process by creating a more positive environment within which to express relational vulnerability, (b) growth of relational acceptance capitalizing both on psychological flexibility (Hayes, Luoma, Bond, Masuda, & Lillis, 2006), and increasing probability of intimate moments (Cordova & Scott, 2001), and (c) increases in behavioral activation and effective relational practices (Jacobson et al., 1996).

In summary, while mental health is not directly targeted by the MC, the current study found that the MC worked to rapidly improve depressive symptoms and diminish the amplitude of depressive symptoms over time. We further found that improvements were strongest for partners who had the most severe depressive symptoms, and that these improvements were largely mediated by improvements in relationship satisfaction. The clinical implication, in keeping with the intention of the MC, is that any time a patient shows up for a Checkup, multiple health domains can be assessed and addressed regardless of the specific target of the checkup. Relationship health checkups provide opportunities to detect, assess, and begin intervention for a variety of related health issues, including both mental and physical health (e.g., anxiety, PTSD, substance abuse, and depression). Furthermore, evidence demonstrates that treating relationship health has beneficial effects on depressive symptoms that appear to be rapid and enduring. Finally, we provided

evidence that treatment diminished the long-term amplitude of depressive symptoms over time, which we suggest is an important outcome measure for disorders that are known to be episodic and chronic.

Limitations

There are several notable study limitations. First, lack of diversity. Future studies of the MC should include more diverse samples regarding race, ethnicity, sexual orientation, and SES. Diversity is an area that too often receives a faint nod within the limitation section, yet merits much more intentional and dedicated research. For example, a relatively high-income sample has access to supportive resources and a lower income sample is subject to more contextual stress, suggesting that future studies consider income as a moderator. We recognize the need for more representative samples and have set the intention to recruit diverse participants and provide access to those who may have previously been unable to participate in research. Also, people who are suffering more debilitating levels of depression may not have been able to participate in a study of this kind, and future research might focus on further reducing the systemic and practical barriers to treatment for people with severe depression.

Conclusion

This study demonstrated that a relationship checkup had the power to significantly improve the depth of suffering for those who experience symptoms of depression. Results indicate that improvements in the relationship appear to strongly influence improvements in depression, suggesting that intimate relationships are perhaps the most significant context for healing. The outcomes are comparable to long-term individual psychotherapy, long-term tertiary couple therapy, and antidepressant medications; and as such, the MC may be an attractive option for people struggling with depression, especially those who do not wish to take medication and those who do not have time and resources for long-term therapy.

REFERENCES

- Arnow, B.A., & Constantino, M.J. (2003). Effectiveness of psychotherapy and combination treatment for chronic depression. *Journal of Clinical Psychology, 59*, 893–905. <https://doi.org/10.1002/jclp.10181>.
- Barbato, A., & D'Avanzo, B. (2008). Efficacy of couple therapy as a treatment for depression: A meta-analysis. *Psychiatric Quarterly, 79*, 121–132. <https://doi.org/10.1007/s1126-008-0968-0>.
- Beach, S.R.H., Katz, J., Kim, S., & Brody, G.H. (2003). Prospective Effects of Marital Satisfaction on Depressive Symptoms in Established Marriages: A Dyadic Model. *Journal of Social and Personal Relationships, 20*, 355–371. <https://doi.org/10.1177/0265407503020003005>.
- Beach, S.R., & O'Leary, K.D. (1992). Treating depression in the context of marital discord: Outcome and predictors of response of marital therapy versus cognitive therapy. *Behavior Therapy, 23*, 507–528. 0.1016/S0005-7894(05)80219-9.
- Bodenmann, G., Plancherel, B., Beach, S.R., Widmer, K., Gabriel, B., Meuwly, N., et al. (2008). Effects of coping-oriented couple therapy on depression: A randomized clinical trial. *Journal of Consulting and Clinical Psychology, 75*, 944–954. <https://doi.org/10.1037/a0013467>.
- Christian, J.L., O'Leary, K.D., & Vivian, D. (1994). Depressive symptomatology in maritally discordant women and men: The role of individual and relationship variables. *Journal of Family Psychology, 8*(1), 32–42. <https://doi.org/10.1037/0893-3200.8.1.32>.
- Cigrang, J.A., Cordova, J.V., Gray, T.D., Najera, E., Hawrilenko, M., Pinkley, C., et al (2016). The Marriage Checkup: Adapting and implementing a brief relationship intervention for military couples. *Cognitive and Behavioral Practice, 23*, 561–570. <https://doi.org/10.1016/j.cbpra.2016.01.002>.
- Cipriani, A., Furikawa, T.A., Salanti, G., Chaimani, A., Atkinson, L.Z., Ogawa, Y., et al. (2018). Comparative efficacy and acceptability of 21 antidepressants drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *Lancet, 391*, 1357–1366. [https://doi.org/10.1016/S0140-6736\(17\)32802-7](https://doi.org/10.1016/S0140-6736(17)32802-7).
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York, NY: Routledge.
- Cohen, S., O'Leary, K.D., & Foran, H. (2010). A randomized clinical trial of a brief, problem-focused couple therapy for depression. *Behavior Therapy, 41*, 433–446. <https://doi.org/10.1016/j.beth.2009.11.004>.
- Cook, W.L., & Kenny, D.A. (2005). The actor-partner interdependence model: A model of bidirectional effects in developmental studies. *International Journal of Behavioral Development, 29*(2), 101–109. <https://doi.org/10.1080/01650250444000405>.

- Cordova, J.V. (2013). *The Marriage Checkup practitioner's guide: Promoting lifelong relationship health*. Washington, DC: American Psychological Association.
- Cordova, J.V., Eubanks Fleming, C.J., Ippolito Morrill, M., Hawrilenko, M., Sollenberger, J.W., Harp, A.G., et al. (2014). The Marriage Checkup: A randomized controlled trial of annual relationship health checkups. *Journal of Consulting and Clinical Psychology, 82*(4), 592–604. <https://doi.org/10.1037/a0037097>.
- Cordova, J.V., & Scott, R. (2001). Intimacy: A behavioral interpretation. *The Behavior Analyst, 24*, 75–86.
- Dimick, J.B., & Ryan, A.M. (2014). Methods for evaluating changes in health care policy: The difference-in-differences approach. *JAMA, 312*(22), 2401–2402. <https://doi.org/10.1001/jama.2014.16153>.
- Doss, B.D., Cicila, L.N., Georgia, E.J., Roddy, M.K., Nowlan, K.M., Benson, L.A., et al. (2016). A randomized controlled trial of the web-based OurRelationship program: Effects on relationship and individual functioning. *Journal of Consulting and Clinical Psychology, 84*(4), 285–296. <https://doi.org/10.1037/ccp0000063>.
- Emanuels-Zuurveen, L., & Emmelkamp, P.M. (1996). Individual behavioural-cognitive therapy vs. marital therapy for depression in maritally distressed couples. *British Journal of Psychiatry, 169*, 181–188. <https://doi.org/10.1192/bjp.169.2.181>.
- Enders, C.K., & Bandalos, D.L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural equation modeling, 8*(3), 430–457. https://doi.org/10.1207/S15328007SEM0803_5.
- Evans, M.D., Hollon, S.D., DeRubeis, R.J., Piasecki, J.M., Grove, W.M., Garvey, M.J., et al. (1992). Differential relapse following cognitive therapy and pharmacotherapy for depression. *Archives of General Psychiatry, 49*(10), 802. <https://doi.org/10.1001/archpsyc.1992.01820100046>.
- Foley, S.H., Rounsaville, B.J., Weissman, M.M., Sholomskas, D., & Chevron, E. (1989). Individual versus conjoint interpersonal psychotherapy for depressed patients with marital disputes. *International Journal of Family Psychiatry, 10*, 29–42.
- Fournier, J.C., DeRubeis, R.J., Hollon, S.D., Dimidjian, S., Amsterdam, J.D., Shelton, R.C., et al. (2010). Antidepressant drug effects and depression severity: A patient-level meta-analysis. *JAMA: The Journal of the American Medical Association, 303*(1), 47–53. <https://doi.org/10.1001/jama.2009.1943>.
- Hayes, S.C., Luoma, J.B., Bond, F.W., Masudo, A., & Lillis, J. (2006). Acceptance and commitment therapy: Model, processes and outcomes. *Behaviour Research and Therapy, 44*(1), 1–25. <https://doi.org/10.1016/j.brat.2005.06.006>.
- Hollon, S.D., DeRubeis, R.J., Shelton, R.C., Amsterdam, J.D., Salomon, R.M., O'Reardon, J.P., et al. (2005). Prevention of relapse following cognitive therapy vs medications in moderate to severe depression. *Archives of General Psychiatry, 62*(4), 417. <https://doi.org/10.1001/archpsyc.62.4.417>.
- Hu, L.T., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal, 6*(1), 1–55. <https://doi.org/10.1080/10705519909540118>.
- Jacobson, N.S., & Christensen, A. (1996). *Acceptance and change in couple therapy: A therapist's guide to transforming relationships*. New York: Norton.
- Jacobson, N.S., Dobson, K., Fruzzetti, A.E., Schmalin, K.B., & Salusky, S. (1991). Marital therapy as a treatment for depression. *Journal of Consulting and Clinical Psychology, 59*, 547–557. <https://doi.org/10.1037/0022-006X.59.4.547>.
- Jacobson, N.S., Dobson, K.S., Truax, P.A., Addis, M.E., Koerner, K., Gollan, J.K., et al. (1996). A component analysis of cognitive-behavioral treatment for depression. *Journal of Consulting and Clinical Psychology, 64*, 295–304. <https://doi.org/10.1037/1522-3736.3.1.323a>.
- Keller, M.B., McCullough, J.P., Klein, D.N., Arnow, B., Dunner, D.L., Gelenberg, A.J., et al. (2000). A comparison of nefazodone, the cognitive behavioral-analysis system of psychotherapy, and their combination for the treatment of chronic depression. *New England Journal of Medicine, 342*(20), 1462–1470. <https://doi.org/10.1056/NeJM200005183422001>.
- Khan, A., Leventhal, R.M., Khan, S.R., & Brown, W.A. (2002). Severity of depression and response to antidepressants and placebo: An analysis of the Food and Drug Administration database. *Journal of Clinical Psychopharmacology, 22*(1), 40–45. <https://doi.org/10.1097/00004714-200202000-00007>.
- Kirsch, I., Deacon, B.J., Huedo-Medina, T.B., Scoboria, A., Moore, T.J., & Johnson, B.T. (2008). Initial severity and antidepressant benefits: A meta-analysis of data submitted to the Food and Drug Administration. *PLoS Med, 5*(2), 45. <https://doi.org/10.1371/journal.pmed.0050045>.
- Klein, D.N., & Santiago, N.J. (2003). Dysthymia and chronic depression: Introduction, classification, risk factors, and course. *Journal of Clinical Psychology, 59*, 807–816. <https://doi.org/10.1002/jclp.10174>.
- Kocsis, J.H. (2003). Pharmacotherapy and chronic depression. *Journal of Clinical Psychology, 59*, 885–892. <https://doi.org/10.1002/jclp.10180>.
- Lambert, N.M., Fincham, F.D., & Stillman, T.F. (2012). Gratitude and depressive symptoms: The role of positive reframing and positive emotion. *Cognition and Emotion, 26*(4), 615–633. <https://doi.org/10.1080/02699931.2011.595393>.

- Leff, J., Alexander, B., Asen, E., Brewin, C.R., Dayson, D., Vearnals, S., et al. (2003). Modes of action of family interventions in depression and schizophrenia: The same or different? *Journal of Family Therapy*, 25, 357–370. <https://doi.org/10.1111/1467-6427.00255>.
- McArdle, J.J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual review of psychology*, 60, 577–605. <https://doi.org/10.1146/annurev.psych.60.110707.163612>.
- McArdle, J.J. (2017). Dynamic structural equation modeling in longitudinal experimental studies. In K. Montfort, J. Oud, & A. Satorra (Eds.), *Longitudinal models in the behavioral and related sciences* (pp. 159–188). New York, NY: Routledge.
- Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, 55(1), 107–122. <https://doi.org/10.1007/BF02294746>.
- Miller, W.R., & Rollnick, S. (2002). *Motivational interviewing: Preparing people for change* (2nd ed.). New York: Guilford Press.
- Morrill, M.I., Fleming, C.E., Harp, A.G., Sollenberger, J.W., Darling, E.V., & Córdova, J.V. (2011). The Marriage Checkup: Increasing access to marital health care. *Family Process*, 50, 471–485. <https://doi.org/10.1111/j.1545-5300.2011.01372.x>.
- Mulrow, C.D., Williams, J.W., Jr, Trivedi, M., Chiquette, E., Aguilar, C., Cornell, J.E., et al. (1998). Treatment of depression – newer pharmacotherapies. *Psychopharmacology Bulletin*, 34(4), 409–795.
- Muthén, L.K., & Muthén, B.O. (1998–2011). *Mplus user's guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.
- National Institute for Health and Care Excellence (2009). Depression in adults: Recognition and management. Retrieved January 20, 2019, from <https://www.nice.org.uk/guidance/cg90>.
- Norton, R. (1983). Measuring marital quality: A critical look at the dependent variable. *Journal of Marriage and the Family*, 45, 141–151. <https://doi.org/10.2307/351302>.
- O'Leary, K.D., & Beach, S.R.H. (1990). Marital therapy: A viable treatment for depression and marital discord. *American Journal of Psychiatry*, 147, 183–186. <https://doi.org/10.1176/ajp.147.2.183>.
- Radloff, L.S. (1977). The CED-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1, 385–401. <https://doi.org/10.1177/014662167700100306>.
- Raykov, T., & Marcoulides, G.A. (2004). Using the delta method for approximate interval estimation of parameter functions in SEM. *Structural Equation Modeling*, 11(4), 621–637.
- Snyder, D.K. (1997). *Marital Satisfaction Inventory, Revised (MSI-R): Manual*. Los Angeles, CA: Western Psychological Services.
- Stark, S., Chernyshenko, O.S., & Drasgow, F. (2006). Detecting differential item functioning with confirmatory factor analysis and item response theory: Toward a unified strategy. *Journal of Applied Psychology*, 91(6), 1292.
- Teichman, Y., Bar-El, Z., Shor, H., Sirota, P., & Elizur, A. (1995). A comparison of two modalities of cognitive therapy (individual and marital) in treating depression. *Psychiatry: Interpersonal and Biological Processes*, 58, 136–148. <https://doi.org/10.1080/00332747.1995.11024721>.
- Thase, Michael, Reynolds, Charles, Frank, Ellen, Simons, Anne, Garamoni, Gregory, McGeary, J., et al. (1994). Response to cognitive behavior therapy in chronic depression. *Journal of Psychotherapy Practice and Research*, 3, 204–214.
- Wald, A. (1945). Sequential Tests of Statistical Hypotheses. *The Annals of Mathematical Statistics*, 16(2), 117–186. [10.1214/aoms/1177731118](https://doi.org/10.1214/aoms/1177731118). <https://projecteuclid.org/euclid.aoms/1177731118>.
- Westen, D., & Morrison, K. (2001). A multidimensional meta-analysis of treatments for depression, panic, and generalized anxiety disorder: An empirical examination of the states of empirically supported therapies. *Journal of Consulting and Clinical Psychology*, 69, 875–899. <https://doi.org/10.1037/0022-006X.69.6.875>.
- Whisman, M.A. (1999). Marital dissatisfaction and psychiatric disorders: Results from the National Comorbidity Survey. *Journal of Abnormal Psychology*, 108, 701–708. <https://doi.org/10.1037/0021-843X.108.4.701>.
- Whisman, M.A. (2001). The association between depression and marital satisfaction. In S.R.H. Beach (Ed.), *Marital and family processes in depression: A scientific foundation for clinical practice* (pp. 3–24). Washington, DC: American Psychological Association. <https://doi.org/10.1037/10350-001>.
- Whisman, M.A. (2007). Marital distress and DSM-IV psychiatric disorders in a population-based survey. *Journal of Abnormal Psychology*, 116, 638–643. <https://doi.org/10.1037/0021-843X.116.3.638>.
- Whisman, M.A., & Beach, S.R.H. (2012). Couple therapy for depression. *Journal of Clinical Psychology*, 68(5), 526–535. <https://doi.org/10.1002/jclp.21857>.
- Williams, J., & MacKinnon, D.P. (2008). Resampling and distribution of the product methods for testing indirect effects in complex models. *Structural Equation Modeling: A Multidisciplinary Journal*, 15(1), 23–51. <https://doi.org/10.1080/10705510701758166>.
- Wittenborn, A.K., Liu, T., Ridenour, T.A., Lachmar, E.M., Mitchell, E.A., & Seedall, R.B. (2018). Randomized controlled trial of emotionally focused couple therapy compared to treatment as usual for depression: Outcomes and mechanisms of change. *Journal of Marital and Family Therapy*, 45, 395–409. <https://doi.org/10.1111/jmft.12350>.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table S1. Fit statistics for intermediate models.

Table S2. Nested comparisons for multigroup models.