

Couple and Family Psychology: Research and Practice

A Comprehensive Meta-Analysis on the Efficacy of Emotionally Focused Couple Therapy

Paul M. Spengler, Nicholas A. Lee, Stephanie A. Wiebe, and Andrea K. Wittenborn

Online First Publication, September 22, 2022. <http://dx.doi.org/10.1037/cfp0000233>

CITATION

Spengler, P. M., Lee, N. A., Wiebe, S. A., & Wittenborn, A. K. (2022, September 22). A Comprehensive Meta-Analysis on the Efficacy of Emotionally Focused Couple Therapy. *Couple and Family Psychology: Research and Practice* Advance online publication. <http://dx.doi.org/10.1037/cfp0000233>

A Comprehensive Meta-Analysis on the Efficacy of Emotionally Focused Couple Therapy

Paul M. Spengler¹, Nicholas A. Lee², Stephanie A. Wiebe³, and Andrea K. Wittenborn⁴

¹ Department of Counseling Psychology, Social Psychology and Counseling, College of Health, Ball State University

² Open Door Health Services, Muncie, Indiana, United States

³ School of Counseling, Psychotherapy and Spirituality, Saint Paul University

⁴ Department of Human Development and Family Studies and Division of Psychiatry and Behavioral Medicine, Michigan State University

Emotionally Focused Couple Therapy (EFT), as developed by Greenberg and Johnson (1988) and Johnson (1996, 2004, 2020), is an evidence-based treatment that works with a variety of presenting concerns (e.g., depression, posttraumatic stress disorder, sexual dissatisfaction) and treatment characteristics, with good maintenance of treatment gains up to 2 years following treatment. In this first comprehensive meta-analysis of EFT, we evaluated the entirety of efficacy outcome research including randomized controlled trial (RCT), quasi-experimental, and dissertation studies. Previous meta-analyses of EFT analyzed solely RCTs (Beasley & Ager, 2019), Johnson's model of EFT combined with Greenberg and Goldman's (2008) conceptually different model (Rathgeber et al., 2018), or EFT embedded and not disentangled from other models of couple therapy (Roddy et al., 2020). Across 20 studies and 332 couples, we found medium to large random effects weighted *d*s for (a) pretest–posttest, $d = .93$, 95% CI [.75, 1.12], (b) EFT versus viable alternative couple interventions, $d = .44$, 95% CI [.03, .85], and (c) pretest to follow-up, $d = .86$, 95% CI [.56, 1.15]. Comprehensive sensitivity and moderator analyses were conducted. This meta-analysis provides support for EFT as an evidence-based couple therapy approach, where 70% of couples will be symptom free at the end of treatment and greater therapist fidelity to the model is associated with stronger couple gains. Limitations and implications are discussed in relation to research, training, and practice of EFT.

Paul M. Spengler  <https://orcid.org/0000-0001-9539-4214>

Stephanie A. Wiebe  <https://orcid.org/0000-0001-5865-4531>

Andrea K. Wittenborn  <https://orcid.org/0000-0003-4608-0635>

The authors have no conflict of interest to disclose.

Portions of this study were presented in August 2016 at the annual meeting of the American Psychological Association in Denver, CO.

Paul M. Spengler played lead role in conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software and writing of original draft, supporting role in funding acquisition and equal role in supervision, validation, visualization and writing

of review and editing. Nicholas A. Lee played lead role in funding acquisition, supporting role in conceptualization, formal analysis, project administration, resources, supervision, visualization, writing of original draft and writing of review and editing and equal role in investigation, methodology and validation. Stephanie A. Wiebe played supporting role in conceptualization, formal analysis, resources, writing of original draft and writing of review and editing. Andrea K. Wittenborn played supporting role in conceptualization, investigation, methodology, validation, visualization, writing of original draft and writing of review and editing.

Correspondence concerning this article should be addressed to Paul M. Spengler, Department of Counseling Psychology, Social Psychology and Counseling, College of Health, Ball State University, Muncie, IN 47306, United States. Email: pspengle@bsu.edu

Public Significance Statement

This empirical synthesis of existing emotionally focused couple therapy (EFT) outcome research provides support for EFT as an effective treatment for couples experiencing significant relationship distress. Of the 330 couples studied, 70% were symptom free at the end of treatment and gains last up to 2 years post treatment.

Keywords: emotionally focused therapy, emotionally focused couple therapy, couple therapy, outcome research, meta-analysis

Supplemental materials: <https://doi.org/10.1037/cfp0000233.supp>

Couple therapy as a field has undergone a great deal of growth particularly in the past two decades. Models for treating couple distress have become increasingly systematized and research has expanded in volume and rigor (Carr, 2018; Lebow et al., 2012). Synthesizing the growing research literature on couple therapy is important because it mitigates biases and limitations inherent in conclusions based on reading of individual studies and efforts at narrative reviews (Shadish & Baldwin, 2003). Findings from meta-analyses reveal moderate-to-large overall effects for couple therapy (Roddy et al., 2020; Shadish & Baldwin, 2003) with clinically significant gains in 65%–70% of treated couples (Halford et al., 2015; Lebow et al., 2012).

Emotionally Focused Couple Therapy

Emotionally focused couple therapy (EFT; Johnson, 1996, 2004, 2020) is one of the modalities with the strongest research support (Sprenkle, 2012; Wiebe & Johnson, 2016) and is considered to be evidence-based (Sexton et al., 2011). EFT—an attachment-based couple therapy built on humanistic and systemic principles—views the negative cycles of interaction typically presented by distressed couples as a reflection of an insecure attachment bond. The goal in EFT is to help partners create a safe and secure bond through the acknowledgement, exploration, and experiential sharing of vulnerable attachment-based emotions and needs. For more information outlining the stages and steps of EFT, see the most recent treatment manual (Johnson, 2020) and research, training, and practice resources posted by the International Centre for Excellence in Emotionally Focused Couple Therapy (ICEEFT; www.iceeft.com).

EFT Research

EFT has accumulated a large body of both process and outcome studies (for review, see Wiebe & Johnson, 2016). Previous meta-analyses of EFT have contributed to establishing EFT as an evidence-based couple therapy, notwithstanding significant limitations in methodology. Meta-analyses consist of analysis of effects from individual studies that are then aggregated and analyzed in an overall analysis, thus the term *meta-analysis*. Johnson et al. (1999) carried out the first empirical synthesis of EFT research, which included a narrative overview of nine EFT studies but meta-analyzed only four that were considered by the authors to be of high-quality randomized controlled trial studies (RCT; Goldman & Greenberg, 1992; Gordon-Walker et al., 1996; James, 1991; Johnson & Greenberg 1985b). They reported a weighted mean effect of $d = 1.31$, a large effect by any convention (Cohen, 1988), with 70%–73% of couples symptom free following treatment. As with many other couple therapy meta-analyses, change in couple relationship satisfaction served as the outcome measure.

More recently, Rathgeber et al. (2018) conducted a meta-analysis of EFT and behavioral couple therapy (BCT). They found a medium effect for the combination of these two models, $g = .60$,¹ with a larger effect for EFT of $g = 0.73$, compared with BCT, $g = 0.53$. At 6-month follow-up, the effect was medium for EFT, $g = 0.66$, and small for BCT, $g = 0.35$ ($g = 0.44$ overall). Rathgeber et al. (2018) assessed moderators, which are basically interaction or

¹ For purposes of comparison of these meta-analyses, Hedges' g is comparable to Cohen's d but is considered to provide correction for bias due to small sample sizes (Hedges & Olkin, 1985).

prediction effects, across both approaches. They found that greater pretest couple distress and couples who had been in longer relationships was associated with better outcomes. A notable limitation of this meta-analysis is that the researchers combined two distinct forms of emotion(ally) focused couple therapy under the acronym EFT that have similarities but are, in fact, different models known respectively as emotionally focused couple therapy (Johnson, 2020) and emotion-focused couple therapy (Greenberg & Goldman, 2008). This oversight resulted in a conflation of the results attributable to the unique characteristics of each model. This is problematic because of significant theoretical differences between these models. Johnson focuses on attachment-related emotions and needs whereas Greenberg and Goldman focus in a more general manner on emotions and personal needs related to self and the relationship. Rathgeber et al. (2018) also isolated their analyses to only RCT studies and, thereby, excluded about 50% of EFT research that is based on quasi-experimental pretest–posttest experimental group only designs.

As with Rathgeber et al. (2018), Beasley and Ager (2019) synthesized only RCT EFT studies and omitted the substantial number of quasi-experimental studies. They reported a surprisingly large overall Hedge's g of 2.09. This is because two peer-reviewed studies included in the analyses arguably should have been identified as statistical outliers (reported effect sizes of $g = 4.46$, Najafi et al., 2015 and $g = 6.25$, Solymani ahmadi et al., 2014²) and were, in fact, eliminated from the present meta-analysis (see Greenhouse & Iyengar, 2009). Although the aim was to include rigorous studies that met RCT standards, some were not peer reviewed and there were no moderation analyses for methodological rigor.

Most recently, Roddy et al. (2020) conducted a meta-analysis of couple therapy studies that included both RCT and quasi-experimental designs. Only nine EFT studies, however, were included, less than 50% of the extant body of research, and these also were not disentangled in the analyses from other couple therapies. For the combined couple therapies, they found a large effect for changes in relationship satisfaction of $g = 1.12$, that is roughly maintained at follow-ups of 6 months to 2 years ($g = .91$). As with Rathgeber et al. (2018), greater couple pretest distress was associated with the greatest couple

gains, which is not surprising given more room for measurable change and regression to the mean. Several potentially important conceptual and methodological moderators, addressed in the current meta-analysis, were not assessed.

These meta-analyses, limitations notwithstanding, are informative and suggest that couple therapy, in general, and EFT in particular produce moderate-to-large improvements for distressed couples in relationship satisfaction. Despite these important contributions, there remains the need for a comprehensive rigorous synthesis of EFT, that is inclusive of RCTs, quasi-experimental studies, and dissertations, without EFT embedded or confused with other couple therapies. While EFT has been extensively researched and has had wide impact on the training and practice of couple therapy, to date a fully comprehensive meta-analysis (CMA) of all existing EFT outcome studies has not been conducted. Our CMA focuses more precisely on a population estimate by including 100% of extant EFT research, not making decisions a priori about the strength or quality of the research, and by analyzing for methodological differences (e.g., RCTs and quasi-experimental studies) that may be associated with the overall effect rather than eliminating studies a priori. In response to this absence, we used meta-analytic techniques to assess the efficacy of EFT and to test an array of relevant moderators that are both (a) standard to meta-analyses, and (b) arise from our observations of couple therapy and EFT literature.

Assessing methodological moderators, such as selection bias, attrition bias, and study quality (Cooper et al., 2009), are important to understanding the strengths and limitations of a body of research. Identifying practice oriented and clinically meaningful moderators, such as therapist experience with EFT, EFT treatment fidelity, and the number or length of EFT sessions, has potential to aid with future model development, training, and practice of EFT. Unlike other meta-analyses of EFT, we implemented a sensitivity analysis strategy (Greenhouse & Iyengar, 2009) by examining factors such as the quality of the data (e.g., inclusion or exclusion of outlier effects, random vs. fixed effects analyses) and other commonly considered threats to validity in meta-analyses (publication bias, funnel plot,

² Solymani ahmadi et al. (2014) is erroneously cited as Ahmadi et al. (2014) by Beasley and Ager (2019).

file drawer analysis). A sensitivity analysis clarifies if the findings are robust to analysis and design choices. Our hope is that the results of this first CMA advances future research, practice, and training of EFT.

Method

Literature Search Process

Prior meta-analyses and reviews on EFT (e.g., Beasley & Ager, 2019; Johnson et al., 1999; Johnson & Wittenborn, 2012; Rathgeber et al., 2018; Roddy et al., 2020; Wiebe & Johnson, 2016; Wood et al., 2005) and the ICEEFT website were examined for EFT outcome studies. Forward and backward searching of identified studies was conducted for all identified studies. Database searches were then performed using PsycINFO, PsycBooks, PsycNET, and ProQuest dissertations and Theses Global. The first outcome study on EFT was published in 1985 by the main authors of the model (Johnson & Greenberg, 1985a). Therefore, our search parameters began with articles published in 1985 and ended with studies dated through January of 2021. The following keywords were used: EFT, EFCT, emotionally focused therapy, and emotionally focused couple therapy. When searching ProQuest dissertations, additional keywords were used: outcome, randomized, efficacy, pretest–posttest, and quasi-experimental. This decision was made after searching with the initial keywords yielded over 6,500 results. If a dissertation was later published, we used the published article, except for Dessaulles et al. (2003) where we used the dissertation (Dessaulles, 1991) to code necessary statistics.

Inclusion and Exclusion Criteria

To be included, a study had to be an *efficacy* study where the outcome of EFT was compared to a wait-list control condition, compared with one or more viable treatments, or assessed in a pretest–posttest only design. The term “efficacy” is customarily contrasted with the use of the term “effectiveness.” Efficacy refers to highly controlled experimental conditions, such as RCTs, whereas effectiveness refers to outcome research in the natural delivery of couple therapy or treatment-as-usual by practitioners. If an outcome

study did not assess couple distress it was not included (e.g., impact of EFT on depression, Alder et al., 2018). Additionally, the EFT model needed to be derived from either the initial Greenberg and Johnson (1988) or Johnson (1996, 2004, 2020) treatment manuals. A study was excluded if it contained missing information needed to calculate effect size estimates, or it was not an efficacy outcome study (e.g., process research, effectiveness study, case application, theoretical or conceptual article). Concerted efforts were made to contact authors for missing data, or reanalysis of their data that would allow for calculation of a study effect size (e.g., Wittenborn et al., 2019). The search process yielded 299 possible studies; 264 were excluded because they did not evaluate the efficacy of EFT, two were excluded because they were statistical outliers (Najafi et al., 2015; Solymani ahmadi et al., 2014), and three were excluded because information needed to compute an effect size was unobtainable. Eight studies shared the same sample of couples; we analyzed only the original study in the pretest–posttest effects. Some studies provided follow-up data in subsequent reports. In total, 19 studies were included in the pretest–posttest analysis and one study was included only in the EFT versus viable treatment analysis, Denton et al. (2012),³ making a total of 20 studies. Ten of the 19 pretest–posttest studies provided follow-up data; three follow-up data sets were obtained from subsequent publications.

Coding Procedures

The search process and coding were conducted by the authors of this study, three psychologists and one doctoral marriage and family therapist. All coders completed advanced training in EFT (4-day externship, year-long core-skills training, and EFT clinical supervision). Three of the coders are certified EFT therapists. All four work in academia and regularly treat couples using EFT. The coding scheme consisted of the following: primary outcomes for treatment, comparison, and control conditions (e.g., *M*, *SD*, *N*); and study, therapist, and

³ We decided to include Denton et al. (2012), but only in the EFT versus viable treatment comparison because EFT was *modified* by combining it with antidepressant medication. This combined model was contrasted with a medication management only treatment condition.

participant characteristics. Study characteristics included assessment of design, exclusion rate, attrition, researcher allegiance, session length, length of overall treatment, fidelity and adherence checks, unit of data analysis, and study quality. Study quality was rated as acceptable, good, and high based on Shadish et al.'s (2002) threats to validity. Cooper (1998) addressed the advantages and disadvantages of using a global rating for study quality; global ratings tend to have better rater agreement with comparable heuristic value. Therapist characteristics included EFT specific training and experience and general therapist experience level. Participant characteristics included mean length of relationship, mean age of couples, initial level of distress, the focus of treatment (e.g., general relationship distress vs. depression in the context of relationship distress), and participant race or ethnicity. In the case of missing information to compute an effect size attempts were made to contact study authors by email and phone. Kappa agreement was calculated using Cohen's κ (Cohen, 1960). Agreement coefficients between the raters ranged from $\kappa = .67$ – 1.0 . According to conventional benchmarks this is high moderate to full agreement (Landis & Koch, 1977). The categories with the lowest interrater agreement were (a) assessment of EFT specific training and experience (prior experience, study specific training, unknown), and (b) type of fidelity checks (bug/mirror supervision, supervision, videotape review, supervision, and videotape review). Disagreements by coders were resolved by consensus (Card, 2012).

Calculation of Effects

The dependent measure for all studies assessed couple satisfaction (e.g., Dyadic Adjustment Scale [DAS]; Spanier, 1976; Revised Dyadic Adjustment Scale [RDAS], Busby et al., 1995). The RDAS is a revised shorter version of the DAS (14 items, 18 fewer than DAS). Both the DAS and RDAS have good reliability and established construct and criterion validity; the DAS and RDAS total scores correlate .97 (Crane et al., 2000). Means and standard deviations or other statistics (e.g., t tests, F tests), along with participant sizes for EFT, comparison, and/or wait-list control groups at pretest, posttest, and follow-up were entered in the software program *CMA Version 2* (Borenstein et al., 2005). We

decided a priori to use random effects weighted d models for three analyses: (a) EFT pretest–posttest, (b) EFT versus a comparison treatment, and (c) EFT pretest to follow-up. A random effects model is most appropriate when the intent is to generalize to diverse research scenarios, compared with a fixed effects model which is more appropriate for closely replicated studies, such as in RCT medication trials (Borenstein et al., 2009). We chose random effects modeling because extant EFT outcome research poses a diverse set of questions, for example, in terms of methodology, rigor, and study of different clinical populations. Pretest–posttest correlations were set at .50. In calculation of the three overall effect sizes, each study effect was weighted by the inverse variance related to sample size (Borenstein et al., 2005). Additionally, we conducted a priori planned mixed-effects moderator and regression analyses when tests of heterogeneity of the effects was equal to or greater than 50% (Higgins et al., 2003). Only one moderator (participant gender) could be assessed using subgroup analyses, otherwise for the overall and the moderator analyses one effect size was analyzed from each study.

Results

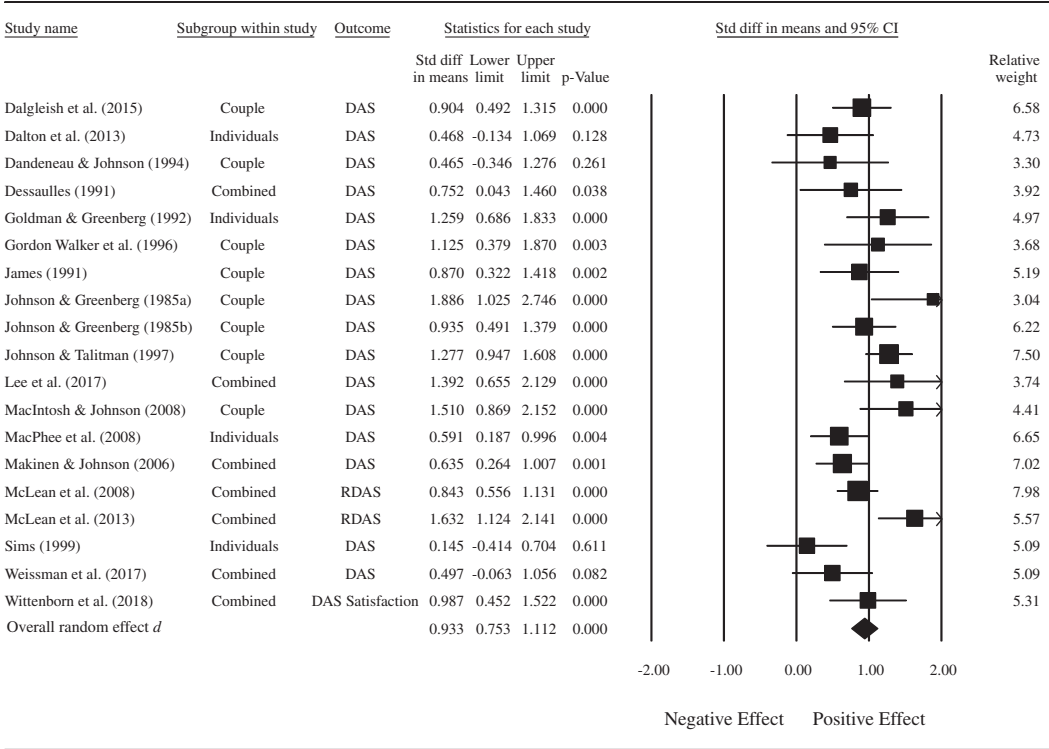
Overall Analyses

EFT Pretest–Posttest

Across 19 studies (9 RCT and 10 pretest–posttest), assessing the use of EFT in the treatment of 308 couples, we found a weighted random effect $d = .93$ reflecting a large treatment effect.⁴ The 95% confidence interval, CI [.75, 1.12], does not include zero meaning this is a statistically significant effect ($p < .001$). The studies, outcome measures, effects, and study moderators along with a forest plot are presented in Figure 1. The forest plot is widely used in meta-analyses and provides a readily understandable and visual representation of meta-analytic findings. All effect sizes were positive and ranged from .15 indicating a small EFT treatment effect to 1.89 reflecting a large impact of EFT. The overall effect is heterogeneous in

⁴ Random effects $d = 1.30$, 95% CI [.95, 1.64], including the two studies removed from the analysis, because they were assessed as outliers; discussed in the sensitivity analyses.

Figure 1
Random Effects (d) Pretest–Posttest EFT Only and Versus Control Group



Note. EFT = emotionally focused couple therapy; DAS = Dyadic Adjustment Scale; RDAS = Revised Dyadic Adjustment Scale. Positive *d* effect indicates positive treatment effect for EFT pretest–posttest or in comparison to control group. The squares represent the *d* effects proportionately sized to inverse-weights used in calculating the overall effect. The lines reflect 95% confidence intervals (CI) for each study. The diamond reflects mean weighted overall EFT effect; width of the diamond represents 95% CI.

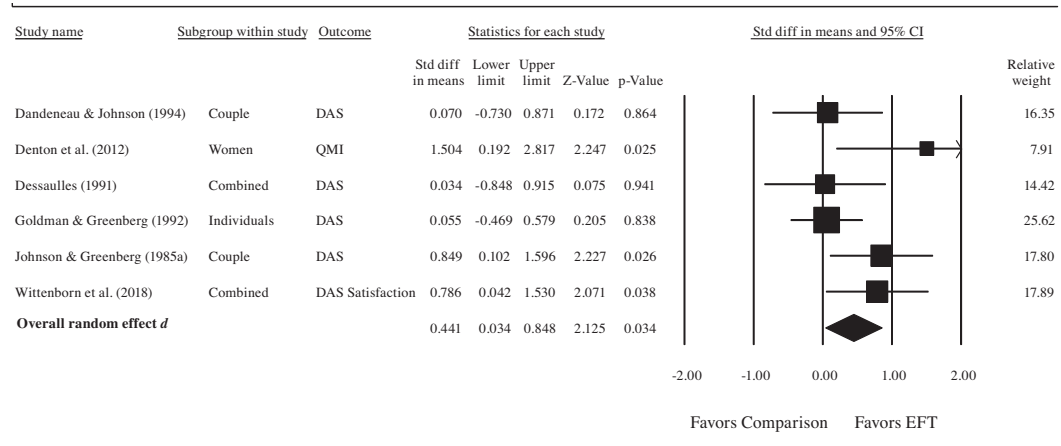
variance reflected by a statistically significant homogeneity index, $Q(18) = 41.69, p = .001$. The *I*-squared index is $I^2 = 56.82$ meaning 57% of the variance between studies is due to heterogeneity, and not chance, which is a moderate degree of heterogeneity (Higgins et al., 2003). This means that it is unlikely that all of the variance between studies is due to sampling error; covariates (or moderators) may account for differences between studies.

EFT Versus Comparison Treatment

Four studies compared EFT with other couple treatments, specifically cognitive marital therapy (Dandeneau & Johnson, 1994), integrated systemic therapy (Goldman & Greenberg, 1992), BCT (Johnson & Greenberg, 1985a), and

treatment as usual couple therapy (Wittenborn et al., 2019). Two studies contrasted EFT alone (Dessaulles, 1991) or EFT combined with medication management (Denton et al., 2012) with antidepressant medication. In both of these studies, medication management was provided to heterosexual women diagnosed with major depressive disorder. The overall weighted random effect $d = 0.44$ reflects that EFT produces stronger treatment effects on couple satisfaction than alternative treatments. The 95% CI [.03, .85] does not include zero reflecting a statistically significant difference in favor of EFT ($p = .03$). All effects were positive ranging from small ($d = .07$) to large ($d = 1.50$). A summary of study characteristics and a forest plot are presented in Figure 2. The overall effect is homogeneous, $Q(5) = 8.13, p = .15$. The *I*-squared index, $I^2 = 38.50$, reflects that 39% of the variance

Figure 2
Random Effects (d) EFT Versus Comparison Treatment



Note. Positive d effect indicates positive treatment effect for EFT versus comparison treatment. All studies compared EFT with a viable alternative treatment. Denton et al. (2012) compared EFT plus medication management with medication management only for heterosexual women with depression. QMI = quality of marriage index; EFT = emotionally focused couple therapy; DAS = Dyadic Adjustment Scale.

is due to chance, which is a low to moderate degree of heterogeneity.

EFT Pretest to Follow-Up

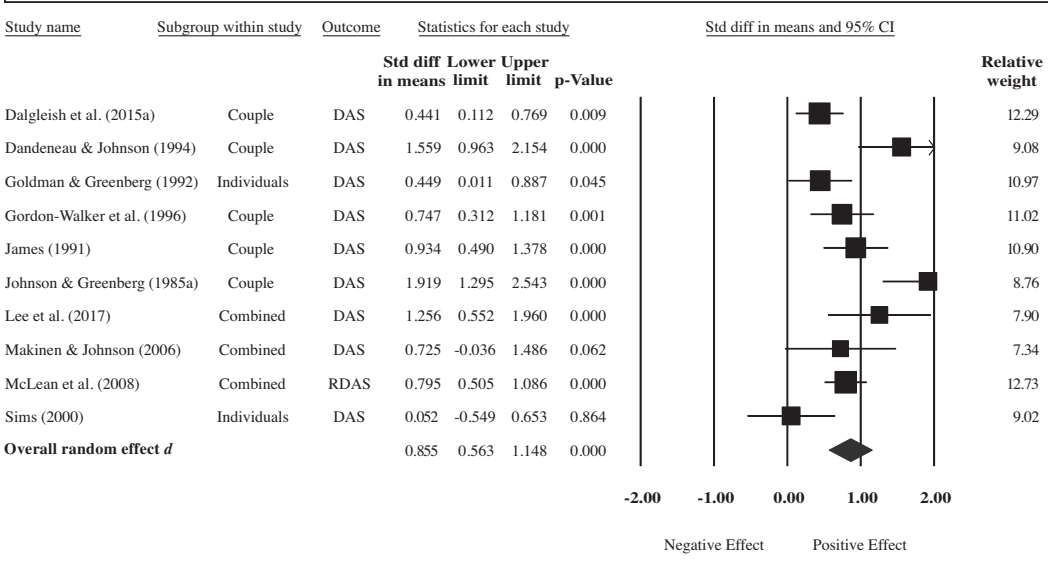
Of the 19 pretest–posttest EFT studies, ten provided follow-up data either in those studies or in subsequent publications. Time of follow-up ranged from 2 to 24 months after couples completed EFT. In each instance, we selected the longest time period available to contrast with pretest scores. A summary of study characteristics and a forest plot for these analyses are presented in Figure 3. The overall weighted random effects $d = .86$ for pretest to follow-up is similar in magnitude to the pretest–posttest $d = .93$, suggesting little decline over time in gains from EFT. This is a statistically significant effect ($p < .001$), reflected by a 95% CI [.56, 1.15] that does not cross zero. All effects were positive ranging from small ($d = .05$) to large ($d = 1.92$) suggesting that there are lasting positive effects of EFT. However, there is considerable heterogeneity between studies, $Q(9) = 33.54$, $p < .001$, with 73% of the variance not due to chance, $I^2 = 73.17$. Therefore, any interpretation of EFT gains being enduring should be viewed tentatively as other factors may account for these results (e.g., attrition bias, time of follow-up).

Sensitivity Analyses

Greenhouse and Iyengar (2009) suggest conducting sensitivity analyses to evaluate the impact of methodological and analysis decisions on meta-analytic results. To this end, we assessed the impact of outlier effects on the overall analyses using the outlier-labeling rule suggested by Tukey (1977) and refined by Hoaglin and Iglewicz (1987). With a lower critical value ($d = -1.18$) and an upper critical value ($d = 3.25$), two outliers were identified in initial pretest–posttest comparisons ($d = 5.72$, Najafi et al., 2015; $d = 6.49$, Solymani ahmadi et al., 2014). Including these outliers resulted in an inflated $d = 1.30$, 95% CI [.95, 1.64]. Removing these outliers resulted in a more centralized $d = .93$, 95% CI [.75, 1.11], that we consider to be the best estimate of the effect of EFT on couple satisfaction. The lower ($d = -2.07$) and upper ($d = 3.13$) critical values for EFT versus comparison treatments identified no outliers. No outliers existed for follow-up analyses using calculated lower ($d = -0.95$) and upper ($d = 2.83$) critical values.

Dandeneau and Johnson (1994) assessed EFT versus cognitive marital Therapy, for relationship enhancement in couples with no marital distress. Because this was the only study with

Figure 3
Random Effects (d) EFT Pretest to Follow-Up



Note. EFT = emotionally focused couple therapy; DAS = Dyadic Adjustment Scale; RDAS = Revised Dyadic Adjustment Scale. Positive *d* effect indicates positive treatment effect for EFT pretest to follow-up. Follow-up for three studies were obtained in subsequent publications: (a) Dalgleish et al. (2015; Wiebe, 2014), (b) Gordon-Walker et al. (1996; Cloutier et al., 2002), and (c) Makinen & Johnson (2006; Halchuk et al., 2010). Original studies are entered in the forest plot.

nondistressed couples, we decided to report effect sizes without Dandeneau and Johnson (1994): (a) EFT pretest–posttest random effects $d = .95$, 95% CI [.77, 1.13], (b) EFT versus a comparison treatment random effects $d = .53$, 95% CI [.05, 1.03], and (c) EFT pretest to follow-up random effects $d = .85$, 95% CI [.57, 1.14]. The only real difference by removing this study was an increase in the effect for EFT versus a comparison treatment ($d = .53$ vs. $d = .44$). We also examined differences between using random effects and fixed effects models. These comparisons show no real differences in the estimated effect sizes: (a) EFT pretest–posttest fixed effects $d = .92$, 95% CI [.81, 1.03], (b) EFT versus a comparison treatment fixed effects $d = .39$, 95% CI [.08, .70], and (c) EFT pretest to follow-up fixed effects $d = .78$, 95% CI [.63, .92].

Moderator Analyses

Moderator analyses are another important step in a sensitivity analysis, especially when 50% or more of the variance is unaccounted for by the

treatment (Higgins et al., 2003). With a magnitude of more than 50%, moderator analyses were conducted for both (a) pretest–posttest, and (b) pretest to follow-up analyses. No moderators were assessed for the EFT versus comparison treatment effect due to homogenous variance.

Categorical Moderator Analyses

Mixed effect moderator analyses were used for categorical variables using a procedure analogous to analysis of variance (ANOVA; Hedges & Olkin, 1985). Mixed effects models are thought to be a more stringent moderator test than fixed effects due to a lower likelihood of Type I errors (Overton, 1998). Each study was treated as the unit of analysis, except for a small number of studies that reported results separately for men and women (Dessaulles, 1991; Lee et al., 2017; Wittenborn et al., 2019), allowing for more powerful gender subgroup analyses. Other EFT studies reported results for additional subgroups (e.g., injured partner, offending partner; Makinen & Johnson, 2006; veterans, partners; Weissman et al., 2017) but were not sufficiently replicated

and we could not meaningfully categorize these for subgroup analyses.

Table 1 presents categorical moderator analyses for the pretest–posttest overall analyses. When reviewing these findings, the between-class effect (Q_B) is equivalent to a one-way ANOVA with levels included below. Moderators that are statistically significant are reflected by a 95% CI that does not cross zero, and differences

between moderators are statistically significant when respective 95% CIs do not overlap (Borman & Grigg, 2009). With the exception of unit of analysis, $Q(2) = 6.32, p = .04$, there were no other statistically significant categorical moderators. Studies that averaged couple scores ($d = 1.10$) were associated with larger treatment effects than studies that analyzed couples' individual scores ($d = .68$). Both effects are statistically significant,

Table 1
Categorical Models for Overall EFT Pretest–Posttest Effects

Moderators and levels	Between class effect		k	Mean weighted effect (d)	95% CI	
	(Q_B)	p			Lower	Upper
Conceptual moderators						
Population	3.30	.35				
Distressed couples			8	.95	.67	1.22
Enrichment couples			1	.47	–.54	1.48
Parents with child needs			4	1.20	.80	1.61
Couple psych symptoms			6	.46	.48	1.12
Researcher allegiance	.35	.55				
EFT Founder			13	.89	.67	1.12
No EFT founder			6	1.01	.69	1.34
Therapist experience EFT	.20	.90				
Prior EFT			8	.89	.61	1.67
Study-specific training			8	.98	.66	1.30
Unknown			3	.97	.50	1.45
Therapist experience	.30	.86				
Graduate degree			8	.88	.59	1.18
Graduate trainee			9	.99	.71	1.28
Mixed			2	.90	.37	1.44
Gender differences	.47	.49				
Men			3	.91	.40	1.41
Women			3	1.17	.62	1.71
Method moderators						
Research design	.07	.80				
Experimental			9	.91	.62	1.20
Quasi-experimental			10	.96	.71	1.13
Study quality	3.78	.29				
Acceptable			4	.82	.41	1.24
Good			7	.88	.60	1.15
Excellent			7	.94	.65	1.24
Inclusion criteria	2.54	.28				
Exclusive			3	1.08	.67	1.50
Typical			14	.94	.73	1.16
Inclusive			2	.47	–.18	1.11
Treatment fidelity	9.25	.06				
None			2	1.06	.61	1.51
Bug/mirror/videotape			1	1.39	.52	2.27
Supervision only			5	.62	.32	.93
Videotape only			1	1.63	.94	2.32
Supervision/videotape			10	.96	.73	1.19
Unit of analysis	6.32	.04				
Averaged couple scores			10	1.10	.88	1.31
Couple's individual scores			8	.68	.43	.93
One partner's scores			1	1.13	.25	2.00

Note. Positive effects reflect EFT efficacy. Gender effects were calculated for $k = 3$ studies using subgroup analyses. 95% CI = confidence interval; EFT = emotionally focused couple therapy.

as confidence intervals do not cross zero, and reflect real comparative differences, as there is no overlap between the confidence intervals. Treatment fidelity approached the convention of $p < .05$ for statistical significance. The largest effects for treatment fidelity were associated with Lee et al. (2017; $d = 1.39$) who reported use of the most rigorous fidelity steps (bug-in-the-ear, live mirror, and videotape review), and McLean et al. (2013; $d = 1.63$) where Sue Johnson supervised treatment fidelity for 25% of the videotaped sessions. Caution should be exercised in drawing conclusions about treatment fidelity and the efficacy of EFT due to the small number of studies ($k = 2$) and overlapping 95% CIs.

The categorical moderator analyses for the pretest to follow-up analyses are similar in findings (see Supplemental Table S1). Few categorical moderators were statistically significant predictors of the follow-up treatment effect, with the exception of therapist specific experience with EFT, $Q(1) = 4.80, p = .03$. Studies that included therapists who received study specific training in EFT ($d = 1.17$) were associated with larger follow-up effects, compared with those that utilized therapists with only prior EFT training ($d = .57$) but no study specific training and feedback. The 95% CI between these two types of EFT specific experience overlap indicating caution is warranted in reaching conclusions of true differences.

Regression Moderator Analyses

Ten continuous moderator variables of the overall pretest–posttest analyses were analyzed using mixed-effects (maximum likelihood) regression analyses (Borenstein et al., 2009).⁵ Of these moderators, three were measures of selection and attrition bias (exclusion rate by researchers, attrition of couples from eligible to pretest, attrition of couples from pre- to post-test), four assessed participant characteristics (age, race, years together, couple pretest DAS/RDAS), and three assessed treatment characteristics (number of EFT sessions, length of EFT sessions, percent of EFT sessions checked for fidelity). Two continuous moderators were assessed for the pretest to follow-up analyses: (a) months to follow-up, and (b) attrition posttest to follow-up.

Selection and Attrition Bias. Ten studies provided rates of researcher exclusion of couples while screening during study recruitment, with

percentages ranging from 0 (Johnson & Greenberg, 1985b) to 94 (Dalglish et al., 2015), $M = 46.4$.⁶ Rate of prestudy exclusion had no relation to the overall effect, $Q(1) = 0.49, p = .53$. Fifteen studies provided information on couples accepted for the study who dropped out from the point of study inclusion to pretest measurement, with percentages ranging from 0 (Dessaullles, 1991; James, 1991; Johnson & Greenberg, 1985a; Lee et al., 2017) to 52 (Goldman & Greenberg, 1992), $M = 13.4$. There was no association between dropout rates of couples prior to pretest and benefit from EFT, $Q(1) = 0.55, p = .46$. All 19 studies provided information on attrition of couples in treatment from pre- to post-test, with a modal dropout percent of zero ($k = 9$) and an upper range of 53% (Weissman et al., 2017), $M = 12.27$. Dropout rates for couples during treatment was also not associated with benefit from EFT, $Q(1) = 1.41, p = .24$.

Participant Characteristics. Sixteen studies reported the mean age of treatment couples, ranging from 36.9 (Walker et al., 1996) to 51.83 years (McLean et al., 2013), $M = 41.33$. Age of couples was unrelated to benefit from EFT, $Q(1) = 0.85, p = .36$. Seven studies provided some or minimal information about participants' race. Percentage of participants who identified as Caucasian ranged from 64.3 (Lee et al., 2017) to 100 (Goldman & Greenberg, 1992) and had no association with gains from EFT, $Q(1) = 0.01, p = .91$. All studies except Lee et al. (2017) reported mean years of couples being together, ranging from 6.9 (Johnson & Greenberg, 1985b) to 22.43 (McLean et al., 2013), $M = 13.03$. Length of couple relationship was unrelated to gains from EFT, $Q(1) = 0.42, p = .52$. With the exception of Wittenborn et al. (2019), who used the DAS Satisfaction subscale, all other studies assessed participants using total scores from the DAS or the RDAS. Scores from the RDAS were converted to DAS total score equivalencies following recommendations by Crane et al. (2000). Initial level of participants' distress did not predict response to EFT, $Q(1) = 0.88, p = .35$.

⁵ Graphs of continuous moderators regressed on standardized differences in mean scores are available upon request.

⁶ All means used in these regression moderator analyses are reported as unweighted means.

Treatment Characteristics. The mean number of EFT sessions in the 19 pretest–posttest studies ranged from 6 (Dandeneau & Johnson, 1994) to 31 (Weissman et al., 2017), $M = 13.89$. Number of sessions was associated with less benefit from treatment, $Q(1) = 5.48, p < .05$. Length of time for sessions was reported in 13 of the 19 pretest–posttest studies; six studies used 60 min, five provided 75 min, and two reported 90 min each session. The length of time for sessions was unrelated to the EFT treatment effect, $Q(1) = 0.92, p = .39$. Percent of sessions checked for treatment fidelity with therapist feedback could be assessed in 18 of the pretest–posttest studies and ranged from 0.00 (Johnson & Greenberg, 1985b; MacIntosh & Johnson, 2008; MacPhee et al., 1995; McLean et al., 2008; Weissman et al., 2017; Wittenborn et al., 2019) to 0.50 (Dandeneau & Johnson, 1994), $M = 18.89$. Higher percentage of sessions checked for treatment fidelity was associated with larger EFT treatment effects, $Q(17) = 6.19, p < .05$.

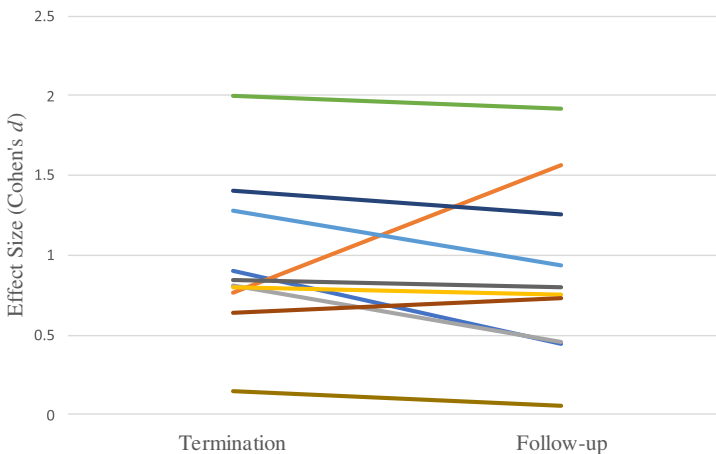
Follow-Up Characteristics. Follow-up data were available for 10 studies reported in either the original study or subsequent publication. Length of follow-up varied from 2 (Johnson & Greenberg, 1985a) to 24 months (e.g., Cloutier et al., 2002). In three studies, more than one time-stamp was reported for follow-up (e.g., 6, 12, 18, 24 months; Wiebe, 2014). Due to nonindependence of these data we elected to assess follow-up

for the data at the longest reported time period. Length of follow-up was associated with a decline in the treatment effect from pretest to time of follow-up, $Q(1) 4.71, p < .05$. Figure 4 provides a visual presentation of change from termination to follow-up. Attrition from termination to follow-up ranged from 0% (Dandeneau & Johnson, 1994; James, 1991; Lee et al., 2017) to 77.78% (Makinen & Johnson, 2006), $M = 25.53$. Higher dropout rates at follow-up was also associated with less benefit for the couples who participated in follow-up, $Q(9) = 7.46, p < .01$.

Publication Bias

For the sensitivity analysis we further analyzed pre- and post-test studies for publication bias. Publication bias occurs when reviewers and editors are, for example, more inclined to publish statistically significant results compared with statistically nonsignificant (Rothstein et al., 2005). A fail-safe analysis (Rosenthal, 1979) resulted in a two-tailed Z-value of 15.56 ($p < .00001$) for the 19 pre- and post-test studies. To obtain a statistically nonsignificant Z-value, under the cutoff of 1.96, a total of 1,180 zero effect EFT studies would need to exist, which seems implausible. A criticism of the fail-safe analysis is the assumption that missing studies would have zero effects, with no consideration of study weights (Becker, 2005). In response to this

Figure 4
Growth From Termination to Follow-Up



Note. See the online article for the color version of this figure.

criticism, we visually inspected a funnel plot with EFT pre- and post-test d effects on the x -axis and a measure of study sample size, or precision (one standard error) in this instance, on the y -axis (see Supplemental Figure S1). Larger studies cluster toward the top and smaller studies toward the bottom of the funnel. Visual inspection reflects no asymmetry suggesting there is no bias to publish EFT studies related to the magnitude of their effect. Egger et al. (1997) proposed, to further test this assumption, using the inverse of the standard error (precision) to predict the EFT pre- and post-test effect which results in a two-tailed $t(17) = 0.51, p = .66$. This finding means the distribution of effect sizes is symmetrical giving further support to there being no evidence of publication bias.

Discussion

In this study we sought to provide a comprehensive review of the outcome research pertaining to the efficacy of EFT (Johnson, 2020). Previous EFT meta-analyses have either included RCTs only, combined different models of couple therapy, or mixed theoretically related, but conceptually different models of couple therapy with the acronym EFT. We solely focused on Sue Johnson's model for EFT given its distinct theoretical and conceptual assumptions compared to, for example, the model of emotion-focused therapy put forth by Greenberg and Goldman (2008). Additionally, our meta-analysis was not limited to RCTs, but included quasi-experimental studies, all pretest–posttest designs, and one dissertation. The overall results show that EFT is a robust treatment for couple distress, that it compares well to a limited sampling of alternative treatments, and that the gains achieved at the end of treatment are sustained up to 2 years with modest decline over time.

Johnson et al. (1999) first reported a d effect of 1.3 after meta-analyzing four existing RCTs of EFT published at that time. We expected a downward correction in the overall d effect given the increase in EFT studies of varying scope and design. Across this body of outcome studies, we found a random effects weighted d of .93 pretest–posttest for improvements in relationship adjustment. Using Cohen's (1988) conventions, this is a large effect that exceeds d effects found in general for individual psychotherapy, ranging on average from d of .70–.80 as noted in broad

research reviews (see Lambert, 2013), and for couple and family therapy, ranging from $d = .59$ –.84 (Dunn & Schwebel, 1995; Shadish & Baldwin, 2003). Interestingly, we found no statistical difference between effects in studies that originated within Johnson's research lab and those from outside. In fact, the overall mean d effect for those studies originating outside of Johnson's lab was higher ($d = 1.01$ vs. .89), tentatively dispelling notions of researcher allegiance or bias. Moreover, we did not find any evidence that the design (RCT vs. quasi-experimental) or study quality moderated this effect, highlighting the importance we placed on analyzing 100% of EFT efficacy research, then analyzing design differences, versus excluding about 50% of studies, a priori, as in previous meta-analyses.

We found a pretreatment to follow-up d effect of .86 from 10 studies for those couples who participated in follow-up assessments at lengths ranging from 2 to 24 months. This is also a large effect per the benchmarks outlined by Cohen (1988) and is on par with the gains made pre- to post-test in EFT, suggesting overall relative stability in couple gains. Conducting follow-up assessment in psychotherapy generally, and couple therapy specifically, has been notoriously hard to do regardless of the model being studied. In perhaps the largest and longest follow-up study of couple therapy, Christensen et al. (2010) followed 134 couples for 5 years after completing treatment with either integrative behavioral couple therapy (IBCT) or traditional behavioral couple therapy (TBCT). At 5 years follow-up they found a pretreatment to follow-up d effect of 1.03 for the IBCT group and .92 for the TBCT group. We urge caution in simply comparing d effects at follow-up between IBCT/TBCT and EFT without embedding this comparison within context. The Christensen et al. (2010) project was a large-scale RCT where therapists received strict training and oversight in deploying their respective treatment protocols. Additionally, couples were paid increasing rates at each time point for their participation incentivizing their likelihood to follow through. The corpus of EFT follow-up research presented in this meta-analysis varies in design, the degree to which EFT fidelity was monitored, and appears to be influenced by moderators such as attrition and therapist fidelity to the model.

In addition to examining the effect of EFT from pretest–posttest, and through follow-up, when

available we examined the differential effect of EFT compared to other interventions in six studies. We found an overall $d = .44$ favoring EFT treatment. This effect was homogenous across studies; meaning, the variability in effects across comparison studies was largely attributable to differences in treatment and not chance. A criticism often cited against differences between treatment effects is unequal comparisons, such as researchers pitting a well-established treatment protocol against a nonbonafide approach. Of the six studies that compared EFT to some other form of treatment, two interventions were pharmacotherapy-based, two were a general form of cognitive interventions, one was integrated-systemic, and the final was treatment as usual. As such, it is important to situate our finding of EFT being more effective to comparison treatments within a context of what kinds of interventions were used. Not all these interventions, in our opinion, rise to the level of what Wampold et al. (1997) called well-established, or bonafide. Future research is needed to address this question by comparing EFT to manualized and well-established couple treatments conducted by clinicians equally trained in each model. Comparing effect sizes obtained from different meta-analyses of couple therapies is problematic due to differences in the way the studies are conducted (e.g., Christensen et al., 2010).

What Accounts for EFT Treatment Differences Across Studies?

Unlike previous meta-analyses we sought to comprehensively examine moderators of the efficacy of EFT and did so in four domains: (a) treatment characteristics, (b) client characteristics, (c) therapist characteristics and (d) methodological differences. Not all EFT meta-analyses assessed for heterogeneity in variance between studies (e.g., Beasley & Ager, 2019; Byrne et al., 2004). Failure to do so can lead to misleading findings and an overabundance of trust in the meaning of a singular index (Spielmanns & Flückiger, 2018). Notably, we found significant variance between studies in the pretest–posttest and follow-up analyses, which makes our confidence in these summary d effects more tentative. Heterogeneity of variance could reflect differences that are clinically significant and/or significant in terms of methodological variables.

Moderators are essentially interaction effects assessed in meta-analytic studies and can be informative for future researchers, practitioners, and trainers. The strongest moderator finding is the association between EFT treatment adherence, or fidelity, and EFT efficacy. Studies that had higher percentages of sessions checked for treatment adherence, with feedback provided to clinicians, were associated with better couple outcomes. Likewise, therapists who received study specific EFT training produced substantially better follow-up outcomes ($d = 1.17$), compared with therapists who only had prior EFT experience but no on-study training ($d = .57$). The method used to monitor treatment adherence, such as bug-in-the-ear with mirror supervision and videotape feedback, or videotape feedback only, only approached the convention for statistical significance ($p = .06$). These findings collectively lend support to the notion that following the EFT model, pursuing EFT training and supervision, and striving for model fidelity will lead to greater therapeutic benefits for couples (Wiebe & Johnson, 2016). Sexton et al. (2011) called for evidence-based treatments to demonstrate a relation between model adherence and treatment efficacy, which is supported by these moderator analyses. Future research should systematically assess for which methods of EFT supervision and feedback are most beneficial to couple outcomes. An extensive body of research exists on EFT supervision and training that could be applied to this question (e.g., see Sandberg, 2011). Likewise, it could be productive to conduct process-outcome research on differences between low and high-fidelity EFT therapists correlated with follow-up outcomes. It could be that high-fidelity therapists are smoother, deeper, more incisive in their delivery of the treatment, or more educational resulting in greater generalization and maintenance of gains made by couples.

It was curious that more EFT sessions provided to couples was associated with less benefit overall. Several variables may be at play here, such as attrition and the need for more sessions for more severely distressed couples. Consequently, we inspected those studies with the highest number of sessions and found, in addition to relationship distress, those couples were presenting with comorbid concerns such as sexual abuse and trauma, posttraumatic stress, and major depressive disorder. It is possible, therefore, that this

finding has less to do with the number of sessions being provided and more to do with the nature of complex distress these couples were experiencing. As any practitioner treating couples can attest to rarely do couples in routine clinical settings present for issues pertaining to relationship concerns alone (Whisman, 2007). Length of follow-up and higher attrition rates at follow-up were also associated with less treatment gains. Further EFT research is needed to fully unpack who is not responding best to EFT and not maintaining gains post treatment. The fact that therapists with more study specific EFT supervision produced better follow-up outcomes seems promising. Perhaps therapists who adhere more closely to the EFT model inoculate couples from future couple distress by being better at experiential and reflective “teaching” of new relationship skills.

The unit of analysis used by researchers, either analyses of averaged couple scores ($d = 1.10$) or analyses of couples’ individual scores ($d = .68$), was a significant methodological moderator. With few exceptions (e.g., Wittenborn et al., 2019), existing EFT outcome studies did not use dyadic methods for data analysis (e.g., nesting individuals within couples) to account for couple interdependence. This issue of unit of analysis created a conundrum for us, as it resulted in studies that analyzed at the individual level being weighted more heavily than warranted. For example, if there were 20 couples and one study analyzed as $N = 20$ and the other analyzed as $N = 40$ that means the second study received more weight in the overall meta-analysis. We recommend that future EFT researchers utilize methodologically and statistically sound dyadic research methods that account for couple interdependence and thereby provide unbiased estimates of treatment effects (for resource on dyadic research, see Wittenborn et al., 2013).

Practical Meaning of Overall EFT Effect

It is important to assess the practical meaning of research findings so they can be easily consumed by practitioners and shared with couples. A d of .93 means that couples who participate in EFT under experimental conditions, for example, with additional supervision, fidelity checks, and other characteristics of outcome research, are on average better off than 81% of couples who do not receive treatment. The overall effect size for

EFT can be converted to a success rate for treated couples (Norcross et al., 2017). For a d effect of .93 this translates to 70% of couples in EFT are successfully treated for their relationship distress. This finding is consistent with the frequently cited finding where Johnson et al. (1999) concluded that 70%–73% of couples are symptom free at the end of EFT.

Limitations and Implications for Future Research

While our study was more comprehensive, and we were able to glean more meaningful moderator analyses, like other couple therapy meta-analyses few statistically significant moderators of the EFT effects were found. This could be an artifact of low statistical power related to the modest sample size of 20 studies with a mean of 14.3 couples per study (for further discussion, see Hedges & Pigott, 2004). Further, because of restrictions in how the data are reported nearly all the moderator analyses were conducted on a between study basis versus the more powerful subgroups or within-studies basis (Wittenborn et al., 2013). This artifact of the EFT research led us to comparisons with lower certainty because the moderators assess each study as a whole and not the specific attribute. Wood and Eagly (2009) argue that the moderator may then be a proxy with other study attributes, thereby reducing statistical conclusion validity.

We agree with Norcross and Cooper (2021) who stated, “It is a false, and to be blunt, misleading presupposition in RCTs that patients are homogeneous” (p. ii). Therefore, one area that could be extended in EFT research is to examine relevant subgroups within studies. We were able to do this only for gender and in only three studies, thereby limiting our ability to assess differences between cisgender heterosexual men’s and women’s responses to EFT. We sought to code other meaningful subgroups (e.g., offending partner vs. nonoffending partner) but too few studies existed. Several subgroups might be systematically explored in accordance with attachment theory and the EFT model, for example, (a) in studies of heterosexual couples men and women given the idea that men are not as emotionally attuned and tend to be withdrawers, (b) withdrawers versus

blamers, and (c) attachment style. Several moderators were not statistically significant (e.g., length of session), and we urge caution in interpreting these findings based on the low statistical power of the moderator analyses.

Like other couple therapy research, the participants in this body of EFT research are predominantly White, middle-class, and moderately to mildly distressed, and are, or are presumed to be, cisgender, heterosexual couples (for a comprehensive review of sexual and gender identity in couple therapy research, see Spengler et al., 2020). The differences between couples who participate in efficacy studies and in naturalistic treatment settings have been widely discussed (e.g., Wright et al., 2007). To make meaningful assumptions about the generalizability of EFT to diverse couples help-seeking in routine practice, EFT researchers need to tackle this issue directly by actively recruiting and conducting studies with more diverse participants. As discussed by Spengler et al. (2020), the implicit or explicit inclusion or exclusion of sexual and gender minority couples, and a general lack of consideration of other sociocultural factors such as race and ethnicity, in couple therapy outcome research risks making these couples invisible. Future EFT researchers are urged to expand tests of the efficacy of the EFT model and to routinely report on self-identified measures of couple characteristics such as sexual, gender, racial, and ethnic identity. In our opinion (Spengler et al., 2020), these are important areas for the future development of research on EFT and, for that matter, all models of couple therapy.

Of course, questions of “transportability” or external validity exist with RCTs and quasi-experimental studies (see Lebow et al., 2012). As with other couple therapies more research is needed on the effectiveness of EFT in naturalistic settings. Mitchell and Spengler (2022) recently conducted a naturalistic study of EFT in routine clinical settings. Across seven therapists and 11 couples being followed for 18 months, Mitchell and Spengler found a statistically significant d effect of .46 for relationship improvement. Such shrinkage in effect size from efficacy to effectiveness studies is common and should lead practitioners to be cautious in assuming the findings from this EFT meta-analysis will generalize to their clinical practice.

Clinical Implications and Applications

Despite these limitations practitioners can conclude that EFT is an efficacious approach to treating couple distress and, with caveats noted, that treatment gains are relatively lasting. Despite working with couples on a regular basis very few practitioners receive substantive training in evidence-based couple treatments beyond a single course or two in graduate school (Lee & Spengler, 2017). Based on this meta-analysis and findings from process and other areas of EFT research (Wiebe & Johnson, 2016), practitioners can feel confident studying and providing EFT to couples. ICEEFT has a training library of master videotapes with real couples in real treatment; a network of consultants and supervisors; training opportunities in large and small groups; and a competency-based certification program aimed at increasing the fidelity with which practitioners practice EFT. Chow et al. (2015) found that therapists who engaged in regular deliberate practice, such as is available in studying EFT, achieve better client outcomes than those who did not. The results of this meta-analysis suggest practitioners can confidently engage in such study and practice of EFT to the benefit of their couples.

References

References with an asterisk indicate studies in the meta-analysis.

- Alder, M. C., Dyer, W. J., Sandberg, J. G., Davis, S. Y., & Holt-Lunstad, J. (2018). Emotionally-focused therapy and treatment as usual comparison groups in decreasing depression: A clinical pilot study. *The American Journal of Family Therapy*, 46(5), 541–555. <https://doi.org/10.1080/01926187.2019.1572478>
- Beasley, C. C., & Ager, R. (2019). Emotionally focused couples therapy: A systematic review of its effectiveness over the past 19 years. *Journal of Evidence-Based Social Work*, 16(2), 144–159. <https://doi.org/10.1080/23761407.2018.1563013>
- Becker, B. J. (2005). The failsafe N or file-drawer number. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment and adjustments* (pp. 111–126). Wiley.
- Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. (2005). *Comprehensive meta-analysis versions 2*. Biostat.

- Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. (2009). *Introduction to meta-analysis*. Wiley. <https://doi.org/10.1002/9780470743386>
- Borman, G. D., & Grigg, J. A. (2009). Visual and narrative interpretation. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 497–520). Sage Publications.
- Busby, D. M., Crane, D. R., Larson, J. H., & Christensen, C. (1995). A revision of the Dyadic Adjustment Scale for use with distressed and nondistressed couples: Construct hierarchy and multidimensional scales. *Journal of Marital and Family Therapy*, 21(3), 289–308. <https://doi.org/10.1111/j.1752-0606.1995.tb00163.x>
- Byrne, M., Carr, A., & Clark, M. (2004). The efficacy of behavioral couples therapy and emotionally focused therapy for couple distress. *Contemporary Family Therapy*, 26(4), 361–387. <https://doi.org/10.1007/s10591-004-0642-9>
- Card, N. (2012). *Applied meta-analysis for social science research*. Guilford.
- Carr, A. (2018). Couple therapy, family therapy and systemic interventions for adult-focused problems: The current evidence base. *Journal of Family Therapy*, 41(4), 492–536. <https://doi.org/10.1111/1467-6427.12225>
- Chow, D. L., Miller, S. D., Seidel, J. A., Kane, R. T., Thornton, J. A., & Andrews, W. P. (2015). The role of deliberate practice in the development of highly effective psychotherapists. *Psychotherapy*, 52(3), 337–345. <https://doi.org/10.1037/pst0000015>
- Christensen, A., Atkins, D. C., Baucom, B., & Yi, J. (2010). Marital status and satisfaction five years following a randomized clinical trial comparing traditional versus integrative behavioral couple therapy. *Journal of Consulting and Clinical Psychology*, 78(2), 225–235. <https://doi.org/10.1037/a0018132>
- *Cloutier, P. F., Manion, I. G., Walker, J. G., & Johnson, S. M. (2002). Emotionally focused interventions for couples with chronically ill children: A 2-year follow-up. *Journal of Marital and Family Therapy*, 28(4), 391–398. <https://doi.org/10.1111/j.1752-0606.2002.tb00364.x>
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46. <https://doi.org/10.1177/001316446002000104>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum.
- Cooper, H. (1998). *Synthesizing research: A guide for literature reviews* (3rd ed.). Sage Publications.
- Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). Sage Publications.
- Crane, D. R., Middleton, K. C., & Bean, R. A. (2000). Establishing criterion scores for the Kansas Marital Satisfaction Scale and the Revised Dyadic Adjustment Scale. *The American Journal of Family Therapy*, 28(1), 53–60. <https://doi.org/10.1080/019261800261815>
- *Dalglish, T. L., Johnson, S. M., Burgess Moser, M., Wiebe, S. A., & Tasca, G. A. (2015). Predicting key change events in emotionally focused couple therapy. *Journal of Marital and Family Therapy*, 41(3), 260–275. <https://doi.org/10.1111/jmft.12101>
- *Dalton, E. J., Greenman, P. S., Classen, C. C., & Johnson, S. M. (2013). Nurturing connections in the aftermath of childhood trauma: A randomized controlled trial of emotionally focused couple therapy for female survivors of childhood abuse. *Couple & Family Psychology*, 2(3), 209–221. <https://doi.org/10.1037/a0032772>
- *Dandeneau, M. L., & Johnson, S. M. (1994). Facilitating intimacy: Interventions and effects. *Journal of Marital and Family Therapy*, 20(1), 17–33. <https://doi.org/10.1111/j.1752-0606.1994.tb01008.x>
- *Denton, W. H., Wittenborn, A. K., & Golden, R. N. (2012). Augmenting antidepressant medication treatment of depressed women with emotionally focused therapy for couples: A randomized pilot study. *Journal of Marital and Family Therapy*, 38(Suppl. 1), 23–38. <https://doi.org/10.1111/j.1752-0606.2012.00291.x>
- *Dessaulles, A. (1991). *The treatment of clinical depression in the context of marital distress* [Doctoral Dissertation, University of Ottawa]. ProQuest Dissertations Publishing.
- *Dessaulles, A., Johnson, S. M., & Denton, W. H. (2003). Emotion-focused therapy for couples in the treatment of depression: A pilot study. *The American Journal of Family Therapy*, 31(5), 345–353. <https://doi.org/10.1080/01926180390232266>
- Dunn, R. L., & Schwebel, A. I. (1995). Meta-analytic review of marital therapy outcome research. *Journal of Family Psychology*, 9(1), 58–68. <https://doi.org/10.1037/0893-3200.9.1.58>
- Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *The BMJ*, 315(7109), 629–634. <https://doi.org/10.1136/bmj.315.7109.629>
- *Goldman, A., & Greenberg, L. (1992). Comparison of integrated systemic and emotionally focused approaches to couples therapy. *Journal of Consulting and Clinical Psychology*, 60(6), 962–969. <https://doi.org/10.1037/0022-006X.60.6.962>
- *Gordon-Walker, J., Johnson, S. M., Manion, I., & Cloutier, P. (1996). Emotionally focused marital intervention for couples with chronically ill children. *Journal of Consulting and Clinical Psychology*, 64(5), 1029–1036. <https://doi.org/10.1037/0022-006X.64.5.1029>
- Greenberg, L. S., & Goldman, R. N. (2008). *Emotion-focused couples therapy*. American Psychological Association. <https://doi.org/10.1037/11750-000>
- Greenberg, L. S., & Johnson, S. M. (1988). *Emotionally focused therapy for couples*. Guilford Press.

- Greenhouse, J. B., & Iyengar, S. (2009). Sensitivity analysis and diagnostics. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 417–434). Sage Publications.
- Halchuk, R. E., Makinen, J. A., & Johnson, S. M. (2010). Resolving attachment injuries in couples using emotionally focused therapy: A three-year follow-up. *Journal of Couple and Relationship Therapy*, 9(1), 31–47. <https://doi.org/10.1080/15332690903473069>
- Halford, W. K., Pepping, C. A., & Petch, J. (2015). The gap between couple therapy research efficacy and practice effectiveness. *Journal of Marital and Family Therapy*, 42(1), 32–44. <https://doi.org/10.1111/jmft.12120>
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Academic Press.
- Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods*, 9(4), 426–445. <https://doi.org/10.1037/1082-989X.9.4.426>
- Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *The BMJ*, 327(7414), 557–560. <https://doi.org/10.1136/bmj.327.7414.557>
- Hoaglin, D. C., & Iglewicz, B. (1987). Fine tuning some resistant rules for outlier labeling. *Journal of the American Statistical Association*, 82(400), 1147–1149. <https://doi.org/10.1080/01621459.1987.10478551>
- *James, P. S. (1991). Effects of a communication training component added to an emotionally focused couples therapy. *Journal of Marital and Family Therapy*, 17(3), 263–275. <https://doi.org/10.1111/j.1752-0606.1991.tb00894.x>
- Johnson, S. M. (1996). *The practice of emotionally focused couple therapy: Creating connection*. Routledge.
- Johnson, S. M. (2004). *The practice of emotionally focused couple therapy: Creating connection* (2nd ed.). Routledge.
- Johnson, S. M. (2020). *The practice of emotionally focused couple therapy: Creating connection* (3rd ed.). Routledge.
- *Johnson, S. M., & Greenberg, L. S. (1985a). Differential effects of experiential and problem-solving interventions in resolving marital conflict. *Journal of Consulting and Clinical Psychology*, 53(2), 175–184. <https://doi.org/10.1037/0022-006X.53.2.175>
- *Johnson, S. M., & Greenberg, L. S. (1985b). Emotionally focused couples therapy: An outcome study. *Journal of Marital and Family Therapy*, 11(3), 313–317. <https://doi.org/10.1111/j.1752-0606.1985.tb00624.x>
- Johnson, S. M., Hunsley, J., Greenberg, L. S., & Schindler, D. (1999). Emotionally focused couples therapy: Status and challenges. *Clinical Psychology: Science and Practice*, 6(1), 67–79. <https://doi.org/10.1093/clipsy.6.1.67>
- *Johnson, S. M., & Talitman, E. (1997). Predictors of success in emotionally focused marital therapy. *Journal of Marital and Family Therapy*, 23(2), 135–152. <https://doi.org/10.1111/j.1752-0606.1997.tb00239.x>
- Johnson, S. M., & Wittenborn, A. K. (2012). New research findings on emotionally focused therapy: Introduction to special section. *Journal of Marital and Family Therapy*, 38(Suppl. 1), 18–22. <https://doi.org/10.1111/j.1752-0606.2012.00292.x>
- Lambert, M. J. (2013). Outcome in psychotherapy: The past and important advances. *Psychotherapy*, 50(1), 42–51. <https://doi.org/10.1037/a0030682>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310>
- Lebow, J. L., Chambers, A. L., Christensen, A., & Johnson, S. M. (2012). Research on the treatment of couple distress. *Journal of Marital and Family Therapy*, 38(1), 145–168. <https://doi.org/10.1111/j.1752-0606.2011.00249.x>
- *Lee, N. A., Furrow, J. L., & Bradley, B. A. (2017). Emotionally focused couple therapy for parents raising a child with an autism spectrum disorder: A pilot study. *Journal of Marital and Family Therapy*, 43(4), 662–673. <https://doi.org/10.1111/jmft.12225>
- Lee, N. A., & Spengler, P. (2017). Training counseling psychologists in couple and family therapy. In J. L. Lebow, A. L. Chambers, & D. C. Breunlin (Eds.), *Encyclopedia of couple and family therapy*. Springer Publishing Company. https://doi.org/10.1007/978-3-319-15877-8_707-1
- *MacIntosh, H. B., & Johnson, S. (2008). Emotionally Focused Therapy for couples and childhood sexual abuse survivors. *Journal of Marital and Family Therapy*, 34(3), 298–315. <https://doi.org/10.1111/j.1752-0606.2008.00074.x>
- *MacPhee, D. C., Johnson, S. M., & Van der Veer, M. M. C. (1995). Low sexual desire in women: The effects of marital therapy. *Journal of Sex & Marital Therapy*, 21(3), 159–182. <https://doi.org/10.1080/00926239508404396>
- *Makinen, J. A., & Johnson, S. M. (2006). Resolving attachment injuries in couples using emotionally focused therapy: Steps toward forgiveness and reconciliation. *Journal of Consulting and Clinical Psychology*, 74(6), 1055–1064. <https://doi.org/10.1037/0022-006X.74.6.1055>
- *McLean, L. M., Jones, J. M., Rydall, A. C., Walsh, A., Esplen, M. J., Zimmermann, C., & Rodin, G. M. (2008). A couples intervention for patients facing advanced cancer and their spouse caregivers: Outcomes of a pilot study. *Psycho-Oncology*, 17(11), 1152–1156. <https://doi.org/10.1002/pon.1319>

- *McLean, L. M., Walton, T., Rodin, G., Espen, M. J., & Jones, J. M. (2013). A couple-based intervention for patients and caregivers facing end-stage cancer: Outcomes of a randomized controlled trial. *Psycho-Oncology*, 22(1), 28–38. <https://doi.org/10.1002/pon.2046>
- Mitchell, A. M., & Spengler, P. M. (2022). *A pilot study of the effectiveness of naturally delivered emotionally focused couple therapy* [Manuscript submitted for publication].
- *Najafi, M., Soleimani, A. A., Ahmadi, K., Javidi, N., & Kamkar, E. H. (2015). The effectiveness of emotionally focused therapy on enhancing marital adjustment and quality of life among infertile couples with marital conflicts. *International Journal of Fertility & Sterility*, 9(2), 238–246. <https://doi.org/10.22074/ijfs.2015.4245>
- Norcross, J. C., & Cooper, M. (2021). *Personalizing psychotherapy: Assessing and accommodating patient preferences*. American Psychological Association. <https://doi.org/10.1037/0000221-000>
- Norcross, J. C., Hogan, T. P., Koocher, G. P., & Maggio, L. A. (2017). *Clinicians' guide to evidence-based practices: Behavioral health and addictions* (2nd ed.). Oxford University Press. <https://doi.org/10.1093/med:psych/9780190621933.001.0001>
- Overton, R. C. (1998). A comparison of fixed-effects and mixed (random effects) models for meta-analysis tests of moderator variable effects. *Psychological Methods*, 3(3), 354–379. <https://doi.org/10.1037/1082-989X.3.3.354>
- Rathgeber, M., Bürkner, P. C., Schiller, E. M., & Holling, H. (2018). The efficacy of emotionally focused couples therapy and behavioral couples therapy: A meta-analysis. *Journal of Marital Therapy*, 45(3), 447–463. <https://doi.org/10.1111/jmft.12336>
- Roddy, M. K., Walsh, L. M., Rothman, K., Hatch, S. G., & Doss, B. D. (2020). Meta-analysis of couple therapy: Effects across outcomes, designs, timeframes, and other moderators. *Journal of Consulting and Clinical Psychology*, 88(7), 583–596. <https://doi.org/10.1037/ccp0000514>
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86(3), 638–641. <https://doi.org/10.1037/0033-2909.86.3.638>
- Rothstein, H. R., Sutton, A. J., & Borenstein, M. (Eds.). (2005). *Publication bias in meta-analysis: Prevention, assessment and adjustments*. Wiley.
- Sandberg, J. G. (2011). Introduction to the special section on learning Emotionally Focused Couples Therapy. *Journal of Marital and Family Therapy*, 37(4), 377–379. <https://doi.org/10.1111/j.1752-0606.2011.00251.x>
- Sexton, T., Gordon, K. C., Gurman, A., Lebow, J., Holtzworth-Munroe, A., & Johnson, S. (2011). Guidelines for classifying evidence-based treatments in couple and family therapy. *Family Process*, 50(3), 377–392. <https://doi.org/10.1111/j.1545-5300.2011.01363.x>
- Shadish, W. R., & Baldwin, S. A. (2003). Meta-analysis of MFT interventions. *Journal of Marital and Family Therapy*, 29(4), 547–570. <https://doi.org/10.1111/j.1752-0606.2003.tb01694.x>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton-Mifflin and Company.
- *Sims, A. E. B. (2000). *Working models of attachment: The impact of emotionally focused marital therapy* [Doctoral Dissertation, University of Ottawa]. ProQuest Dissertation Publishing.
- *Solymani ahmadi, F., Zarei, E., & Fallahchai, S. R. (2014). The effectiveness of emotionally-focused couple therapy in resolution of marital conflicts between the couples who visited the consultation centers. *Journal of Education and Management Studies*, 4(1), 118–123.
- Spanier, G. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and Family*, 38(1), 15–28. <https://doi.org/10.2307/350547>
- Spengler, E. S., DeVore, E. N., Spengler, P. M., & Lee, N. A. (2020). What does “couple” mean in couple therapy outcome research? A systematic review of implicit and explicit, inclusion and exclusion of gender and sexual minority individuals and identities. *Journal of Marital and Family Therapy*, 46(2), 240–255. <https://doi.org/10.1111/jmft.12415>
- Spielmans, G. I., & Flückiger, C. (2018). Moderators in psychotherapy meta-analysis. *Psychotherapy Research*, 28(3), 333–346. <https://doi.org/10.1080/10503307.2017.1422214>
- Sprenkle, D. H. (2012). Intervention research in couple and family therapy: A methodological and substantive review and an introduction to the special issue. *Journal of Marital and Family Therapy*, 38(1), 3–29. <https://doi.org/10.1111/j.1752-0606.2011.00271.x>
- Tukey, J. W. (1977). *Exploratory data analysis*. Addison Wesley.
- *Walker, J. G., Johnson, S., Manion, I., & Cloutier, P. (1996). Emotionally focused marital intervention for couples with chronically ill children. *Journal of Consulting and Clinical Psychology*, 64(5), 1029–1036. <https://doi.org/10.1037/0022-006X.64.5.1029>
- Wampold, B. E., Mondin, G. W., Moody, M., Stich, M., Benson, K., & Ahn, H. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, “all must have prizes.” *Psychological Bulletin*, 122(3), 203–215. <https://doi.org/10.1037/0033-2909.122.3.203>

- *Weissman, N., Batten, S. V., Rheem, K. D., Wiebe, S. A., Pasillas, R. M., Potts, W., Barone, M., Brown, C. H., & Dixon, L. B. (2017). The effectiveness of emotionally focused couples therapy with veterans with PTSD: A pilot study. *Journal of Couple & Relationship Therapy*, 17(1), 25–41. <https://doi.org/10.1080/15332691.2017.1285261>
- Whisman, M. A. (2007). Marital distress and DSM-IV psychiatric disorders in a population-based national survey. *Journal of Abnormal Psychology*, 116(3), 638–643. <https://doi.org/10.1037/0021-843X.116.3.638>
- *Wiebe, S. A. (2014). *A two-year longitudinal study of emotionally focused therapy for couples: Maintenance and predictors of relationship change* [Doctoral Dissertation, University of Ottawa].
- Wiebe, S. A., & Johnson, S. M. (2016). A review of the research in emotionally focused therapy for couples. *Family Process*, 55(3), 390–407. <https://doi.org/10.1111/famp.12229>
- Wittenborn, A. K., Dolbin-MacNab, M. L., & Keiley, M. K. (2013). Dyadic research in marriage and family therapy: Methodological considerations. *Journal of Marital and Family Therapy*, 39(1), 5–16. <https://doi.org/10.1111/j.1752-0606.2012.00306.x>
- *Wittenborn, A. K., Liu, T., Ridenour, T. A., Lachmar, E. M., Mitchell, E. A., & Seedall, R. B. (2019). 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(3), 395–409. <https://doi.org/10.1111/jmft.12350>
- Wood, N. D., Crane, D. R., Schaalje, B., & Law, D. D. (2005). What works for whom: A meta-analytic review of marital and couples therapy in reference to marital distress. *The American Journal of Family Therapy*, 33(4), 273–287. <https://doi.org/10.1080/01926180590962147>
- Wood, W., & Eagly, A. H. (2009). Advantages of certainty and uncertainty. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 455–472). Sage Publications.
- Wright, J., Sabourmin, S., Mondor, J., McDuff, P., & Mamodhoussen, S. (2007). The clinical representativeness of couple therapy outcome research. *Family Process*, 46(3), 301–316. <https://doi.org/10.1111/j.1545-5300.2007.00213.x>

Received July 7, 2021

Revision received June 23, 2022

Accepted July 8, 2022 ■