Routledge
Taylor & Francis Group

What Works for Whom: A Meta-Analytic Review of Marital and Couples Therapy in Reference to Marital Distress

NATHAN D. WOOD

Utah Valley State College, Orem, Utah, USA

D. RUSSELL CRANE AND G. BRUCE SCHAALJE

Brigham Young University, Provo, Utah, USA

DAVID D. LAW

Utah State University, Uintah Basin, Utah, USA

Due to methodological limitations, past meta-analytic research was not able to identify which treatment was most effective for specific marital distress levels. By converting pre- and post-test scores from marital research into equivalent Dyadic Adjustment Scale scores, it was possible to isolate mild, moderate, and severe levels of marital distress. Results show that Emotionally Focused Therapy is significantly more effective than isolated Behavioral Marital Therapy interventions for the treatment of moderate marital distress. Future directions of outcome research are also discussed.

Over the years, researchers in the field of Marriage and Family Therapy (MFT) have attempted to validate the profession's existence through outcome studies and meta-analyses. The results of these studies did exactly what they were designed to do, they took a wide-angled snapshot of the field and everyone looked good from a distance. All meta-analyses to date have shown positive results when MFT has been compared to no treatment (e.g., Butler & Wampler, 1999; Dunn & Schwebel, 1995; Hahlweg & Markman, 1988; Hazelrigg, Cooper, & Borduin, 1987; Shaddish, 1993, Wampler, 1982).

Because of these meta-analyses, we can claim that what we do matters and that we do make a difference.

These studies then focused on comparing individual treatment models against each other. Much like tightly focusing on a wide-angled photograph, the results gave a blurry, but general impression of the details. Hence the conclusions that have been reached are also general. The most common conclusion that has been reached is that different therapeutic approaches are essentially equal in their effectiveness. The phrase from *Alice in Wonderland* "all have won, and all must have prizes" inspired Luborsky, Singer, and Luborsky (1975) to dub these general findings as the "Dodo bird effect."

One reason for the Dodo bird effect in marital therapy was that all previous meta-analyses were not methodologically designed beforehand to analyze differences between marital therapy treatments. Previous research also made no attempt to account for varying degrees of marital health. Meta-analyses need to isolate marital outcome studies and account for marital distress levels to have the best chance of finding differences in marital treatments.

Dunn and Schwebel (1995) took the first step and only included marital therapy outcome studies in their meta-analysis, but neglected distress level. This enabled them to more accurately look at treatment effectiveness of marital therapy by using an apparently homogeneous sample of studies. They found that no particular theoretical approach was more effective than another, the apparent Dodo bird effect once again. However, Dunn & Schwebel (1995) assumed that all measures of marital distress are created equal, going as far as equating observational measures with pencil and paper measures of marital distress. It is important to note that all other previous meta-analyses have also made this assumption.

The use of multiple measures of marital distress/satisfaction in meta-analyses could have a diluting effect on finding the most effective treatment. To include studies with different measurements of marital satisfaction assumes that all the instruments are measuring the same construct, in the same way. Research by Crane, Allgood, Larson, and Griffin (1990) show that this assumption is erroneous and potentially misleading.

Before the study by Crane et al. (1990) was published, couples were considered distressed if they scored below 100 on the Marital Assessment Test (MAT: Locke & Wallace, 1959), or below 92 (Baucom & Hoffman, 1996) or 97 (Jacobson, Schmaling, & Holtzworth-Munroe, 1987) on the Dyadic Adjustment Scale (DAS: Spanier, 1976). Crane et al. (1990) argued that the MAT score of 100 was comparable to DAS scores of 107, not 92 or 97.

Following Crane et al. (1990), DAS scores of 85–92 (traditionally accepted scores for mild distress) are actually more comparable to MAT scores of 72–81 (accepted scores in the range of moderately distressed couples). If a meta-analyst defined marital distress according to traditionally accepted

cutoff scores, they would be combining studies of moderately distressed couples along with those studies of mildly distressed couples.

Shaddish et al. (1995) stated "[as] long as there is such enormous diversity of measurement characteristics, it is difficult to see how we will be able to generate more coherent analyses of treatment effects over studies" (p. 354). They assert that with a standard measure of marital distress it would be far easier for those who review the MFT literature to tell if "differences between behavior and nonbehavior therapy … would disappear when the same outcomes were measured, or if they would still persist" (p. 354). Given all the productive work to date, the field is still lacking understanding of which therapeutic approach, or theory, may be the most appropriate for differing levels of marital distress.

The first step toward identifying which treatment best fits with a corresponding level of marital distress is to standardize the operationalization of marital distress. Crane, Allgood, Larson, and Griffin (1990) made the initial steps toward standardizing the main measures of marital satisfaction used in marital therapy outcome research. In a sample of 302 individuals, multiple regression analysis was used to estimate the linear relationship between the DAS, MAT, and the Revised Marital Adjustment Test (RMAT; Kimmel & Van Der Veen, 1974). This study enabled clinician and researchers to convert scores back and forth between measures.

Later, Crane, Middleton, and Bean (2000) added the Kansas Marital Satisfaction Scale (KMSS; Schumm, Nichols, Schectman, & Grisby, 1993; Schumm, Paff-Bergen, Hatch, Obiorah et al., 1986) and the Revised Dyadic Adjustment Scale (RDAS; Busby, Crane, Larson & Christensen, 1995) to the mix. It is now possible to convert scores between the DAS, MAT, RMAT, KMSS, and RDAS.

Now marital assessment scores included in meta-analyses can be converted into a single metric from which differing levels of distress can be determined. The DAS provides the solution to the meta-analytic problem of a standardized measure of marital satisfaction (Shadish et al., 1995). It is now possible to explore treatment effectiveness by varying levels of marital distress in a more meaningful way utilizing a meta-analytic format.

METHOD

Inclusionary Criteria

Studies were included in the current meta-analysis if the following criteria were met: (a) The study focused specifically on conjoint treatment of marital distress. Studies that included conjoint treatment of other disorders were excluded from the current study. (b) Treatment involved both spouses. (c) Measures used in the studies were one of the following: DAS, MAT, RMAT, KMSS, or the RDAS. (d) The studies were published between 1963 and 2002.

Procedure

The initial search for relevant literature was in electronic databases such as PSYCHINFO, ERIC, SOCIOLOGICAL ABSTRACTS, Social Science Citation Index (SSCI), and Dissertation Abstracts International. Key words used as search terms in finding relevant articles included but were not limited to: marital distress, outcome study, marital therapy, couples therapy, conjoint therapy, marital satisfaction, and marital conflict. Once relevant literature was identified, reference lists and bibliographies were examined to identify more literature. Review articles and meta-analyses were also searched for literature that fit the inclusionary criteria.

Articles were coded by advanced undergraduate and graduate students who had a thorough understanding of statistical and methodological issues (contact the first author for copy of the code book). Groups of 2–3 students were initially given the same articles for coding to establish inter-rater reliability. Consistency in the coding was validated by a visual comparison between coding sheets. This process continued until each code sheet was identical to the others. Once inter-rater reliability was established, each coder received unique articles.

Non-primary literature such as unpublished dissertations were included because they generate a more realistic picture of treatment effectiveness (Lipsey & Wilson, 2001; Rosenthal, 1998; Wampler & Serovich, 1996). Of the 23 studies coded for this meta-analysis, 20 peer-reviewed articles were coded along with 3 dissertations (See Table 1 for articles included in the analysis). From the 23 studies included, 41 treatment groups were identified. Seven treatment groups fell within the mildly distressed range, 33 treatment groups were classified as moderately distressed, and only one group fell into the severely distressed range.

Couples' scores were coded and used in the analysis rather than husbands' or wives' individual scores. Treatment approach was coded in two ways. The first was to code the approach as it was labeled in the paper itself. The second approach following Dunn & Schwebel (1995) was to affiliate the main thrust of the treatment with previously established models such as Behavioral Marital Therapy (BMT), Emotionally Focused Therapy (EFT), and others. The approaches that did not fit solidly into a single category, such as behavioral marital therapy with emotional validation interventions added, were coded as "Mixed" (e.g., Halford, Sanders, & Behrens, 1993). Other studies contained only individual interventions of BMT, such as problem solving only, or communication training only (e.g., Jacobson, 1984). These individual interventions from BMT were coded as "BMT Components."

Studies were coded into two groups, true experimental and quazi-experimental designs. Following previous meta-analytic research, true experimental designs had random assignment to group, presence of a control group, and a clear treatment protocol (Dunn & Schwebel, 1995; Carrol &

TABLE 1 Studies Included in the Meta-Analysis

Study	Models included
Azrin et al. (1980)	BMT
Baucom (1982)	BMT Components, BMT
Baucom (1986)	BMT, Others
Baucom, Sayers, & Sher (1990)	BMT, Mixed
Beach & Broderick (1983)	BMT
Bennun (1985)	BMT, Others
Bornstein et al. (1983)	BMT
Carlton (1978)*	EFT
Follingstad & Haynes (1981)	BMT
Halford, Sanders, & Behrens (1993)	BMT, Mixed
Hazelrigg, Cooper, & Borduin (1987)	EFT
Huber & Milstein (1985)	BMT, Others
Jacobson (1978)	BMT, Mixed
Jacobson (1984)	BMT Components
James (1991)	Mixed
Johnson & Greenberg (1985a)	BMT, EFT
Johnson & Greenberg (1985b)	EFT
Goldman & Greenberg (1992)	EFT
Hudson (1978)*	BMT
Liberman et al. (1976)	BMT
Margolin & Weiss (1978)	BMT Components
Simms (1999)*	EFT
Wilson, Bornstein, & Wilson (1988)	BMT, Others

*Studies were unpublished dissertations.

Doherty, in press; Shaddish et al., 1993). Quazi-experimental designs have similar features to true experimental designs with the exception of random assignment to group.

Special care was taken to ensure that each study was independent of all other studies. The same sample may have been used for several different studies and therefore results from these studies could not be considered independent. If dependent studies were identified, only one study of the set was used. Study independence is an important issue as resultant effect sizes are to be used for traditional significance testing which holds data independence as a primary assumption.

Preliminary Analysis

Initially, all the MAT scores in the coded literature were converted to DAS scores (see Crane et al., 1990). No studies were found that used the RMAT, KMSS, or RDAS. Standard deviations were converted from MAT (sMAT) to DAS (sDAS) by the following formula:

$$sDAS = sMAT * .605$$

The conversion of scores and standard deviations was possible because of the linear relationship between MAT and DAS values (Crane et al., 1990).

Additional computations were performed on those studies for which only husband and wife data were coded. For those with husband and wife scores, the two were added and divided by two. For missing data, couple's standard deviations were calculated by using the following formula.

$$\text{Couple } sd = \sqrt{\frac{1}{4}\left(s_b^2 + s_w^2 + 2rs_bs_w\right)}$$

The formula takes into account the correlation between husband and wife's scores and was derived from statistical theory. The higher the correlation coefficient used, the larger the estimate of the couple's standard deviation. As standard deviations increase the resultant effect sizes calculated decrease, or become more conservative.

Various studies have reported the correlation of husband's and wife's scores to fall between .5 and .73 (Crane, Soderquist, & Frank, 1995; Oppenheim, Wambboldt, Gavin, Renouf, & Emde, 1996). When the correlation is placed in the aforementioned formula, smaller correlations will generate more liberal effect sizes while larger correlations are associated with more conservative effect sizes. A correlation of .70 was chosen from the empirically supported range of correlations in order to generate conservative estimates.

Where studies did not report husband, wife, or couple standard deviations, standard deviations were substituted by a standard deviation from a similar sample in another study that used the same assessment instrument (Lipsey & Wilson, 2001).

Standardized mean gain (ES_{sg}) and mean difference (ES_{sm}) effect sizes were calculated and a Q-test was then performed on the data. Q-tests are based on a Chi-square distribution with k−1 degrees of freedom. Significant values imply a heterogeneous distribution. The Q-test indicated that the both forms of standardized effect sizes represented a heterogeneous population of studies (p < .05).

Given that some couple data and standard deviations had been calculated (i.e., converted from MAT to DAS scores) rather than directly measured, it was hypothesized that there might be a systematic difference between directly measured data versus calculated data causing the heterogeneity. Measurement error could be one explanation of the difference between DAS scores and converted DAS scores. Measurement error is the difference between a subject's "true score," or the score they would receive under perfect conditions, and their actual score. Additional error was possibly introduced when a score with existing measurement error was converted from the MAT scale to the DAS scale.

With these two potential sources of error in converted DAS scores, additional heterogeneity tests were performed on effect sizes generated from original DAS scores verses effect sizes from converted DAS scores. Heterogeneity was still present in both groups suggesting that the heterogeneity present was not from converting scores to the DAS scale. While the conversion of scores to the DAS scale did not explain the heterogeneity, there is still the potential for error to influence the final results.

Possible error in the converted scores was controlled for statistically by including a dummy variable in regression analysis. The dummy variable in this case was the assessment instrument administered in the study, (0 = DAS, 1 = MAT). The presence of this variable ("Assessment") in a regression model eliminated the possibility of any difference found between treatment models to be explained by the conversion of scores.

Where there were no apparent fixed effects, or systematic differences between the studies that contributed to the sample heterogeneity, it could be safely assumed that random effects were present. Random effects are unknown variables that can contribute to the overall variability in effect sizes beyond sampling error (Lipsey & Wilson, 2001).

Therefore, a mixed model weighted linear regression was utilized in this study. The mixed model aspect took into account the random effects and adjusted the effect sizes accordingly. This was in addition to weighting each effect size by the inverse variance to take into account the sample size of each individual effect size.

RESULTS

Of the treatment groups found in the initial analysis, 17% (n = 7) scored in the mildly distressed range (DAS 96–107). Eighty percent (n = 33) of the groups included in this meta-analysis were moderately distressed groups whose pre-test DAS scores ranged from 80–95.9. Only one group fell below the DAS cutoff for severe marital distress (DAS < 80). Due to the low sample size of the mildly and severely distressed groups, the main focus of the results and discussion will be on moderately distressed couples.

Standardized effect sizes do not lend themselves to intuitive interpretation. They refer directly to the pooled standard deviation of reference group (e.g., control group in the case of mean differences and pre-test/post-test scores in reference to treatment gains). For example, a mean difference effect size of .50 means the treatment group performed one half of a pooled standard deviation better than the control group. A mean difference effect size of one means the treatment group performed one pooled standard deviation above controls.

Mean Gain

Mean gain effect sizes (ES_{sg}) refer to standardized pre-post test differences of each treatment group. Moderate distress at pre-test translates into DAS scores between 80 and 95.99. Each treatment modality included in this analysis had at least four effect sizes available for analysis. The Standardized Mean Gain ES_{sg} for moderately distressed groups was used at the dependent measure in Table 2.

Table 2 shows a regression comparing all theoretical approaches to BMT Components after controlling for the assessment instrument used at pretest. The variable "constant" represents BMT Components. The choice of the comparison group was arbitrary, any theoretical model could have been the comparison group. Coefficients for the theoretical models represent the magnitude of the ES_{sg} that needs to be added to, or subtracted from, the ES_{sg} of BMT Components.

Emotionally Focused Therapy had the highest ES_{sg} while BMT Components had the lowest ES_{sg}. These data did not show whether or not each approach was effective. It shows that there was not any specific approach that was significantly more predictive of treatment gains than BMT Components. This was also accounting for the influence of the other treatments, and the marital assessment used at pretest.

"Other" approaches came close to being significantly different from BMT Components ($p = .07$) requiring that an additional .69 be added to the ES_{sg} of BMT Components in order for BMT Components to be equal with "Other approaches." The assessment instrument used at pretest significantly predicted ES_{sg} scores. If the MAT was used in a study, the effect size of *any* treatment model would need to be reduced by .71 ($p < 0.01$). The significance of the "Assessment" variable may be a reflection of measurement error introduced

TABLE 2 Summary of Random Effects Regression Analysis for Treatment of Moderate Marital Distress, Mean Gains (DAS = 80–95.9, n = 33)

Variable	B	SE B	Beta
Constant (BMT Comp)	1.3626**	0.2852	0.0000
BMT	0.4284	0.2849	0.2750
EFT	0.5476	0.3798	0.2647
Mixed	0.0936	0.3744	0.0465
Other	0.6958	0.3848	0.2847
Assessment	−.7140**	0.2551	−0.4674

Constant = Isolated components of BMT; BMT = Behavioral Marital Therapy; EFT = Emotionally Focused Therapy; Mixed = Approaches that combined elements of EFT and BMT; Others = Cognitive Behavioral Therapy, "Systemic Therapy"; Assessment = Categorical variable coded 0 = DAS, 1 = MAT; Dependent Variable = ES_{sg}; Method of Moments Random Effects Variance Component, v = .2344; Q = 17.0665, p < .01 (Overall Model); R^2 = .2983; *p < .05; **p < .01.

TABLE 3 Summary of Random Effects Regression Analysis for Treatment of Moderate Marital Distress, Mean Difference (DAS = 80–95.9, n = 33)

Variable	B	SE B	Beta
Constant (EFT)	1.0079	0.5477	0.0000
BMT	−0.6577	0.4836	−0.3587
BMT Comp	−1.1513*	0.5549	−0.5410
Mixed	−0.8884	0.5092	−0.3848
Other	−0.4048	0.7703	−0.0958
Protocol	1.0667	0.5827	0.3221
Assessment	−0.7691*	0.3826	−0.4103

Constant = Emotionally Focused Therapy; BMT = Behavioral Marital Therapy; Mixed = Approaches that combined elements of EFT and BMT; BMT Comp = Isolated components of BMT; Others = Cognitive Behavioral Therapy, "Systemic Therapy"; Protocol = Categorical variable to indicate a clear treatment protocol; Assessment = Categorical variable coded 0 = DAS, 1 = MAT; Dependent Variable = ES_{sm}; Method of Moments Random Effects Variance Component, v = .25747; Q = 18.6671, p < .01 (Overall Model) R^2 = .4922; *p < .05, **p < .01.

by converting MAT scores into DAS scores. It may also be indicative of the assessment instrument used in the studies.

Mean Difference

Mean difference effect sizes (ES_{sm}) are the effect sizes commonly reported in other meta-analyses (i.e., Cohen's *d*). Table 3 shows the mixed model regression results for treatment of moderate marital distress with ES_{sm} as the dependent variable. The overall model significantly predicted 49% of the variance in treatment vs. control comparisons (Q = 18.67, p < .05). The constant in Table 3 represents EFT. It is important to note the presence of the "protocol" variable (p = .067). This was a dichotomous variable indicating whether or not there was a clear treatment protocol (e.g., treatment manual) that was followed in the study. The coefficients returned for "Assessment" and "Protocol" remained the same no matter the comparison group. Therefore it can be said that when clear protocols are followed, an additional 1.067 can be added to the effect size of any treatment used.

The assessment instrument used at pretest ("Assessment") was also significantly predictive of ES_{sm} (p < 0.05). This indicates that if the MAT was used in the study, the effect size of *any* theoretical approach needs to be adjusted by −.77. This may or may not be a statistical anomaly because of the conversion of MAT scores into DAS scores. In depth explanations behind the significance of the "Assessment" variable are present in the discussion section.

Once the variables "protocol" and "test" have been controlled for, EFT performed significantly better than BMT Components in reference to their

respective control groups. All things being equal, effects sizes from BMT Components need to be reduced by 1.15 in comparison to EFT treatments ($p < 0.05$). Mixed models approached being significantly different (i.e., worse) than EFT ($p = 0.081$).

Analysis of mean gain and mean difference effect sizes for mild distress show no significant differences between treatment approaches. No statistical procedures were able to be performed with severely distressed couples as there was only one group.

DISCUSSION

Limitations

One of the fundamental weaknesses in this study was the small number of studies available for consideration. The relatively low number of groups included in the analysis lowers the statistical power of the current results and increases the odds of saying there are no differences in treatments when in fact there may be. Saying that there are no differences in treatment in this situation is not the same as claiming the Dodo bird effect. In fact, in some instances, the results of this study throw doubt onto the Dodo bird effect.

The best example of this was in the analysis of mean gains (ES_{sg}) in the treatment of moderate distress. Table 2 shows obvious differences between full theoretical approaches and isolated BMT Components. The lack of additional statistically significant differences in treatment models may be a Type II error showing non-significant results when there may actually be a difference present.

More replications need to be performed on EFT, Mixed, CMT, and CBMT before any differences can be determined in terms of treatment gains. The low sample size is due to two main reasons. The first was the narrow focus of this study. The second major reason of smaller sample size is the lack of established outcome measures in the field.

The relative lack of replication, besides that of BMT and EFT, may be more of a comment on the state of academia and what is considered publishable material. Replications rarely show anything new and therefore may not be as readily published. Beyond these systemic barriers, clinical outcome research is expensive and very time consuming which also add to the relative lack of pure replication. For the field of Marriage and Family Therapy to thrive, there must be more sophistication and standardization in the way we measure and treat marital and family issues.

The best example of an approach that was excluded from this analysis is Insight Oriented Marital Therapy (IOMT; Snyder & Wills, 1989). Studies of IOMT were excluded due to the fact that the dependent measures used in their study could not be converted to DAS scores. It is wise to explore the impact of treatment on different aspects of relational and individual functioning.

However, to exclude established measurements of marital distress makes it difficult to gain a better understanding of treatment through meta-analysis. Ideally, the same measure needs to be used across studies. Statistical conversions are possible, as shown by this study. However, conversions between scores at the scale of meta-analysis can add unwanted variability that may not be necessarily present if all studies used the same instrument.

The work of Crane et al. (1990) started the process of attempting to empirically show where the cutoffs were for all three distress levels rather than statistical speculation. If the field is to use a common metric such as the DAS, or another instrument, clear criteria must be set forth to distinguish between distressed and non-distressed populations.

The best example of the importance of setting criteria came in the relative absence of severely distressed couples. This may not be as much the fault of researchers as it is the population. Attrition rates may be higher with severely distressed couples. Their motives in coming to therapy may also be more disguised. For example, one partner may already be emotionally divorced from the relationship, affairs, or decreased commitment to the marriage may all effect attrition and what is considered a "successful outcome."

An additional confounding factor with severely distress couples is how the field of marriage and family therapy defines successful outcomes, or clinical significance. In instances of physical, emotional, and/or sexual abuse, a successful outcome may be the termination of the relationship. Unfortunately, these types of outcomes are difficult to measure, especially if the abuse was not disclosed in the first place.

More research needs to be done to identify clear cutoffs for mild, moderate, and severe distress. The results from this study clearly show that once distress level is identified, the Dodo bird hypothesis can be tentatively put on the endangered species list. It is in no way extinct yet, as more research must be done. Specifically, more replication needs to be performed with distress level included in the design.

Are All Assessments Created Equal?

The argument for a single measurement of marital distress is buoyed by the statistically significant presence of "Assessment" variable included in the analyses. This variable was included to control for any measurement or statistical errors in converting scores from MAT to DAS. The fact that it was statistically significant in every regression model for moderately distressed couples raises questions.

It simply may be a reflection of error introduced by the conversion of scores and nothing more. Another explanation may relate to the year of the study used. The DAS has been used much more in recent marital research than the MAT. It could be argued that recent research is more

methodologically stringent and would reflect higher effect sizes on average that past research that used the MAT.

Yet, the impact of the "Assessment" variable in many analyses required the effect size of some theoretical models be reduced by half. This would imply that the DAS is more predictive of higher effect sizes than the MAT. Following this conclusion would call into question previous meta-analytic work that included effect sizes based on DAS scores because the overall average effect size generated from the meta-analysis would be higher than otherwise would be expected if there were no differences between measures.

Another assessment implication of the current study is who is considered to be mildly distressed. The mildly distressed group in this the study would be considered to be an "enrichment" population as their pre-test scores are close to (at least a 96 on the DAS) or greater than the traditionally accepted cutoff score (i.e., DAS = 97). Yet there was no indication from the authors of those studies that they were studying an enrichment population.

Clinical Implications

Clinical conclusions for mildly distressed couples are very tentative due to the low sample size. The data show that any treatment would be helpful to some extent. More research needs to be done once a clear definition of what constitutes a mildly distressed couple.

Treatment for moderately distressed couples should have a clear treatment plan that is followed. While EFT is showing significant results with moderately distressed couples, other full treatment models were also significant. Treatment that delivered only isolated components of BMT had treatment gains substantially lower than other full model treatments. "Other" treatment models (CMT, CBMT, and BMT Group) also approached being significantly different from BMT Components ($p = 0.07$). This was after controlling for the assessment instrument used in the studies. More specifically it can be stated that holding all other things equal, "Other" treatment approaches were associated with an additional 9-point jump on DAS scores from pre-test to post-test beyond the treatment gains of BMT Components.

The difference between EFT and BMT Components in reference to control groups is intuitive as EFT offers a comprehensive treatment plan where the BMT Component studies were focused on a single technique or intervention. A group of interventions tied together have more potential for benefit than a single intervention. This would be especially true for a moderately distressed population. It is difficult to say whether including more studies in the moderately distressed groups would confirm or disprove the differences between EFT and the Mixed approaches.

The evidence of a clear treatment protocol approaching significance in the treatment of moderate marital distress is more intuitive. The results

suggest that all other things being equal, a clear plan of treatment and follow through with that plan, is associated with a 17-point DAS difference between treatment and control groups at post-test.

FINAL THOUGHTS

This study was unique in many ways. It was the first to standardize outcome measures across studies. The presence of a standard dependent measure across studies made it possible to compare treatment gains of the treatment groups across multiple studies. Treatment gains are more reflective of treatment effectiveness as they focus solely on the magnitude of change of the treatment group. Until the field as a whole adopts a standard dependent measure, these results represent the best attempt at comparing realized treatment gains.

This study was also first in attempting to measure treatment effectiveness by differing levels of marital distress. The data from this study more conclusively showed that mild, moderate, and severely distressed couples should not be aggregated. Differences in treatment approaches appeared once distress levels were operationalized and isolated. Through these data, the initial attempts to match presenting complaint to treatment have been initiated.

In reference to what works for whom, any intervention for mildly distressed couples is better than no intervention. No one intervention stood out against any other for mild distress. Moderately distressed couples should receive a full treatment model rather than isolated components or interventions. Emotionally Focused Therapy stood out in terms of treatment versus control comparisons when compared to isolated BMT Components. "Other" approaches stood out in terms of treatment gains when compared to isolated BMT Components. Only BMT was used to treat a severely distressed group, but to great success.

REFERENCES

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*, 289–308.

Butler, M. H., & Wampler, K. S. (1999). A meta-analytical update of research on the couple communication program. *American Journal of Family Therapy, 27*, 223–237.

Carrol, J. S., & Doherty, W. J. (2003). Evaluating the Effectiveness of Premarital Prevention Programs: A Meta-Analytic Review of Outcome Research. *Family Relations: Interdisciplinary Journal of Applied Family Studies, 52*, 105–118.

Crane, D. R., Allgood, S. M., Larson, J. H., & Griffin, W. (1990). Assessing marital quality with distressed and nondistressed couples: A comparison and equivalency

table for three frequently used measures. *Journal of Marriage and the Family, 52*, 87–93.

Crane, D. R., Middleton, K. & Bean, R. (2000). Establishing criterion scores for the Kansas Marital Satisfaction Scale and the Revised Dyadic Adjustment Scale. *American Journal of Family Therapy, 28*, 53–60.

Crane, D. R., Soderquist, J. N., & Frank, R. L. (1995). Predicting divorce at marital therapy intake: A preliminary model. *American Journal of Family Therapy, 23*(3), 227–236.

Dunn, R. L., & Schwebel, A. I. (1995). Meta-analytic review of marital therapy outcome research. *Journal of Family Psychology, 9*, 58–68.

Gurman, A. S., & Kniskern, D. P. (1978). Research on marital and family therapy: Program, perspective, and prospect. In S. Garfield & A. Bergin (Eds.), *Handbook of psychotherapy and behavior change; An empirical analysis* (pp. 820–821). New York: Wiley.

Hahlweg K., & Markman, H. J. (1988). Effectiveness of behavioral marital therapy: Empirical status of behavioral techniques in preventing and alleviating marital distress. *Journal of Consulting and Clinical Psychology, 56*, 440–447.

Hazelrigg, M. D., Cooper, H. M., & Borduin, C. M. (1987). Evaluating the effectiveness of family therapies; An integrative review and analysis. *Psychological Bulletin, 101*, 388–395.

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy, 15*, 336–352.

Jacobson, N. S., Follette, W. C., Revenstorf, D., Boucom, D. H., Hahlweg, K., & Margolin, G. (1984). Variability in outcome and clinical significance of behavioral marital therapy: A reanalysis of outcome data. *Journal of Consulting and Clinical Psychology, 52*, 497–504.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12–19.

Lipsey, M. W., & Wilson, D. B. (2001). Practical meta-analysis. Thousand Oaks, CA: Sage.

Locke, H. J., & Wallace, K. M. (1959). Short marital adjustment and prediction tests: Their reliability and validity. *Marriage and Family Living, 21*, 251–255.

Luborksy, L., Singer, B., & Luborksy, L. (1975). Comparative studies of psychotherapies: Is it true that "Everybody has won and all must have prizes?" *Clinical Psychology: Science and Practice, 2*, 106–109.

Oppenheim, D., Wambboldt, F., Gavin, L. A., Renouf, A. G., & Emde, R. N. (1996). Couples' co-construction of the story of their child's birth: Associations with marital adaptation. *Journal of Narrative and Life History, 6*, 1–21.

Schumm, W. R., Nichols, C. W., Schectman, K. L., & Grigsby, C. C. (1983). Characteristics of responses to the Kansas Marital Satisfaction Scale by a sample of 84 married mothers. *Psychological Reports, 53*, 567–572.

Schumm, W. R., Paff-Bergen, L. A., Hatch, R. C., Obiorah, F. C. et al. (1986). Concurrent and discriminant validity of the Kansas Marital Satisfaction Scale. *Journal of Marriage & the Family, 48*, 381–387.

Shaddish, W. R. (2002). The effects of marriage and family interventions: What can we learn from meta-analyses? In D. Sprenkle (Ed.), *Effectiveness research in marriage and family therapy*. Washington, D.C.: American Association for Marriage and Family Therapy.

Shaddish, W. R. (1992). Do family and marital psychotherapies change what people do? A meta-analysis of behavioral outcomes. In T. D. Cook, H. M. Cooper, D. S. Cordray, H. Hartmann, L. V. Hedges, R. J. Light, T. A. Louis, & F. Mosteller (Eds.), *Meta-analysis for explanation: A casebook* (pp. 129–208). New York: Sage.

Shaddish, W. R., Montgomery, L. M., Wilson, P., Wilson, M. R., Bright, I., & Okwumabua, T. (1993). Effects of family and marital psychotherapies: A meta-analysis. *Journal of Consulting and Clinical Psychology, 61*, 992–1002.

Shadish, W. R., Ragsdale, K., Glaser, R. R. & Montgomery, L. M. (1995). The efficacy and effectiveness of marital and family therapy: A perspective from meta-analysis. *Journal of Marital and Family Therapy, 21*, 345–360.

Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and the Family, 38*, 15–28.

Wampler, K. S. (1982). The effectiveness of the Minnesota Couple Communication Program: A review of research. *Journal of Marital and Family Therapy, 8*, 345–355.

Wampler, K. S., & Serovich, J. M. (1996). Meta-analysis in family therapy research. In D. H. Sprenkle, & S. M. Moon (Eds.), *Research methods in family therapy* (pp. 286–307). New York: Guilford.