



Published in final edited form as:

Behav Ther. 2012 March ; 43(1): 77–87. doi:10.1016/j.beth.2010.12.006.

Exploring Programmatic Moderators of the Effectiveness of Marriage and Relationship Education Programs: A Meta-Analytic Study

Alan J. Hawkins,
Brigham Young University

Scott M. Stanley,
University of Denver

Victoria L. Blanchard, and
State University of New York at Albany

Michael Albright
Brigham Young University

Abstract

This study uses meta-analytic methods to explore programmatic moderators or common factors of the effectiveness of marriage and relationship education (MRE) programs. We coded 148 evaluation reports for potential programmatic factors that were associated with stronger intervention effects, although the range of factors we could code was limited by the lack of details in the reports. Overall, we found a positive effect for program dosage: moderate-dosage programs (9–20 contact hours) were associated with stronger effects compared to low-dosage programs (1–8 contact hours). A programmatic emphasis on communication skills was associated with stronger effects on couple communication outcomes, but this difference did not reach statistical significance for the relationship quality/satisfaction outcome. There was no evidence that institutionalized MRE programs (formal manuals, ongoing presence, formal instructor training, multiple evaluations) were associated with stronger effects. Similarly, there was little evidence of differences in program setting (university/laboratory vs. religious). We discuss possible explanations for these findings and implications for program design and evaluation.

Keywords

common factors; marriage and relationship education; meta-analysis

Systematic exploration of common factors in marital and family therapy effectiveness has been gaining momentum (Sprenkle, Davis, & Lebow, 2009). But this approach to improving preventative efforts in marriage and relationship education (MRE) has lagged behind. Meta-analytic work over the past few years generally has established the efficacy of MRE (Blanchard, Hawkins, Baldwin, & Fawcett, 2009; Fawcett, Hawkins, Blanchard, & Carroll,

© 2010 Published by Elsevier Ltd on behalf of Association for Behavioral and Cognitive Therapies.

Address correspondence regarding this article to the first author: 2050 JFSB, Brigham Young University, Provo, UT 84602-6723; hawkinsa@byu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

2010; Hawkins, Blanchard, Baldwin, & Fawcett, 2008; Hawkins & Fackrell, 2010). And while these studies have explored various methodological moderators of program effects, they have not given in-depth attention to potential programmatic moderators of MRE outcomes. With a rigorous analysis of programmatic moderators or common factors, we hope to provide MRE practitioners greater guidance for improving programs and highlight needs for improvement.

Accordingly, in this meta-analytic study we focus on potential programmatic moderators of MRE effectiveness. A handful of MRE evaluation studies were designed to test the effects of specific programmatic elements on program outcomes (e.g., self-guided vs. classroom format, see Duncan, Steed, & Needham, 2009). And some scholars have provided conceptual analyses of programmatic features that should produce stronger effects (Halford, Markman, Kline, & Stanley, 2003). While these efforts are valuable, a thorough analysis of common factors in MRE is advanced by looking at the full body of evaluation work using meta-analytic methods.

Increasingly, meta-analysis has been used to search for programmatic moderators of family life education intervention effects (Lundhal, Risser, & Lovejoy, 2006; MacLeod & Nelson, 2000; Nowak & Heinrichs, 2008; Piquart & Teubert, 2010a; Piquart & Teubert, 2010b). A prime example of this approach is the study by Kaminski and her colleagues (Kaminski, Valle, Filene, & Boyle, 2008) of programmatic factors associated with parent training program effectiveness as measured by positive parenting behaviors and children's externalizing behaviors. These researchers analyzed 77 evaluation studies to find that larger effects were associated with, for instance, an emphasis on increasing positive parent-child interactions and emotional communication skills. Weaker effects were associated with a program emphasis on promotion of children's cognitive or social skills.

These meta-analytic studies of programmatic factors in family life interventions provide a few clues about potential common factors to explore in MRE. We also used the Comprehensive Framework for Marriage Education—a set of concepts to help relationship educators think more thoroughly, systematically, and creatively about their craft—to identify other potential common factors (Hawkins, Carroll, Doherty, & Willoughby, 2004). Higher program dosage (contact hours) predicts stronger effects in interventions for new parents (Piquart & Teubert, 2010a, 2010b), studies of premarital education (Stanley, Amato, Johnson, & Markman, 2006), and programs for the prevention of child maltreatment (MacLeod & Nelson, 2000). An earlier meta-analysis of MRE programs also found that moderate intensity programs yielded stronger effects than low-intensity programs (Hawkins et al., 2008). The current study updates the earlier one with more recent studies and a more in-depth analysis. MRE program length ranges from 1 to 120 hours, with a median of about 12 hours. We hypothesized that stronger doses would be associated with larger effects in our study. Also, different program content emphases have been associated with differential effects (Kaminski et al., 2008). Communication skills training is emphasized in many MRE programs due to the influential cognitive-behavioral line of research showing the importance of communication skills to subsequent relationship quality and stability (Gottman & Silver, 1994; Markman, Rhoades, Stanley, Ragan, & Whitton, 2010). Many other programs, however, emphasize alignment of couple expectations and specific knowledge about marriage and healthy relationships (e.g., Schulz, Cowan, & Cowan, 2006). Still other programs emphasize improving relationship virtues such as forgiveness and empathy (e.g., Ripley & Worthington, 2002). We hypothesized that an emphasis on communication skills would be associated with stronger effects in our study. We were able to code nearly all of the evaluation studies for dosage and content emphasis.

In addition, we examined the “institutional status” of a program. That is, we hypothesized that institutionalized MRE programs—programs that use formal manuals, require formal instructor training, have an ongoing presence in the field, and have had multiple evaluations—would be associated with stronger effects in our study. In the MRE field, there are a number of well-known, institutionalized programs (e.g., Prevention and Relationship Enhancement Program, Markman, Stanley, & Blumberg, 2010) but many evaluation studies report on lesser-known programs that may not have an ongoing presence in the field and have only been evaluated once. Also, we were able to code for the setting of the program. The most common program setting in evaluation studies is a university “laboratory,” classroom, or mental health clinic associated with a university. But the most common field setting for MRE is a church. We hypothesized that university/laboratory settings would be associated with stronger effects because the program instructors would likely have greater training than instructors in religious or other community settings. Pinquart and Teubert (2010a, 2010b) found that new-parent intervention programs were more effective when they used more highly trained instructors. However, we note that one study (Stanley et al., 2001) directly compared MRE instruction with trained mental health professionals in a university setting to instruction by trained religious leaders in a religious setting and found roughly equivalent results.

Unfortunately, our list of potential common factors that could be coded is short because study reports often provided minimal programmatic details. For instance, one of the strongest common factors found in MFT intervention is the therapeutic alliance, including the emotional bond between client and therapist (Sprenkle, Davis, & Lebow, 2009). Only a few MRE evaluation studies ask program participants to report on their bond with the instructor(s) in evaluation questions such as, “How much did you like the instructor and feel a connection with him/her?” With only a small percentage of studies reporting on the “pedagogic alliance,” meta-analysis cannot yield a helpful perspective. Similarly, the impact of providing more supportive components, such as mentoring or other services, has not been investigated much. MacLeod and Nelson (2000) did find that child maltreatment prevention programs with social support components were more effective, but again, it is still rare for evaluated MRE programs to use mentoring as a part of the intervention. Consequently, we were unable to provide a fair test of some important potential moderators. Our study, then, is able to address only a handful of potential common factors. Nevertheless, we believe this initial effort to explore programmatic common factors in MRE yields interesting early clues and can be a model for further work.

Method

In this meta-analytic study, we coded 148 reports to begin to search for programmatic moderators or common factors of the effectiveness of MRE programs. Some of these reports examined more than one intervention condition (against a control group or against a different treatment condition), generating multiple studies within a report (Lipsey & Wilson, 2001). The two most common generic outcomes evaluated in these studies were relationship quality or satisfaction and some form of couple communication; other measured outcomes were infrequent. Given the theoretical importance of these two outcomes and their commonality across studies, they are the focus of this study. We coded a wide array of communication constructs but in analyses not reported here (Blanchard, 2008) we found that more fine-grained analyses of specific constructs yielded little new information; thus, in this study we aggregated all communication outcomes into a single global construct which allowed for more statistical power to investigate moderator effects. (When coding for the large array of communication outcomes, we took care in coding whether an effect was positive, such as a decrease in negative problem-solving strategies.) Most studies used a common, standardized assessment of relationship quality such as the Dyadic Adjustment

Scale or the Marital Adjustment Test. From the 148 reports, there were 166 studies available to examine for relationship quality and 168 studies available to examine for couple communication.

Literature Search

We identified studies from 1975, when serious MRE evaluation research began, through 2009 by reviewing the reference lists of previous MRE meta-analyses (e.g., Butler & Wampler, 1999; Carroll & Doherty, 2003; Giblin, Sprinkle, & Sheehan, 1985; Hight, 2000) and the actual search list from a recent MRE meta-analysis (Reardon-Anderson, Stagner, Macomber, & Murray, 2005). We also conducted searches with PsychInfo and Dissertation Abstracts International to find more recent work and to search for studies missed by earlier meta-analytic work. Also, we contacted many researchers and practitioners over a three-year period to find unpublished reports. Although we did not do a specific search for studies in non-English languages, we did come across five such studies in our search. We employed translators to help us with the coding of these studies.

Selection and Inclusion Criteria

Psychoeducational couple intervention—All studies assessed the effects of a psychoeducational intervention designed to improve couple relationship quality and/or communication skills. We included studies of couple interventions. A few studies evaluated relationship literacy programs targeted to youth or young adults rather than couples; we excluded these studies from our analyses to keep the focus on couples. Therapeutic interventions were excluded in order to provide a clear picture of the effects of educational intervention. However, we note that a few studies reported that their samples included a significant proportion of distressed couples as well as couples seeking preventive services (e.g., Cummings, Faircloth, Mitchell, Cummings, & Schermerhorn, 2008).

Reporting of outcome data—Included studies had to report effects using quantitative methods that could produce an effect size. Some quantitative studies did not report some data necessary to calculate an effect size. We succeeded in “rehabilitating” a limited number of these by following recommendations outlined in Lipsey and Wilson (2001). Yet, six published and five unpublished studies were excluded from analyses because rehabilitation efforts or contacting authors for more information did not yield sufficient data to code effect sizes.

Study design—We included both experimental and quasi-experimental studies. Some studies included a comparison group that received a minimal “intervention” such as a small brochure, but intervention groups beyond this level of treatment were coded as separate studies. In some cases, then, what may have been designed and reported as one experimental study was coded as multiple one-group pre/post studies (e.g., Halford, Sanders, & Behrens, 2001).

In addition, we included a large group of studies using one-group pre/post designs in our investigation as a supplement to analyses of controlled studies. These studies also may yield clues to programmatic common factors of MRE effectiveness even though they are subject to more internal validity concerns. We analyzed these studies separately from control-group studies, both because the effect size statistic is computed differently and to see if they replicated the pattern of findings in control-group studies. Our interest in this study is not so much the exact magnitude of an effect size, which has been examined in other recent meta-analytic studies (see introduction) but whether certain programmatic factors produce stronger or weaker effects. These one-group pre/post studies generally were conducted with otherwise sound methods, so ignoring them would have excluded an important body of

evaluation work and potentially limited our understanding of common factors. Included in the one-group pre/post studies are a number of reports that compared one MRE intervention to another with two independent samples. In these cases, we coded each treatment separately as a one-group/pre-post study.

Publication status—We included both published and unpublished studies to control for potential publication bias. Studies with non-significant results are less likely to be submitted and accepted for publication, thus upwardly biasing effect sizes (Lipsey & Wilson, 2001). Of the 148 reports analyzed in this study, nearly half (69) were unpublished reports, most of which were doctoral dissertations. Previous meta-analytic work, however, found only weak evidence of potential publication bias in this body of work (Hawkins et al., 2008).

Variable Coding

Two trained coders coded every study. After separately coding, the two coders compared answers. Although coders agreed more than 90% of the time, when there were discrepancies, coders sought clarification from the study text until they reached agreement. Thus, we did not compute inter-coder reliability; rather we used coder discrepancies as a stimulus for deeper investigation into the study to ascertain the correct code.

We explored possible differences due to four potential programmatic moderators: (a) dosage (measured as contact hours: low [1–8 contact hours] vs. moderate [9–20 hours] vs. high [21+ hours]; also measured as number of sessions); (b) institutionalization status (institutionalized vs. not; see previous definition); (c) setting (university/laboratory vs. religious; other community settings were too infrequent in the body of work to justify separate analyses); and (d) primary content emphasis (communication skills vs. expectations alignment and healthy relationship knowledge vs. relationship virtues enhancement). The first three of these program moderators were relatively straightforward to code, but program primary content emphasis was more challenging because many MRE programs have multiple components. When this was the case and when the report did not provide adequate clues, we took additional steps to code this moderator, including examining first-hand published curricula, when available, and even contacting program providers for their opinion.

Computation of Effect Sizes

We computed effect sizes with Comprehensive Meta Analysis II (Biostat, 2006). Standardized mean group differences were calculated for control-group studies. The standardized mean change score was computed for one-group/pre-post studies. Each effect size was weighted by the inverse variance (squared standard error) to account for the precision of the effect size estimates. Hedges' (1981) correction for small sample size bias was used because many studies had small sample sizes. Because previous meta-analytic work found little evidence of deterioration (or gain) of effects over the first year after intervention (Blanchard et al., 2009; Hawkins et al., 2008), and because we found limited evidence in analyses here of significant differences between post-test and follow-up effects, we combined immediate post-assessment and follow-up effects. (We coded the follow-up effect closest to 12 months post intervention, although 3- and 6-month follow-ups were the norm.) While examining follow-up effects would be a more rigorous evaluation, about 40% of studies did not assess follow-up effects, so our pool of studies to explore programmatic common factors would have been much smaller. We report the random effects results. Meta-analytic experts now recommend random effects estimates as standard practice (Shadish & Baldwin, 2003) because they allow for the possibility that differences in effect sizes from study to study are associated not only with participant-level sampling error but also with variations in study methods and interventions.

A precise effect size calculation for one-group/pre-post studies requires the pre/post correlation between the outcomes, information that was seldom reported. Often in these circumstances, meta-analysts reasonably estimate the correlation to be .50, which we did in this study. In a meta-analysis of parenting education interventions, Nowak and Heinrich (2008) reconstructed a reasonable pre/post correlation from other statistical information in reports and found an average correlation of .54.

Finally, because we examined effects for two different outcomes—communication skills and relationship quality/satisfaction—in two different study designs—control-group and one-group/pre-post—our analyses generated a set of four effect size statistics rather than a single effect size. Consequently, we try to interpret a pattern of results, giving somewhat greater weight to effects from control-group studies when inconsistencies emerge in the pattern.

Results

Program Participants

The large majority of MRE program participants were involved in marital enrichment programs (83%); 17% were involved in premarital education for engaged couples. Sample modal age was between 30–35. Sample modal education was “some college.” Only a small number of studies had significant numbers of lower-income and non-White participants, even though studies with more disadvantaged and diverse samples have increased in recent years (Hawkins & Fackrell, 2009). We revisit the implications of this sample homogeneity issue in our conclusion.

Our analyses of common factors of MRE effectiveness are displayed in Table 1.

Program Dosage

Our hypothesis for dosage was generally confirmed. The pattern of results for dosage indicates that less than 9 hours of contact hours in MRE programs may be insufficient to produce significant effects. Looking at control-group studies, moderate dosage programs—the modal group—had significantly larger effects than low-dosage programs ($Q_{com} = 8.12, p < .01; Q_{rq} = 7.44, p < .01$). Moreover, moderate-dosage programs had larger effects than high-dosage programs, although these differences did not reach statistical significance possibly due to the small number of high-dosage-program studies ($Q_{com} = 1.65, ns; Q_{rq} = 2.50, ns$). This pattern, then, suggests a curvilinear relationship between dosage and outcome, with moderate dosage being optimal. The one-group/pre-post studies, however, suggest a linear relationship between dosage and outcomes, with moderate-dosage programs showing larger effects than low-dosage programs ($Q_{com} = 7.30, p < .01; Q_{rq} = 9.10, p < .001$) and high-dosage program effects larger than moderate-dosage program effects, though not quite significantly for communication skills ($Q_{com} = 1.89, ns; Q_{rq} = 9.00, p < .01$).

In addition to contact hours, we also coded for number of sessions. Lower-dosage programs tended to be offered as one- or two-session interventions. Still, single-session programs (usually done on a weekend) yielded significant, moderate effects (for controlled studies, $d_{comm} = .356, p < .01, k = 16; d_{rq} = .342, p < .05, k = 25$). However, the largest effects were from programs with 10 or more sessions (usually spread out over 10+ weeks) (for controlled studies, $d_{comm} = .661, p < .05, k = 8; d_{rq} = .583, p < .01, k = 15$).

Program Primary Content Emphasis

The pattern of results for primary curricular content emphasis provides mixed evidence for our hypothesis that an emphasis on communication skills training produces stronger effects. Looking at control-group studies, communication skills curricula—the modal group—had

significantly larger effects than expectations alignment curricula on the communication skills outcome ($Q_{com} = 7.19, p < .01$). This result would be expected because the instruction is directly aligned with the measured outcome (i.e., teaching to the test). A more stringent test comes when looking at the relationship quality/satisfaction outcome, which is not directly aligned. In this case, communication skills curricula had somewhat larger effects than expectation alignment curricula, but the difference was not significant ($Q_{com} = 1.78, ns$). There were no differences between communication skills curricula and the small number of virtues curricula. A similar pattern emerges when looking at the one-group/pre-post studies. Communication skills curricula had significantly larger effects than expectations alignment curricula on the communication skills outcome ($Q_{com} = 13.69, p < .001$). Again, this result would be expected because the program instruction is directly aligned with the measured outcome. When looking at the relationship quality outcome, which is not directly aligned, communication skills curricula had somewhat larger effects than expectation alignment curricula, but again, the difference was not significant ($Q_{com} = 2.23, ns$). Only a couple of one-group/pre-post studies evaluated virtues curricula, so comparisons with these programs were unreliable, despite the large effect sizes. Overall, curricula that emphasize communication skills are indeed associated with larger outcomes for communication compared to curricula that emphasize expectations alignment and healthy relationship knowledge, but this difference is smaller and non-significant when looking at the relationship quality outcome.

Program Institutionalization

Contrary to our hypothesis, the pattern of results for program institutionalization provide no evidence that institutionalized programs produce stronger effects than the myriad non-institutionalized programs (for control-group studies: $Q_{com} = 1.12, ns$; $Q_{rq} = 0.22, ns$; for one-group/pre-post studies: $Q_{com} = 0.14, ns$; $Q_{rq} = 3.54, ns$). (Institutionalized programs are well-known programs with an ongoing presence in the field, manualized curricula, formal instructor training, and multiple outcome evaluations.)

Program Setting

Our hypothesis that programs delivered in university/laboratory settings would be associated with stronger effects than programs delivered in the field in religious settings received only weak support from the pattern of findings. When looking at control-group studies, programs in university/laboratory settings had significantly larger effects than those in religious settings for the communication skills outcome ($Q_{com} = 6.75, p < .01$). In contrast, there was no difference between these settings for the relationship quality outcome ($Q_{rq} = 0.31, ns$). Moreover, when looking at one-group/pre-post studies, programs delivered in religious settings were somewhat stronger, although these differences were not statistically significant ($Q_{com} = 1.65, ns$; $Q_{rq} = 1.12, ns$). Thus, overall, programs delivered in university/laboratory settings do not appear to have much of an advantage over programs delivered in religious settings.

Discussion

This study used meta-analytic methods to search for programmatic common factors of MRE effectiveness. The lack of program detail provided in study reports and the general lack of pedagogical diversity in MRE programs limited the range of potential moderators we could examine. Thus, our study is only an initial attempt to understand common factors in MRE. Additional meta-analytic approaches using more creative and fine-grained coding of MRE programs may improve our understanding. Careful conceptual analyses of programmatic moderators will be valuable, as well. Also, studies designed to test the effects of specific programmatic factors on program outcomes are a crucial way to illuminate best practices.

Despite our study limitations, we did confirm that program dosage (instructional contact hours) was a significant moderator of program effects, similar to meta-analytic studies of other family life education programs (MacLeod & Nelson, 2000; Piquart & Teubert, 2010a, 2010b). We found that a moderate dosage of 9–20 hours produced significant, larger effects compared to low-dosage programs. However, high-dosage programs of more than 20 hours did not produce significant program effects, suggesting that a moderate dosage is optimal. One-group/pre-post studies suggested that high-dosage programs yield even larger effects, but this was not the case when looking only at control-group studies. Giving greater weight to control-group study findings, and noting that Stanley and his colleagues (2006) found in a cross-sectional survey that relationship satisfaction did not continue to increase with premarital education of more than 20 hours, we conclude that a moderate dosage of 9–20 hours yields the largest effects for MRE programs. However, we should be cautious interpreting the meaning of results with high-dosage programs because, in some cases, higher doses are given because the program targets couples at more risk who may need more services, thus confounding dosage with prior risk. A dosage effect could best be evaluated in studies where different dose levels of an intervention were set, a priori, in randomized, controlled trials with high completion rates across dose conditions (so that selection effects do not compromise the interpretation of the dosage moderator).

Nevertheless, the modal dosage in MRE programs is about 12 hours, and this may be about right, at least for white, middle-class, relatively non-distressed couples typical of MRE participants. This is fortuitous in that it may be harder to recruit volunteers and retain participants for programs that demand higher levels of involvement. Larger doses may be important for more disadvantaged and distressed couples who are increasingly accessing MRE programs. There are examples of programs that include large doses of MRE (30+ hours) and that successfully deliver these large doses to couples with significant risk characteristics (see Cowan, Cowan, Pruett, Pruett, & Wong, 2009; also see results for the Oklahoma City site in the Building Strong Families 15-month impact report; Wood, McConnell, Moore, Clarkwest, & Hsueh, 2010).

One explanation for the ineffectiveness of lower-dosage programs might be that it takes a certain amount of time for new skills, behaviors, and attitudes to congeal. Low-dosage programs tended to be offered as one- or two-session interventions, whereas moderate- and high-dosage programs were almost always broken up into several sessions of shorter duration but spread over weeks or months. Spreading the period of intervention over a longer period may allow more time for new skills and behaviors to set. Further, smaller doses that occur only in one session may have a disadvantage of not having another, upcoming time point when the couples come back in with the implicit (or explicit) expectancy that they will have practiced some of the strategies taught in the prior session. And indeed, we found that programs spread over 10+ sessions yielded the largest effects. This finding seems similar to findings in the Piquart and Teubert (2010b) study: new-parenting interventions that lasted 3–6 months were associated with stronger effects than programs lasting 3 months or less. However, new-parenting interventions that lasted longer than 6 months were the least effective in their study. There may be a point at which longer interventions yield no greater return or may even be counterproductive (Olds, Sadler, & Kitzman, 2007). (Although, again, this can be confounded with characteristics of the sample receiving the services.) When considering dosage, however, it should be noted that our analyses suggested single-session (weekend) programs still produced significant, albeit more modest effects. Thus, we believe there is a role for single-session programs to play in the field, especially if they attract more participants than programs with many sessions.

We also explored primary content emphasis as a potential common factor. In a meta-analytic study of common factors of parenting education effectiveness, Kaminski and her colleagues

(2008) found that programs that emphasize teaching parents effective emotional communication skills with their children produced stronger positive child outcome effects. And in the MRE field, the strong cognitive-behavioral influence places a premium on effective communication and problem-solving skills for forming and sustaining healthy marriages and relationships. Thus, we expected programs that emphasize communication skills would produce larger effects than programs that emphasize alignment of couple expectations. This was the case, not surprisingly, when looking at the communication skills outcome. But this may be an unfair comparison because other kinds of programs do not place as much emphasis on communication skills. A fairer comparison may be with the more generic outcome of relationship quality/satisfaction. When looking at this outcome, communication skills programs had larger effects than other programs, but these effects did not reach statistical significance. Thus, programs that emphasize communication skills have obvious advantages in terms of increasing positive and decreasing negative communication behaviors. But this advantage does not show up reliably in reports of relationship quality. Of course, improvements in relationship quality may take time to show up as effective communication behaviors are maintained; relatively few studies follow participants for a lengthy period of time (Blanchard et al., 2009). In addition, perhaps more fine-grained coding and analyses of what kinds of communication and problem-solving skills are emphasized would produce more differentiated results. One MRE study found that an emphasis on positive communication and friendship-building behavior produced larger effects on conflict management skills than an emphasis on decreasing negative communication (Gottman, Ryan, Swanson, & Swanson, 2005).

Also, it would be valuable if programs could be examined on other important outcomes such as commitment, sacrifice, forgiveness, and fairness, which are positive constructs or virtues that may be just as important to relationship quality and stability as communication skills (Fincham, Stanley, & Beach, 2007; Fowers, 2000). Unfortunately, only a small number of MRE programs directly target these elements of healthy relationships and even fewer evaluations have assessed them. Future research should remedy this problem.

We hypothesized that institutionalized MRE programs would be associated with stronger effects compared to non-institutionalized programs. That is, a number of programs have a sustained presence in the field, use formal manuals, formally train educators to deliver the program, have invested in multiple outcome evaluations studies, and make ongoing adjustments to improve the program. Such controls and investments would be expected to yield better outcomes than programs with fewer controls and investments. Similarly, programs delivered in a university/laboratory setting usually use highly trained mental health professionals as instructors and likely have greater fidelity to the program design. Thus, they may be more effective. However, we found little evidence that institutionalized programs and programs delivered in university/laboratory settings were associated with stronger effects than non-institutionalized programs and programs delivered in religious settings. We did find that control-group studies of programs delivered in university/laboratory settings were associated with larger effects for communication skills, but this is likely due to an outcome measurement issue. That is, studies in university/laboratory settings were more likely to use observational measures of communication instead of self-report methods, and observational measures in MRE studies yield significantly larger effects (Blanchard et al., 2009; Fawcett et al., 2010).

Of course, one explanation for this counterintuitive set of findings is that MRE program developers overestimate the importance of the carefully designed features of their programs. That is, beyond sufficient dosage, the program particulars may be less important than non-specific factors related to participating in MRE or the beliefs, attitudes, and motivations that participants themselves bring to the intervention. Here the body of research to date provides

little help. Seldom do evaluators consider participants' personal attitudes and motivations for involvement in the program. Only a handful of studies examine distress levels of couples entering MRE (Blanchard et al., 2009). Nor do they often examine other personal factors such as personality, interpersonal awareness, or IQ. A reasonable implication of our results is that program evaluators should design studies that give more attention to participants' personal beliefs, attitudes, distress levels, and characteristics as potential moderators of program effects. MRE programs may be more effective for people with a specific profile of personal beliefs, attitudes, and characteristics, and these personal factors may override programmatic elements.

There are other ways that personal and interpersonal factors may be affected by MRE regardless of programmatic elements. Regardless of the content of a specific curriculum, the act of going to a workshop with one's partner may have an effect on commitment. Stanley has argued that there is an important emblematic role of culturally sanctioned rituals of deepening commitment (e.g., getting engaged) but that these emblems are diminishing (Stanley, 2010). In this frame, Stanley suggests that going to MRE may be a mutually reinforcing signal of commitment in a culture increasingly devoid of emblems of commitment. Behaviors that reflect future intention, and that are observable, should reinforce commitment between partners, which in turn, should confirm the security of the relationship (Stanley, Rhoades, & Whitton, in press). Regardless of the MRE curriculum, just showing up might demonstrate a willingness to step up when extra commitment and sacrifice is needed in a relationship.

We were unable to provide evidence here that the institutionalized programs produce uniformly superior results to other less formal programs. Admittedly, meta-analytic methods are coarse rather than fine-grained, in order to make comparisons across studies. And recently there has been substantial evolution in MRE approaches and models among most of the major programs. It does seem the case that the more institutionalized models receive regular refinement in methods, systems of delivery, and dissemination models based on ongoing experience and research. Such refinements may produce better effects over time in this field as program developers continue to improve usability, access, change models, and customer satisfaction for the services they provide. Still, the field is populated with many non-institutionalized MRE programs, and at this point, the evidence is that they too are effective interventions.

We acknowledge that our finding of no consistent difference between more and less formalized programs appears to diverge from the meta-analytic findings reported by Pinquart and Teubert (2010a, 2010b) who found that more highly educated and formally trained program instructors of new-parent education programs produced stronger effects. We can only speculate briefly about possible reasons for these divergent findings. One possibility might be that the transition to parenthood is a particularly novel and unsettling life course transition such that couples are more attuned to signals about the expertise and training of the instructor at such a time. In contrast, MRE is often available by clergy, lay leaders, and other community leaders wherein no particular expertise may be sought other than caring and reasonable competence.

We urge some caution in interpreting the general lack of differences found in our moderator analyses. We analyzed immediate post-treatment effects in this study due to the paucity of long-term follow-up effects in the body of evaluation studies. Differences could emerge if there are delayed effects, and a few studies have documented emergent treatment effects several years after program participation (Halford, Sanders, & Behrens, 2001; Markman, Renick, Floyd, Stanley, & Clements, 1993).

Importantly, MRE evaluation research has only begun to investigate directly the “pedagogic alliance”—the perceived positive connection between a participant and instructor—as a moderator of program outcomes. Some studies have tested whether participants’ consumer satisfaction with the program moderated outcomes (Hawkins, Fawcett, Carroll, & Gilliland, 2006), but this would be a very rough proxy for the pedagogic alliance. A recent study (Higginbotham & Myler, 2010) with couples in stepfamilies found that quality facilitation and instruction were more important to participants’ ratings of the program than whether participants and facilitators had similar demographic characteristics. One recent evaluation study (Owen, Rhoades, Stanley, & Markman, in press), however, directly addressed pedagogic alliance and found that the facilitator/participant working alliance accounted for significant variability in participant reports of relationship satisfaction and observed measures of positive and negative communication. Given the importance of the therapeutic alliance in MFT research (Sprenkle et al., 2009), this factor deserves more attention in future MRE evaluations.

We conclude this meta-analysis with a meta-observation. There is a programmatic homogeneity in the body of MRE work. That is, overall, programs pursue their intervention goals in much the same way. They teach much the same thing (communication skills dominate), deliver it much the same way (i.e., didactic classrooms) in only a few settings (universities/laboratories and churches), and measure the same generic outcomes (i.e., communication skills and relationship quality/satisfaction). The field is not so advanced that a convergence has taken place on clear best practices for marriage and relationship education. Instead, there seems to be a kind of “group-think” problem in the field. As a result, we see a need for experimenting with more divergent approaches to MRE programming to learn what works best and for whom. Program developers and evaluators should still be exploring different things to teach, taught in diverse ways and settings, and examining their effects on a wider range of important relational outcomes, all with an expectation that the personal attitudes and characteristics that participants bring to MRE will moderate program effects.

For instance, MRE curricula seem to assume a solid commitment to the relationship exists among participants, and thus provide skills and knowledge to build on this relational foundation. Yet recent scholarship suggests that commitment should not be taken for granted in MRE, perhaps especially for couples who are cohabiting or who cohabited before a decision to marry (Rhoades, Stanley, & Markman, 2009). Also, although didactic, classroom education will probably always have a role in the field and the social integration and support elements may help to produce stronger effects, still our increasingly online society will demand that effective relationship education be available online in ways that inform, engage, entertain, and motivate—and probably in more flexible, customized, and self-guided ways and perhaps in smaller doses (Halford, Moore, Wilson, Dyer, & Aarrugia, 2004). (For instance, see www.PowerofTwo.org. Other programs are moving in the same direction.) The field will need to adjust to these kinds of demands and evaluators will need to study them. Further, over the past decade the field has expanded to provide MRE opportunities to more disadvantaged, at-risk, and diverse couples (Hawkins & Fackrell, 2010). While this has produced some programmatic changes, more are needed. The increasing diversity of program participants needs to be matched by a growing diversity of programmatic approaches to helping couples form and sustain healthy marriages and relationships. More creativity and greater experimentation with different programmatic approaches is needed to advance our understanding of common factors of MRE effectiveness.

References

(A list of references of coded studies included in this meta-analysis is available upon request from the first author.)

- Biostat. Comprehensive Meta-Analysis (Version no. 2.2). Englewood, NJ: 2006.
- Blanchard, VL. Unpublished masters thesis. Brigham Young University; Provo, Utah: 2008. Does marriage and relationship education improve couples' communication? A meta-analytic study.
- Blanchard VL, Hawkins AJ, Baldwin SA, Fawcett EB. Investigating the effects of marriage and relationship education on couples' communication skills: A meta-analytic study. *Journal of Family Psychology*. 2009; 23:203–214. [PubMed: 19364214]
- Butler MH, Wampler KS. A meta-analytic update of research on the Couple Communication program. *The American Journal of Family Therapy*. 1999; 27:223–237.
- Carroll JS, Doherty WJ. Evaluating the effectiveness of premarital prevention programs: A meta-analytic review of outcome research. *Family Relations*. 2003; 52:105–118.
- Cowan PA, Cowan CP, Pruett MA, Pruett KD, Wong JJ. Promoting fathers' engagement with children: Preventive interventions for low-income families. *Journal of Marriage & Family*. 2009; 71:663–679.
- Cummings EM, Faircloth WB, Mitchell PM, Cummings JS, Schermerhorn AC. Evaluating a brief prevention program for improving marital conflict in community families. *Journal of Family Psychology*. 2008; 22:193–202. [PubMed: 18410206]
- Duncan SF, Steed A, Needham CM. A comparison evaluation study of web-based and traditional marriage and relationship education. *Journal of Couple & Relationship Therapy*. 2009; 8:162–180.
- Fawcett EB, Hawkins AJ, Blanchard VL, Carroll JS. Do premarital education programs really work? A meta-analytic study. *Family Relations*. 2010; 59:232–239.
- Fincham FD, Stanley SM, Beach SRH. Transformative processes in marriage: An analysis of emerging trends. *Journal of Marriage & Family*. 2007; 69:275–292. [PubMed: 19890466]
- Fowers, BJ. *Beyond the myth of marital happiness*. San Francisco: Jossey-Bass; 2000.
- Giblin P, Sprenkle DH, Sheehan R. Enrichment outcome research: A meta-analysis of premarital, marital and family interventions. *Journal of Marital & Family Therapy*. 1985; 11(3):257–271.
- Gottman J, Ryan K, Swanson C, Swanson K. Proximal change experiments with couples: A methodology for empirically building a science of effective interventions for changing couples' interaction. *Journal of Family Communication*. 2005; 5:163–190.
- Gottman, J.; Silver, N. *Why marriages succeed or fail*. New York: Simon & Schuster; 1994.
- Halford WK, Sanders MR, Behrens BC. Can skills training prevent relationship problems in at-risk couples? Four-year effects of a behavioral relationship education program. *Journal of Family Psychology*. 2001; 15(4):750–768. [PubMed: 11770479]
- Halford WK, Markman HJ, Kline GH, Stanley SM. Best practice in couple relationship education. *Journal of Marital & Family Therapy*. 2003; 29:385–406. [PubMed: 12870411]
- Halford WK, Moore EM, Wilson KL, Dyer C, Farrugia C. Benefits of a flexible delivery relationship education: An evaluation of the Couple CARE program. *Family Relations*. 2004; 53:469–476.
- Halford WK, Sanders MR, Behrens BC. Can skills training prevent relationship problems in at-risk couples? Four-year effects of a behavioral relationship education program. *Journal of Family Psychology*. 2001; 15(4):750–768. [PubMed: 11770479]
- Hawkins AJ, Fackrell TA. Does couple education for low-income couples work? A meta-analytic study of emerging research. *Journal of Couple & Relationship Therapy*. 2010; 9:181–191.
- Hawkins AJ, Blanchard VL, Baldwin SA, Fawcett EB. Does marriage and relationship education work? A meta-analytic study. *Journal of Consulting & Clinical Psychology*. 2008; 76:723–734. [PubMed: 18837590]
- Hawkins AJ, Carroll JS, Doherty WJ, Willoughby B. A comprehensive framework for marriage education. *Family Relations*. 2004; 53:547–558.
- Hawkins AJ, Fawcett EB, Carroll JS, Gilliland TT. The Marriage Moments program for couples transitioning to parenthood: Divergent conclusions from formative and outcome evaluation data. *Journal of Family Psychology*. 2006; 20:561–570. [PubMed: 17176190]

- Hedges LV. Distribution theory for Glasser's estimator of effect size and related estimators. *Journal of Educational Statistics*. 1981; 6:107–128.
- Higginbotham BJ, Myler C. The influence of facilitator and facilitation characteristics on participants' ratings of stepfamily education. *Family Relations*. 2010; 59:74–86.
- Hight TL. Do the rich get richer? A meta-analysis of the methodological and substantive moderators of couple enrichment. *Dissertation Abstracts International*. 2000; 61:3278B.
- Kaminski JW, Valle LA, Filene JH, Boyle CL. A meta-analytic review of components associated with parent training program effectiveness. *Journal of Abnormal Child Psychology*. 2008; 36:567–589. [PubMed: 18205039]
- Lipsey, MW.; Wilson, DB. *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications; 2001.
- Lundhal B, Risser HJ, Lovejoy MC. A meta-analysis of parent training: Moderators and follow-up effects. *Clinical Psychology Review*. 2006; 26:86–104. [PubMed: 16280191]
- MacLeod J, Nelson G. Programs for the promotion of family wellness and the prevention of child maltreatment: A meta-analytic review. *Child Abuse & Neglect*. 2000; 24:1127–1149. [PubMed: 11057701]
- Markman HJ, Renick MJ, Floyd FF, Stanley SM, Clements M. Preventing marital distress through communication and conflict management training: A four and five year follow-up. *Journal of Consulting and Clinical Psychology*. 1993; 61:70–77. [PubMed: 8450110]
- Markman HJ, Rhoades GK, Stanley SM, Ragan E, Whitton S. The premarital communication roots of marital distress: The first five years of marriage. *Journal of Family Psychology*. 2010; 24:289–298. [PubMed: 20545402]
- Markman, HJ.; Stanley, SM.; Blumberg, SL. *Fighting for your marriage*. 3. San Francisco: Jossey-Bass; 2010.
- Nowak C, Heinrichs N. A comprehensive meta-analysis of Triple-P Positive Parenting program using hierarchical linear modeling: Effectiveness and moderating variables. *Clinical Child & Family Psychology Review*. 2008; 11(3):114–144. [PubMed: 18509758]
- Olds DL, Sadler L, Kitzman H. Programs for parents of infants and toddlers: Recent evidence from randomized trials. *Journal of Child Psychology & Psychiatry*. 2007; 48:355–391. [PubMed: 17355402]
- Owen JJ, Rhoades GK, Stanley SM, Markman HJ. The role of leaders' working alliance in premarital education. *Journal of Family Psychology*. (in press).
- Pinquart M, Teubert D. A meta-analytic study of couple interventions during the transition to parenthood. *Family Relations*. 2010a; 59:221–231.
- Pinquart M, Teubert D. Effects on parenting education with expectant and new parents: A meta-analysis. *Journal of Family Psychology*. 2010b; 24:316–327. [PubMed: 20545405]
- Reardon-Anderson, J.; Stagner, M.; Macomber, JE.; Murray, J. *Systematic review of the impact of marriage and relationship programs*. Washington D.C: Urban Institute; 2005. Retrieved February 5, 2006, from <http://www.urban.org/url.cfm?ID=411142>
- Rhoades GK, Stanley SM, Markman HJ. Working with cohabitation in relationship education and therapy. *Journal of Couple & Relationship Therapy*. 2009; 8:95–112. [PubMed: 20613887]
- Ripley JS, Worthington EL Jr. Hope-focused and forgiveness-based group interventions to promote marital enrichment. *Journal of Counseling & Development*. 2002; 80:452–463.
- Schulz MS, Cowan CP, Cowan PA. Promoting Healthy Beginnings: A randomized controlled trial of a preventive intervention to preserve marital quality during the transition to parenthood. *Journal of Consulting & Clinical Psychology*. 2006; 74:20–31. [PubMed: 16551140]
- Shadish WR, Baldwin SA. Meta-analysis of MFT interventions. *Journal of Marital & Family Therapy*. 2003; 29:547–570. [PubMed: 14593694]
- Sprenkle, DH.; Davis, S.; Lebow, J. *Common factors in couple and family therapy: The overlooked foundation for effective practice*. New York: Guilford; 2009.
- Stanley, SM. What is it with Men and Commitment, Anyway?. Working paper based on a keynote address to the 6th Annual Smart Marriages Conference in 2002 in Washington DC; 2010. Retrieved from http://www.prepinc.com/main/docs/scottscorner/Men_and_Commitment_Stanley_Update.pdf

- Stanley SM, Amato PR, Johnson CA, Markman HJ. Premarital education, marital quality, and marital stability: Findings from a large, random household survey. *Journal of Family Psychology*. 2006; 20:117–126. [PubMed: 16569096]
- Stanley SM, Markman HJ, Prado LM, Olmos-Gallo PA, Tonelli L, St Peters M, Leber BD, Bobulinski M, Cordova A, Whitton SW. Community-based premarital prevention: Clergy and lay leaders on the front lines. *Family Relations*. 2001; 50:67–76.
- Stanley SM, Rhoades GK, Whitton SW. Commitment and the securing of romantic attachment. *Journal of Family Theory and Review*. (in press).
- Wood, RC.; McConnell, S.; Moore, Q.; Clarkwest, A.; Hsueh, J. Strengthening unmarried parents' relationships: The early impacts of Building Strong Families. Washington D.C: Administration for Children and Families, Office of Planning, Research, and Evaluation; 2010. The Building Strong Families Project. Retrieved on May 27, 2010, from http://www.acf.hhs.gov/programs/opre/strengthen/build_fam/reports/unmarried_parents/15_impact_exec_summ.pdf

Table 1
Effect Sizes by Programmatic Moderators for (A) Control-Group and (B) 1-Group/Pre-Post Studies.

Programmatic Moderator	Communication Skills				Relationship Quality/Satisfaction					
	k	d	C.I. low	C.I. high	z	k	d	C.I. low	C.I. high	z
A. Control-group Studies										
Dose:										
Low (1-8 hrs.)	21	0.100	-0.069	0.268	1.16 ns	35	0.099	-0.034	0.223	1.46 ns
Moderate (9-20 hrs.)	70	0.405	.280	0.530	6.33***	64	0.380	0.229	0.532	4.92***
High (21+ hrs.)	7	0.012	-0.574	0.599	0.041 ns	10	0.088	-0.241	0.417	0.53 ns
Primary Content Emphasis:										
Communication Skills	75	0.397	0.264	0.531	5.84**	69	0.297	0.146	0.448	3.86***
Expectations/Knowledge	17	0.105	-0.062	0.272	1.24 ns	24	0.139	-0.039	0.316	1.53 ns
Virtues/Motivations	5	0.285	-0.066	0.635	1.59 ns	12	0.162	-0.047	0.371	1.52 ns
Institutionalization:										
Yes	53	0.382	0.246	0.517	5.52***	56	0.229	0.084	0.374	3.10**
No	45	0.266	0.101	0.431	3.16**	55	0.279	0.133	0.426	3.73***
Setting:										
University/Laboratory	59	0.461	0.302	0.621	5.66***	58	0.266	0.125	0.406	3.70***
Religious	20	0.108	-0.105	0.321	0.99 ns	28	0.190	-0.038	0.417	1.64 ns
B. 1-group/Pre-Post Studies										
Dose:										
Low (1-8 hrs.)	8	0.279	0.129	0.429	3.64***	11	0.219	0.186	0.251	13.20***
Moderate (9-20 hrs.)	41	0.567	0.422	0.713	7.64***	35	0.433	0.298	0.569	6.26***
High (21+ hrs.)	20	0.891	0.446	1.337	3.92***	11	1.249	0.734	1.765	4.75***
Primary Content Emphasis:										
(Q = 15.00***)										

Programmatic Moderator	Communication Skills				Relationship Quality/Satisfaction				z
	k	d	C.I. low	C.I. high	k	d	C.I. low	C.I. high	
Communication Skills	54	0.607	0.500	0.714	34	0.487	0.361	0.612	7.60***
Expectations/Knowledge	8	0.212	0.033	0.392	16	0.325	0.154	0.496	3.72***
Virtues/Motivations	2	2.620	-1.007	6.247	4	1.481	-0.996	3.958	1.17 ns
Institutionalization:									(Q = 3.54 ns)
Yes	38	0.640	0.500	0.779	25	0.376	0.255	0.497	6.10***
No	31	0.576	0.281	0.872	32	0.631	0.395	0.868	5.23***
Setting:									(Q = 1.12 ns)
University/Laboratory	40	0.554	0.401	0.707	23	0.421	0.246	0.595	4.71***
Religious	13	1.061	0.302	1.819	15	0.706	0.206	1.205	2.77**

k = number studies; d = effect size; C.I. = confidence interval for d; z = significance test for d; Q = test of group differences; ns = p ≥ .05

* = p < .05

** = p < .01

*** = p < .001.