

## EFFECT OF CLIENT FEEDBACK ON COUPLE PSYCHOTHERAPY OUTCOMES

ROBERT J. REESE,  
MICHAEL D. TOLAND, AND  
NORAH C. SLONE  
*University of Kentucky*

LARRY A. NORSWORTHY  
*Abilene Christian University*

*Using outcome data to monitor the progress of treatment and the therapeutic alliance, also known as “client feedback” or “patient-focused research,” has yielded impressive results in individual psychotherapy. Client feedback has demonstrated reductions in premature terminations and improved psychotherapy outcomes. However, little research has been conducted using this paradigm with couples receiving therapy. The purpose of this study was to investigate whether the effectiveness of client feedback would extend to couple therapy. Results from a randomized couple clinical trial conducted in a naturalistic setting indicated that couples in a client feedback condition demonstrated statistically significantly more improvement compared with couples receiving treatment as usual and that improvement occurred more rapidly. Also, 4 times as many couples in the feedback condition reported clinically significant change by the end of treatment.*

**Keywords:** couple therapy, patient-focused research, feedback, psychotherapeutic outcomes

Continuous assessment, or client feedback, is a method of tracking client progress across psychotherapy that allows for clinicians to monitor whether progress is being made in treatment. If clients are not making progress as expected, therapists have the opportunity to modify or adjust treatment as necessary. The American Psychological Association (APA) Division 29 Task Force for Empirically Supported Relationships has made several recommendations to help clients achieve positive outcomes in psychotherapy. In a conclusion based on these recommendations, Ackerman et al. (2001) stated, “Practitioners are encouraged to routinely monitor patients’ responses to the therapy relationship and ongoing treatment. Such monitoring leads to increasing opportunities to repair alliance ruptures, to improve the relationship, to modify technical strategies, and to avoid premature termination” (p. 496).

A growing body of research investigating the effects of client feedback in psychotherapy has yielded encouraging results (e.g., Lambert et al., 2001, 2002; Reese, Norsworthy, & Rowlands, 2009; Whipple et al., 2003). A meta-analysis of three previous client feedback studies conducted by Lambert and colleagues (2003) found an overall effect size (Cohen’s *d*) of 0.39 for clients in a feedback condition identified as deteriorating (declined in treatment by about half of a standard deviation) compared with clients deteriorating in a no-feedback condition. Client feedback has consistently been found to benefit clients identified as at risk for terminating prematurely (i.e., clients who do not improve or deteriorate early in treatment).

The research, however, has found conflicting results on whether client feedback works for *all* clients, not just those who are not progressing as

---

Robert J. Reese, Michael D. Toland, and Norah C. Slone, Department of Educational, School, & Counseling Psychology, University of Kentucky; and Larry A. Norsworthy, Department of Psychology, Abilene Christian University.

Correspondence regarding this article should be addressed to Robert J. Reese, PhD, Department of Educational, School, & Counseling Psychology, University of Kentucky, Lexington, KY 40506-0017. E-mail: jeff.reese@uky.edu

expected (Harmon et al., 2007; Lambert, Harmon, Slade, Whipple, & Hawkins, 2005; Reese et al., 2009). Conclusions are difficult because studies that have found feedback to benefit all clients (Harmon et al., 2007; Hawkins, Lambert, Vermeersch, Slade, & Tuttle, 2004; Reese et al., 2009) have implemented client feedback differently from studies that have found improvement only for clients identified as at risk for terminating prematurely (Lambert et al., 2002; Whipple et al., 2003). For example, Whipple et al. (2003) found that only clients not on track benefited from feedback; however, the feedback data were shared only with the therapist. In contrast, Hawkins et al. (2004) found that providing feedback was beneficial for all clients when it was provided to both the therapist and client. In turn, Harmon et al. (2007) showed that feedback benefited all clients, but providing feedback to both therapist and client did not lead to increased effectiveness. More recent research (Reese et al., 2009) has found that client feedback was beneficial for all clients when compared with a treatment as usual (TAU) condition. In this study, feedback was provided for both therapist and client and included a measure to monitor the therapeutic alliance every session. Research is needed to further address the processes by which feedback is most effective, but the continuous assessment literature has consistently established that feedback is beneficial for improving psychotherapy outcomes, especially for clients at risk for dropping out of treatment.

The rationale for continuous assessment is based in part on research that has demonstrated that, in the aggregate, clients who benefit from therapy demonstrate improvement sooner rather than later in treatment (e.g., Howard, Kopta, Krause, & Orlinsky, 1986; Lutz, Martinovich, & Howard, 1999). Monitoring outcome early in treatment increases the likelihood of identifying clients who are not progressing as expected. Another predictor of effective psychotherapy is having a strong therapeutic alliance (Horvath & Symonds, 1991). As was stated earlier, the APA Division 29 Task Force recommends the ongoing monitoring of both outcome and the therapeutic alliance. Monitoring the therapeutic alliance has been found to be a statistically significant predictor of positive outcomes (Harmon et al., 2007). Therapists can quickly and directly respond to problems with the alliance when alerted (Lambert et al., 2002). As with monitoring treatment out-

comes, therapists can more readily intervene and attend to disconnects in the therapeutic relationship. Harmon et al. (2007) and Whipple et al. (2003) have found that adding clinical support tools, including a measure of the therapeutic alliance, yielded incremental effectiveness for at-risk clients when compared with just tracking outcome.

The research on client feedback is impressive but has focused almost exclusively on individual therapy. Couples and individuals experience similar barriers to positive psychotherapy treatment outcome, including deterioration, premature termination, and ruptured therapeutic alliances (Snyder, Castellani, & Whisman, 2006). Meta-analytic studies reported an overall effect size (Cohen's  $d$ ) for couple therapy ranging from 0.61 (Shadish et al., 1993) to 0.84 (Shadish & Baldwin, 2003). According to recent findings (Snyder et al., 2006), couples that receive therapy are approximately 80% better off than couples that do not receive treatment, which is comparable to effect sizes seen in the individual psychotherapy literature (Lambert & Ogles, 2004; Wampold, 2001). Shadish et al. (1993) compared the effect sizes for couple therapy studies with those for individual therapy studies and found a nonsignificant difference between the effect sizes ( $d = -0.05$ ,  $SE = 0.12$ ,  $n = 6$ ). The outcome literature provides substantial evidence that both individual and couple psychotherapy are effective forms of treatment.

The psychotherapy outcome literature for couples, like the individual literature, has also demonstrated that several approaches are effective. For example, Snyder et al. (2006) reported that both emotion-focused couple therapy and behavioral couple therapy have yielded impressive results in multiple clinical trials. Shadish and Baldwin's (2005) meta-analytic findings suggest that couples in treatment with behavioral couple therapy were 72% better off than couples in a control condition. Gollan and Jacobson (2002) demonstrated effective couple therapy using emotion-focused couple therapy, demonstrating recovery rates of 70–73% with a weighted mean effect size of 1.31 in comparison to a waitlist control. Although these therapies have evidenced greater effectiveness when compared with no-treatment conditions, the literature suggests that when these approaches are directly compared with one another, evidence has yet to demonstrate one approach being superior to another (Shadish &

Baldwin, 2003; Shadish et al., 1993; Snyder et al., 2006).

One factor that makes studying couple therapy outcome difficult is that therapy is both an individual and shared experience for each partner. Outcome is affected by the influence of each partner's readiness to change and level of distress (Isakson et al., 2006; Tambling & Johnson, 2008). An individual within the partnership may have significantly different views of the partnership itself, the therapy experience, and the therapist. A continuous feedback system used in couple therapy may allow both researchers and clinicians to better understand how individuals respond in couple psychotherapy at both the individual and couple level.

Only one study was identified that has used client feedback with couples and therapists while in therapy. Anker, Duncan, and Sparks (2009) used the Partners for Change Outcome Management System (PCOMS; Duncan, Miller, & Sparks, 2004) with a sample of 205 White Euro-Scandinavian heterosexual couples. PCOMS consists of two brief measures that are used to track client progress in therapy during each session. The Outcome Rating Scale (ORS; Miller, Duncan, Brown, Sparks, & Claud, 2003) consists of four items and measures client outcome, and the Session Rating Scale (SRS; Duncan et al., 2003) also consists of four items and measures the therapeutic alliance. The ORS is administered at the beginning of each session and the SRS is administered and scored at the end of each session.

The results of each scale were administered, scored, and discussed every session. Participants in the study presented with a broad range of relationship issues, including communication, jealousy/conflict, and coping with partner's physical or psychological problems (Anker et al., 2009). Couples were assigned to one of 10 therapists and randomly assigned to a treatment condition: feedback or TAU. Couples in the feedback condition reported statistically significantly higher residual ORS scores than couples in the TAU condition, yielding an effect size of  $d = 0.50$ , which is considered large when comparing the differences between treatments (Wampold, 2001). Four times as many couples in the feedback condition experienced clinically significant change (i.e., change beyond the standard error of measure that includes starting treatment below the clinical cut score and finishing treatment above the clinical cut score). Couples in the feed-

back condition also reported higher levels of marital satisfaction at posttreatment, and a greater percentage of marriages were intact at follow-up when compared with marriages in the TAU condition. These findings for PCOMS are consistent with previous studies that focused on individual therapy (Miller, Duncan, Brown, Sorrell, & Chalk, 2006; Reese et al., 2009).

Replication of the Anker et al. (2009) findings is necessary for further evidence that continuous assessment, and specifically PCOMS, works with couples. The purpose of our study was to replicate the results of Anker et al. with a sample from the United States. The current study focused on the effectiveness of using PCOMS with couples in psychotherapy as compared with a TAU control condition. We had three major hypotheses. First, we hypothesized that couples in the feedback condition would demonstrate better outcomes than those in the TAU condition as measured by the ORS after controlling for pre-ORS scores. Second, we hypothesized that couples in the feedback condition would improve more quickly (i.e., in fewer sessions) than couples in the TAU condition. Third, we hypothesized that more couples in the feedback condition would meet the criteria for clinical significance at posttreatment than would couples in the TAU condition.

## Method

### *Participants*

*Clients.* Clients were 46 heterosexual couples ( $N = 92$ ) that received couple therapy during the course of an academic year at a graduate training clinic for a marriage and family therapy master's program. There were 55 possible couples, but nine couples (3 = feedback condition, 6 = TAU) did not return for a second session for reasons unknown. The mean pretreatment ORS score for those that did not return for a second session (23.36) was almost identical to those included in the study (23.62). Seventy-four percent of the sample was Caucasian ( $n = 68$ ), 4.3% African American ( $n = 4$ ), 16.3% Hispanic/Latino ( $n = 15$ ), 3.2% multiracial ( $n = 3$ ), and 2.2% ( $n = 2$ ) did not indicate ethnicity. The mean age was 30.18 years ( $SD = 9.71$ ), with ages ranging from 19 to 56 years. The primary reason couples sought counseling included relationship distress ( $n = 36$  couples; marital discord, communication, parenting, divorce, separation, extra-

marital affairs, sexual difficulties), individual distress affecting the relationship ( $n = 4$  couples; pornography addiction, depression, anxiety, sexual abuse), and relationship enhancement ( $n = 6$  couples; premarital, relationship enhancement).

*Therapists.* All of the 261 sessions at the training clinic for marriage and family therapy were provided by 13 second-year practicum students (7 women and 6 men; 10 = Caucasian, 2 = African American, 1 = Hispanic) enrolled in an American Association for Marriage and Family Therapy–approved program. Practicum students received weekly individual and group supervision. All of the couple sessions were video-recorded for supervision purposes. One therapist met with the couples for a 50-min session typically on a weekly basis. There were no session limits, and the sessions did not follow a particular treatment format or protocol. Theoretical orientations of the student therapists across both treatment conditions were all grounded in a general family systems framework, using a variety of approaches including solution-focused, narrative/postmodern, and strategic therapy. The median number of couples seen by each therapist was three, ranging from one to eight couples. Therapists met with couples on average for 5.91 sessions ( $Mdn = 5$ ), ranging from two to 17 sessions.

### Measures

*ORS.* The ORS is a four-item, self-report measure that is designed to evaluate session-to-session progress made in therapy. Using a visual analog scale, clients rate their level of psychological distress on items adapted from the three areas of the Outcomes Questionnaire 45 (OQ45; Lambert et al., 1996). Specifically, clients respond to how they are doing *individually* (personal well-being), *socially* (work, school, friendships), *interpersonally* (family, close relationships), and *overall* (general sense of well-being). Clients make a mark on each of the four analog scales that are 10 cm in length, with marks near the left end of the scale indicating lower distress and marks near the right end of the scale indicating higher distress. A ruler or template is then used to measure the distance from the left end of the scale to the client's mark. The score is recorded for each item to the nearest millimeter and then all are summed, for a total score ranging from 0 to 40. Lower scores reflect more distress.

The internal consistency estimated with the ORS (first session) for the current sample was .88, 95% CI [.84, .92]. Anker et al. (2009) reported an internal consistency coefficient alpha estimate of .93 with 410 individuals participating in couple therapy. Reese et al. (2009) have found similar reliability estimates. Evidence of concurrent validity for scores derived from the ORS is based on Pearson correlations with scores on other established outcome measures, including the Symptom Checklist–90—Revised (Derogatis, 1992;  $r = -.57$ ; Reese, Norsworthy, & Rowlands, 2006), the Clinical Outcomes in Routine Evaluation (Barkham et al., 2001;  $r = .67$ ; Miller & Duncan, 2004), and the OQ45 ( $r = .59$ ; Miller et al., 2003).

*PCOMS.* All therapists and supervisors in the feedback condition attended a 1-hr training session that covered the rationale for using PCOMS and how to administer, score, and interpret the ORS and SRS. The feedback condition used the protocol as outlined in the scoring and administration manual for PCOMS (Miller & Duncan, 2004). Each client was administered the ORS at the beginning of every couple session with the therapist present. After completing the ORS (approximately 1 min), the therapist scored the items in the session. The total score was charted on a graph that indicated each client's progress across treatment. Because this study was conducted with couples, one chart was used that showed the individual progress of each partner. Therapists used the data within the session as they saw fit, but the manual provides guidelines for how to intervene with clients who fall into the following four categories:

- **No change.** For a client who has not shown reliable change (a gain of 5 points) after three sessions, therapists are directed to address the therapeutic alliance and the course of treatment. If the client has not demonstrated reliable improvement after six sessions, the manual suggests consultation, supervision, or staffing.
- **Deteriorating.** Clients in this category (a decrease of 5 points since entering treatment) are considered to be at risk for terminating prematurely or having a poor outcome. Therapists are directed to discuss possible reasons for the drop in score, review the SRS items with the client to assess the therapeutic alliance, or consider changing the treatment approach, fre-



quency, mode, or even therapist if no improvement is noted after three sessions.

- **Reliable change.** Treatment is going accordingly (evidenced by a gain of at least 5 points since beginning therapy). Therapists are advised to reinforce changes and to continue treatment until progress begins to plateau, whereupon a therapist should consider reducing the frequency of sessions.
- **Clinically significant change.** The client may no longer be struggling with issues that led to seeking therapy. Clinically significant change is defined by a client beginning treatment below the clinical cut score of 25, improving at least 5 points since starting therapy, and having a total score in the nonclinical range (25 or above). Therapists are advised to consolidate changes, anticipate potential setbacks, and consider reducing the frequency of sessions.

The SRS was administered to each client and again scored by the therapist (approximately 1 min) toward the end of the session. If the total score was below 36 or any one of the items was below 9, the therapist followed up and asked about the reason for the lower scores. The total score was then charted on a graph for the corresponding session. Again, scores on the SRS for each partner in the couple were recorded on one graph. The SRS was used as part of the feedback process for PCOMS, but the data were not included in the analyses for the current study.

### Data Analysis

We applied multilevel modeling (MLM; Hox, 2002), also referred to as hierarchical linear modeling (Raudenbush & Bryk, 2002), to answer the first two primary hypotheses (for a gentle introduction to MLM, see Peugh, 2010). In general, multilevel data tend to result when data are naturally nested data structures (e.g., clients nested within therapists, therapists nested within a counseling center, repeated observations nested within clients, who are then nested within therapists). The issue with nested data structures is that the traditional assumption of independence of observations is violated, which is necessary for traditional techniques such as analysis of variance (Peugh, 2010). Ignoring this issue will result in biased parameter estimates (i.e., means, variances, and covariances) and increase Type I error rates. In this study, the data structure is naturally

nested wherein each client is nested within a couple, which is then nested within a therapist. This means that the ORS scores of partners within the same couple are likely to be more correlated than ORS scores for partners in different couples. In the language of MLM, each client is perceived as a Level-1 unit and couples are seen as a Level-2 unit. Similarly, ORS scores of couples within the same therapist are likely to be more correlated than ORS scores for couples working with different therapists. As a result, the language of MLM would consider therapist to be a Level-3 unit.

For the first research hypothesis, we predicted that couples in the feedback condition would demonstrate better outcomes than those in the TAU condition as measured by the ORS after controlling for pre-ORS scores. This means a two-level cross-sectional multilevel model was needed to address this hypothesis. The Level-2 predictor is feedback condition (FEEDBACK; 1 = feedback condition; 0 = TAU condition) and the Level-1 predictor or covariate is pre-ORS scores. Because the primary interest is in the Level-2 predictor, FEEDBACK, pre-ORS scores were grand mean centered by subtracting each client's score from the overall mean pre-ORS score ( $M_{\text{pre-ORS}}$ ; for details on centering predictors in MLM, see Enders & Tofighi, 2007).

The multilevel model used to address the first research hypothesis or explain variation in ORS scores is

$$Y_{ij} = \gamma_{00} + \gamma_{01}(\text{FEEDBACK}_j) + \gamma_{10}(\text{pre-ORS}_{ij} - M_{\text{pre-ORS}}) + \mu_{0j} + r_{ij}, \quad (1)$$

where  $Y_{ij}$  is the post-ORS score for client  $i$  in couple  $j$ ;  $\gamma_{00}$  is a fixed effect reflecting the overall mean post-ORS for couples in the TAU condition after controlling for pre-ORS scores;  $\gamma_{01}$  is a fixed effect reflecting mean difference between couples in the TAU and feedback conditions after controlling for pre-ORS scores (i.e., a positive difference would mean that couples in the feedback condition had a higher mean post-ORS than couples in the TAU condition after controlling for pre-ORS scores);  $\gamma_{10}$  is a fixed effect or covariate reflecting the slope between pre- and post-ORS scores after controlling for FEEDBACK;  $\mu_{0j}$  is a Level-2 random couple effect or the deviation of couple  $j$  from the overall mean post-ORS for couples after controlling for pre-ORS scores; and  $r_{ij}$  is a Level-1 random client effect or client  $ij$ 's difference in

post-ORS score from the overall mean post-ORS for couples after controlling for pre-ORS scores.

In MLM, the random effects are estimated as variances such that  $\sigma_{\text{Couple}}^2$  and  $\sigma_{\text{Client}}^2$  capture the intercept variances in ORS scores at the couple and client levels, respectively. Conceptually,  $\sigma_{\text{Couple}}^2$  measures the variation in mean post-ORS scores across couples that is not due to feedback condition and is similar to  $MS_{\text{Between}}$  in analysis of covariance (ANCOVA). Similarly,  $\sigma_{\text{Client}}^2$  is the average variance in individual clients' scores within couples after accounting for pre-ORS scores and feedback condition and is like  $MS_{\text{Within}}$  in ANCOVA.

For the second research hypothesis, we predicted that couples in the feedback condition would improve more quickly (i.e., in fewer sessions) than couples in the TAU condition. This means that a three-level multilevel growth model was needed to address this hypothesis. In this model, repeated observations or time represented Level 1, which are nested within each client (Level 2), which are then nested within each couple (Level 3). The Level-3 predictor in this model is feedback condition (FEEDBACK; as previously defined) and Level-1 predictors are the time measure as a linear function (SESSION) and nonlinear function (SESSION<sup>2</sup>). The nonlinear function of time allows the model to capture the curvature in the ORS growth patterns, which is more realistic than assuming all couples' ORS scores grow in a linear manner.

The multilevel model used to address the second research hypothesis or explain variation in ORS scores over sessions is

$$\begin{aligned}
 Y_{ij} = & \gamma_{000} + \gamma_{100}(\text{SESSION}_{ij}) \\
 & + \gamma_{200}(\text{SESSION}_{ij}^2) + \gamma_{001}(\text{FEEDBACK}_j) \\
 & + \gamma_{101}(\text{FEEDBACK}_j)(\text{SESSION}_{ij}) \\
 & + \gamma_{201}(\text{FEEDBACK}_j)(\text{SESSION}_{ij}^2) + \mu_{00j} \\
 & + r_{0ij} + e_{ij}, \quad (2)
 \end{aligned}$$

where  $Y_{ij}$  is the ORS score at session  $t$  for client  $i$  in couple  $j$ ;  $\gamma_{000}$  is a fixed effect reflecting the overall average couple mean ORS at the start of therapy for those in the TAU condition (centered at Session 1 such that substituting 0 for SESSION reflects the effect of the treatment at the first session);  $\gamma_{100}$  is the overall average linear growth rate between adjacent sessions for those in the TAU condition (i.e., the expected linear change

in ORS for a one-session increment in time);  $\gamma_{200}$  is the quadratic or curvature growth rate between adjacent sessions for those in the TAU condition (i.e., the expected nonlinear change in ORS for a one-session increment in time);  $\gamma_{001}$  is the mean difference between couples in the TAU and feedback conditions at the start of therapy;  $\gamma_{101}$  is the average linear slope difference between couples in the TAU and feedback conditions (i.e., a positive value would mean that couples in the feedback condition improved faster in ORS scores than those in the TAU condition);  $\gamma_{201}$  is the difference in curvature growth rates between couples in the TAU and feedback conditions (i.e., a positive value would mean that couples in the feedback condition have more positive curvature growth rates than couples in the TAU condition);  $\mu_{00j}$  is a Level-3 random couple effect or the deviation of each couple mean from the overall initial couple ORS mean;  $r_{0ij}$  is a Level-2 random client effect or the deviation of client  $ij$ 's ORS score from the overall initial couple ORS mean; and  $e_{ij}$  is a Level-1 random client effect or client  $ij$ 's residual error at session  $t$  (this error term reflects the difference between each client's predicted and observed ORS score).

Similar to the random effects estimated in the multilevel model for the first hypothesis, the random effects in growth models are estimated as variances such that  $\sigma_{\text{Couple}}^2$ ,  $\sigma_{\text{Client}}^2$ , and  $\sigma_{\text{Error}}^2$  estimate the intercept variances in ORS scores at the couple-level, client-level, and repeated observations level, respectively. Conceptually,  $\sigma_{\text{Couple}}^2$  measures the variance in mean ORS scores across couples that is not due to feedback condition,  $\sigma_{\text{Client}}^2$  is the average variance in individual clients' scores within couples that is not due to feedback condition (i.e., individual differences between clients), and  $\sigma_{\text{Error}}^2$  captures within-person variation in ORS scores (i.e., the variability of an individual client's score around her or his mean ORS score).

For all MLM analyses, predictors or covariates were added to each of these basic models at the appropriate level (client level or couple level). To compare the statistical fit of competing models, we used the  $-2$  log-likelihood ( $-2LL$ ) value or deviance statistic from two nested models and found the difference in deviance estimates. Models are nested when one model is a subset of the larger statistical model. Also, the deviance is a measure of fit, and the higher the deviance, the poorer the fit of the model to the sample data. The

difference in the deviances tests the null hypothesis that two models do not have statistically significantly different model fits to the sample data. A rejection of this null hypothesis indicates that the model with more estimated parameters fits the sample data better than a model with fewer estimated parameters. The difference in the deviance statistics is  $\chi^2$  distributed with degrees of freedom equal to the difference in parameters estimated by two nested models. The test of two models' deviances is often referred to in the statistical literature as a likelihood ratio test. All MLM analyses were conducted with Proc Mixed in SAS Version 9.2 using maximum likelihood estimation and the Satterthwaite degrees of freedom method.

**Results**

*Preliminary Analyses*

Although our data are inherently nested within therapists at the highest level, initial MLM analyses indicated that therapist fixed effects (i.e., feedback condition vs. TAU) could not be directly estimated at the therapist level because of the limited number of therapists used in this study ( $n = 13$ ), but therapist-level variance at the intercept ( $\sigma^2_{\text{Therapist}}$ ) could be estimated (i.e., the variation in mean ORS scores across therapists). It is important to note that ignoring the variability at the therapist level results in this variability being pushed into the variance estimates at other levels of the multilevel model, which are then ultimately biased. Therefore, we estimated the  $\sigma^2_{\text{Therapist}}$  for each of our models. As a result, we used a three-level multilevel model to analyze the nested structure of our data, clients nested within couples, to address the first research hypothesis. Similarly, we used a four-level multilevel growth model (session within client within couple within therapist). However, each of these models is estimating a single variance component at the therapist level and described as models with one less level (i.e., ignoring therapist fixed effects because they could not be estimated).

*Did PCOMS Produce Differences in Outcomes for Couples?*

Descriptive statistics show that clients in the feedback condition improved 8.58 points compared with the 3.64-point improvement by clients

in the TAU condition (see Table 1 for pre- and post-ORS mean treatment scores, standard deviations, and effect sizes within each condition). To evaluate the first hypothesis, we first estimated a model like that shown in Equation 1 except we included only the covariate pre-ORS scores (grand mean centered;  $M_{\text{pre-ORS}} = 23.62$ ) in the model (covariate-only model in Table 2). The covariate-only model was estimated as a baseline model as is typically done in traditional hierarchical regression analyses to determine the incremental improvement of one model to the next. The covariate-only model (see column 1 of Table 2) suggests a statistically significant positive slope ( $\gamma_{10} = 0.26, p < .001$ ) between pre-ORS scores and post-ORS scores across clients. This means that scores improved from pre-ORS to post-ORS, while the average post-ORS for a client with an average pre-ORS score was 30.17 ( $\gamma_{00}$ ). The standardized mean effect size from pre- to post-ORS was 0.71 ( $[30.17 - 23.62]/9.21$ ), indicating that clients improved by almost three fourths of a standard deviation from pre- to post-ORS.

One way to understand the overall utility of the covariate-only model is to compute a global pseudo- $R^2$  effect size statistic, like multiple  $R^2$  in regression. This is done by correlating and squaring the predicted ORS scores for each participant, using the fixed effects parameters for the covariate-only model with the observed ORS scores for each client. The global pseudo- $R^2 = .13$ , which means that 13% of the variation in ORS scores can be explained by knowing the pre-ORS scores (see Peugh, 2010). To understand the specific amount of variability explained at a level, we computed a local pseudo- $r^2$  statistic, which is similar to a semipartial  $r^2$  statistic in traditional regression. As such, the estimated proportion of variance between couples explained by

TABLE 1. Pretest and Posttest Mean Outcome Rating Scale (ORS) Scores and Effect Sizes for the Client Feedback and Treatment as Usual Conditions

Measure	Client feedback ( $n = 54$ )		Treatment as usual ( $n = 38$ )	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Pre-ORS score	23.34	9.15	24.03	9.47
Post-ORS score	31.92	7.15	27.67	9.53
Standardized effect size	0.94		0.38	

Note. Standardized effect size =  $(M_{\text{post}} - M_{\text{pre}})/SD_{\text{pre}}$ .

TABLE 2. Fixed and Random Effect Estimates for Multilevel Models Predicting Postoutcome Rating Scale (ORS) Scores

Parameter	Covariate-only model	ANCOVA
Fixed effects (regression coefficients)		
Intercept: Mean post-ORS ( $\gamma_{00}$ )	30.17*** (0.99)	27.56*** (1.46)
Client pre-ORS ( $\gamma_{10}$ )	0.26** (0.09)	0.27** (0.09)
Feedback ( $\gamma_{01}$ )		4.44* (1.9)
Random effects (regression variances)		
Client intercept variance ( $\sigma_{Client}^2$ )	33.27*** (7.04)	33.48*** (7.1)
Couple intercept variance ( $\sigma_{Couple}^2$ )	27.51** (10.31)	22.52** (9.35)
Therapist intercept variance ( $\sigma_{Therapist}^2$ )	1.14	0.99
Standardized effect size	0.71 <sup>a</sup>	0.48 <sup>b</sup>

Note. Standard errors are in parentheses. Client pre-ORS = client’s initial ORS score grand mean centered; Feedback = type of feedback condition (0 = treatment as usual; 1 = feedback).

<sup>a</sup> Standardized effect size =  $(M_{post} - M_{pre})/SD_{pre}$ .

<sup>b</sup> Standardized effect size = 
$$\frac{\gamma_{01}}{\sqrt{\frac{(n_{TAU} - 1)s_{TAU(post-ORS)}^2 + (n_{Feedback} - 1)s_{Feedback(post-ORS)}^2}{N - 2}}}$$

\*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

the covariate-only model with pre-ORS is 0.27 (i.e.,  $[37.63 - 27.51]/37.63$ ). This means that 27% of the between-couples variance in post-ORS scores is accounted for by knowing the pre-ORS scores. Moreover, the intraclass correlation for couple ( $ICC_{Couple} = \sigma_{Couple}^2 / [\sigma_{Client}^2 + \sigma_{Couple}^2 + \sigma_{Therapist}^2]$ ) was .44 (which had been .53). This means that 44% of the variance in post-ORS scores is due to pre-ORS scores. The  $ICC_{Therapist}$  was .02, meaning that 2% of the variability in post-ORS scores (after controlling for pre-ORS scores) was attributed to therapists, which is within the range of other naturalistic therapist effects (see Anker et al., 2009; Baldwin, Berkeljon, Atkins, Olsen, & Nielsen, 2009).

For the second model we added treatment condition (FEEDBACK) as a predictor to the former model (see Equation 1). The second model can be conceptually thought of as a multilevel ANCOVA model (see column 2 of Table 2). The  $-2LL$ s for the covariate-only and ANCOVA models were 629.6 and 624.4, respectively. The difference in fit between these two models was statistically significant,  $\chi^2(1) = 5.2, p = .02$ . Results from this approach suggest that including a model with a treatment effect for feedback while controlling for pre-ORS scores improves the overall fit of the model to the data versus including only pre-ORS scores. This means that couples in the feedback condition scored on average 4.44 ( $\gamma_{00}$ ) ORS

points higher than couples in the TAU condition after controlling for pre-ORS scores (see last column of Table 2). The standardized mean effect size between couples’ ORS scores after controlling for pre-ORS scores was 0.54 (see formula at bottom of Table 2; U.S. Department of Education: Institute of Education Sciences, 2008, p. 43). This means that the feedback condition scored just over half a standard deviation higher on post-ORS scores than the TAU condition after controlling for pre-ORS scores, which is within the range of other naturalistic therapist effects (see Anker et al., 2009).

When we examine the ANCOVA model random effects, we see that 39.5% of the variance in post-ORS scores remains between couples. Moreover, comparing the  $ICC_{Couple}$  from both models (covariate only vs. ANCOVA), the proportion of variance between couples explained by the ANCOVA model is  $(27.51 - 22.52)/27.51 = .18$  or 18%. That is, 18% more between-couples variance in post-ORS scores is explained by knowing the type of feedback condition. The global pseudo- $R^2$  effect size statistic for the ANCOVA model was .19.

### Preliminary Growth Curve Analyses

In the MLM literature, it is recommended that a minimum of four time points be used to specify



a quadratic (or nonlinear) group model (Willett, Singer, & Martin, 1998). By adding a nonlinear component to the model, researchers can increase precision in estimates of change (Muthén, 1999; Muthén & Curran, 1997). However, couples varied in the number of sessions attended. Results from preliminary analyses identified an outlier couple (17 sessions) as an influential case and was subsequently removed from all subsequent multilevel growth curve analyses. Although the minimum recommended number of sessions for a quadratic growth model is four, we chose to include all couples with at least two sessions because models including all couples did not differ from models including couples who attended a minimum of four sessions. Moreover, including couples with fewer than four sessions increases the generality of the results to naturalistic settings and helps maintain adequate statistical power.

*Growth Curves for Feedback and TAU Conditions*

To evaluate the second hypothesis, we first estimated a model like that shown in Equation 2 except that we did not include feedback condition (FEEDBACK) in the growth model (see unconditional growth model in Table 2). The unconditional growth model estimates an intercept or average starting mean ORS score, linear growth rate, and nonlinear (quadratic) growth rate across couples. These results indicate that all couples start with an average ORS score of 24.46 and have statistically significant linear ( $\gamma_{100} = 1.87$ ,

$p < .001$ ) and quadratic ( $\gamma_{100} = -0.08$ ,  $p < .001$ ) growth rates (see Table 3). The global pseudo- $R^2$  effect size statistic for the unconditional model was .09, which suggests that 9% of the variation in ORS scores can be explained by knowing linear change and quadratic change.

To evaluate whether growth rates vary between couples receiving TAU versus feedback during couple therapy (Hypothesis 2; Equation 2), we added treatment condition (FEEDBACK) to the growth model (conditional growth model; see Table 3). The  $-2LLs$  for the unconditional and conditional growth models were 2,684.8 and 2,675.9, respectively. The difference in fit between these two models was statistically significant,  $\chi^2(3) = 8.9$ ,  $p = .053$ , indicating that couples in the feedback condition improved more quickly than couples in the TAU condition. The global pseudo- $R^2$  effect size statistic for the conditional model was .10. Inspection of the conditional growth model results specifically shows that clients receiving feedback during couple therapy have a statistically significant different linear growth rate compared with those in the TAU condition ( $\gamma_{101} = 1.5$ ,  $p = .02$ ,  $d = 0.81$ , i.e.,  $d = [\text{effect}(\text{time})]/SD_{\text{pre}}$ , where time = 5 and  $SD_{\text{pre}} = 9.31$ ; for more details, see Feingold, 2009, p. 7). Because the session numbers varied across couples, the value of time was set to 5 to reflect the median number of sessions attended. However, there was not a statistically significant difference in the conditions' quadratic growth rates ( $\gamma_{201} = -0.11$ ,  $p = .14$ ). A depiction of the difference in these growth rates up to five sessions is presented in Figure 1. We chose to stop

TABLE 3. Fixed and Random Effect Estimates for Multilevel Growth Models for Outcome Rating Scale (ORS) Scores

Parameter	Unconditional growth model	Conditional growth model
Fixed effects (regression coefficients)		
Intercept: Mean ORS ( $\gamma_{000}$ )	24.46*** (1.07)	24.53*** (1.65)
Session ( $\gamma_{100}$ )	1.87*** (0.28)	0.91 (0.56)
Session <sup>2</sup> ( $\gamma_{200}$ )	-0.08** (0.03)	-0.01 (0.07)
Feedback ( $\gamma_{001}$ )		-0.001 (2.16)
Feedback $\times$ Session ( $\gamma_{101}$ )		1.5* (0.65)
Feedback $\times$ Session <sup>2</sup> ( $\gamma_{201}$ )		-0.11 (0.07)
Random effects (regression variances)		
Error variance ( $\sigma_{\text{Error}}^2$ )	32.53*** (2.62)	31.68*** (2.55)
Client intercept variance ( $\sigma_{\text{Client}}^2$ )	10.01** (4.14)	10.3** (4.16)
Couple intercept variance ( $\sigma_{\text{Couple}}^2$ )	24.48** (9.82)	23.66** (9.68)
Therapist intercept variance ( $\sigma_{\text{Therapist}}^2$ )	10.7	10.78

Note. Standard errors are in parentheses. Session = session number centered at Session 1; Feedback = type of feedback condition (0 = treatment as usual; 1 = feedback).

\*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

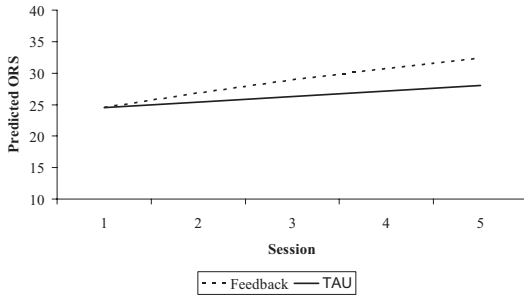


FIGURE 1. Average growth curves across five sessions for the feedback (dashed line) and treatment as usual (TAU; solid line) conditions.

at five sessions given that the sample median number of sessions per couple was five. Figure 1 depicts the quicker improvement in ORS scores for the feedback condition over the TAU condition after five sessions.

Clinical Significance

Observing the number of clients who incur clinically significant change across treatment has become a common way to assess psychotherapy outcome (Lambert, Hansen, & Bauer, 2008). Jacobson and Truax (1991) developed formulas to evaluate change in therapy using the terms *reliable change* and *clinically significant change* to denote meaningful change in therapy. Reliable change is simply the increase or decrease in a client’s score on an outcome measure that exceeds the measurement error for the instrument. For the ORS, the amount of change needed to incur reliable change is 5 or more points. A decrease of 5 or more points is termed *deterioration*. Clinically significant change occurs when a client has reliable change (gain of 5 points) and

the client finishes therapy above an established cut score that separates a clinical from nonclinical population. The cut score for the ORS is 25. The reliable change index and cut score for the ORS were based on two samples from a community mental health center (Miller et al., 2003) and a residential alcohol and drug treatment center (Miller, Mee-Lee, Plum, & Hubble, 2005).

More clients in the feedback condition, both at the individual and couple level, completed treatment having obtained reliable (gain of 5+ points) and clinically significant (gain of 5+ points and crossing the clinical threshold) change when compared with clients in the TAU condition (see Table 4). Approximately 65% of clients at the individual level reported reliable or clinical change in the feedback condition compared with approximately 31.6% in the TAU condition. At the couple level (only couples where both partners met the criteria were included), 44.4% of the couples in the feedback condition compared with 15.8% of the couples in the TAU condition reported reliable or clinically significant change. Four times as many couples in the feedback condition were categorized as obtaining clinically significant change. It is important to note that only 15 couples in the feedback condition and 11 couples in the TAU condition were eligible to achieve clinical significance (both partners had pre-ORS < 25); if this is considered, then 53.3% of the eligible couples in the feedback condition achieved clinical significance and 18.2% achieved clinical significance in the TAU condition. Chi-square analyses indicated that the differences in the outcome classifications across treatment conditions were statistically significant at both the individual,  $\chi^2(3, N = 92) = 10.42$ ,

TABLE 4. Individuals and Couples That Achieved Clinical Significance or Reliable Change in the Client Feedback and Treatment as Usual Conditions

Classification	Individuals				Couples			
	Client feedback (n = 54)		Treatment as usual (n = 38)		Client feedback (n = 27)		Treatment as usual (n = 19)	
	n	%	n	%	n	%	n	%
1. Deteriorated	4	7.4	4	10.4	1	3.7	1	5.3
2. No change	15	27.8	22	57.9	3	11.1	9	47.4
3. Reliable change	9	16.7	2	5.3	4	14.8	1	5.3
4. Clinically significant change	26	48.1	10	26.3	8	29.6	2	10.5
5. Not classified					11	40.7	6	31.6

Note. Couples were classified only if both partners completed treatment in the same category.

$p = .02$ , and couple,  $\chi^2(3, N = 46) = 8.18, p = .04$ , levels.

## Discussion

An impressive amount of research has accumulated that supports the efficacy of using continuous outcome assessment (i.e., client feedback) in individual psychotherapy. Little research has been conducted to evaluate whether using client feedback in couple therapy would yield similar results. Our study investigated whether the benefits of using a client feedback system, PCOMS (Duncan et al., 2004), would extend to couples in therapy. Results indicated that couples randomly assigned to a feedback condition experienced statistically significant more improvement than those in the TAU condition and also improved more quickly as evidenced by a steeper growth curve. Couples in the feedback condition were also more likely to incur reliable and clinically significant change. The results of this study are comparable with the studies that have used PCOMS with individuals to address outcome (Miller et al., 2005; Reese et al., 2009) and with the only other couple therapy study using PCOMS (Anker et al., 2009).

In our study, couples in the feedback condition experienced treatment gains more than double of those in the TAU condition on the ORS (8.58 points vs. 3.64 points). Findings from Anker et al. (2009) found almost identical treatment gains when comparing feedback and TAU conditions (8.3 points vs. 3.11 points). Effect sizes for the feedback condition in both studies were also found to be large ( $d > 0.8$ ; Cohen, 1992). In addition, the effect size for the difference between the feedback condition and TAU condition ( $d = 0.48$ ) was also similar to the Anker et al. (2009) effect size ( $d = 0.50$ ).

When observing treatment effectiveness from a clinical significance perspective, clients in the feedback condition were more likely to experience clinically significant change (48.1%) compared with those in the TAU condition (26.3%). Similar to the Anker et al. (2009) study, we found that 4 times as many couples in the feedback condition were classified as having incurred clinically significant change when compared with those in the TAU condition. The results for the feedback condition are comparable to Shadish and Baldwin's (2003) summary of meta-analytic studies that reported 40–50% of clients in couple

therapy achieved clinical significance. At the couple level, approximately 30% of the couples in the feedback condition that completed treatment were classified as clinically significant compared with only 10.5% in the TAU condition. These rates are much lower than those reported by Christensen et al. (2004), who found that 52% of couples that received a form of behavior therapy reported clinically significant change. This comparison is problematic, however, on the surface. The couples in the Christensen et al. study were chronic and more distressed and also attended more sessions (22.9 sessions vs. 5.9 sessions) than the couples in our sample. When couples that were ineligible for clinical significance (pre-ORS scores  $>25$ ) are removed from our sample, the rate of achieving clinical significance is 53.3% for couples in the feedback condition, which is comparable to the Christensen et al. study.

The differences between pre- and post-ORS scores for couples are similar to client feedback studies that used PCOMS compared with TAU for individual psychotherapy. Although the current differences in outcome are slightly smaller than the 10.8-point gain found in Miller et al. (2005) and the 12.69- and 10.83-point gains found in two samples by Reese et al. (2009), the difference in scores between treatment conditions (at least double) is similar.

Although the results of our study closely resemble the Anker et al. (2009) couple study, there are two differences of note. First, the therapists in our sample were all graduate trainees and the therapists in the Anker et al. study were licensed professionals. Second, the trainees in our study received much less training (1 hr vs. 8 hr). The trainees in our study, however, did have supervisors who were able to provide continued instruction and discuss couple progress and the ORS and SRS measures. Our study provides evidence that client feedback is useful for therapy trainees who provide couple therapy.

## Limitations of Our Study

There are multiple limitations of our study that warrant mentioning. First, many clients in the feedback condition had missing session data. We were not able to discern a pattern for the missing data other than from anecdotal evidence from therapists that indicated both logistical issues (“I forgot” or “I did not bring copies of the measures

to the session”) and clinical reasons (“It did not feel necessary every week” or “The couple had a crisis and it did not feel appropriate to use”). All clients had ORS and SRS data for at least half of their sessions, but the consequences of these missing data may have led to underestimating the effects of the feedback intervention. Concerns for this limitation are tempered by the similar results in our study compared with those in the Anker et al. (2009) study. Future research should investigate the influence of administering a continuous assessment system every session compared with every second or third session. Such a study would address the potential differential effects for some clinics that administer continuous assessment systems every few sessions rather than every session.

A second limitation is that we did not use multiple outcome measures, such as marital satisfaction or couple distress. We are also unable to extrapolate the results of our study to assume that larger treatment gains in the feedback condition resulted in more couples remaining together when compared with the couples in the TAU condition. The findings of our study, however, are very similar to those in the Anker et al. (2009) study, which provide evidence of the validity for the ORS in couple therapy. Treatment gains as measured by the ORS in their study also showed treatment gains on an established measure of marital adjustment, the Locke–Wallace Marital Adjustment Test (Locke & Wallace, 1959). Couples who had better outcomes as measured by the ORS were also more likely to remain together at a 6-month follow-up.

A third limitation is the lack of consistently monitored treatment integrity. The appealing qualities of PCOMS are that it is easy to implement and provides the therapist with latitude as to how to best integrate the measures into treatment. Although conducting a study in a naturalistic setting is a strength, a weakness is being uncertain of the differences in how PCOMS was implemented. The effects of feedback may have been underestimated. Anecdotally, therapists reported differing levels of allegiance to using PCOMS. This concern is tempered by the fact that therapists received weekly supervision and supervisors were encouraged to ensure protocol compliance by noting the use of PCOMS in video-recorded sessions (all sessions were recorded) and through verbal reminders in supervision. Also, the therapists in the TAU condition

were aware of PCOMS, and some expressed frustration with not being able to use it with their clients. It is possible that they may have been applying components of the system verbally with their clients. This possibility, however, was not monitored or evaluated.

A fourth concern is that only therapist trainees were used, thereby limiting the generalizability of the results to therapists with more experience. More experienced therapists may have used PCOMS more effectively and demonstrated larger treatment gains, or conversely, the lack of experience may have heightened the demand characteristics and led to an overestimate of treatment effects. For example, couples were aware that their therapist was a trainee being evaluated and perhaps did not want to negatively influence the student’s grade. We do not believe, however, that this is a large concern. A previous study we conducted (Reese et al., 2009) did not show treatment outcome differences between licensed, professional staff and trainees. In addition, the Anker et al. (2009) couple study found similar treatment outcomes with experienced therapists.

#### *Future Research and Conclusions*

The use of PCOMS with both individuals and couples appears to have much promise, but more research is needed to clarify the variables and mechanisms of change associated with the positive outcomes found in studies using PCOMS. There is little understanding of why PCOMS leads to better outcomes, and until these processes are better understood, the confidence one can attribute to the specific effects of PCOMS is limited. Continuous assessment originally was designed to identify clients not progressing as expected. It would logically follow that being able to identify clients early in treatment who are not improving would afford the therapist the opportunity to alter treatment. However, PCOMS has been found to work with all clients, including those progressing as expected. Most of the research using the OQ45 has found that continuous assessment is more beneficial for clients not progressing as expected (e.g., Whipple et al., 2003). PCOMS differs from Lambert and colleagues’ (1996) signal system because PCOMS uses a measure of the therapeutic relationship every session, whereas the Lambert et al. system uses an alliance measure when a client is not improving as expected. Does this difference matter? The



authors of PCOMS (Duncan et al., 2004) have opined that having access to weekly feedback regarding the relationship may serve to heighten attention and focus on the therapeutic alliance and promote active collaboration. Future research should attempt to isolate the contribution of the SRS to the effectiveness of PCOMS.

PCOMS may also be effective because seeing the measures weekly creates expectancy effects regarding improvement. A second possibility is that cognitive dissonance plays a role in reporting improved outcome, and a good therapeutic relationship may increase the likelihood that clients feel better and have a good therapeutic relationship. Another possibility is that seeing one's graphed progress promotes improvement; there is an established body of literature in psychology that points to the importance of receiving feedback for promoting behavioral change (e.g., Alvero, Bucklin, & Austin, 2001). Using PCOMS may stimulate such behavioral changes for both clients and therapists. Studies that manipulate the manner in which therapist and client observe the data would help address these issues. Similar studies have been conducted with the OQ45 (e.g., Harmon et al., 2007; Hawkins et al., 2004), but as mentioned earlier, the findings have been mixed.

Future research should also investigate whether demand characteristics influence scores on the ORS and SRS given that the measures are completed and then directly discussed with the therapist. The SRS, in particular, appears susceptible to social desirability because the client is evaluating the quality of the session in the presence of the therapist. It is possible that the demand characteristics are heightened in couple therapy given the added presence of a partner, which makes the results even more "public." Perhaps simply the process of asking the client to evaluate the session is pivotal because it overtly communicates that the therapist values the client's input and the scores from the measure are less important. To evaluate the role of social desirability, a study could have three randomized treatment conditions use PCOMS. One condition would use PCOMS as typically prescribed, the second condition would complete the SRS without the therapist present and see the results after the client has left, and the third group would complete the SRS without the therapist having access to the results. Differences in the scores and outcome could be assessed.

With the increased need to demonstrate psychotherapy's utility due to forces such as managed care and third-party reimbursement, measuring the progress of treatment as it occurs has become an important area of study with exciting results. Ongoing client feedback has been found to help avoid premature termination and meet the needs of clients in a more effective, efficient manner. Overall, the results of this study indicate that using client feedback is a useful approach with couples that received treatment at a graduate training clinic and are consistent with the findings from previous client feedback studies focused on individual therapy. More research needs to be conducted, but PCOMS appears to hold much promise for use with couples given its ease of use and encouraging results.

## References

- ACKERMAN, S., SMITH, B., BEUTLER, L. E., GELSO, C. J., GOLDFRIED, M. R., HILL, C., & RAINER, J. (2001). Empirically supported therapy relationships: Conclusions and recommendations of the Division 29 Task Force. *Psychotherapy Theory, Research, Practice, Training, 38*, 495–497. doi:10.1037/0033-3204.38.4.495
- ALVERO, A. M., BUCKLIN, B. R., & AUSTIN, J. (2001). An objective review of the effectiveness and essential characteristics of performance feedback in organizational settings. *Journal of Organizational Behavior Management, 21*, 3–29. doi:10.1300/J075v21n01\_02
- ANKER, M. G., DUNCAN, B. L., & SPARKS, J. A. (2009). Using client feedback to improve couple therapy outcomes: A randomized clinical trial in a naturalistic setting. *Journal of Consulting and Clinical Psychology, 77*, 693–704. doi:10.1037/a0016062
- BALDWIN, S. A., BERKELJON, A., ATKINS, D. C., OLSEN, J. A., & NIELSEN, S. L. (2009). Rates of change in naturalistic psychotherapy: Contrasting dose-effect and good-enough level models of change. *Journal of Consulting and Clinical Psychology, 77*, 203–211. doi:10.1037/a0015235
- BARKHAM, M., MARGISON, F., LEACH, C., LUCOCK, M., MELLOR-CLARK, J., EVANS, C., & MCGRATH, G. (2001). Service profiling and outcome benchmarking using the CORE-OM: Toward practice-based evidence in the psychological therapies. *Journal of Consulting and Clinical Psychology, 69*, 184–196. doi:10.1037/0022-006X.69.2.184
- CHRISTENSEN, A., ATKINS, D. C., BERNIS, S., WHEELER, J., BAUCOM, D., & SIMPSON, L. E. (2004). Traditional versus integrative behavioral couple therapy for significantly and chronically distressed married couples. *Journal of Consulting and Clinical Psychology, 72*, 176–191. doi:10.1037/0022-006X.72.2.176
- COHEN, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159. doi:10.1037/0033-2909.112.1.155
- DEROGATIS, L. R. (1992). *Administration, scoring and procedures manual II*. Baltimore: Clinical Psychometric Research.

- DUNCAN, B. L., MILLER, S. D., REYNOLDS, L., SPARKS, J., CLAUD, D., BROWN, J., & JOHNSON, L. D. (2003). The Session Rating Scale: Psychometric properties of a "working" alliance scale. *Journal of Brief Therapy, 3*, 3–12. Retrieved from <http://www.journalbrieftherapy.com/>
- DUNCAN, B. L., MILLER, S. D., & SPARKS, J. A. (2004). *The heroic client: A revolutionary way to improve effectiveness through client-directed, outcome-informed therapy*. San Francisco: Jossey-Bass.
- ENDERS, C. K., & TOFIGHI, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods, 12*, 121–138. doi:10.1037/1082-989X.12.2.121
- FEINGOLD, A. (2009). Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis. *Psychological Methods, 14*, 43–53. doi:10.1037/a0014699
- GOLLAN, K. K., & JACOBSON, N. S. (2002). Developments in couple therapy research. In H. A. Liddle, D. A. Santisteban, R. F. Levant, & J. H. Bray (Eds.), *Family psychology: Science-based interventions* (pp. 105–122). Washington, DC: American Psychological Association.
- HARMON, C. S., LAMBERT, M. J., SMART, D. M., HAWKINS, E., NIELSEN, S. L., SLADE, K., & LUTZ, W. (2007). Enhancing outcome for potential treatment failures: Therapist–client feedback and clinical support tools. *Psychotherapy Research, 17*, 379–392. doi:10.1080/10503300600702331
- HAWKINS, E. J., LAMBERT, M. J., VERMEERSCH, D. A., SLADE, K. L., & TUTTLE, K. C. (2004). The therapeutic effects of providing patient progress information. *Psychotherapy Research, 14*, 308–327. doi:10.1093/ptr/kph027
- HORVATH, A. O., & SYMONDS, B. D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology, 38*, 139–149. doi:10.1037/0022-0167.38.2.139
- HOWARD, K. L., KOPTA, S. M., KRAUSE, M. S., & ORLINKSY, D. E. (1986). The dose-effect relationship in psychotherapy. *American Psychologist, 41*, 159–164. doi:10.1037/0003-066X.41.2.159
- HOX, J. (2002). *Multilevel analysis techniques and applications*. Mahwah, NJ: Erlbaum.
- ISAKSON, R. L., HAWKINS, E. J., HARMON, S. C., SLADE, K., MARTINEZ, J. S., & LAMBERT, M. J. (2006). Assessing couple therapy as a treatment for individual distress: When is referral to couple therapy contraindicated? *Contemporary Family Therapy, 28*, 313–322. doi:10.1007/s10591-006-9008-9
- 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. doi:10.1037/0022-006X.59.1.12
- LAMBERT, M. J., HANSEN, N. B., & BAUER, S. (2008). Assessing the clinical significance of outcome results. In A. M. Nezu & C. M. Nezu (Eds.), *Evidence-based outcome research: A practical guide to conducting randomized controlled trials for psychosocial interventions* (pp. 359–378). New York: Oxford University Press.
- LAMBERT, M. J., HANSEN, N. B., UMPHRESS, V., LUNNEN, K., OKISHI, J., BURLINGAME, G., & REISINGER, C. (1996). *Administration and scoring manual for the OQ 45.2*. Stevenson, MD: American Professional Credentialing Services.
- LAMBERT, M. J., HARMON, C., SLADE, K., WHIPPLE, J. L., & HAWKINS, E. J. (2005). Providing feedback to psychotherapists on their patients' progress: Clinical results and practice suggestions. *Journal of Clinical Psychology, 61*, 165–174. doi:10.1002/jclp.20113
- LAMBERT, M. J., & OGLES, B. (2004). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 139–193). New York: Wiley.
- LAMBERT, M. J., WHIPPLE, J. L., HAWKINS, E. J., VERMEERSCH, D. A., NIELSEN, S. L., & SMART, D. W. (2003). Is it time for clinicians to routinely track patient outcome? A meta-analysis. *Clinical Psychology, 10*, 288–301. doi:10.1093/clipsy/bpg025
- LAMBERT, M. J., WHIPPLE, J. L., SMART, D. W., VERMEERSCH, D. A., NIELSEN, S. L., & HAWKINS, E. J. (2001). The effects of providing therapists with feedback on patient progress during psychotherapy: Are outcomes enhanced? *Psychotherapy Research, 11*, 49–68. doi:10.1093/ptr/11.1.49
- LAMBERT, M. J., WHIPPLE, J. L., VERMEERSCH, D. A., SMART, D. W., HAWKINS, E. J., NIELSEN, S. L., & GOATES, M. (2002). Enhancing psychotherapy outcomes via providing feedback on client progress: A replication. *Clinical Psychology and Psychotherapy, 9*, 91–103. doi:10.1002/cpp.324
- LOCKE, H. J., & WALLACE, K. M. (1959). Short marital-adjustment and prediction tests: Their reliability and validity. *Marriage and Family Living, 21*, 251–255. Retrieved from <http://www.jstor.org>
- LUTZ, W., MARTINOVICH, Z., & HOWARD, K. (1999). Patient profiling: An application of random coefficient regression models to depicting the response of a patient to outpatient psychotherapy. *Journal of Consulting and Clinical Psychology, 67*, 571–577. doi:10.1037/0022-006X.67.4.571
- MILLER, S. D., & DUNCAN, B. L. (2004). *The Outcome and Session Rating Scales: Administration and scoring manual*. Ft. Lauderdale, FL: Author.
- MILLER, S. D., DUNCAN, B. L., BROWN, J., SORRELL, R., & CHALK, B. (2006). Using outcome to inform and improve treatment outcomes. *Journal of Brief Therapy, 5*, 5–22. Retrieved from <http://www.journalbrieftherapy.com/>
- MILLER, S. D., DUNCAN, B. L., BROWN, K., SPARKS, J., & CLAUD, D. (2003). The Outcome Rating Scale: A preliminary study of the reliability, validity, and feasibility of a brief visual analog measure. *Journal of Brief Therapy, 2*, 91–100. Retrieved from <http://www.journalbrieftherapy.com/>
- MILLER, S. D., MEE-LEE, D., PLUM, W., & HUBBLE, M. A. (2005). Making treatment count: Client-directed, outcome informed clinical work with problem drinkers. In J. L. Lebow (Ed.), *Handbook of clinical family therapy* (pp. 281–308). Hoboken, NJ: Wiley.
- MUTHEN, B. O. (1999, October 29). *Re: Timepoints* [Online discussion]. Retrieved from <http://www.statmodel.com/discussion/messages/14/20.html#POST16727>
- MUTHEN, B. O., & CURRAN, P. J. (1997). General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation. *Psychological Methods, 2*, 371–402. doi:10.1037/1082-989X.2.4.371

- PEUGH, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology, 48*, 85–112. doi:10.1016/j.jsp.2009.09.002
- RAUDENBUSH, S. W., & BRYK, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (6th ed.). Newbury Park, CA: Sage.
- REESE, R. J., NORSWORTHY, L. A., & ROWLANDS, S. (2006, August). *Does a popular client feedback model improve psychotherapy outcome?* Poster session presented at the annual meeting of the American Psychological Association, New Orleans, LA.
- REESE, R. J., NORSWORTHY, L., & ROWLANDS, S. (2009). Does a continuous feedback model improve psychotherapy outcomes? *Psychotherapy Theory, Research, Practice, Training, 46*, 418–431. doi:10.1037/a0017901
- SHADISH, W. R., & BALDWIN, S. A. (2003). Meta-analysis of MFT interventions. *Journal of Marital and Family Therapy, 29*, 547–570. doi:10.1111/j.1752-0606.2003.tb01694.x
- SHADISH, W. R., & BALDWIN, S. A. (2005). Effects of behavioral marital therapy: A meta-analysis of randomized controlled trials. *Journal of Consulting and Clinical Psychology, 73*, 6–14. doi:10.1037/0022-006X.73.1.6
- SHADISH, W. R., MONTGOMERY, L., WILSON, P., WILSON, M., BRIGHT, I., & OKWUMABUA, T. (1993). The effects of family and marital psychotherapies: A meta-analysis. *Journal of Consulting and Clinical Psychology, 61*, 992–1002. doi:10.1037/0022-006X.61.6.992
- SNYDER, D. K., CASTELLANI, A. M., & WHISMAN, M. A. (2006). Current status and future directions in couple therapy. *Annual Review of Psychology, 57*, 317–344. doi:10.1146/annurev.psych.56.091103.070154
- TAMBLING, R. B., & JOHNSON, L. N. (2008). The relationship between stages of change and outcome in couple therapy. *American Journal of Family Therapy, 36*, 229–241. doi:10.1080/01926180701290941
- U.S. Department of Education: Institute of Education Sciences. (2008). *What works clearing house: Procedures and standards handbook* (version 2.0). Retrieved from [http://ies.Ed.gov/ncee/wwc/pdf/wwc\\_procedures\\_v2\\_standards\\_handbook.pdf](http://ies.Ed.gov/ncee/wwc/pdf/wwc_procedures_v2_standards_handbook.pdf)
- WAMPOLD, B. E. (2001). *The great psychotherapy debate: Models, methods, and findings*. Hillsdale, NJ: Erlbaum.
- WHIPPLE, J. L., LAMBERT, M. J., VERMEERSCH, D. A., SMART, D. W., NIELSEN, S. L., & HAWKINS, E. J. (2003). Improving the effects of psychotherapy: The use of early identification of treatment failure and problem-solving strategies in routine practice. *Journal of Counseling Psychology, 50*, 59–68. doi:10.1037/0022-0167.50.1.59
- WILLETT, J. B., SINGER, J. D., & MARTIN, N. C. (1998). The design and analysis of longitudinal studies of development and psychopathology in context: Statistical models and methodological recommendations. *Development and Psychopathology, 10*, 395–426. doi:10.1017/S0954579498001667