

DOES A CONTINUOUS FEEDBACK SYSTEM IMPROVE PSYCHOTHERAPY OUTCOME?

ROBERT J. REESE
University of Kentucky

LARRY A. NORSWORTHY AND
STEVE R. ROWLANDS
Abilene Christian University

Using outcome data on a continual basis to monitor treatment progress has been identified as a way to enhance psychotherapy outcome. The purpose of this study was to investigate the use of a continuous feedback assessment system, the Partners for Change Outcome Management System (PCOMS; Miller & Duncan, 2004). Findings from 2 client samples that attended individual therapy at a university counseling center (N = 74) or a graduate training clinic (N = 74) indicated that clients who used PCOMS with their therapists (feedback condition) demonstrated statistically significant treatment gains when compared to clients receiving treatment as usual (no-feedback condition). Clients using PCOMS were also more likely to experience reliable change and in fewer sessions. A survival analysis demonstrated that approximately 50% of the clients in the feedback condition demonstrated reliable change after the 7th (graduate training clinic) or 9th session (university counseling center). Further find-

ings, limitations of the study and ideas for future research are discussed.

Keywords: psychotherapy, counseling effectiveness, client progress and outcome, feedback information

Research has suggested that outcome assessment conducted on a continuous basis can lead to increased effectiveness, particularly for identifying clients who are not progressing in therapy as expected (e.g., Beutler, 2001; Lambert, Hansen, & Finch, 2001; Lueger et al., 2001). The research findings are so compelling that the American Psychological Association's (APA) Division 29 Task Force on Empirically Supported Relationships advised practitioners "to routinely monitor patients' responses to the therapy relationship and ongoing treatment. Such monitoring leads to increased opportunities to repair alliance ruptures, to improve the relationship, and to avoid premature termination" (Ackerman et al., 2001, p. 496). Continuous feedback provided by clients can help therapists determine when therapy is not effective, identify client deterioration, and provide opportunities to make the necessary changes to the treatment or approach (Lambert et al., 2001). In addition, Brown and Jones (2005) stated that continuous client feedback can assist in improving the therapeutic relationship and in making referrals when appropriate.

Lambert and his colleagues at Brigham Young University have accumulated an impressive body of evidence for the utility of tracking outcome across treatment (e.g., Lambert, Whipple, et al., 2001; Whipple et al., 2003). Lambert, Whipple, et al.'s (2001) system is based on using the Outcome Questionnaire 45 (OQ45; Lambert et al., 1996) that is completed by the client to measure weekly outcome and to provide feedback to the therapist on progress or lack thereof. Scores are tracked using a "signal system" that identifies if a

Robert J. Reese, Department of Educational, School, and Counseling Psychology, University of Kentucky; Larry Norsworthy, Department of Psychology, Abilene Christian University; and Steve Rowlands, University Counseling Center, Abilene Christian University.

Correspondence regarding this article should be addressed to Robert J. Reese, PhD, Department of Educational, School, and Counseling Psychology, 235 Dickey Hall, University of Kentucky, Lexington, KY 40506-0017. E-mail: jeff.reese@uky.edu

client is progressing as expected. If not, the therapist is to intervene by changing the course of treatment and/or assessing the reason(s) for the lack of progress. A meta-analysis (Lambert, Whipple, & Hawkins, 2003) that summarized three of Lambert and colleagues' previous studies found a medium effect size of .39 across the studies when comparing the treatment gains of clients identified as deteriorating (had declined, on average, half of a standard deviation) who were in the feedback group (therapists were provided feedback) versus the no-feedback group. Using continuous assessment to identify clients that are not benefiting from therapy has consistently been found to increase the likelihood of "turning things around" in therapy.

Research on continuous assessment has been extended to investigate how to maximize the use of outcome data. Examples include assessing whether sharing outcome assessment results with therapists and clients is more effective than sharing the results with only therapists (Harmon et al., 2007; Hawkins, Lambert, Vermeersch, Slade, & Tuttle, 2004), incorporating measures of the therapeutic alliance, stages of change, and social support to increase effectiveness (Whipple et al., 2003), and assessing if using continuous outcome data is beneficial for all clients instead of only clients identified as "not-on-track" (Lambert et al., 2003).

Two studies (Harmon et al., 2007; Hawkins et al., 2004) have examined whether providing data on treatment progress to both therapist and client influences effectiveness. Hawkins et al. (2004) found that providing feedback data on treatment progress to both clients and therapists was associated with statistically significant gains in treatment outcome. However, Harmon et al. (2007) failed to replicate these results, finding no incremental effectiveness by allowing the client to see the feedback results. Studies by Whipple et al. (2003) and Harmon et al. found that adding measures of the therapeutic alliance, motivation to change, and perceived social support for clients identified as not-on-track via continuous assessment demonstrated incremental effectiveness over just using continuous feedback alone. Deterioration rates were reduced from 21% to 8%. Successful outcome rates increased from 21% to 50%.

Percevic, Lambert, and Kordy (2004) found that most continuous assessment studies indicated providing continuous feedback to therapists

has no effect on clients with predicted positive outcome. In addition, the collective results from four large-scale studies suggest that measuring, monitoring, and predicting treatment failure enhances treatment outcome for clients who do not have high likelihood of positive outcomes but yield little impact for other clients (Lambert, Harmon, Slade, Whipple, & Hawkins, 2005). More recently, however, Harmon et al. (2007) found that using continuous assessment was helpful to all clients, although those not expected to make progress from receiving treatment benefited more. Taken as a whole, there appears to be solid evidence that regularly monitoring client progress increases the likelihood of a client staying with treatment and having a positive treatment outcome.

Building on the extant psychotherapy outcome literature, Miller and Duncan (2004) developed a feedback system called the Partners for Change Outcome Management System (PCOMS) that uses two brief measures to track outcome and the counseling relationship in every session. The system can be used in individual, couples, family, or group therapy formats. Much of the system's appeal is that the scales used to measure outcome and the counseling relationship are much shorter than traditional outcome and therapeutic alliance measures. The Outcome Rating Scale (ORS; Miller & Duncan, 2000) and the Session Rating Scale (SRS; Miller, Duncan, & Johnson, 2000) are both four-item measures developed to track outcome and the therapeutic alliance, respectively. The proposed advantage is that the measures' brevity makes implementation by clinicians more likely. Although other factors (e.g., training and treatment allegiance) impact compliance with using outcome measures, there is some evidence that suggests that the amount of time required to complete measures does matter. Miller, Duncan, Brown, Sparks, and Claud (2003) reported that compliance rates for the ORS and SRS at one site was 86% compared to 25% at another site using the OQ45.

The development of PCOMS was based on Lambert Whipple, et al. (2001) continuous assessment system using the OQ45. In addition to the discrepancy in length of the measures, other substantial and potentially important differences exist. First, PCOMS is viewed as part of the therapy process. During sessions the therapist and client examine the feedback data together. Although there is research on clients seeing their

OQ45 outcome data (e.g., Hawkins et al., 2004) a comprehensive literature review did not reveal any studies that examined the process of therapist and client going over feedback data collaboratively. Second, PCOMS assesses the therapeutic relationship every session, with every client. Research on using the OQ45 (e.g., Whipple et al., 2003) has investigated the impact of assessing the therapeutic relationship, but only when there is a lack of progress in treatment. Duncan et al. (2003) argued that one advantage of assessing the relationship every session, particularly early in treatment, is that it allows for immediate response within the session.

Initial results from research by Duncan and Miller on PCOMS have been positive; finding that its use has resulted in fewer premature terminations and increased effectiveness. For example, in a study that had 6,424 clients who received services through a telephone-based employee assistance program, Miller, Duncan, Brown, Sorrell, and Chalk (2006) found that effect sizes increased from .37 to .79 when their system was implemented. All measures were administered over the telephone. Miller and Duncan also reported that their measures generate reliable and valid scores (Duncan et al., 2003; Miller et al., 2003). Although the number of clinicians using this feedback system has been increasing, little research has been conducted to replicate their findings. In addition, much of their evidence cited is based on samples that received services via the telephone.

We sought to replicate their research by conducting two studies that utilized PCOMS with clients in a university counseling center (Study 1) and a community-based graduate training clinic (Study 2). Specifically, we had three hypotheses. First, we hypothesized that clients in a feedback condition (used PCOMS) would exhibit greater pre/postresidual treatment gains on ORS scores compared to clients in a no-feedback condition that did not use PCOMS. Second, we hypothesized that more clients in a feedback condition would experience reliable change than clients in a no-feedback condition as measured by the ORS. Third, we posited that clients in the feedback condition would demonstrate reliable change more quickly (i.e., have a steeper dose-response curve) than those assigned to a no-feedback condition.

Method

Participants

Clients. Study 1 was composed of clients ($N = 74$) that received individual therapy at a university counseling center (UCC) on the campus of a small-to-medium, private southwestern university during the course of an academic year. The UCC serves enrolled university students. Initially the study included 131 participants, but 57 were excluded because they either did not return for a second session ($N = 24$; pretreatment ORS, $M = 19.24$, $SD = 8.90$) or did not comply with the treatment protocol ($N = 33$; pretreatment ORS, $M = 21.83$, $SD = 10.05$). Noncompliance occurred in two ways: participants in the feedback condition not completing the ORS and SRS in at least 50% of the sessions ($N = 5$; pretreatment ORS, $M = 21.65$; $SD = 8.68$) or the participants in the no-feedback condition failing to complete the posttest measures ($N = 28$; pretreatment ORS, $M = 21.86$, $SD = 10.30$). Such a compliance/completion rate is consistent with other continuous assessment studies that are therapist dependent (Whipple et al., 2003). An analysis of variance (ANOVA) did not find pretreatment functioning differences between those included in the study (ORS, $M = 19.93$; $SD = 8.50$) and the two excluded groups, $F(2, 128) = 0.23$, $p > .05$.

The final client sample consisted of 53 women and 18 men; three participants did not indicate their sex. The majority of the sample was White (78.4%), 4.1% was African American, 2.7% Asian American, 6.8% Hispanic/Latino, and 5.4% were international students. There were two participants who did not indicate ethnic/racial origin. The mean age was 20.17 years ($SD = 1.90$), with ages ranging from 18 to 27.

Although roughly half of the participants were originally randomly assigned to the feedback ($N = 60$) and no-feedback conditions ($N = 53$), 50 participants were in the final feedback group and 24 were in the no-feedback group. This disparity is due to the ease of having posttest data from the feedback group because data were collected every session coupled with the difficulty of getting therapists in the no-feedback group to remind participants to complete the posttest measure. The pretest measures of those in the no-feedback group who completed posttest data

were not statistically significantly different from those who failed to do so, $t(48) = .30, p > .05$.

Study 2 was composed of clients ($N = 74$) that received individual therapy over the course of an academic year at a graduate training clinic for a marriage and family therapy master's program (MFC). The MFC is located on the same campus of the UCC. Practicum students provide individual, couples, and family therapy services based on a sliding-scale fee for clients from the surrounding community. Only clients that received individual therapy were included in the study. Initially the study had 96 clients (52 = feedback condition, 44 = no-feedback condition), but 22 were excluded because they did not return for a second session ($N = 8$; pretreatment ORS, $M = 14.53, SD = 5.68$), complete the ORS and SRS measures consistently in the feedback condition ($N = 4$; pretreatment ORS, $M = 16.48, SD = 3.93$) or complete a measure at posttreatment in the no-feedback condition ($N = 10$; pretreatment ORS, $M = 18.71, SD = 8.20$). Pretreatment functioning mean comparisons of those included in the study compared to the two groups of participants excluded (i.e., attended only one session or did not complete the measures as directed) were not statistically significant, $F(2, 93) = 1.04, p > .05$.

The final client sample consisted of 51 women and 21 men; two clients did not indicate sex. The majority of the sample was White (79.6%), 3.7% was African American, 14.6% was Hispanic/Latino, and 2.1% did not indicate ethnicity/race. The mean age for clients was 32.96 ($SD = 12.32$) with ages ranging from 18 to 69. The ORS was administered every session for Study 2 to facilitate the collection of post-treatment data for the no-feedback condition. This change appeared to have been marginally helpful, with 45 clients in the feedback condition and 29 in the no-feedback condition. The mean pretreatment ORS score for those that did not complete the ORS at posttreatment (19.78) in the no-feedback condition was nearly identical to those that did (19.64).

Therapists. The therapists in Study 1 were comprised of both professional staff and practicum students at the UCC. There were five professional staff members, all master's level practitioners with a mean of 8 years of experience, and five second year practicum students (second or third practicum) enrolled either in a master's counseling or clinical psychology program. Staff

members provided 391 (72.41%) of the 540 sessions. For Study 2, all of the 446 sessions at the MFC were provided by 17 second-year practicum students enrolled in a master's marriage and family therapy program. Practicum students at both sites received weekly individual supervision. Theoretical orientations of the therapists in both studies consisted of cognitive-behavioral, family systems, solution-focused, or an integrated/eclectic approach.

Measures

ORS. The ORS (Miller & Duncan, 2000) consists of four items that are measured using a visual analog scale. The items were adapted from the three areas of the OQ45 (Lambert et al., 1996). Specifically, clients respond to how they are doing Socially (work, school, friendships), Interpersonally (family, close relationships), and Individually (personal well-being). An Overall (general sense of well-being) score is also obtained. Clients make a hash mark on each of the four analog scales that are 10 cm in length, with scores on the left side of the scale indicating lower functioning and scores on the right indicating higher functioning. Using a ruler to measure the distance from the left end of the scale to the client's hash mark, the score is recorded for each item. The scores are then totaled, ranging from 0 to 40.

Using a sample of 34,790 participants, a clinical cut-off score of 25 was determined (77th percentile for a nontreatment sample), meaning that clients who score below 25 are more typically found to benefit from therapy, whereas those scoring above 25 are more consistent with a nonclinical population and less likely to improve in psychotherapy (Miller & Duncan, 2004). Miller et al. (2003) also found that the ORS discriminates well among clients and nonclients.

Initial research has indicated that the ORS generates reliable scores among individuals who receive therapy in a community mental health center. Miller et al. (2003) conducted a psychometric study and reported an internal consistency coefficient of .93. Test-retest reliability from the first to second session was .60. The internal consistency for the ORS for the two current samples was .88 and .84, respectively. The test-retest reliability from the first to second session was .51 in Study 1 and .72 for Study 2. However, Vermeersch, Whipple, and Lambert (2004) reported

that it is likely that test–retest coefficients will be attenuated for outcome measures that are designed to be sensitive to change, particularly from the first repeat administration.

Evidence for construct validity (also from the Miller et al. study) found a correlation coefficient of .59 between the ORS and OQ45. Miller et al. (2003) also provided further evidence for construct validity because client gains across therapy were demonstrated. Lambert et al. (1996) stated that evidence for construct validity can be established by showing that scores differ from those obtained at the beginning of treatment.

SRS. The SRS (Miller et al., 2000) consists of four items that are measured via a visual analog scale. Based on Bordin's (1979) pantheoretical definition of the therapeutic alliance and an inclusion of the client's theory of change, the scale assesses the therapeutic relationship ("I felt heard, understood, and respected"), goals and topics covered in therapy ("We worked on or talked about what I wanted to work on or talk about"), the approach used in therapy ("The therapist's approach is a good fit for me"), and the overall rating of the session ("Overall, today's session was right for me"). Clients make a hash mark on each of the four analog scales that are 10 cm in length, with scores to the left of the scale indicating less satisfaction and scores on the right indicating higher satisfaction for each item. Once again, a ruler is used to measure the distance from the left end of the scale to the hash mark. The individual items are then recorded and totaled, ranging from 0 to 40. A clinical cut-score of 36, or if any one item is below a 9, is used to denote when there is/are problem(s) with the therapeutic alliance. Initial research has indicated the SRS generates reliable and valid scores. Duncan et al. (2003) found that with a sample of 337 community mental agency clients, the SRS had a coefficient alpha of .88 and possessed a correlation coefficient of .48 with the Helping Alliance Questionnaire–II (HAQ–II; Luborsky et al., 1996). Test–retest reliabilities averaged .74 across the first six sessions with the SRS compared to .69 for the HAQ–II. Internal consistency estimates for the current samples were .88 (Study 1) and .90 (Study 2). The SRS test–retest coefficient from Session 1 to Session 2 was .66 (Study 1) and .54 (Study 2), which is comparable to Miller et al.'s (2003) finding of .60.

Miller and Duncan (2004) found that increases on the SRS during the course of treatment were

statistically significantly associated with better outcome. When compared to clients who did not use the SRS, clients who used the SRS were three more times likely to attend their next session and experienced more change during treatment. The cut-off score of 36 was derived from a sample of 15,000 clients of whom only 24% scored below 36 and were "at a statistically greater risk for dropping out of or experiencing a negative or null outcome from treatment" (Miller and Duncan, p. 14).

PCOMS

This study followed the protocol as outlined in the scoring and administration manual for PCOMS (Miller & Duncan, 2004). A client is administered the ORS at the beginning of every session with the therapist present. After completing the ORS (approximately one minute), the therapist scores the items with a ruler (or computer software is now available for administration and scoring) and totals up the items. The items are then charted on an ORS graph that indicates a client's progress, or lack thereof, across the course of treatment. A composite score below 25 indicates that a client has a level of distress consistent with people typically found in therapy. The scores can be used to frame content or to give a therapist an area to focus on in session. Discretion is given to the therapist to decide how to best integrate the scores within a given session. However, general guidelines are provided for how to proceed with clients that do not improve (less than a gain of 5 points), "deteriorate" during therapy (scores go down at least 5 points), have "reliable change" (a gain of 5 or more points) or demonstrate "clinically significant improvement" (i.e., demonstrating at least a 5 point gain and traversing the ORS cut-score of 25 during treatment). We have included a brief description of how to proceed with clients in each category (see Miller & Duncan, 2004, for a complete description).

No change. For a client that 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 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, frequency, mode, or even therapist if no improvement is noted after three sessions.

Reliable change. Treatment is going accordingly. Therapists are advised to reinforce changes and to continue treatment until progress begins to plateau, then a therapist should consider reducing the frequency of sessions.

Clinically significant change. The client is likely no longer struggling with issues that led to seeking therapy. Therapists are advised to consolidate changes, anticipate potential setbacks, and to consider reducing the frequency of sessions.

Toward the end of every session, the SRS is administered to the client and again scored by the therapist (approximately one minute). If the total score is below 36 or one of the items is below 9, the therapist intervenes and inquires about the reason for the lower scores. The total score is then charted on a graph for the corresponding session.

Procedure

Study 1. Clients for an academic year at the UCC were assigned by the director to either the feedback group or no-feedback condition via a randomized block design to help control for therapist effects. All of the clients were new clients at the UCC, not having received services there previously. Approximately half of the new clients assigned to therapists were in the feedback group and the other half were in the no-feedback group. Clients in the feedback condition completed the ORS at the beginning of each session and the SRS at the end of each session. Participants in the no-feedback condition completed the ORS at the beginning and end of treatment. The SRS was not administered to the no-feedback condition. There was a concern that exposing the clients to the items might unduly influence their perceptions/expectations of treatment, leading to a possible deviation from a "treatment as usual" paradigm. If a client had not completed treatment by the end of the academic year, the participant's

last ORS score was used as the posttreatment measure for the feedback condition and those in the no-feedback condition completed a post-ORS measure.

Study 2. This study took place the next academic year after Study 1. Clients at the MFC also used PCOMS but had two deviations from the Study 1 protocol. First, clients in the no-feedback condition completed the ORS at the beginning of each session, rather than just at the beginning and end of treatment, to help increase compliance with collecting posttreatment data. Doing this also allowed for comparison of outcome with clients not progressing in treatment, a proposed advantage of continually monitoring client outcome, and to compare dose response curves of both the feedback and no-feedback groups. The ORS results were not seen or scored by the therapist or shared with clients in the no-feedback condition. Second, therapists, rather than clients, were randomly assigned to the feedback and no-feedback conditions because the graduate faculty over the MFC felt that it would be too cumbersome and confusing for beginning practicum students to deviate from their normal treatment paradigm by alternating between the two conditions.

Therapists for both studies were trained to administer, score, and provide feedback to clients via the training manual provided for the ORS and SRS (Miller & Duncan, 2004). The first author of the current study conducted a 1-hr training session for the therapists and practicum supervisors. A summary handout was also provided to each therapist as a reminder of how to follow the protocol if needed. Two case studies were provided in the training to facilitate application of PCOMS. In the feedback condition, progress was tracked, charted, and discussed with the client every session. The no-feedback condition received treatment as usual and did not utilize PCOMS. For both studies, copies of ORS/SRS measures were made by therapists and placed in a collection box for the measures to be rescored before being entered into a database to ensure scoring accuracy. Any scoring errors were relayed to the therapist to correct the original copies kept in the client's file.

Results

Pre- and posttest ORS mean total scores and standard deviations for each of the treatment conditions in both studies can be observed in Table 1.

TABLE 1. Means and Standard Deviations of the ORS for the Feedback and No-Feedback Conditions

ORS total	Feedback		No feedback	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Study 1				
Pre	18.59	7.60	22.71	9.70
Post	31.28	6.63	29.53	7.26
Study 2				
Pre	18.68	10.39	19.64	6.46
Post	29.51	9.58	24.33	7.51

Note. ORS = Outcome Rating Scale.

To assess if pretreatment ORS mean scores were different for the feedback and no-feedback conditions, independent samples *t* tests were computed and found that the pretreatment mean differences were not statistically significant for either Study 1, $t(72) = 1.99, p > .05$, or Study 2, $t(72) = -0.49, p > .05$. This indicates that the initial random assignment appears to have been effective in creating equivalent groups for both samples. The mean SRS total scores for the feedback condition were both at the upper end of the continuum and had little variability (Study 1: $M = 35.94, SD = 4.22, \text{range} = 20.90$; Study 2: $M = 37.09, SD = 3.79, \text{range} = 14.00$). Participants in the feedback condition generally felt favorable about the alliance with their therapist.

Individuals in each study’s client feedback condition reported more treatment gains on the ORS compared to the no-feedback condition. The client feedback groups reported mean treatment gains of 12.69 (Study 1) and 10.84 (Study 2) points whereas the no-feedback groups reported mean treatment gains of 6.82 and 5.04, respectively. For Study 1, a repeated-measures ANOVA indicated that therapy gains were statistically significant across both groups, $F(1, 72) = 60.32, p = .00, \eta^2 = .46$, but the interaction between the treatment condition and time (pre-post) on the ORS total score indicated that those in the feedback condition had statistically significant more change than the no-feedback condition, $F(1, 72) = 5.46, p < .05, \eta^2 = .07$. For Study 2, a repeated-measures ANOVA with therapist added as a covariate (because therapists were assigned to either the feedback or no-feedback condition) indicated that therapy gains were statistically significant for all clients, $F(1, 72) = 22.76, p = .00, \eta^2 = .24$. The interaction between treatment condition and time also found that those in the feedback group experienced sta-

tistically significant more change than those in the no-feedback group, $F(1, 72) = 7.51, p < .01, \eta^2 = .10$. The therapist covariate was not statistically significant, $F(1, 72) = 1.10, p > .05, \eta^2 = .01$. Using a Cohen’s *d* to compute an effect size as is typically found in psychotherapy outcome studies that compare treatments, medium to large effect sizes were found for both Study 1 ($d = .54$) and Study 2 ($d = .49$).

Although clients in the feedback condition demonstrated larger treatment gains, they did not attend statistically significantly more sessions on average than the no-feedback condition in Study 1 (6.27 vs. 5.66), $t(72) = 0.51, p > .05$, or Study 2 (8.02 vs. 5.79), $t(72) = 1.74, p > .05$. We found it interesting that in Study 1 professional staff and practicum students had equivocal pre-post ORS treatment outcomes for clients that were seen in the feedback group, $F(1, 48) = .00, p > .05, \eta^2 = .00$, and for all clients irrespective of treatment condition, $F(1, 72) = .03, p > .05, \eta^2 = .00$.

Another common way to assess psychotherapy outcome is to view the number of clients who incur clinically significant change (Lambert, Hansen, & Bauer, 2008). Posited by Jacobson and Truax (1991), there are two criteria for establishing clinically significant, or meaningful, change in psychotherapy. The first criterion, “reliable change,” is the increase or decrease of a client’s score on an outcome measure that exceeds the measurement error for the instrument. The second criterion, “clinical significance,” requires reliable change and that the client started treatment in the clinical range and concluded treatment in the nonclinical range based on an established cut-score. Jacobson and Truax’s formulas were used to establish a reliable change index (RCI) of 5 points and a cut-score of 25 for the ORS that was based on two studies (Miller et al., 2003; Miller, Mee-Lee, & Plum, 2005) that used samples from a community mental health and a residential alcohol and drug treatment center, respectively. Specifically, reliable change is denoted by a 5-point increase indicating “improvement,” whereas a 5-point decrease is considered to indicate “deterioration.”

The less stringent criterion of reliable change was used in this study, because 28.4% of the university counseling center sample began treatment in the nonclinical range, and other researchers have suggested that using reliable change was appropriate for university counseling centers

given the likelihood that the population would generally report less distress (Snell, Mallinckrodt, Hill, & Lambert, 2001). In Study 1 (see Table 2), the continuous feedback condition incurred reliable change on the ORS more frequently when compared to the no-feedback condition (80% vs. 54.2%). A chi-square analysis found a statistically significant difference between the feedback and no-feedback groups, $\chi^2(1, N = 74) = 5.32, p < .05$. Very few clients in both groups reported deteriorating during treatment. In Study 2, a higher percentage of feedback condition participants (66.67%) also incurred reliable change compared to clients in the no-feedback condition (41.40%). A chi-square analysis also found a statistically significant difference, $\chi^2(1, N = 74) = 4.60, p < .05$. As in Study 1, few clients demonstrated deterioration across treatment.

We compared clients in both the feedback condition ($n = 16$) and no-feedback condition ($n = 11$) that were identified as not progressing (NP) to evaluate if feedback was helpful for clients at risk for poor outcome. This comparison could only be made in Study 2, because clients in the no-feedback condition completed the ORS every session. Consistent with the administration and scoring manual, a NP client was identified as having improved less than 5 points on the ORS after three sessions as (Miller & Duncan, 2004). The NP clients in the feedback condition showed larger treatment gains (6.06 vs. 2.48 points) at the end of treatment than the no-feedback condition. However, a repeated-measures ANOVA did not find statistically significant pre/posttreatment scores between the groups, $F(1, 25) = 1.59, p >$

.05, $\eta^2 = .07$. However, please note that the effect size of .07 is comparable to the effect size of .10 for the entire sample; the lack of statistical significance is likely a function of a small sample size. The NP clients in both the feedback and no-feedback conditions attended nearly the same number of sessions, 6.9 and 5.9 sessions, respectively.

One way of analyzing the dose-response curve is to assess when clients achieve reliable change as defined earlier. To do so, a survival analysis was computed, a nonparametric statistic commonly used with longitudinal data that provides an estimate of the percentage of clients that will demonstrate reliable change from session-to-session. The possibility exists that a client could incur a 5-point improvement and then regress in latter sessions. For this analysis, reliable change was only noted when there was no subsequent regression before treatment ended. First, an analysis using Cox regression was computed with therapist (professional staff vs. practicum students) selected as a covariate to see if multiple-survival curves needed to be computed. The overall goodness of fit chi-square value was statistically significant for the regression model, $\chi^2(2, N = 354) = 6.49, p < .05$. Therapist was statistically significant ($\beta = .434, \text{Wald} = 6.18, df = 3, p < .05$), indicating that clients assigned to professional staff demonstrated improvement more quickly than clients paired with a practicum student. A Kaplan–Meier survival analysis was then conducted to create separate survival curves for professional staff and practicum students to estimate the median number of sessions needed to acquire reliable change (see Figure 1). Data for clients that had not achieved reliable change were censored. The survival analysis found that 51% of the clients in the feedback condition were estimated to achieve reliable change after a median of nine sessions. For clients of professional staff, 50% of the clients were estimated to achieve reliable change after a median of seven sessions and clients of practicum students were estimated to take a median of 12 sessions.

A survival analysis was also conducted for Study 2 to assess the median number of sessions estimated for clients to obtain reliable change. Reliable change was used as the criterion to maintain consistency with Study 1 for comparison purposes. Four clients were removed from the analysis (two from each condition) because their initial ORS scores were above 35 and made

TABLE 2. Percentage of Clients in Feedback and No-Feedback Conditions Who Achieved Reliable Change at End of Treatment

Classification	Feedback		No feedback	
	<i>n</i>	%	<i>n</i>	%
Study 1				
Deteriorated	2	4.00	3	12.50
No change	8	16.00	8	33.30
Reliable change	40	80.00	13	54.20
Study 2				
Deteriorated	2	4.44	1	3.44
No change	13	28.89	16	55.16
Reliable change	30	66.67	12	41.40

Note. $\chi^2 = 4.60, p < .05$ (Study 1), $\chi^2 = 16.67, p < .01$ (Study 2).

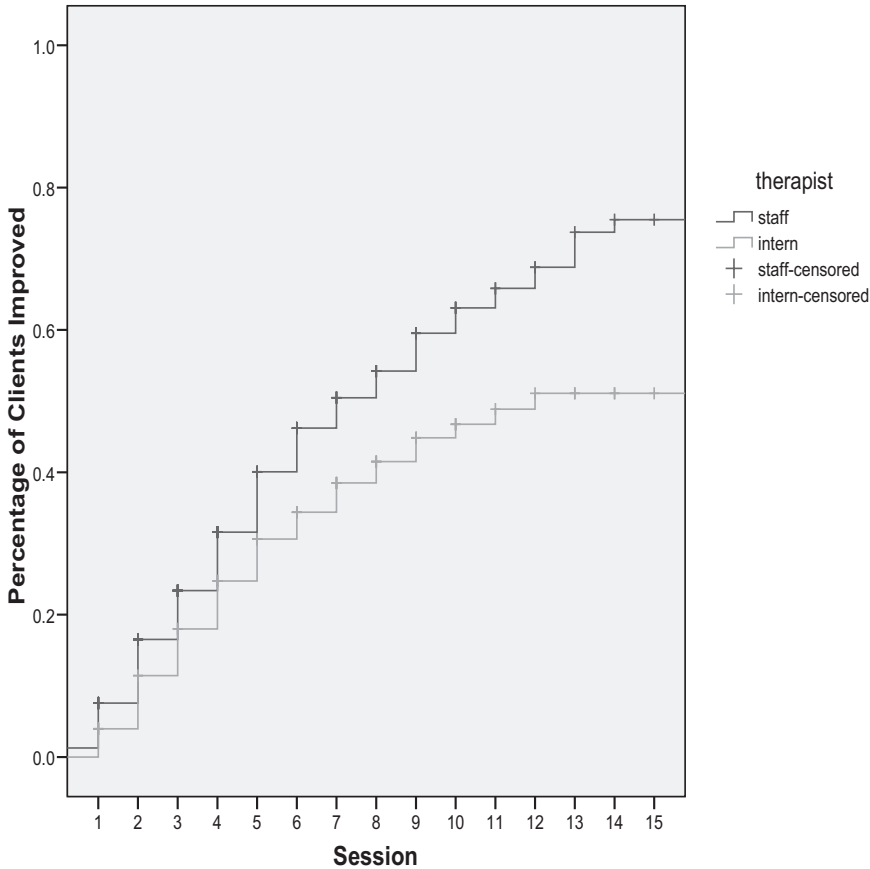


FIGURE 1. Study 1 survival plots of reliable change rates for clients of professional staff and practicum students in the feedback condition by session.

it impossible to achieve reliable change. A Cox regression model was computed to assess if there were differences in the recovery rates for the feedback and no-feedback conditions. The overall goodness of fit chi-square value was statistically significant, $\chi^2(1, N = 296), = 5.59, p < .05$, indicating that those in the both conditions achieved reliable change at different rates. A Kaplan–Meier survival analysis was conducted to view the survival curves for the feedback and no-feedback groups separately (see Figure 2). The survival analysis found that 56% of the clients in the feedback condition were estimated to achieve reliable change after a median of 7 sessions whereas 52% of the clients in the no-feedback condition were estimated to achieve reliable change after a median of 10 sessions.

Discussion

Two studies were conducted to evaluate an increasingly used continuous assessment system, PCOMS (Miller & Duncan, 2004), the first using a sample of psychotherapy clients in a university counseling center and the second a sample of psychotherapy clients in a community-based graduate training clinic. In general, both studies replicated the positive findings that Miller and Duncan reported in other studies (Miller et al., 2003; Miller, Duncan, Sorrell, & Brown, 2005). The results indicated that clients in the feedback condition (i.e., clients that completed an outcome and alliance measure every session and reviewed these results in session), reported more change than those in the no-feedback condition (i.e., received therapy in a treatment as usual format).

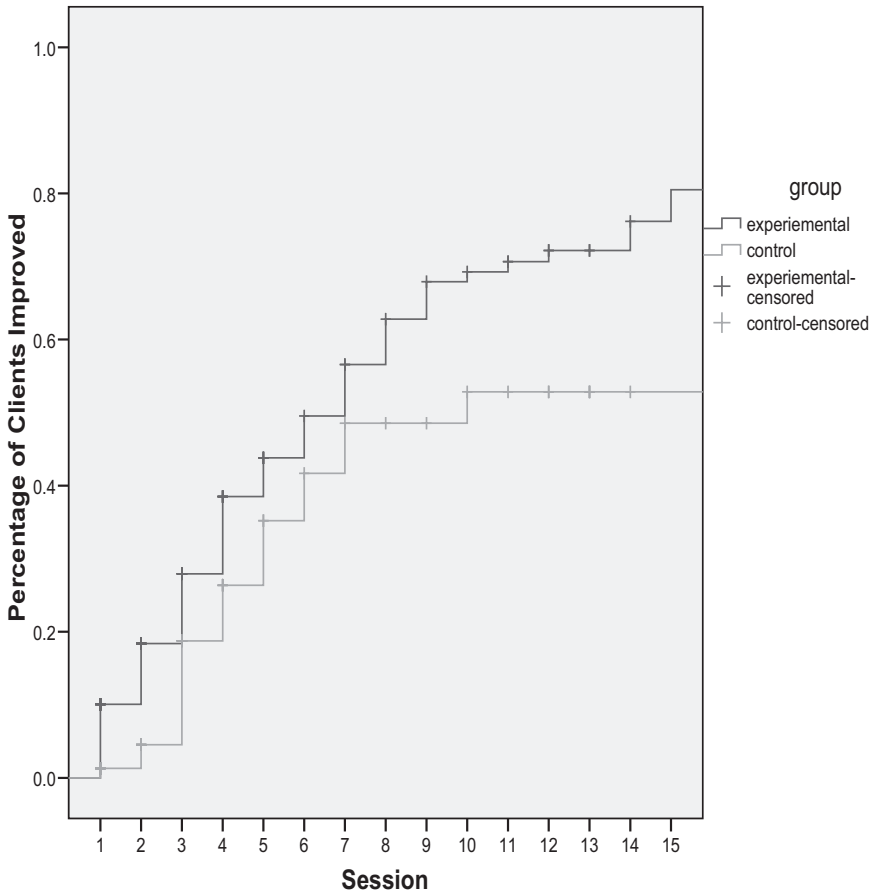


FIGURE 2. Study 2 survival plots of client reliable change rates for the feedback and no-feedback conditions by session.

The majority of clients in the feedback group evidenced reliable change by the end of treatment. Last, a survival analysis in Study 2 demonstrated that clients in the feedback condition were estimated to achieve reliable change in fewer sessions than those in the no-feedback condition.

Improved Outcome

Individuals in both the feedback and no-feedback conditions showed statistically significant improvement on pre/postmeasures of the ORS total score. In both studies the feedback group showed roughly twice as much improvement as the no-feedback group (12.69 points vs. 6.82 points in Study 1; 10.83 vs. 4.69 points in Study 2). This amount of improvement is similar to the 10.8 point gain that Miller, Duncan, et al. (2005) reported from a study

of those who completed a treatment program for alcohol and substance use.

Medium to large effect sizes were found in both studies ($d = .54$ and $.49$). Wampold et al. (1997) found in a meta-analytic study that effect sizes for compared psychotherapy treatments did not exceed $.21$. However, Whipple et al. (2003) found a larger effect size ($d = .70$) when comparing a continuous assessment system to a control group using the OQ45 for clients identified as at-risk for terminating prematurely or having a poor treatment outcome. When examining outcome from a reliable change perspective, the results are just as impressive. There were many more clients in the feedback condition that reported reliable change at the end of treatment compared to the no-feedback condition (80% vs. 54.2% in Study 1; 66.67% vs. 41.4% in Study 2).

One of the biggest advantages proposed for using continuous assessment is that therapists can more readily identify clients not progressing in treatment. If identified early, a therapist can intervene and assess why the client is not improving before the client terminates prematurely or has a negative outcome in therapy. In Study 2, the results indicated that clients in the feedback group who were not progressing by the third session demonstrated greater treatment gains than those in the no-feedback group. However, the difference between the mean number of sessions attended was nearly identical (five or six sessions). This may be due to the pretreatment means for the feedback group being much higher (25.06 vs. 20.06); therefore, perhaps requiring fewer sessions.

These findings seem consistent with previous research; using continuous outcome assessment appears to lead to better treatment outcomes for those that are not-on-track early in treatment. However, when comparing feedback and no-feedback pre/post-ORS treatment gains the effect size for clients not progressing was similar to the effect size for the entire sample. Previous research has been mixed on this (Harmon et al., 2007; Lambert et al., 2005) but has generally found that clients not progressing early in therapy benefit more from tracking outcome. This study provides evidence that all clients, not just those projected to do poorly, benefit from using a continuous assessment system. PCOMS, however, is implemented differently than other continuous assessment systems. These differences are discussed further in this section.

Dose-response curve. In Study 1, half of the feedback group was estimated to have met the criterion for reliable change after a median of nine sessions. We found it interesting that clients assigned to professional staff were more likely to evidence reliable change sooner (Session 7) compared to clients assigned to a practicum student (Session 12). However, practicum students were just as effective as professional staff when observing pre/post-ORS treatment gains in the feedback condition and across treatment conditions. The results seem to indicate that ultimate outcome is equivalent but clients paired with professional staff improved more quickly. This finding did not appear to be replicated in Study 2. The dose-response finding with the MFC sample was similar to the survival curve for clients seen by professional staff at the UCC, with 54% of the

participants in the feedback condition estimated to demonstrate reliable change after a median of seven sessions compared to 50% of the participants in the no-feedback condition estimated to require a median of 10 sessions.

The dose-response curve for both studies appears to be consistent with other outcome research utilizing survival analysis. For example, Wolgast, Lambert, and Puschner (2003) found that it took an estimated 10 sessions for 51% of a sample of 788 university counseling clients to evidence and maintain reliable change as measured by the OQ45. In another example, Anderson and Lambert (2001) found 50% of clients in a university-training clinic were estimated to evidence reliable change on the OQ45 after nine sessions.

Limitations of the Study

There are several limitations in both studies. The largest concern is the number of clients excluded. Of the original 237 possible participants across both studies, 89 participants could not be included. The biggest problem was the number of participants in the no-feedback condition that did not complete the posttreatment ORS ($n = 34$). An attempt was made to correct the difficulty with collecting posttreatment measures by having participants in Study 2 complete the ORS every session. Although this led to some improvement, it was still problematic. Frequent reminders were sent on a monthly basis, but appeared to have little impact with the no-feedback group. Once again, the pretreatment difference between those in the no-feedback condition that did or did not complete the posttreatment measure was not statistically significant.

It is important to note that half of those excluded in both studies did not return for a second session ($n = 45$). The possibility exists that some did not return because using PCOMS was not appealing. Because of the concern that eliminating data of these clients might bias or skew the data favorably, analyses were rerun with those who attended one session using the pretreatment score as the posttreatment score. None of the analyses revealed differences that would have influenced the findings and conclusions of the study.

Another limitation is the large number of clients that had missing session data. A decision was made to exclude participants that had not

completed the ORS and SRS for at least half of their sessions for both studies ($n = 10$). The decision was made to limit the possibility of underestimating the intervention effect; however, 28 participants in the feedback condition (29.47% across both studies) still had at least one session with missing data (i.e., the ORS and SRS was not given). The impact of not using PCOMS every session is unknown; it is plausible that inclusion of data with sessions skipped led to underestimating the treatment effects. However, the differences in treatment outcome gains were not different for this group compared to participants with no missing data. Future research could compare tracking outcome every second or third session to every session, particularly given that some clinics and university counseling centers already use continuous assessment systems in such a manner.

A related limitation was not monitoring treatment integrity. No manipulation checks were done to assess how well PCOMS was implemented. Completing the measures is only part of the system; it was unknown if the measures were discussed and implemented within session appropriately. In addition, participant and therapist dyads who complied with the study protocol may have been different from dyads that did not comply. Therapists complied with some clients but not other clients. Reasons cited by therapists were: "I forgot" or "I was too busy" or "The client could not wait." It seems reasonable that client characteristics influenced therapist compliance, as well as general therapist attitudes toward using the system. Resistance to using the system was not perceived to be a general problem, but it certainly did occur at some level. Conversely, some of the therapists expressed frustration at having a useful tool at their disposal but not being able to use it with certain clients in the no-feedback condition, particularly clients they felt were not progressing. It may be that some therapists were applying the system verbally with these clients.

Future Study

Given the positive results, continued replication and extension of research using PCOMS is warranted. Four suggestions are provided for consideration. First, a current limitation of PCOMS is that little research exists that addresses effectiveness with individuals from di-

verse cultural and ethnic/racial backgrounds. For example, PCOMS assumes a collaborative, client-directed process but this approach may be less preferred with clients from cultures that emphasize deference to professionals.

Second, future research should focus on why PCOMS has been found effective for all clients not just those identified at risk for terminating prematurely. Research on monitoring outcome throughout treatment (Lambert et al., 2005) has generally indicated that outcome is only enhanced for those who are projected to do poorly in treatment. However, it is not understood why clients progressing as expected would benefit. Therapists, in this case, would not appear to have a need to change or alter anything. A possible study would be to compare PCOMS to other continuous assessment systems. PCOMS assesses the therapeutic relationship with all clients, but Lambert and colleagues' signal system uses a measure of the therapeutic relationship only after a client is identified as deteriorating or not progressing as expected (e.g., Whipple et al., 2003). Both acknowledge the importance of the therapeutic relationship in relation to outcome (Horvath & Bedi, 2002), but does the ability to discuss problems with the therapeutic alliance immediately with clients, rather than retrospectively, matter?

Third, future study should also consider the potential influence of demand characteristics or social desirability inflating the scores due to completing the measures in the therapist's presence and then discussing the scores with the therapist. This appears more likely to be a problem for the SRS than the ORS. Many clients do hide things from their therapist, but they are more likely to withhold an immediate negative reaction to the therapist or session than to hide or misrepresent their level of distress (Farber, 2003). An additional possibility is that seeing the measures consistently may create an expectancy effect that improvement should occur. Conversely, having access to weekly feedback regarding the relationship may serve to heighten attention and focus on the therapeutic alliance and promote active collaboration. Yet another possibility is that having a visual prompt may also make a difference. It is well-established that receiving feedback on performance can promote positive behavior change (e.g., Alvero, Bucklin, & Austin, 2001). An example perhaps analogous to therapy is that fre-

quent weighing has been found to promote weight loss for dieters (Wing & Hill, 2001).

Last, it has been suggested that utilizing client outcome data may be beneficial to clinical training and supervision (Worthen & Lambert, 2007). Specifically, this would involve taking the outcome data provided by clients and utilizing that information within clinical supervision. Worthen and Lambert (2007) proposed that using client outcome data would facilitate the supervisor's ability to provide specific and critical feedback to trainees. Hoffman, Hill, Holmes, and Freitas (2005) noted that almost all supervisors withhold feedback regarding performance, although specific feedback is considered a good marker of supervision (Lehrman-Waterman & Ladany, 2001). Research could assess whether using PCOMS in supervision helps supervisees provide more specific and critical feedback.

With the increased need to demonstrate psychotherapy's utility due to such forces as managed care and third-party reimbursement, measuring the progress of treatment as it occurs has become an emerging area of study with exciting results. Ongoing feedback has been found to prevent premature termination and to help meet the needs of clients in a more effective, efficient manner. Overall, the results of this study indicated that the PCOMS approach of providing outcome feedback on a client's progress and the counseling relationship is a useful approach and is consistent with findings by the developers. Although more research certainly needs to be conducted, this system appears to hold promise given its ease of use and encouraging results.

References

- ACKERMAN, S. J., BENJAMIN, L. S., BEUTLER, L. E., GELSO, C. J., GOLDFRIED, M. R., HILL, C., et al. (2001). Empirically supported therapy relationships: Conclusions and recommendations of the Division 29 Task Force. *Psychotherapy: Theory, Research, Practice, Training*, 38, 495–497.
- ALVERO, A. M., BUCKLIN, B. R., & AUSTIN, J. (2001). An objective review of the effectiveness and essential characteristics of performance feedback in organizational settings (1985–1998). *Journal of Organizational Behavior Management*, 21, 3–30.
- ANDERSON, E. W., & LAMBERT, M. J. (2001). A survival analysis of clinically significant change in outpatient psychotherapy. *Journal of Clinical Psychology*, 57, 857–888.
- BEUTLER, L. E. (2001). Comparisons among quality assurance systems: From outcome assessment to clinical utility. *Journal of Consulting and Clinical Psychology*, 69, 197–204.
- BORDIN, E. S. (1979). The generalizability of the psychoanalytic concept of working alliance. *Psychotherapy: Theory, Research and Practice*, 16, 252–260.
- BROWN, G. S., & JONES, E. R. (2005). Implementation of a feedback system in a managed care environment: What are patients teaching us? *Journal of Clinical Psychology*, 61, 187–198.
- DUNCAN, B. L., MILLER, S. D., SPARKS, J. A., CLAUD, D. A., REYNOLDS, L. R., BROWN, J. B., et al. (2003). The Session Rating Scale: Preliminary psychometric properties of a "working alliance" measure. *Journal of Brief Therapy*, 3, 3–12.
- FARBER, B. A. (2003). Patient self-disclosure: A review of the research. *Journal of Clinical Psychology*, 59, 589–600.
- HARMON, S. C., LAMBERT, M. J., SMART, D. M., HAWKINS, E., NIELSEN, S., SLADE, K., et al. (2007). Enhancing outcome for potential treatment failures: Therapist-client feedback and clinical support tools. *Psychotherapy Research*, 17, 379–392.
- HAWKINS, E. J., LAMBERT, M. J., VERMEERSCH, D. A., SLADE, K. L., & TUTTLE, K. C. (2004). The therapeutic effects of providing patient progress information to therapists and patients. *Psychotherapy Research*, 14, 308–327.
- HOFFMAN, M. A., HILL, C. E., HOLMES, S. E., & FREITAS, G. F. (2005). Supervisor perspective on the process and outcome of giving easy, difficult, or no feedback to supervisees. *Journal of Counseling Psychology*, 52, 3–13.
- HORVATH, A. O., & BEDI, R. P. (2002). The alliance. In J. C. Norcross (Ed.), *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients* (pp. 37–70). New York, NY: Oxford University Press.
- 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.
- LAMBERT, M. J., BURLINGAME, G. M., UMPHRESS, V., HANSEN, N. B., VERMEERSCH, D. A., CLOUSE, G. C., & YANCHAR, S. C. (1996). The reliability and validity of the Outcome Questionnaire. *Clinical Psychology and Psychotherapy*, 3, 249–258.
- 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., & FINCH, A. E. (2001). Patient-focused research: Using patient outcome data to enhance treatment effects. *Journal of Consulting and Clinical Psychology*, 69, 159–172.
- LAMBERT, M. J., HANSEN, N. B., UMPHRESS, V., LUNNEN, K., OKISHI, J., BURLINGAME, G. M., et al. (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 re-

- sults and practice suggestions. *Journal of Clinical Psychology*, *61*, 165–174.
- LAMBERT, M. J., WHIPPLE, J. L., HAWKINS, E. J., VERMEERSCH, D. A., NIELSEN, S. L., & SMART, D. (2001). Is it time for clinicians to routinely track patient outcome? A meta-analysis. *Journal of Clinical Psychology: Science and Practice*, *10*, 288–301.
- 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.
- LEHRMAN-WATERMAN, D., & LADANY, N. (2001). Development and validation of the Evaluation Process Within Supervision Inventory. *Journal of Counseling Psychology*, *48*, 168–177.
- LUBORSKY, L., BARBER, J., SIQUELAND, L., JOHNSON, S., NAJAVITS, L., FRANK, A., et al. (1996). The Helping Alliance Questionnaire (HAQ-II): Psychometric properties. *The Journal of Psychotherapy Practice and Research*, *5*, 260–271.
- LUEGER, R. J., HOWARD, K. I., MARTINOVICH, Z., LUTZ, W., ANDERSON, E. E., & GRISSOM, G. (2001). Assessing treatment progress of individual patients using expected treatment response models. *Journal of Consulting and Clinical Psychology*, *69*, 150–158.
- MILLER, S. D., & DUNCAN, B. L. (2000). *The Outcome Rating Scale*. Chicago, IL: Authors.
- MILLER, S. D., & DUNCAN, B. L. (2004). *The Outcome and Session Rating Scales: Administration and scoring manual*. Chicago, IL: Authors.
- MILLER, S. D., DUNCAN, B. L., BROWN, J., SORRELL, R., & CHALK, M. (2006). Using formal client feedback to improve outcome and retention: Making ongoing, real-time assessment feasible. *Journal of Brief Therapy*, *5*, 5–22.
- MILLER, S. D., DUNCAN, B. L., BROWN, J., 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*(2), 91–100.
- MILLER, S. D., DUNCAN, B. L., & JOHNSON, L. D. (2000). *The Session Rating Scale 3.0*. Chicago, IL: Authors.
- MILLER, S. D., MEE-LEE, D., & PLUM, W. (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.
- MILLER, S. L., DUNCAN, B. L., SORRELL, R., & BROWN, G. S. (2005). The partners for change outcome management system. *Journal of Clinical Psychology*, *61*, 199–208.
- PERCEVIC, R., LAMBERT, M. J., & KORDY, H. (2004). Computer-supported monitoring of patient treatment response. *Psychotherapy Research*, *16*, 364–373.
- SNELL, M. N., MALLINCKRODT, B., HILL, R. D., & LAMBERT, M. J. (2001). Predicting counseling center clients' response to counseling: A 1-year follow-up. *Journal of Counseling Psychology*, *48*, 463–473.
- VERMEERSCH, D. A., WHIPPLE, J. L., & LAMBERT, M. J. (2004). Outcome questionnaire: Is it sensitive to changes in counseling center clients? *Journal of Counseling Psychology*, *5*, 38–49.
- WAMPOLD, B. E., MONDIN, G. W., MOODY, M., STICH, F., BENSON, K., & AHN, H. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, "all must have prizes." *Psychological Bulletin*, *122*, 203–215.
- 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 and problem-solving strategies in routine practice. *Journal of Counseling Psychology*, *50*, 59–68.
- WING, R. R., & HILL, J. O. (2001). Successful weight loss maintenance. *Annual Review of Nutrition*, *21*, 323–341.
- WOLGAST, B. W., LAMBERT, M. J., & PUSCHNER, M. J. (2003). The dose-response relationship at a college counseling center: Implications for setting session limits. *Journal of College Student Psychotherapy*, *18*, 15–29.
- WORTHEN, V. E., & LAMBERT, M. J. (2007). Outcome-oriented supervision: Advantages of adding systematic client tracking to supportive consultations. *Counseling and Psychotherapy Research*, *7*, 48–53.