The Partners for Change Outcome Management System

▼

Scott D. Miller and Barry L. Duncan
Institute for the Study of Therapeutic Change

▼

Ryan Sorrell
Resources for Living

▼

George S. Brown
Center for Clinical Informatics

A number of systems provide feedback regarding client progress and experience of the therapeutic alliance to clinicians. Available evidence indicates that access to such data improves retention and outcome for clients most at risk for treatment failure. Over the last several years, the team at the Institute for the Study of Therapeutic Change has worked to develop an outcome management system that not only provides valid and reliable feedback, but also is as user-friendly as possible for therapists and consumers. In this article, we describe the system and summarize current research findings. © 2004 Wiley Periodicals, Inc. J Clin Psychol/In Session 61: 199–208, 2005.

Keywords: psychotherapy; therapeutic alliance; outcome measurement

> What can be done with fewer means is done in vain with many.
> —William of Ockham

More than any previous time in the history of the field, policy makers and payers are stridently insisting that therapists and the systems of care in which they operate must "deliver the goods." Accountability is the watchword of the day, and "return on investment" the guiding metric. Like it or not, psychotherapy has become a commodity and those footing the bill want proof of the effectiveness and value of the product being purchased.

A special thanks to Mark Hubble, Ph.D. for his valuable editorial review and suggestions.
Correspondence concerning this article should be addressed to: Scott D. Miller, Ph.D., Institute for the Study of Therapeutic Change, P.O. Box 578264, Chicago, IL 60657-8264; e-mail: scottdmiller@talkingcure.com.

JCLP/In Session, Vol. 61(2), 199–208 (2005)

Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jclp.20111

Lest there be any confusion, interest in outcome is neither exclusive to mental health nor limited to payers and policy makers. Rather, the emphasis on outcome is part of a worldwide phenomenon. Even consumers of treatment are demanding proof of results. Indeed, although stigma, lack of knowledge, and concerns about the length of treatment are frequently offered as explanations for not seeking therapy, a significantly larger number of potential consumers identify low confidence in the outcome of services as the major deterrent (APA, 1998). Clearly, in the mental health arena, outcomes are "on the line."

Over the last several years our team at the Institute for the Study of Therapeutic Change (ISTC) has been working to develop and implement a system for both monitoring and improving the effectiveness of treatment (Duncan, Miller, & Sparks, 2004). The approach builds in a major way on two key findings from previous research: first, the general trajectory of change in successful therapy is predictable (e.g., Howard, Moras, Brill, Martinovich, & Lutz, 1996); second, measures of client progress and experience of the therapeutic alliance can be used to "determine the appropriateness of the current treatment . . . [and] the need for further treatment . . . [and to] prompt a clinical consultation for patients who [are] not progressing at expected rates" (Howard et al., 1996, p. 1063). With regard to the latter, Whipple and colleagues (2003) found that clients at risk for a negative outcome were less likely to deteriorate, more likely to stay in treatment longer, and *twice as likely* to achieve a clinically significant change when their therapists had access to outcome *and* alliance information.

In addition to establishing a system that is valid and reliable, a major goal of our efforts at ISTC has been to make the process of collecting and using outcome data as user-friendly as possible for both therapists and consumers. As is news to no clinician on the front lines of treatment, the number of forms, authorizations, and other oversight procedures has exploded in recent years. Few therapists have the time or resources to devote to the repeated administration, scoring, and interpretation of lengthy structured interviews or standardized measures. The majority of practitioners in one study did not consider any measure or combination of measures that required more than 5 minutes per session to complete, score, and interpret practical (Brown, Dreis, & Nace, 1999).

After experimenting with a number of outcome and alliance measures in a variety of treatment contexts, we found that similar tolerance levels (5 minutes) apply to consumers. Clients, we have consistently found, quickly tire of measures that lack obvious face validity, require more than a few minutes to complete, or use time spent with the clinician. Low compliance rates are the most frequent result. Worse yet, given the time pressures operating in real-world clinical settings, is the failure to use whatever data or feedback the lengthier outcome tools make available. In primary health-care settings—where treatment contact is very limited or when behavioral health services are delivered via the Internet or telephone—the need for feasible outcome management tools becomes ever more apparent.

In the sections that follow, we describe the Partners for Change Outcome Management System (PCOMS), specifically emphasizing the feedback process. Information regarding the development and validation of the alliance and outcome measures employed in the system is presented first, along with options for administration and interpretation. A summary of research on the impact of PCOMS in real-world clinical settings follows. The article concludes with a discussion of ongoing research and limitations of the system.

The Measures

In PCOMS, treatment outcomes are assessed via the Outcome Rating Scale (ORS), a four-item self-report instrument. Completing and scoring the scale take less than

1 minute—and the test is available in both written and oral forms in several different languages. The ORS was developed as a brief alternative to the Outcome Questionnaire-45 (OQ-45)—a widely used and well-validated measure (Lambert et al., 1996). Both measures were designed to assess change in three areas of client functioning generally considered valid indicators of treatment progress: individual (or symptomatic) functioning, interpersonal relationships, and social role performance (work adjustment, quality of life). We have found the ORS to have adequate reliability and validity (Miller, Duncan, Brown, Sparks, & Claud, 2003).

Therapeutic alliance is assessed via the Session Rating Scale 3.0 (SRS), a four-item client-completed measure. As does the ORS, the SRS requires less than 1 minute to complete and score, and it, too, is available in several different languages. Items on the scale reflect the classical definition of the alliance first stated by Bordin (1979) and a related construct termed the *client's theory of change* (Duncan, Miller, & Sparks, 2004). As such, the scale assesses four interacting elements: the quality of the relational bond, as well as the degree of agreement between the client and therapist on the goals, methods, and overall approach of therapy. As with the ORS, we have found the SRS has adequate validity and reliability (Duncan, Miller, Reynolds, et al., 2004). Both scales may be downloaded and used free of charge at www.talkingcure.com.

Options for the administration and interpretation of the SRS and ORS run from simple to complex. With regard to administration, the practitioner working in private practice can readily adopt a single-subject design—graphing and discussing the measures with the individual client at each session. The general pattern of alliance formation and change in treatment established by existing outcome research can then be used as a guide to interpretation of the results (Duncan, Miller, & Sparks, 2004). As an example, a score of 36 or below on the SRS can be utilized to warn of a potential problem in the alliance, as it falls below the 25th percentile. Because research indicates that clients frequently drop out of treatment *before* discussing problems in the relationship, a therapist would be well advised to use the opportunity afforded by the scale to open discussion and work to restore the alliance. The same general strategy could be followed for the ORS. Given research showing that the majority of change in treatment occurs earlier rather than later (e.g., Brown, Dreis, & Nace, 1999; Howard, Moras, Brill, Martinovich, & Lutz, 1996), an absence of improvement in the first handful of visits could serve as a warning to the therapist, signaling the need for opening a dialogue with the client about the nature of treatment.

If clinicians prefer, these tasks can be accomplished through an automated Windows-based system that provides "real-time" warnings to therapists when an individual client's ratings of either the alliance or the outcome fall significantly outside established norms (Miller, Duncan, Brown, Sorrell, & Chalk, in press). As an example of the feedback that therapists receive when a particular client's outcomes fell outside the expected parameters, consider Figure 1.

The dotted line represents the expected trajectory of change for clients whose total score at intake on the ORS is 10. Consistent with previous research and methodology, trajectories of change in this system are derived via linear regression and provide a visual representation of the relationship between ORS scores at intake and those of subsequent administrations. Bands corresponding to the 25th (light gray) and 10th (black) percentiles mark the distribution of actual scores *below* the expected trajectory over time. The horizontal dashed-dotted line at 25 represents the clinical cutoff score for the ORS. Scores falling above the line are characteristic of individuals who are not seeking treatment, and scores below are similar to those of people who are in treatment and likely to improve (Duncan, Miller, Reynolds, et al., 2004). The remaining solid line designates the client's actual score from session to session.

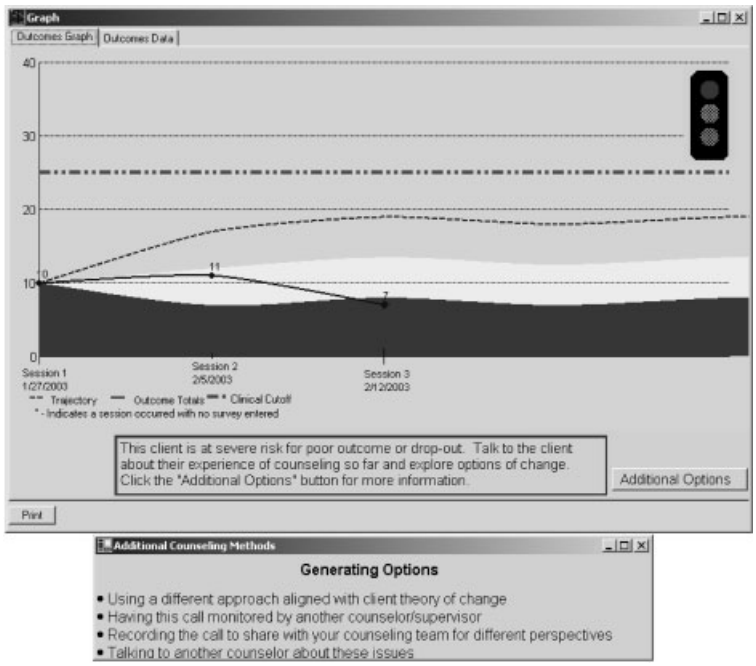

Figure 1. SIGNAL outcome feedback screen.

As seen in Figure 1, this particular client's score at the second session falls below the 25th percentile. By session 3 the score has fallen even further, landing in the black area representing the 10th percentile in the distribution of actual scores. As a result, the therapist receives a *red* signal, warning of the potential for premature dropout and an increased risk for a negative or null outcome should therapy continue unchanged. An option button provides suggestions for addressing the problem: (1) talking with the client about problems in the alliance, (2) changing the type and amount of treatment being offered, and (3) recommending consultation or supervision. Client feedback regarding the alliance is presented in a similar fashion at the end of each visit.

Recently, we launched a Web-based version of PCOMS: www.treatmentoutcomes online.com (TOOL). Briefly, TOOL provides the precision and reliability without the extensive work and expense associated with development of an automated outcome management system. Independent practitioners, agencies, and behavioral health-care entities can monitor caseloads, receive feedback on clients at risk for a negative or null outcome, and even determine overall effectiveness relative to national norms. Other features make it possible for consumers to follow their progress over the course of treatment as well as enables supervisors and larger systems of care to identify cases requiring review or clinicians in need of consultation or further training.

The Research

In this section we review the research reported to date on PCOMS. A number of questions regarding the system have been explored, including: (1) the overall feasibility of the system; (2) its impact on retention in and outcome from treatment; and (3) its ability to identify reliable differences in effectiveness among clinicians. In addition, for the first time we summarize the application of PCOMS in the area of substance abuse.

Feasibility

As noted in the introduction, feasibility is a critical issue in outcome management. Two studies have found high rates of compliance among therapists using the ORS and SRS in treatment (Miller, Duncan, Brown, et al., 2003; Miller, Duncan, Brown, et al., in press). The same studies note the lack of client complaints and the ease with which the scales can be integrated into treatment. Of course, neither the ORS nor the SRS offers the same breadth of assessment as longer scales (e.g., the absence of "critical items" related to suicide or alcohol and drug use, greater detail about the alliance). At the same time, a measure that is unused is useless regardless of its strengths. Moreover, few would argue for the utility of more time-consuming measures in services provided over the telephone or Internet. In the real and evolving world of clinical practice, finding the right set of tools for a particular setting means striking a balance among the competing demands of validity, reliability, and feasibility.

Impact on Retention and Outcome

One study has examined the impact of PCOMS on client retention and outcome (Miller, Duncan, Brown, et al., in press). Our team compared outcome and alliance data from a 6-month baseline period to data gathered after the implementation of PCOMS. The sample was composed of 6,424 culturally and economically diverse clients served by an international employee assistance program. The presenting problems were comparable to those seen in a typical mental health clinic, including anxiety, depression, substance abuse, work and family complaints, as well as chronic mental and physical health problems (Miller, Duncan, Brown et al., 2003). As shown in Figure 2, the availability of real-time outcome feedback resulted in a substantial improvement in effectiveness. The overall effect size of treatment more than doubled from the baseline period to the final evaluation phase (baseline $ES = .37$ versus final phase $ES = .79$; $p < .001$). Consistent with research on other outcome management systems, the improvement in outcome was realized without any attempt to organize, systematize, or otherwise control the treatment process. Nor were the therapists in the study trained in any new therapeutic modalities, treatment techniques, or diagnostic procedures.

We also found that access to alliance feedback had a beneficial impact on outcomes and retention rates (Miller, Duncan, Brown, et al., in press). For reasons unknown, during the baseline period in the study, 20% of the patients who had ORS scores at intake did not

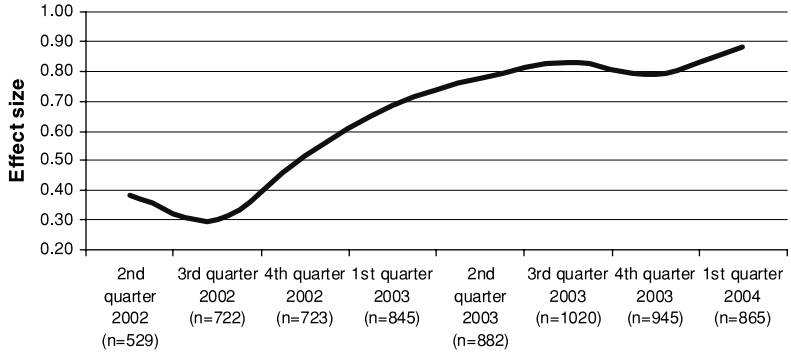

Figure 2. Improvement in effect size after feedback.

have SRS scores for that visit. Interestingly, such cases were three times *less* likely to have an additional session than those for whom alliance data were present (6% versus 19%, respectively). Failure to complete the SRS was also associated with less change on the ORS at the end of treatment. Among clients who remained with the same therapists throughout treatment, those who completed the SRS at intake averaged 3.3 points more change (residualized gain score) than those who did not ($p < .01$; two-tailed *t*-test). By the final evaluation period, utilization rates for the SRS had improved so much that failure to complete the measure was no longer predictive of dropout after the first session. Even in this phase, however, failure to complete the SRS was associated with less change by the end of treatment (mean residualized change score = 1, $p < .05$; two-tailed *t*-test).

Differences in Effectiveness Among Clinicians

Research provides substantial evidence of differences in effectiveness among clinicians and treatment settings. Indeed, conservative estimates indicate that between 6% and 9% of the variance in outcomes is attributable to therapist effects. The setting in which treatment is conducted accounts for 3%–4% (Wampold, 2001; Miller, Duncan, & Hubble, 2002). The ability to identify reliable differences in outcome among therapists and among treatment settings has obvious implications for training, supervision, and quality assurance—perhaps even for marketing and reimbursement (Lambert, 2002).

As an example of using the PCOMS to identify differences in effectiveness, consider data on 22 therapists who participated in one of our studies (Miller, Duncan, Brown et al., 2004). Figure 3 shows the distribution of outcomes for each clinician who had at least 30 completed cases during the baseline period. The solid horizontal line represents the average effect size of all therapists in the sample. Basically, a therapist can be considered statistically "above average" at a 90% confidence level when the bottom end of his or her personal range falls above the average effect size for the agency as a whole. Conversely, when the top end of a clinician's personal range falls entirely below the average effect size, he or she can be considered statistically "below average" at the 90% confidence level.

Unfortunately, little is known at present about the cause(s) of the differences between clinicians. Nor do we know whether anything can be done to close the gap between more and less effective clinicians (e.g., distillation of effective practices by studying the most

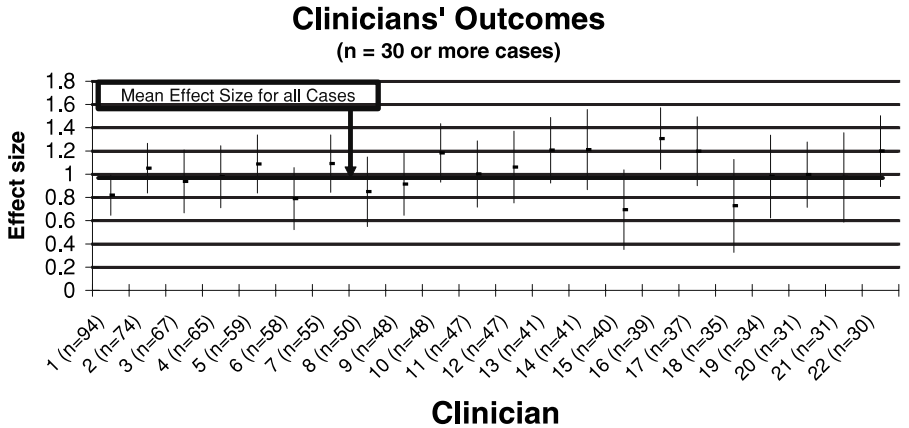

Figure 3. Distribution of clinician effectiveness.

effective therapists, additional supervision or training). Interestingly, however, despite documentation of tremendous improvements in cases at risk for a negative outcome, Lambert (2003, personal communication) has not found that therapist ability to identify failing cases improves with time and feedback. Rather, from year to year, the number of at-risk warnings a given clinician receives remains constant.

In an e-mail to the first author dated July 3, 2003, Lambert said:

> "The question is—have therapists learned anything from having gotten feedback? Or, do the gains disappear when feedback disappears? About the same question. We found that there is little improvement from year to year even though therapists have gotten feedback on half their cases for over three years. It appears to us that therapists do not learn how to detect failing cases. Remember that in our studies the feedback has no effect on cases who are progressing as expected—only the signal alarm cases profit from feedback."

If confirmed, such findings, when taken in combination with the weak historical link between training and outcome in psychotherapy (Lambert & Ogles, 2004), underscore the need to make feedback a routine part of clinical practice. After all, ongoing and systematic evaluation of outcome has the potential of offering real-time protection to consumers and payers. Perhaps instead of empirically supported *therapies*, consumers could have access to empirically validated *therapists*.

Partners for Change Outcome Management System and Substance Abuse

Recently, the PCOMS has been applied to clients who seek treatment for problems related to alcohol and other substance use. Data are currently being gathered in a number of treatment contexts, including residential treatment centers, intensive outpatient, outpatient, and telephonic and Internet-based treatment contexts. Here we report results from one program, New Start™, an alcohol and drug service offered by an international employee assistance program.

Outcome and alliance scores were gathered for 160 clients who participated in the program between January 1, 2003, and June 1, 2004. Briefly, the treatment program was designed to help people who have drug or alcohol problems maintain their employment. Entry into the program can be either voluntary or mandated by the employer. Form and content of the program are similar in most respects to those of traditional alcohol and drug services, with a specific focus on skill building and the development of "sober-friendly" support networks.

Several findings have implications for the use of feedback in the treatment of clients who have drug and alcohol problems. First, the average ORS intake score of participants in the sample is significantly higher (showing *less* distress) than scores reported in two studies of the general mental health population (24.1 versus 19.6 and 19.1, respectively; Miller, Duncan, Brown et al., 2003; Miller, Duncan, Brown, et al., in press). In addition, the general trajectory of change for the New Start™ sample is quite different from that reported for the general mental health population. General mental health clients tended to become worse with treatment when their initial score on the ORS fell above the clinical cutoff of 25 points. In contrast, the average substance-abusing client improved regardless of his or her intake score.

A second difference between the substance abuse sample and general mental health clients was found in the rate and total amount of change over time. Consistent with findings from other studies of substance abuse treatment (e.g., Cooney, Babor, DiClemente, & Del Boca, 2003), longer contact resulted in better outcomes for the participants in New

Start™. By contrast, data for a general mental health sample typically show little or no gain in outcomes after the first handful of visits.

Third, one cluster of findings related to client engagement should be mentioned. In contrast to other studies documenting improved retention rates when clients who have alcohol and drug problems are mandated into care (Martin, et al., 2003), in the current sample no differences in attrition were found between voluntary and involuntary clients. At the same time, clients who completed the program—whether voluntary or mandated—averaged significantly more change than those who dropped out (10.8 versus 7.4 points of change, $p < .05$). Interestingly, however, only the mandated clients who ended treatment unsuccessfully (e.g., dropout, positive urine screen result, termination from work) scored above the clinical cutoff at intake (see Figure 4). This group was also the only one whose change scores did not significantly differ from intake to last recorded session.

Although awaiting further exploration, one potential explanation for these findings may be a mismatch between the treatment program and certain clients' stage of change (Prochaska, 1999). People who are mandated *and* score above the clinical cutoff are indicating that they are not distressed about their situation. As a result, they may be said to fit the precontemplation stage of change. A growing body of research indicates that therapists may need to exercise greater skill in gauging the level and intensity of treatment in order to keep such clients engaged (Prochaska & Norcross, 2002). Whatever the case proves to be, feedback via a valid, reliable, and feasible outcome tool is an essential first step.

Assessment of the alliance also proved important for substance-abusing clients. Similar to findings reported for a general mental health population, failure to seek feedback regarding the alliance at the first session was significantly higher among unsuccessful clients in general and mandated clients in particular (84% for the successful versus 63% for unsuccessful; 70% for the voluntary, unsuccessful; and 53% for the mandated, unsuccessful). A number of potential explanations for the discrepancy exist; for example, the therapist believed the client was too angry to ask about the alliance or believed the session did not go well and opted not to complete the measure. Recall that we previously found that outcome and retention rates improved when utilization of the alliance measure at intake increased—a finding that suggests that clinicians would be well served by making extra efforts to collect an alliance measure in all cases.

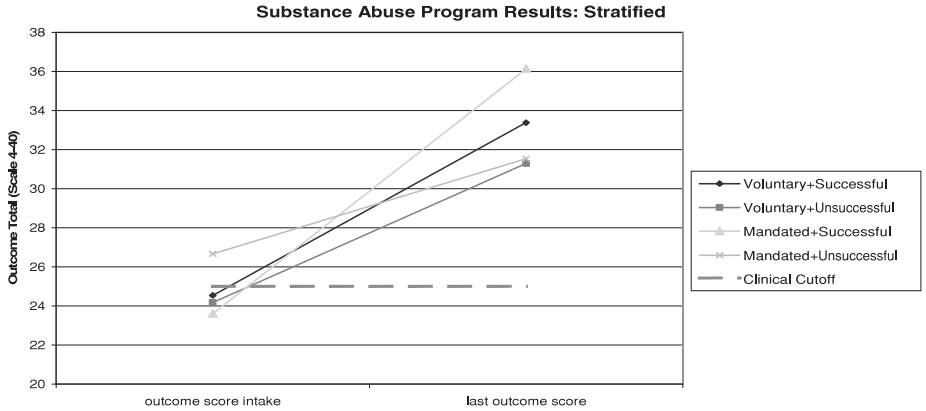

Figure 4. Attrition rates stratified by entry status and outcome.

Clinical Issues

Research on PCOMS adds to the growing literature documenting the salutary impact of feedback on retention and outcome in therapy. To summarize briefly, available research shows that access to the client's experience of progress and alliance as measured by the ORS and SRS can as much as double the effect size of treatment and simultaneously improve client retention and cost-effectiveness. Moreover, the brevity of the two measures that form the basis of the system makes clear that the method for obtaining client feedback need not be complicated or time-consuming.

The findings on PCOMS are limited by a number of factors. First is the reliance on client self-report measures. In addition, evaluation of outcome and alliance via the ORS and SRS is far from comprehensive. Nor does PCOMS contain multiple perspectives (e.g., therapists, outside judges, objective criteria). Both measures, however, are similar in nature to those being used in patient-focused as opposed to traditional efficacy types of research studies.

Caution should also be exercised when generalizing beyond the samples employed in the studies of PCOMS thus far. For example, although the sample did not differ in any known way from traditional mental health clients seen in outpatient settings, the largest study on PCOMS published to date is based on clinical services delivered over the telephone (Miller, Duncan, & Brown, et al., in press). Additionally, the findings reported in this article for alcohol and drug treatment services, although intriguing, are based on a relatively small sample and await confirmation. On a positive note, research on feedback derived from the ORS and SRS is currently under way for a variety of different populations (e.g., alcohol and drug, disease management, children, and adults) and treatment settings (primary care, residential treatment, outpatient community mental health, college counseling center).

Select References/Recommended Readings

American Psychological Association. (1998). Communicating the value of psychology to the public. Washington, DC: American Psychological Association.

Bordin, E.S. (1979). The generalizability of the psychoanalytic concept of the working alliance. Psychotherapy, 16, 252–260.

Brown, J., Dreis, S., & Nace, D.K. (1999). What really makes a difference in psychotherapy outcome? Why does managed care want to know? In M.A. Hubble, B.L. Duncan, & S.D. Miller (Eds.), The heart and soul of change: What works in therapy (pp. 389–406). Washington, DC: American Psychological Association.

Cooney, N., Babor, T., DiClemente, C., & Del Boca, F. (2003). Clinical and scientific implications of Project MATCH. In T. Babor & F. Del Boca (Eds.), Treatment matching in alcoholism (pp. 222–237). Oxford: Oxford University Press.

Duncan, B.L., Miller, S.D., Reynolds, L., Sparks, J., Claud, D., Brown, J., et al. (2004). The session rating scale: Psychometric properties of a "working" alliance scale. Journal of Brief Therapy, 3(1).

Duncan, B.L., Miller, S.D., & Sparks, J. (2004). The heroic client: Principles of client-directed, outcome-informed therapy (Rev. ed.). San Francisco: Jossey-Bass.

Howard, K.I., Moras, K., Brill, P.L., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. American Psychologist, 51, 1059–1064.

Lambert, M.J. (2002, June). Improving psychotherapy outcome for poorly responding clients via measuring, monitoring, and feedback. Presentation at the First Bi-Annual Heart and Soul of Change Conference. Toronto, Ontario.

Lambert, M.J., Hansen, N.B., Umphress, V., Lunnen, K., Okiishi, 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., & Hill, C.E. (1994). Assessing psychotherapy outcomes and processes. In A.E. Bergin & S.L. Garfield (Eds.), Bergin and Garfield's handbook of psychotherapy and behavior change (5th ed., pp. 72–113). New York: John Wiley.

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.

Martin, B., Clapp, L., Bridgeford, D., Aponsah, A., Bialkowski, D., Lyons, L., et al. (2003). Compliance to supervised disulfiram therapy: A comparison of voluntary and court-ordered patients. American Journal of Addiction, 12, 137–143.

Miller, S.D., Duncan, B.L., Brown, J., Sorrell, R., & Chalk, M.B. (in press). Using outcome to inform and improve treatment outcomes. Journal of Brief Therapy.

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., & Hubble, M.A. (2002). Client-directed, outcome-informed clinical work. In F.W. Kaslow & J. Lebow (Eds.), Comprehensive handbook of psychotherapy. Integrative/Eclectic (Vol. 4, pp. 185–212). New York: Wiley.

Prochaska, J.O. (1999). How do people change, and how can we change to help many more people? In M.A. Hubble, B.L. Duncan, & S.D. Miller (Eds.), The heart and soul of change: What works in therapy (pp. 227–258). Washington, DC: American Psychological Association Press.

Prochaska, J.O., & Norcross, J.C. (2002). Stages of change. Psychotherapy, 38, 443–448.

Wampold, B.E. (2001). The great psychotherapy debate: Models, methods, and findings. Hillsdale, NJ: Lawrence 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 and problem-solving strategies in routine practice. Journal of Counseling Psychology, 50, 59–68.