

# Exploring the Contribution of Supervisors to Intervention Outcomes

Jennifer L. Callahan  
University of North Texas

Christina M. Almstrom, Joshua K. Swift,  
and Susan E. Borja  
Oklahoma State University

Christopher J. Heath  
University of North Texas

Supervision has been identified as perhaps the most important mechanism for developing competencies in therapists in training. However, there is scant research on the effect of supervisors on client outcome. Moreover, in a relatively recent review of the existing literature, significant methodological concerns were raised that attenuate the interpretability of the existing reported findings. The current study therefore sought to address such methodological concerns. Crosstabulation of supervisors by client outcome categories (i.e., recovered, reliably improved, no reliable change, deteriorated) indicated that supervisors are significantly related to client outcome, generating a moderate effect. Training programs are therefore encouraged to routinely track client outcomes as an objective indicator of quality supervision. Additional implications for training programs and suggestions for future research are discussed.

*Keywords:* supervision, training, client outcomes, supervisor competencies, practicum

It has recently been reported that a very high rate of premature termination, over 75%, is observed in the training clinic (Callahan, Aubuchon-Endsley, Borja, & Swift, in press). In contrast, the rate of premature termination in other outpatient settings is typically reported to fall between 40 and 60% (e.g., Clarkin & Levy, 2003;

Garfield, 1994; Wierzbicki & Pekarik, 1993). In addition, established research has revealed that client improvements occur more slowly with comparatively fewer successful outcomes in training clinics compared with other outpatient settings (Callahan & Hynan, 2005; Kadera, Lambert, & Andrews, 1996). Such adverse outcomes are somewhat surprising, given that it appears the process of change has been found to be similar in the training clinic setting to those reported in other outpatient settings (Callahan, Swift, & Hynan, 2006).

Better understanding of these outcomes is important for at least two reasons. First, such understanding might elucidate potential ways in which psychotherapy training may be improved. Second, training clinics carry an important responsibility for protecting a vulnerable population; the financially disadvantaged. In a recent survey of social justice activities within psychology training clinics, it was noted that 50% of training clinics provide some pro bono services, with 2 of the 28 clinics sampled providing exclusively free services (Lewis, 2008). Aside from the two free service clinics, all of the surveyed training clinics reported using a generous sliding fee scale. In those settings, 14% of clients are seen pro bono, with an average fee across paying clients of \$15.00. Clearly, training clinics serve an important role of responsibility to provide both quality training and, simultaneously, quality treatment to financially disadvantaged clients. Thus, supervision to trainee clinicians serves to ensure quality in clinical services as well as in training.

Supervision has long been recognized as an important factor in the training process and has been identified as perhaps the most important mechanism for developing competencies in therapists in training (Falender et al., 2004; Stoltenberg, 2005). The effects of supervision on factors such as the trainee's in session behavior, the trainee's interpersonal skills, the trainee's attitude toward the client, the trainee's ability to implement a particular treatment, and

---

JENNIFER L. CALLAHAN received her PhD in clinical psychology from the University of Wisconsin-Milwaukee. She is an assistant professor in psychology at the University of North Texas and holds board certification in clinical psychology. Her research interests span the assessment and psychotherapy domains, with emphasis on competency development and training in these areas.

CHRISTINA M. ALMSTROM is a PhD candidate in clinical psychology at Oklahoma State University. She is currently completing her internship at the Oklahoma State University Health Sciences Center. Her research interests include gender stereotyping and benevolent sexism, evidence-based treatments using cognitive behavioral therapy, and evolutionary psychology.

JOSHUA K. SWIFT is a doctoral candidate in the clinical psychology program at Oklahoma State University. His research interests include psychotherapy processes and outcomes with a specific focus on examining the client's role in treatment.

SUSAN E. BORJA received her MS in clinical psychology from Oklahoma State University. She is currently on internship at the Medical University of South Carolina/Charleston Veterans Affairs Medical Center Consortium in the Traumatic Stress Track. Her research interests are focused on traumatic stress, including precursors and various outcomes.

CHRISTOPHER J. HEATH received his master's degree in psychology from the University of Memphis. He is now a doctoral student in the clinical psychology program at the University of North Texas. His interests are in psychotherapy treatment evaluations and outcomes, as well as research design and statistical analyses.

CORRESPONDENCE CONCERNING THIS ARTICLE should be addressed to Jennifer L. Callahan, Department of Psychology, University of North Texas, Denton, TX 76203-1280. E-mail: jennifercallahan@unt.edu

the trainee's ability to show empathy have been well documented (Holloway & Neufeldt, 1995; Lambert & Arnold, 1987).

However, while much research exists on the effects of supervisors on trainees, there has been a continual, but mostly unheard, call for examination of the effect of supervisors on client outcome (Ellis, Ladany, Kregel, & Schult, 1996; Holloway & Neufeldt, 1995; Lambert & Arnold, 1987). Holloway and Carroll (1996), in particular, called this issue into sharp focus by chastising researchers for focusing too much on the impact of supervision on trainee development at the expense of public protection, drawing an analogy to "viewing parenthood solely for the enrichment of parents" (p. 54).

In response to these calls for examination, Freitas (2002) reviewed the small existing body of literature for the preceding two decades and identified 10 studies that addressed, at least in part, the relationship between supervision and client outcome. The review concluded that supervisors did appear to impact client outcome, but also found serious methodological problems attenuating the interpretability of reported findings (e.g., outcome measures with poor or unknown psychometric properties). As a result, five methodological suggestions were made for subsequent research.

First, Freitas (2002) noted that psychometrically sound measures should be used to quantify client outcome and minimize Type I or Type II errors. Second, it was encouraged that individuals being supervised have similar training backgrounds, be at similar levels of training, and provide services to a fairly uniform clientele. Third, random assignment of clients to trainees and/or conditions was recommended. Fourth, obtaining measures of client outcome from multiple informants was encouraged. Fifth, use of therapists no longer in training was encouraged so that a no supervision condition could be created and contrasted with a supervision condition.

Although such recommendations highlight pertinent methodological limitations in the existing literature, the recommendations vary in terms of practicality. Conducting such research within a preinternship training clinic setting allows for inclusion of a sample of therapists with similar training backgrounds and is relatively easily accomplished. Similarly, at first glance, use of psychometrically sound outcome measures also seems easily addressed; however, this is only true for outcome measures completed by clients themselves. That is, there is lack of other-informant outcome measures available at present. In addition, while use of therapists who are no longer in training might allow for a clear experimental design, by creating supervision versus no supervision conditions, the generalizability of such a study to the training clinic setting with trainee therapists would be compromised.

The present study therefore sought to address several of the recommendations made by Freitas (2002), more fully described in the methods section of this article, while remaining consistent with the typical training clinic setting so that the findings may be replicable in other training clinics and the practical implications more easily discerned. Despite the greater rigor in methodology, it was hypothesized that a significant effect of supervisors on client outcome would still be observed. Because the scant literature indicates only nonspecific supervisor effects on client outcome, we did not specifically hypothesize supervisor characteristics that might be related to outcome. However, we were aware of clinical lore that some variables might be relevant (e.g., time since grad-

uation, licensure status, attainment of board certification, etc.) for exploratory purposes.

## Method

### Participants

Archival data from 76 discharged clients (61% female, 39% male) in a South-Central training clinic were utilized. The average client was 29.25 years of age (range = 18–67 years) and of Caucasian ethnicity (91%). Profiles on the Symptom Checklist-90, Revised (Derogatis, 1992) indicated that a range of common clinical presentations was represented in the sample, with a mean score on the global severity index (GSI) of 1.11 ( $SD = 0.70$ ). Scores on the Beck Depression Inventory-II (BDI-II; Beck, Steer, & Brown, 1996) were also comparable to normative clinical populations ( $M = 18.36$ ,  $SD = 11.37$ ; range = 0–44). All clients completed a course of treatment that was at least three sessions in length during a 5-year training period. The average course of treatment for these clients was 17.89 sessions long, with a median of 11.00 sessions and mode of 8.00 sessions.

Although not direct participants, characterization of the trainees is salient to interpreting and appreciating the potential generalizability of findings. All trainee clinicians ( $n = 40$ ) were enrolled in a clinical psychology, scientist-practitioner model doctoral program. The training clinic operates year round and emphasizes cognitive-behavior therapy. Each trainee clinician was supervised by a member of the tenure-track clinical program faculty (PhD) and received 1 hour of individual supervision and 2 hours of group supervision on a weekly basis. All trainee clinicians in this sample were at the preinternship level. The typical trainee clinician in the program is 26.57 ( $SD = 5.53$ ) years of age, female (63%), Caucasian (70.4%), and holds a bachelor's degree (87.5%) upon admittance to the program.

The supervisors represented all academic ranks and also served as research mentors and course instructors for the core courses required of all students in the clinical program. Supervisors were responsible for the same group of trainees for 12-month cycles. Each trainee group was comprised so that an approximately even number of students from each cohort were enrolled on each supervision team (i.e., a vertical team approach). A total of 3 or 4 teams, ranging in size from 5 to 8 students total, were created each year depending upon how many students were enrolled. Of the 5 to 8 students on the team, typically 2 served as therapists (i.e., other students were either very early in training and preparing to serve as therapists or more advanced and providing services in other settings). The clinic director constructed the supervision teams. Each student rotated to a new team annually and multiyear monitoring ensured that students were not assigned to the same supervisor twice. Moreover, supervisors did not self-select therapy cases.

### Measures

All clients were administered the Beck Depression Inventory-II (BDI-II; Beck, Steer, & Brown, 1996) as a standard part of the intake process and again at termination. The BDI-II is a 21-item, self-report measure of common depressive symptoms. According to the measures' authors, internal reliability and test-retest reli-

ability with outpatient samples is excellent and scores correlate highly with other established measures of depression (e.g., Hamilton Rating Scale for Depression,  $r = .47$ ).

The clients were also administered the Symptom Checklist-90, Revised (Derogatis, 1992) at the intake session. The 90-item checklist is a self-report measure of physical and psychological symptoms commonly seen by outpatients seeking psychological help. Clients rate each of the 90 symptoms on a 4-point scale, indicating that the symptom *bothered them not at all* (1), *a little* (2), *quite a bit* (3), or *extremely* (4).

As a precaution, in case it became necessary to control for therapists in the analysis of supervisor effect on outcome, archival ratings from clients in the style of the shortened Counselor Rating Form (CRF-S; Corrigan & Schmidt, 1983) were also included. Clients rated their trainee therapist on single word attributes (i.e., warm, likable, friendly, expert, reliable, sociable, prepared, sincere, skillful, trustworthy) using a 7-point, Likert-type scale with qualitative anchors at 1 (*not very*) and 7 (*very*). The CRF-S has been found to highly correlate with the 36-item Counselor Rating Form (Barak & LaCrosse, 1975) and hold up well under factor analysis (Corrigan & Schmidt, 1983). Items were summed to produce a total Therapist Attributes score. Composite internal reliability of this measure was good ( $\alpha = .92$ ) in this sample.

### Procedure

All clients completed the BDI-II and Symptom Checklist-90 at the time of intake, and completed the BDI-II again at termination. To allow for adequate exposure to the therapists' personal attributes, following the third session clients were routinely administered the measure of Therapist Attributes. All clients and their resultant data were treated in accordance with the *Ethical Principles of Psychologists and Code of Conduct* (American Psychological Association, 2002). This study was also approved and conducted in compliance with the university internal review board.

### Results

The data set represented 76 adult training clinic clients receiving psychotherapy services from 40 trainee therapists and 9 supervisors. No more than two cases per trainee therapist were associated with any one supervisor, and all supervisors were associated with at least four cases in the data set.

Client intake scores on the BDI-II revealed no significant differences between supervisors. BDI-II scores at intake and termination were significantly correlated ( $r = .23$ ;  $p = .03$ ). However, correlations between the Therapist Attributes total score with intake and termination BDI-II scores were not statistically significant. Moreover, analysis of variance revealed no significant differences in Therapist Attributes total score among supervisors.

Next, client outcome was categorically coded by examining the change in BDI-II scores from intake to termination. Using the reliable change index method (Jacobson & Truax, 1991), clients were considered to have reliably changed from intake if the difference between the intake and termination scores exceeded the standard error of change associated with the BDI-II. Thus, to be considered reliably changed, the BDI-II score needed to change at least 7.67 points from intake to termination. With that determination, clients were then assigned to one of four outcome categories:

- Recovered: the client's score had reliably changed from Session 1 and had moved from the clinical range at intake into the normal range at termination ( $n = 36$ ; 47.4%).
- Reliably improved: the client demonstrated reliable improvement from intake, but not enough to move into the normal range at termination ( $n = 5$ ; 6.6%).
- No change: no reliable change in score from intake was observed at termination ( $n = 30$ ; 39.5%).
- Deteriorated: the client's score was reliably worse at termination in comparison to their intake score ( $n = 5$ ; 6.6%).

Crosstabulation of supervisors by client outcome categories (see Table 1) revealed a significant Pearson  $\chi^2$  ( $\chi^2 = 47.72$ ;  $p = .003$ ), yielding a moderate effect size (Cramer's  $V = .46$ ).

Next, an analysis examining the amount of variance in outcome that may be accounted for by Therapist Attributes and Supervisors was conducted. For this analysis, only the two largest outcome categories (no reliable change vs. recovered) were included. An additional predictor variable, intake BDI-II score, was included in this analysis because a significant point biserial correlation was found between outcome status (no reliable change/recovered) and scores on the intake BDI-II ( $r_{pb} = 0.61$ ;  $p < .01$ ).

Thus, intake BDI-II scores were entered in step 1 of a logistical regression analysis. In step 2, Therapist Attributes scores were entered. Finally, the categorical Supervisor variable was entered in step 3. This analysis revealed a significant main effect for intake BDI-II scores,  $\chi^2(1, N = 66) = 24.76$ ,  $p < .01$ . As an approximation to ordinal least squares  $R^2$  (Nagelkerke, 1991), the Nagelkerke's  $R^2$  strength of this association model accounts for approximately 44.5% of the variance in outcome indicating that initial client severity substantially predicts treatment course outcome. Therapist Attributes did not produce a significant main effect,  $\chi^2(1, N = 66) = .72$ ,  $ns$ , and accounted for only an additional 1.1% of the variance. Finally, although Supervisor only neared significance  $\chi^2(1, N = 66) = 12.64$ ,  $p = .08$ , it accounted for 16.4% of the variance in outcome.

Visual inspection of the crosstabulation (see Table 1) suggested that Supervisor 1 and Supervisor 5 were most dissimilar in outcomes. In comparing these two supervisors, no significant differences were found with respect to clients' intake BDI-II scores or

Table 1  
Crosstabulation of Supervisors by Client Outcome Category

| Supervisor | Outcome category   |                       |                   |           |
|------------|--------------------|-----------------------|-------------------|-----------|
|            | No reliable change | Reliably deteriorated | Reliably improved | Recovered |
| 1          | 0                  | 0                     | 0                 | 5         |
| 2          | 8                  | 4                     | 1                 | 5         |
| 3          | 1                  | 0                     | 0                 | 4         |
| 4          | 4                  | 0                     | 2                 | 9         |
| 5          | 12                 | 0                     | 0                 | 4         |
| 6          | 3                  | 0                     | 0                 | 2         |
| 7          | 1                  | 0                     | 0                 | 3         |
| 8          | 0                  | 1                     | 2                 | 1         |
| 9          | 1                  | 0                     | 0                 | 3         |

Therapist Attribute ratings. Furthermore, both supervisors were licensed and of the same gender. Supervisor 1 held lower academic rank and supervised fewer cases because of being hired late into the data capture window, which might indicate comparatively less experience in supervision; however, she was the only Supervisor in the sample that held American Board of Professional Psychology (ABPP) certification, which could arguably indicate greater relevant expertise.

## Discussion

Client outcome has been referred to as the “acid test” of good supervision (Ellis & Ladany, 1997, p. 485), although there is scant research supporting this position. Based on the few studies that have been completed, it has generally been found that supervision does have a positive effect on client outcome (e.g., Holloway & Neufeldt, 1995), but such studies have been criticized for significant methodological problems (Freitas, 2002). The present study addressed many of the common methodological shortcomings by using a psychometrically sound outcome measure and a sample of trainees at the same level, from the same training background, and providing services with an orientation congruent with supervisors. Although a strictly randomized design was not employed, cases were not self-selected by supervisors, and no therapist was represented more than twice in any supervisor’s data. Thus, the natural procedures of the training clinic approximated randomization. Despite the greater rigor in design, in the present study it was observed that supervisors did have a significant impact on client outcome, generating a moderate effect. In addition, partitioning of the variance indicated that supervisors may account for approximately 16% of the variance in outcome beyond that accounted for by the client’s initial severity and the treating therapist’s attributes. Clearly, a client’s initial severity level remains highly predictive of treatment outcome; a finding well established in the literature (e.g., Elkin et al., 1995; McLellan et al., 1983). However, the contribution of supervisors should not be overlooked. The 16% of variance in outcome associated with supervisors in this study exceeds the amount of variance commonly associated with specific treatment interventions (e.g., 8%; Wampold, 2001).

In the introduction, it was noted that training clinics serve at least two important responsibilities. First, they serve an important training role with students and, second, they also provide services to financially disadvantaged persons. The finding that supervisors play a significant role in client outcome impacts both of these responsibilities. Faculty supervisors are not just teachers, they are also performing a clinical service; examination of both teaching evaluations and client outcomes associated with a supervisor is therefore recommended. Such an approach could result in improved client outcomes, thereby promoting the clinic’s reputation and the university’s relationship with the community. Moreover, it provides an objective mechanism for monitoring the quality of supervision provided to trainee therapists that is independent of standard, subjective course evaluations. As a result, Directors of Clinical Training (DCTs) are encouraged to work with Clinic Directors in tracking client outcomes associated with supervisors and then using that information, in conjunction with teaching evaluations, to determine who continues to serve in supervisory roles.

Despite the aforementioned methodological problems, a handful of studies have examined specific components of supervision that may lead to positive outcomes and these studies suggest some possible future directions for research. For example, Steinhelber, Patterson, Cliffe, and LeGoullon (1984) found that while the amount of supervision did not relate to client change for 237 outpatient clients, congruence of theoretical orientation between the supervisor and trainee did. By design, the current study reflected congruent trainee-supervisor orientations and this might be one reason why a larger effect of supervisors was observed than is typically reported in the existing literature. In a more eclectic program, the effect of supervisors on client outcomes might not be as apparent as seen in our investigation.

Another useful future inquiry might be to examine supervisor variables that might be associated with better client outcomes. In the present study, the sample size was too small to effectively address such follow-up inquiries. However, a few ideas for future research were generated. The supervisor associated with the best client outcomes was the only one with ABPP status. The supervisor associated with the worst client outcomes was licensed, but most of the postdoctoral hours required for licensure represented time spent supervising trainees. Examination of these types of variables is therefore encouraged.

Although the present study served as a more methodologically sound investigation of the impact of supervisors on client outcome, several limitations remain. Dodenhoff (1981) found that a strong positive trainee emotional response to the supervisor related to an improved client outcome. In the present study, no ratings of trainee clinician’s emotional response to their supervisor were available in the archival data. However, it is unclear whether such ratings would have truly been helpful or whether they would have contributed other methodological problems. For example, trainees’ emotional responses to supervisors might be state dependent and vary between cases or times of year (e.g., early in training year vs. late; before vs. after semester grading, etc.).

Similarly, Harkness and Hensley (1991) found that “client-focused” supervision resulted in more client improvement when compared with supervision focused on administrative case management for 161 outpatient clients. Further analyzing the same data set, Harkness (1995) found that supervisory problem solving and supervisor empathy led to improved clinical outcomes. In the present study, information on the content of supervision was not available, but it may be that some aspects of supervision contributed differentially to client outcomes. The observation that Supervisor 1 and Supervisor 5 were more similar than dissimilar, at face value, while the client outcomes associated with them appear disparate suggests they were indeed doing something different in supervision. It would be particularly interesting to examine whether the aspects that positively impact client outcomes also positively impact trainee development of competencies.

Finally, Kivlighan, Angelone, and Swafford (1991) found that live supervision resulted in better client outcomes compared to videotaped supervision. In the present study, no supervisor typically engaged in live supervision. An interesting future study would be to determine whether the already moderate effect size might be even larger in a sample using exclusively live supervision. Given the relatively low rate of positive client outcomes in training clinics (e.g., Callahan & Hynan, 2005; Kadera et al., 1996), such a study appears particularly important.



We appreciate that, at present, other outcome measures are preferred to the BDI-II (e.g., Outcome Questionnaire [OQ] 45.2; Lambert et al., 1996). However, such preferences have reached consensus only relatively recently with the BDI-II being one of the previously preferred measures for the purpose of tracking broad and general outcomes. The scope of this investigation spanned a lengthy period of time (in order to generate enough cases to have adequate power in analyses); thus, the only measure consistently available was the BDI-II. We consider this to be an acceptable measure for the purpose of the investigation because of the BDI-II being well documented as reflecting broad and general symptoms of distress. For example, the correlation between the OQ total score and the BDI-II is reported to be .7959 by the OQ authors (Lambert et al., 1996).

A few aspects of this study are worth highlighting as important considerations when speculating as to the generalizability of this study's findings. Not all of the clients in this study began treatment in the "clinical range." Our thinking is that in order to characterize the clients as clinical or nonclinical, we would have to establish a "clinical" cut score for the BDI-II. Such an approach would be inconsistent with the typical usage of the BDI-II, which is conventionally expressed in terms of level of symptom severity (mild, moderate, etc.; Beck et al., 1994). In addition, some clients seek treatment in the absence of measurable symptoms. In light of the tendency for scores at the ends of a distribution to move toward the mean, one might speculate that such individuals may be at heightened risk for deterioration during treatment. The removal of cases at this end of the distribution of attained scores could therefore be problematic. Given the high rate of deterioration reported in training clinics (e.g., Callahan & Hynan, 2005) we felt it important to retain deterioration as a consideration. The focus was not limited to identifying supervisor contributions to positive change; we also included deterioration as an outcome to which supervisors might contribute. Second, all trainees were at the preinternship level and results may not generalize to settings more typical later in training (e.g., outlying practicum or internship) when less close supervision is perhaps more common.

The finding that Therapist Attributes were not significantly related to client outcome is in contrast to at least some previous literature (e.g., Martin, Garske, & Davis, 2000). However, the current study included only cases that attended at least 3 sessions of treatment, in order to allow for adequate exposure to the therapist before completing the Therapist Attributes ratings. It is possible that clients that felt more strongly negative about their trainee therapist discontinued treatment prior to the third session of treatment. In the current study, although the range of scores on the Therapist Attributes rating (48–70) indicates some variability, the mean (65.32) does suggest that clients were generally highly pleased with their therapist.

This issue also brings into question whether premature terminations should be represented in the outcome data to capture the full range of possible client outcomes. Although not necessarily obvious, in this study premature terminations that occurred at any point beyond the third session actually were represented in the outcome classification. Our outcome classification includes premature terminations by utilizing the suggested definition of premature termination provided by Hatchett and Park (2003). In responding to the shortcomings found in the assorted existing definitions of premature termination, Hatchett and Park suggested,

and empirically supported, a method for conceptualizing dropout based on the criteria of attaining clinically significant change (CSC). That is, those who do not attain CSC prior to termination are considered to have discontinued treatment prematurely. In the present study, we defined "recovered" clients as those that attain CSC. By Hatchett and Park's definition, all other clients could be said to have prematurely terminated treatment. We could use a different method of defining premature termination and then include attrition as a distinct variable in analyses. However, in light of Hatchett and Park's work, doing so would result in undesirable multicollinearity. It is important to note though that clients who prematurely terminated treatment prior to the third session were not reflected in the data because they did not have the opportunity to provide any ratings of Therapist Attributes. It remains possible that Therapist Attributes are more salient to whether a client discontinues very early in treatment (i.e., prior to the third session). A potentially useful future study would be to elicit clients' ratings of therapist attributes at different point in the course of therapy to determine whether Therapist Attributes play a more significant role very early in treatment with the significance of Supervisors emerging later in treatment.

## References

- American Psychological Association (2002). *Ethical principles of psychologists and code of conduct*. Washington, DC: Author.
- Barak, A., & LaCrosse, M. B. (1975). Multidimensional perception of counselor behavior. *Journal of Counseling Psychology*, 22, 471–476.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *The BDI-II manual*. London: Harcourt Brace.
- Callahan, J. L., Aubuchon-Endsley, N., Borja, S. E., & Swift, J. K. (in press). Pretreatment expectancies and premature termination in a training clinic environment. *Training and Education in Professional Psychology*.
- Callahan, J. L., & Hynan, M. T. (2005). Models of psychotherapy outcome: Are they applicable in training clinics? *Psychological Services*, 2, 65–69.
- Callahan, J. L., Swift, J. K., & Hynan, M. T. (2006). Test of the phase model of psychotherapy in a training clinic. *Psychological Services*, 3, 129–136.
- Clarkin, J. F., & Levy, K. N. (2003). Influence of client variables on psychotherapy. In M. Lambert (Ed.), *Handbook of psychotherapy and behavior change* (5th ed.). New York: Wiley.
- Corrigan, J. D., & Schmidt, L. D. (1983). Development and validation of revisions in the Counselor Rating Form. *Journal of Counseling Psychology*, 30, 64–75.
- Derogatis, L. R. (1992). *SCL-90-R: Administration, scoring & procedures manual-II* (2nd ed.). Towson, MD: Clinical Psychometric Research, Inc.
- Dodenhoff, J. T. (1981). Interpersonal attraction and direct-indirect supervisor influence as predictors of counselor trainee effectiveness. *Journal of Counseling Psychology*, 28, 47–52.
- Elkin, I., Gibbons, R. D., Shea, M. T., Sotsky, S., Watkins, J., & Pilkonis, P. (1995). Initial severity and differential treatment outcome in the National Institute of Health Treatment of Depression Collaborative Research Program. *Journal of Clinical and Consulting Psychology*, 63, 841–846.
- Ellis, M. V., & Ladany, N. (1997). Inferences concerning supervisees and clients in clinical supervision: An integrative review. In C. E. Watkins (Ed.), *Handbook of psychotherapy supervision* (pp. 447–507). New York: Wiley.
- Ellis, M. V., Ladany, N., Krenzel, M., & Schult, D. (1996). Clinical supervision research from 1981 to 1993: A methodological critique. *Journal of Counseling Psychology*, 43, 35–50.

- Falender, C. A., Cornish, J. A. E., Goodyear, R., Hatcher, R., Kaslow, N. J., Leventhal, G., et al. (2004). Defining competencies in psychology supervision: A consensus statement. *Journal of Clinical Psychology, 60*, 771–785.
- Freitas, G. J. (2002). The impact of psychotherapy supervision on client outcome: A critical examination of 2 decades of research. *Psychotherapy: Theory, Research, Practice, Training, 39*, 354–367.
- Garfield, S. L. (1994). Research on client variables in psychotherapy. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 190–228). New York: Wiley.
- Harkness, D. (1995). The art of helping in supervised practice: Skills, relationships, and outcomes. *The Clinical Supervisor, 13*, 63–76.
- Harkness, D., & Hensley, H. (1991). Changing the focus of social work supervision: Effects on client satisfaction and generalized contentment. *Social Work, 36*, 506–512.
- Hatchett, G. T., & Park, H. L. (2003). Comparison of four operational definitions of premature termination. *Psychotherapy: Theory, Research, Practice, Training, 40*, 226–231.
- Holloway, E., & Carroll, M. (1996). Reaction to the special section on supervision research: Comment on Ellis et al. (1996), Ladany et al. (1996), Neufeldt et al. (1996), and Worthen and McNeill (1996). *Journal of Counseling Psychology, 43*, 51–55.
- Holloway, E. L., & Neufeldt, S. A. (1995). Supervision: Its contributions to treatment efficacy. *Journal of Consulting and Clinical Psychology, 63*, 207–213.
- 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.
- Kadera, S. W., Lambert, M. J., & Andrews, A. A. (1996). How much therapy is really enough? A session-by-session analysis of the psychotherapy dose-effect relationship. *Psychotherapy: Research and Practice, 5*, 1–21.
- Kivlighan, D. M., Angelone, E. O., & Swafford, K. G. (1991). Live supervision in individual psychotherapy: Effects on therapist's intention use and client's evaluation of session effect and working alliance. *Journal of Counseling Psychology, 22*, 489–495.
- Lambert, M. J., & Arnold, R. C. (1987). Research and the supervisory process. *Professional Psychology: Research and Practice, 18*, 217–224.
- Lambert, M. J., Hansen, N. B., Umphress, V., Lunnen, K., Okiishi, J., & Burlingame, G. M. (1996). *Administration and scoring manual for the OQ-45.2*. Stevenson, MD: American Professional Credentialing Services LLC.
- Lewis, B. (2008, February 18). Social Justice Activity Survey Psychology Training Clinic Version. Message posted to Association of Directors of Psychology Training Clinics electronic mailing list, archived at [http://groups.google.com/group/ADPTC/browse\\_thread/thread/e5f1fe099ae4742c/ee383e39f2c71f05#ee383e39f2c71f05](http://groups.google.com/group/ADPTC/browse_thread/thread/e5f1fe099ae4742c/ee383e39f2c71f05#ee383e39f2c71f05)
- Martin, D. J., Garske, J. P., & Davis, M. K. (2000). Relation of the therapeutic alliance with outcome and other variables: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 68*, 438–450.
- McLellan, A. T., Luborsky, L., Woody, G. E., Druley, K. A., & O'Brien, C. P. (1983). Predicting response to alcohol and drug abuse treatments: Role of psychiatry severity. *Archives of General Psychiatry, 40*, 620–625.
- Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika, 78*, 691–692.
- Steinhebel, J., Patterson, V., Cliffe, K., & LeGoullon, M. (1984). An investigation of some relationships between psychotherapy supervision and patient change. *Journal of Clinical Psychology, 40*, 1346–1353.
- Stoltenberg, C. D. (2005). Enhancing professional competence through developmental approaches to supervision. *American Psychologist, 60*, 857–864.
- Wampold, B. E. (2001). *The great psychotherapy debate: Models, methods, and findings*. Mahwah, NJ: Erlbaum.
- Wierzbicki, M., & Pekarik, G. (1993). A meta-analysis of psychotherapy dropout. *Professional Psychology: Research and Practice, 24*, 190–195.

Received September 24, 2007

Revision received September 2, 2008

Accepted September 25, 2008 ■

### E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at <http://notify.apa.org/> and you will be notified by e-mail when issues of interest to you become available!