

# The Evidence Base for Evidence-Based Mental Health Treatments: Four Continuing Controversies

Peter E. Nathan, PhD

Substantial disagreement continues to divide mental health professionals on the evidence base that underlies evidence-based practices. This article considers four of the most hotly debated unresolved issues underlying this controversy; suggests the nature of evidence that might permit their resolution; and speculates on the consequences that might follow a failure in efforts at resolution. Resolution of these issues is crucial to the future of evidence-based practices. [*Brief Treatment and Crisis Intervention* 4:243–254 (2004)]

KEY WORDS: evidence-based treatments, empirically supported treatments, psychotherapy outcome studies, psychotherapy, efficacy vs. effectiveness, common factors vs. treatment factors, “dodo bird” effect, art vs. science.

---

Evidence-based practices—professional interventions based on reliable and valid empirical findings—now figure prominently in the work of most mental health practitioners (Callaghan, 2001; Nathan, 2003). Many believe we should prefer evidence-based treatments over those without empirical support because the former have been systematically compared to alternative treatments by appropriate and powerful methods and, for that reason, should provide us assurance of superior efficacy (Barlow, 1996).

Managed care organizations, third-party reimbursers, and state and federal agencies have come increasingly to expect psychologists, psychiatrists, and social workers whenever possible to employ practices with empirical support (Institute of Medicine, 2002; U.S. General Accounting Office, 2002). As a consequence, many psychiatric residents and psychology and social work students receive training that emphasizes these practices, as mandated by their respective professions (Borkovec, Echemendia, Ragusea, & Ruiz, 2001; Daniels, Alva, & Olivares, 2002). Practice guidelines incorporating evidence-based treatments have been put forth by the American Psychological Association and the American Psychiatric Association, as well as by the U.S. Department of Veterans Affairs and the U.S. Agency for Health

---

From the University of Iowa.

Contact author: Peter E. Nathan, Department of Psychology, College of Liberal Arts and Sciences, University of Iowa, E119 Seashore Hall, Iowa City, IA 52242.  
E-mail: peter-nathan@uiowa.edu.

DOI: 10.1093/brief-treatment/mhh021

Brief Treatment and Crisis Intervention Vol. 4 No. 3, © Oxford University Press 2004; all rights reserved.

Care Policy and Research. And more and more patients have come to expect their therapists to know and use empirically supported treatments whenever possible (Callaghan, 2001).

Despite these developments, substantial disagreement continues to divide mental health professionals on the strength and legitimacy of the evidence base that underlies evidence-based practices (Boisvert & Faust, 2003; Deegear & Lawson, 2003). This article considers four of the most hotly debated unresolved issues underlying this controversy; suggests the nature of evidence that might permit their resolution; and speculates on the consequences that might follow a failure in efforts at resolution. Resolution of these issues is essential to the future of evidence-based practices. If some or all of them can finally be resolved, agreement by most mental health professionals on the worth of evidence-based practices would seem to be assured. By the same token, if few or none can be settled, the momentum toward evidence-based practices will surely slow and ultimately stop.

### **Issue 1: Does the Efficacy Model or the Effectiveness Model Yield the Most Valid Picture of Psychotherapy Outcomes?**

A great deal of controversy continues over which of two psychotherapy outcome research models—the efficacy model or the effectiveness model—best captures the most crucial differences among therapy techniques and procedures and can hence be relied on to provide the most accurate picture of therapy outcomes. We refer to the efficacy model when we describe most carefully controlled, time-limited psychotherapy outcome research—much of it involving random assignment of patients to treatments (done largely by psychotherapy researchers in laboratory or other controlled settings), using therapists intensively trained to provide the experimental treatment and using

psychotherapy patients carefully selected diagnostically to receive it. By contrast, when we refer to the effectiveness model, we describe psychotherapy research done in real-world clinical settings, utilizing clinicians in their usual treatment settings doing the kind of psychotherapy they customarily do with the patients who customarily come to see them.

Psychotherapy research that employs the efficacy model is concerned above all with replication, because replicated psychotherapy outcome data are more likely to be valid. Efficacy studies contain a number of research elements that are not generally included in effectiveness studies. These elements include the following:

1. *randomized assignment of patients to treatment and comparison groups*, so that any differences in outcomes between or among the groups reflect differences in the efficacy of the treatments rather than any systematic bias in patient assignment to groups;
2. *inclusion of an active comparison treatment*, in preference to a no-treatment control, because comparing the experimental treatment to a treatment with demonstrated effectiveness is the strongest test of the efficacy of the experimental treatment;
3. *documentation of the delivery of treatments as planned*, to ensure that the treatments whose effectiveness is being compared are being delivered as designed and with fidelity;
4. *reliance on multiple outcome measures*, so that as many relevant behavioral changes attributable to the treatments as possible will be assessed;
5. *appropriate length of treatment follow-up*, so that the “staying power” of the experimental and comparison treatments can be compared.

Although these key methodological features of the efficacy model are necessary, they are not sufficient to prove that an experimental treatment has shown empirical support. The experimental treatment must also yield significantly better outcomes for significantly more patients than comparison treatments. Such findings, moreover, ought to be replicated by more than a single team of investigators (Nathan, 2003).

Effectiveness research, by contrast, is concerned above all with the feasibility of treatments in real-world settings. Effectiveness studies incorporate persons in need of treatment, regardless of diagnosis, comorbid psychopathology, or duration of illness. Because they are usually active practitioners in research, rather than participants, therapists in effectiveness studies are not usually trained to deliver, and do not generally expect to be asked to provide, a specific experimental or comparison treatment that requires the extensive training that is a hallmark of efficacy studies. Clinical considerations, rather than the demands of a research protocol, largely dictate choice of treatment method, as well as its frequency, duration, and means of outcome assessment. Although assignment of patients to treatments in effectiveness studies may be randomized, disguising (“blinding”) the treatment to which the patient has been assigned is rarely feasible. Outcome assessments are often broadly defined and may include such “soft” indexes as changes in degree of disability, quality of life, or personality, rather than efficacy studies’ preference for targeted evaluations of symptoms by means of structured interviews.

Barlow drew a useful and concise distinction between efficacy studies and effectiveness studies in 1996: efficacy studies yield “a systematic evaluation of the intervention in a controlled clinical research context. Considerations relevant to the internal validity of these conclusions are usually highlighted”

(p. 1051); by contrast, effectiveness studies explore “the applicability and feasibility of the intervention in the local setting where the treatment is delivered” and are designed to “determine the generalizability of an intervention with established efficacy” (p. 1055).

Most of the research that has led to identification of evidence-based treatments to this time has been done according to the efficacy model. As a result, critics of the evidence underlying evidence-based treatments have claimed that efficacy studies do not reflect therapy outcomes “in the real world” (Garfield, 1996; Seligman, 1996; Westen & Morrison, 2001; Westen, Morrison, & Thompson-Brenner, in press). While supporters of the efficacy model have mounted a vigorous defense of the treatment model (Hollon, 1996; Jacobson & Christensen, 1996; Wilson, 1995), the question of which of the two models provides the most valid picture of psychotherapy outcomes remains unresolved. Until it is, the evidence base of evidence-based treatments will remain suspect.

The efficacy model and the effectiveness model represent quite different approaches to studying behavior change. Because neither model by itself appears to capture the entirety of what makes a treatment effective, clinical researchers have begun to try to integrate the two approaches in the design of psychotherapy research in the effort to gather the most broadly based empirical support for the treatments being evaluated. For instance, Norquist, Lebowitz, and Hyman (the latter the director of NIMH at the time) acknowledged that “the intrinsic efficacy of an intervention (either pharmacological or psychotherapeutic) . . . is not usually informative for treatment practice in the community” (1999, p. 1). They proposed that NIMH—in consultation with basic scientists, advocates, and other federal agencies—bridge the gap between regulatory (efficacy) and public health (effectiveness) models

to assess outcomes. Their proposal incorporated experimental and observational work, albeit after each received methodological changes. While their ideas were clearly in the process of evolution at the time they made the proposal, Norquist and his colleagues specified the following:

Research designs must permit a loosening of exclusion criteria to allow for enrollment of people with different levels of disease severity and comorbid conditions. In addition, treatment settings must be more diversified to allow for a range of providers from primary care, managed care settings to tertiary care, academic centers. Outcome measures will need to incorporate domains that are important to consumers, families and policy makers (e.g., performance, disability, cost, resource use, etc.). . . . In addition, specific areas of treatment intervention research need to be launched (e.g., rehabilitation research) and reinvigorated (e.g., psychosocial intervention research). (Norquist et al., 1999, p. 6)

With similar goals in mind, Klein and Smith (1999) proposed the development of “dedicated, multi-site efficacy/effectiveness clinics” (p. 1) to address the problems posed by the conflicting demands of internal and external validity in efficacy and effectiveness studies. These special clinics would also be structured to examine such understudied treatment issues as compliance, comorbidity, refractory illness, and withdrawal syndromes, as well as adjunctive and maintenance treatments. The clinics would emphasize careful, reliable, documented studies on process and outcome. They would also promote development of outcome norms for well-defined populations on such variables as diagnosis, economic status, history, and comorbidity. The proximate goal would be to generate, across cooperating clinics, “a large volume of well-delineated patients (who) could

be treated and studied who may have high comorbidity with medical, psychiatric, and substance abuse conditions” (1999, p. 5). The distal goals would be, first, to develop benchmarks for expected treatment outcomes for these distinct groups of patients by means of normative sampling; and, second, to serve as hypothesis-generating therapeutic endeavors.

Both the NIMH proposal and Klein and Smith’s plan to resolve the efficacy–effectiveness issue are long on enthusiasm, problem identification, and aspirations for change but are a good deal short on concrete details on design, methodology, and statistical analysis. This relative lack of substance leaves the reader who appreciates some of the problems attendant on integrating efficacy and effectiveness studies uncertain to what extent integration can actually be affected.

If the continuing efficacy/effectiveness controversy cannot be resolved, then

1. treatments already established as evidence based by efficacy studies may begin to be discounted, and the legitimacy of the evidence-based psychotherapy movement may begin to be questioned; and
2. advocates for reliance on “clinical judgment,” rather than on empirical data, in choosing treatments may start to be listened to more seriously.

## **Issue 2: Which Contributes the Most Variance to Psychotherapy Outcomes, Common Factors or Treatment Factors?**

The controversy over Issue 1—the relative validity of efficacy versus effectiveness studies in assessing treatment outcomes—came to a head with the publication of practice guidelines by the American Psychiatric Association (1993, 1994, 1995, 1996, 1997) and the Division

of Clinical Psychology of the American Psychological Association (Chambless et al., 1996, 1998; Task Force on Promotion and Dissemination of Psychological Procedures, 1995). However, the appearance of these guidelines also led critics of evidence-based practices to step up the amplitude of a related continuing controversy: whether treatment factors or common factors account for more of the variance associated with treatment outcomes—that is, whether differences in psychotherapy outcomes are more strongly associated with specific types or schools of psychotherapy (as many psychotherapists believe) or with therapist, patient, and therapy process variables common to all psychological treatments (as many clinical researchers claim). Parenthetically, if the so-called common factors turn out to be most important, that would explain the “dodo bird” effect, the continuing controversy discussed next, as Issue 3.

Treatment factors refer to the array of therapeutic behaviors and techniques a therapist is taught and must learn as he or she acquires the skills appropriate to the practice of a specific intervention. For example, a therapist treating an anxiety disorder from the cognitive-behavioral perspective must learn such techniques as constructing an anxiety hierarchy, with the patient’s cooperation; exposing the patient to that hierarchy, both in the office and in vivo (in real-life situations); and helping the patient develop alternative behaviors that decrease his or her having to face anxiety-provoking situations in their full intensity. By contrast, common factors refer to patient attributes such as age, gender, and personality; therapist attributes such as interpersonal and social skills; and therapeutic process factors such as the nature of the relationship between therapist and patient that may influence the outcomes of many therapies.

Through the years, a number of well-respected psychotherapy researchers have con-

cluded that common factors account for a substantial amount of the treatment outcome variance. In an influential 1994 article, Lambert and Bergin identified therapist, patient, and therapeutic process as independent sources of common factors. Therapist variables thought to affect therapy outcomes (regardless of the kind of therapy techniques the therapist uses) range from the therapist’s demographic characteristics and sociocultural background to subjective factors such as values, attitudes and beliefs. Beutler, Machado, and Neufeldt (1994) have written that therapist variables reflecting behaviors specific to the therapeutic relationship—including the therapist’s professional background, style and choice of interventions—may exert the most powerful effects on therapy outcomes.

In contrast to the voluminous data attesting to the impact of therapist variables on therapy outcomes, patient variables have failed to demonstrate a robust relationship to outcome variables (Nathan, Stuart, & Dolan, 2000). Consider, for example, the well-known National Institute of Mental Health Treatment of Depression Collaborative Research Program (NIHM-TDCRP; Elkin et al., 1989), a comparative study of treatments for depression. In that study, no single patient variable correlated significantly with outcome. More recently, Project MATCH failed to identify relationships between patient-treatment matches and outcomes of treatment for alcohol abuse and dependence (Project MATCH Research Group, 1997).

Therapeutic process variables—factors influencing therapists’ reactions to patients’ behavior and attitudes, and vice versa—have also been claimed to affect therapy outcomes. To this end, Orlinsky and Howard (1986) concluded that process variables—which they believed also included the strength of the therapeutic bond, the skillfulness with which interventions are undertaken, and the duration of the treatment relationship—all positively

affect outcomes. Nonetheless, critics of process research have continued to emphasize the difficulties associated with the reliable collection of process data (e.g., Stiles & Shapiro, 1989).

In an extensive review of outcome research data, Lambert (1992) concluded that about 30% of psychotherapy outcome variance is attributable to therapist variables, prominently including therapist empathy, warmth, and acceptance of the patient. Others—including Svartberg, Seltzer, and Stiles (1998); and Horvath and Symonds (1991), following Strupp (1973)—have stressed the central role of the therapeutic alliance in determining outcomes. At the same time, a sizable core of mental health professionals and researchers—prominently including the developers of mental health practice guidelines—continue to assert that choice of therapy technique plays a significant role in determining therapy outcomes (Nathan, 1998). If the continuing common factors—techniques controversy cannot be resolved, programs that train psychotherapists may shift their training emphases from specific therapeutic techniques to generic therapist skills.

Managed-care organizations may more aggressively seek the lowest-cost providers, especially those without extensive professional education and training who can nonetheless claim to have acquired the therapist attributes that have been linked to positive outcomes.

### **Issue 3: The “Dodo Bird” Effect—Are Most Psychosocial Treatments Equally Effective?**

The “dodo bird” effect, so designated by Saul Rosenzweig in a prescient 1936 article on common factors in diverse psychotherapies, is named after the well-known race in Lewis Carroll’s *Alice in Wonderland*. At the conclusion of that contest, the dodo bird declares that since the racers, including Alice, had failed to keep to

the racecourse, it was impossible to know who had won; and so, “everyone has won, and all must have prizes.” Thus, the “dodo bird” effect in psychotherapy research refers to findings that indicate that there are few or no meaningful differences among psychotherapies in effectiveness and that, accordingly, outcomes of therapy don’t really depend on the kind of therapy patients.

Acknowledging their debt to Rosenzweig for his initial reference to the dodo bird, Luborsky, Singer, and Luborsky (1976) titled their review chapter in an edited volume on psychotherapy evaluation “Comparative Studies of Psychotherapies: Is It True That ‘Everyone Has Won and All Must Have Prizes?’” Luborsky and his colleagues compared outcomes from group and individual psychotherapy, time-limited and open-ended psychotherapy, and client-centered and psychodynamic therapy, concluding that “most comparative studies of different forms of psychotherapy found insignificant differences in proportions of patients who improved by the end of psychotherapy” (1976, p. 12). Adopting a system to grade methodological quality that extended from A (“the main principles of research design were satisfied”) to E (“the deficiencies were so serious that the results were not included”), the authors reviewed “only studies in which some attention was paid to the main requirements of controlled comparative research.” These “requirements” included a number of elements of efficacy research listed at the beginning of our consideration of Issue 1, including random assignment of subjects, suitable controls, and blind-outcome assessment. Missing features included careful specification of the form of treatment, an appropriately extended follow-up period, and follow-up assessments that focused on specific, rather than more general, outcome behaviors. These authors concluded their observations with a plea for more adequate

research methodologies for examining, weighing, and comparing treatment efficacy. Furthermore, they made the following statement:

Our title implies what I think many of us believe—that all psychotherapies produce some benefits for some patients. What we do not know is whether there are some psychotherapies which produce significantly better results and whether certain psychotherapies are especially well suited to certain patients. (Luborsky et al., 1976, p. 3)

More recently, the “dodo bird” effect has been taken to refer to efficacy comparisons among psychotherapies by means of meta-analysis that have failed to find differences in efficacy. While Luborsky et al. (1976) were the first contemporary researchers to use this term for this phenomenon, Smith and Glass (1977); Stiles, Shapiro, and Elliott (1986); Wampold and his colleagues (1997); and a number of other researchers subsequently adopted the same position—that most psychotherapies are effective in inducing behavior change but that they do not differ in efficacy. However, a number of behavioral researchers—including Krasner (1971), Bergin and Suinn (1975), Rachman and Wilson (1980), Nathan, Stuart, and Dolan (2000), and Chambless and Ollendick (2001)—have vigorously disputed this position. They point to data from a number of randomized clinical trials that strongly suggest that some psychosocial treatments—most of them behavioral or cognitive behavioral—do appear to yield significantly better outcomes than do other psychotherapies. They also cite data that find substantial problems with meta-analysis as a means of comparing the differential efficacy of psychotherapies (e.g., Cris-Christoph, 1997; Howard, Krause, Saunders, & Kopta, 1997). As noted, the drafters of practice guidelines by the American Psychiatric Association, the American Psychological Association, and the Mental

Health and Behavioral Science Service of the Veterans Administration, among others, have taken a similar position in recommending some treatments over other treatments on the basis of outcome studies.

If the continuing controversy regarding the “dodo bird” effect cannot be resolved, then

1. the influence of practice guidelines, which depend largely on professionals’ agreeing that some treatments are more effective than others, may diminish; and
2. programs that train mental health professionals in psychosocial treatments may decide to de-emphasize training in specific techniques in favor of training in common therapist factors such as empathy, warmth, and unconditional positive regard.

#### **Issue 4: Teaching, Learning, and Doing Therapy—Art or Science?**

The issue of art versus science in the mental health disciplines underlies all three issues discussed so far. The art-versus-science contrast also has a specific and somewhat different set of referents, which will be briefly reviewed at this point.

Although most of us believe that medical practice has always been heavily influenced by scientific data, at least as science has been conceived through the years, that assumption does not appear to be valid. In fact, the roots of evidence-based medical practice (EBMP) are now usually traced to a statement published in the *Journal of the American Medical Association* a little over a decade ago. In 1992, Gordon Guyatt, who was an internist in Hamilton, Ontario, led a group of physicians who advocated evidence-based medical practice over medicine as an art. The article “ignited a debate

about power, ethics, and responsibility in medicine that is now threatening to radically change the experience of health care" (Patterson, 2002). The change in medicine over the past decade and more, then, has been away from intuition—that is, away from the art in medicine—and more toward the science in medicine. Nonetheless, divisions among physicians over the value of intuition continue: "The EBMP position is that there are reliable, validated data, and then there are data that aren't reliable and validated, and that's really what matters" (Patterson, 2002).

The parallels are obvious between evidence-based medical practice and evidence-based mental health practice—in particular, their respective histories, current statuses, and member reactions (both popular and professional). What advocates for evidence-based treatments in psychology, psychiatry, and social work currently hear from their critics would be all too familiar to what advocates for evidence-based medicine have heard over the past decade, just as what advocates are saying to their critics to justify evidence-based mental health practice would be familiar to physicians who support evidence-based medicine.

This issue has a particularly lengthy and important history in clinical psychology. It dates back more than 60 years, to the era of the explosive growth of clinical psychology during and after World War II, during a time when it was transformed from a small, primarily academic discipline to one of the core mental health professions. In 1942, social psychologist Theodore Sarbin predicted, on the basis of some of his own data, that actuarial prediction methods ("science") would ultimately be able to outperform humans ("art") along a variety of judgment dimensions. And in 1954, Paul Meehl, who was to become one of the towering figures in clinical psychology over the next several decades, published a book that summarized his data confirming the consistent superiority of

statistical prediction ("science") over clinical prediction ("art"). Although many advocates for art over science in clinical psychology since then have been heard from, support from behavioral scientists in favor of Sarbin's and Meehl's positions has been strong and consistent (e.g., Goldberg, 1965; Grove et al., 2000; Sines, 1971).

Campaign statements from two recent candidates for the presidency of the American Psychological Association epitomize the continuing difference in views on the matter:

*Art:* The present body of scientific evidence is not sufficiently developed to serve as the sole foundation for practice. . . . We (need to) bring to bear . . . what is often referred to as clinical judgment. (Ron Levant, 2003)

*Science:* We are facing a crisis, once again, within APA. In the past year, while campaigning for president, I have become aware of some very strong opposition to the definition of psychology as a discipline built on an emerging and developing scientific foundation. (Larry Beutler, 2003)

Notwithstanding opposition to evidence-based practice by some mental health professionals, increasing efforts are being expended to require mental health practitioners to follow practice guidelines, and practice guidelines are becoming more prescriptive. American Psychological Association accreditation criteria now specify that these treatments must be taught to graduate students in approved programs. The new Department of Veterans Affairs guidelines for treatment of serious mental illness, substance abuse and dependence, and posttraumatic stress disorder are evidence based and quite prescriptive. Demands from health management organizations are increasingly being felt for evidence of efficacy before additional psychosocial treatments are authorized.

If the continuing controversy over the respective roles of art and science in mental health practice cannot be resolved, then

1. the mental health professions may return to an earlier training-and-practice model, when science was seen as substantially less central to the clinical enterprise than it is today and when art (“intuition,” “clinical judgment,” and the like) was more important; and
2. patients may choose practitioners on the basis of the eloquence of their artistry rather than on the validity of their science.

## Summary and Conclusions

The well-known complexity of the psychotherapy process—in particular, the range of variables in the relationship between patient and therapist that are capable of influencing therapy outcomes—has led advocates for, and critics of, evidence-based treatments to the present stalemate. On the one hand, advocates point to voluminous data attesting to the power of efficacy studies to differentiate among therapies in outcome; to the association of specific kinds of treatment, most often cognitive-behavioral, with positive outcomes; and to the failure of “art” through history in its struggle with science for primacy. On the other hand, critics have no problem citing research in each instance that supports their view that these four questions cannot be resolved nearly as simply and expeditiously as the advocates would like.

While the most attractive solution to this deadlock is to suggest that partisans on each side of this issue tone down their rhetoric until enough data have been gathered to resolve each of these four crucial questions, that solution is

unlikely to satisfy either side. Instead, I anticipate that the advocates and the critics of evidence-based treatments will continue their noisy dialogue until the court of public opinion makes its decision, pro or con. When and how that will come is anybody’s guess. Meanwhile, those of us quite convinced of the wisdom of our positions will continue to work hard to defend them.

## References

- American Psychiatric Association. (1993). Practice guideline for the treatment of major depressive disorder in adults. *American Journal of Psychiatry*, *150* (4, Suppl.), 1–26.
- American Psychiatric Association. (1994). Practice guideline for the treatment of patients with bipolar disorder. *American Journal of Psychiatry*, *151* (12, Suppl.), 1–36.
- American Psychiatric Association. (1995). Practice guideline for the treatment of patients with substance use disorders: Alcohol, cocaine, opioids. *American Journal of Psychiatry*, *152* (11, Suppl.), 1–59.
- American Psychiatric Association. (1996). Practice guideline for the treatment of patients with nicotine dependence. *American Journal of Psychiatry*, *153* (10, Suppl.), 1–31.
- American Psychiatric Association. (1997). Practice guideline for the treatment of patients with schizophrenia. *American Journal of Psychiatry*, *154* (4, Suppl.), 1–63.
- Barlow, D. H. (1996). Health care policy, psychotherapy research, and the future of psychotherapy. *American Psychologist*, *51*, 1050–1058.
- Bergin, A. E., & Suinn, R. M. (1975). Individual psychotherapy and behavior therapy. In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual review of psychology* (Vol. 26, pp. 509–556). Palo Alto, CA: Annual Reviews.
- Beutler, L. E., Machado, P. P. P., & Neufeldt, S. A. (1994). Therapist variables. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy*

- and behavior change (4th ed., pp. 229–269). New York: Wiley.
- Boisvert, C. M., & Faust, D. (2003). Leading researchers' consensus on psychotherapy research findings: Implications for the teaching and conduct of psychotherapy. *Professional Psychology: Research and Practice, 34*, 508–513.
- Borkovec, T. D., Echemendia, R. J., Ragusea, S. A., & Ruiz, M. (2001). The Pennsylvania Practice Research Network and possibilities for clinically meaningful and scientifically rigorous psychotherapy effectiveness research. *Clinical Psychology: Science and Practice, 8*, 155–167.
- Callaghan, G. M. (2001). Demonstrating clinical effectiveness for individual practitioners and clinics. *Professional Psychology: Research and Practice, 32*, 289–297.
- Chambless, D. L., Baker, M. J., Baucom, D. H., Beutler, L. E., Calhoun, K. S., Crits-Christoph, P., et al. (1998). Update on empirically validated therapies, II. *Clinical Psychologist, 51*, 3–16.
- Chambless, D. L., & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. In S. T. Fiske, D. L. Schacter, & C. Zahn-Waxler (Eds.), *Annual review of psychology* (Vol. 52, pp. 685–716). Palo Alto, CA: Annual Reviews.
- Chambless, D. L., Sanderson, W. C., Shoham, V., Bennett Johnson, S., Pope, K. S., Crits-Christoph, P., et al. (1996). An update on empirically validated therapies. *Clinical Psychologist, 49*, 5–18.
- Crits-Christoph, P. (1997). Limitations of the dodo bird verdict and the role of clinical trials in psychotherapy research: Comment on Wampold et al. (1997). *Psychological Bulletin, 122*, 216–220.
- Daniels, J. A., Alva, L. A., & Olivares, S. (2002). Graduate training for managed care: A national survey of psychology and social work programs. *Professional Psychology: Research and Practice, 33*, 587–590.
- Deegear, J., & Lawson, D. M. (2003). The utility of empirically supported treatments. *Professional Psychology: Research and Practice, 34*, 271–277.
- Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., et al. (1989). National Institute of Mental Health Treatment of Depression Collaborative Research Program: General effectiveness of treatments. *Archives of General Psychiatry, 46*, 971–982.
- Garfield, S. L. (1996). Some problems associated with “validated” forms of psychotherapy. *Clinical Psychology: Science and Practice, 3*, 218–229.
- Goldberg, L. R. (1965). Diagnosticians vs. diagnostic signs: The diagnosis of psychosis and neurosis from the MMPI. *Psychological Monographs, 79*(9, Whole No. 602).
- Grove, W. M., Zald, D. H., Hallberg, A. M., Lebow, B., Snitz, E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment, 12*, 19–30.
- Hollon, S. D. (1996). The efficacy and effectiveness of psychotherapy relative to medications. *American Psychologist, 51*, 1025–1030.
- Horvath, A. O., & Symonds, D. B. (1991). Relationship between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology, 38*, 139–149.
- Howard, K. L., Krause, M. S., Saunders, S. M., & Kopta, S. M. (1997). Trials and tribulations in the meta-analysis of treatment differences: Comments on Wampold et al. (1997). *Psychological Bulletin, 122*, 221–225.
- Institute of Medicine. (2002). *Medical innovation in the changing healthcare marketplace: Conference summary*. Washington, DC: National Academy Press.
- Jacobson, N. S., & Christensen, A. (1996). Studying the effectiveness of psychotherapy: How well do clinical trials do the job? *American Psychologist, 51*, 1031–1039.
- Klein, D. F., & Smith, L. B. (1999). Organizational requirements for effective clinical effectiveness studies. *Prevention & Treatment, 2*, Article 0002a. Retrieved May 1, 2004, from <http://journals.apa.org/prevention/volume2/pre0020002a.html>
- Krasner, L. (1971). Behavior therapy. In P. H. Mussen & M. R. Rosenzweig (Eds.), *Annual review of psychology* (Vol. 22, pp. 483–532). Palo Alto, CA: Annual Reviews.
- Lambert, M. J., (1992). Psychotherapy outcome research: Implications for integrative and

- eclectic therapies. In J. C. Norcross & M. R. Goldfried (Eds.), *Handbook of psychotherapy integration* (pp. 94–129). New York: Basic Books.
- Lambert, M. J., & Bergin, A. E. (1994). The effectiveness of psychotherapy. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 143–189). New York: Wiley.
- Luborsky, L., Singer, B., & Luborsky, L. (1976). Comparative studies of psychotherapies: Is it true that “everybody has won and all must have prizes”? In R. L. Spitzer & D. F. Klein (Eds.), *Evaluation of psychological therapies* (pp. 3–22). Baltimore: Johns Hopkins University Press.
- Meehl, P. E. (1954). *Clinical vs. statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis: University of Minnesota Press.
- Nathan, P. E. (1998). Practice guidelines: Not yet ideal. *American Psychologist*, *53*, 290–299.
- Nathan, P. E. (2003). *Empirically-supported treatments, public policy, and public practice*. Workshop at the annual meeting of the New Mexico Psychological Association, Albuquerque, October 11, 2003.
- Nathan, P. E., Stuart, S. P., & Dolan, S. L. (2000). Research on psychotherapy efficacy and effectiveness: Between Scylla and Charybdis? *Psychological Bulletin*, *126*, 964–981.
- Norquist, G., Lebowitz, B., & Hyman, S. (1999). Expanding the frontier of treatment research. *Prevention & Treatment*, *2*, Article 0001a. Retrieved May 1, 2004, from <http://journals.apa.org/prevention/volume2/pre0020001a.html>
- Orlinsky, D. E., & Howard, K. I. (1986). Process and outcome in psychotherapy. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (3rd ed., pp. 311–381). New York: Wiley.
- Patterson, K. (2002, May 5). What doctors don't know (almost everything). *New York Times Magazine*, 74–77.
- Project MATCH Research Group. (1997). Matching alcoholism treatment to client heterogeneity: Project MATCH post-drinking outcomes. *Journal of Studies on Alcohol*, *58*, 7–29.
- Rachman, S., & Wilson, G. T. (1980). *The effects of the psychological therapies*. Oxford: Pergamon Press.
- Rosenzweig, S. (1936). Some implicit common factors in diverse methods in psychotherapy. *American Journal of Orthopsychiatry*, *6*, 412–415.
- Sarbin, T. R. (1942). A contribution to the study of actuarial and individual methods of predictions. *American Journal of Sociology*, *48*, 593–602.
- Seligman, M. E. P. (1996). Science as an ally of practice. *American Psychologist*, *51*, 1072–1079.
- Sines, J. O. (1971). Actuarial versus clinical prediction in psychopathology. *British Journal of Psychiatry*, *116*, 129–144.
- Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, *32*, 752–760.
- Stiles, W. B., & Shapiro, D. A. (1989). Abuse of the drug metaphor in psychotherapy process-outcome research. *Clinical Psychology Review*, *9*, 521–543.
- Stiles, W. B., Shapiro, D. A., & Elliott, R. (1986). Are all psychotherapies equivalent? *American Psychologist*, *41*, 165–180.
- Strupp, H. H. (1973). *Psychotherapy: Clinical, research, and theoretical issues*. New York: Jason Aronson.
- Svartberg, M., Seltzer, M. H., & Stiles, T. C. (1998). The effects of common and specific factors in short-term anxiety-provoking psychotherapy. *Journal of Nervous and Mental Disease*, *186*, 691–696.
- Task Force on Promotion and Dissemination of Psychological Procedures. (1995). Training in and dissemination of empirically-validated psychological treatments: Report and recommendations. *Clinical Psychologist*, *48*, 3–23.
- U.S. General Accounting Office. (2002). *Community health centers: Adapting to changing health care environment key to continued success* (GAO/HEHS-00-39). Washington, DC: U.S. Government Printing Office.

- Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H.-N. (1997). A meta-analysis of outcome studies comparing *bona fide* psychotherapies: Empirically, "all must have prizes." *Psychological Bulletin*, *122*, 203–215.
- Westen, D., & Morrison, K. (2001). A multidimensional meta-analysis of treatments for depression, panic, and generalized anxiety disorder: An empirical examination of the status of empirically supported therapies. *Journal of Consulting and Clinical Psychology*, *60*, 875–899.
- Westen, D., Morrison, K., & Thompson-Brenner, H. (in press). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin*.
- Wilson, G. T. (1995). Empirically supported treatments as a basis for clinical practice: Problems and prospects. In S. C. Hayes, V. M. Follette, R. M. Dawes, & K. E. Grady (Eds.), *Scientific standards of psychological practice: Issues and recommendations* (pp. 163–196). Reno, NV: Context Press.