# The Science and Practice of Empirically Supported Treatments

**JAMES D. HERBERT**
*Drexel University*

Despite impressive gains over the past three decades in the development and evaluation of empirically supported psychotherapies, such treatments are not used widely by front-line practicing clinicians. In an attempt to address this science-practice gap, efforts have turned recently to constructing lists of empirically supported treatments (ESTs) and disseminating information about these treatments to professionals and the public. This effort has been met with criticism, however, by both practitioners, on one hand, and psychotherapy researchers on the other. The current procedures for identifying ESTs are critically reviewed, and recommendations are offered to improve the scientific viability of the process. It is argued that lists of ESTs are viewed most productively as one step toward the development of best practice guidelines.

***Keywords:*** *empirically supported treatments, ESTs, psychotherapy dissemination, evidence-based medicine, psychotherapy*

**Despite important advances** in the development and empirical evaluation of psychotherapies over the past three decades, the chasm between clinical practice and the scientific literature is as wide as ever. In 1993, the Division of Clinical Psychology of the American Psychological Association established the Task Force on the Promotion and Dissemination of Psychological Procedures with the goal of establishing criteria for identifying empirically supported treatments (ESTs), constructing lists of such treatments, and disseminating this information to various stakeholders. This effort was spawned in large measure by concerns over the gap between science and practice with respect to

---

AUTHOR'S NOTE: Correspondence concerning this article should be addressed to James D. Herbert, Department of Psychology, Drexel University, Mail Stop 988, 245 N. 15th St., Philadelphia, PA 19102-1192; e-mail: james.herbert@drexel.edu.

BEHAVIOR MODIFICATION, Vol. 27 No. 3, July 2003 412-430
DOI: 10.1177/0145445503253836

psychotherapy. The Task Force, now known as the standing Committee on Science and Practice (CSP), published its initial report in 1995, with updates in 1996 and 1998 (Chambless et al., 1996, 1998; Task Force, 1995). Although not official Task Force publications, several other recent articles published by members of the CSP outline the committee's work to date and discuss future directions (Chambless & Hollon, 1998; Chambless & Ollendick, 2001; Sanderson, this issue; Weisz, Hawley, Pilkonis, Woody, & Follette, 2000). The effort to identify and promote ESTs is consistent with the broader evidence-based medicine movement initially developed in the United Kingdom (Gambrill, 1999; Sackett, Richardson, Rosenberg, & Haynes, 1997).

Although on initial consideration the idea of identifying lists of ESTs would seem to be uncontroversial, the project has in fact proved quite contentious. Most criticisms have centered on three broad themes: (a) concerns about the impact of lists of ESTs on practicing clinicians; (b) the rejection of empirical epistemologies in general, and randomized clinical trials in particular, as legitimate methods for evaluating psychotherapies; and (c) the scientific status of the specific procedures for identifying ESTs. Many clinicians object to the effort on the grounds that it oversimplifies the treatment process and undermines clinical judgment and the autonomy of the practicing professional. Several observers—including both clinicians and psychotherapy researchers—have discussed practitioners' wariness of treatment manuals, which many view as overly restrictive, inadequate to capture the myriad nuances of psychotherapy, and not applicable to the complex cases that are typically seen in most settings (Addis, Wade, & Hatgis, 1999; Garfield, 1996; Lambert, 1998; Parloff, 1998; Raw, 1993; Silverman, 1996). Indeed, many of the reservations about ESTs among practitioners stem directly from their disdain for treatment manuals. Some practitioners have also voiced concerns about potential negative economic consequences of listing ESTs, especially in relation to the further demise of long-term psychotherapy, as well as the focus of ESTs on symptom reduction rather than psychological growth and self-actualization (Fensterheim & Raw, 1996). As discussed below, many practicing psychotherapists reject the notion that experimental scientific studies can possibly provide adequate tests of psychotherapies as they are implemented in the community.

On the other end of the ideological spectrum, many scientifically minded clinical psychologists have also been critical of the work of the CSP. Like their practitioner colleagues, some clinical scientists also object to the very idea of lists of ESTs on the grounds that they oversimplify inherently complex phenomena (Nezu, 1996). In addition, as discussed below, psychotherapy researchers have raised a number of criticisms of the movement, the most common and important of which concern the criteria established by the CSP for identifying ESTs.

## SHOULD WE EVEN TRY TO LIST ESTs?

These concerns raise a fundamental question: Should we establish lists of ESTs? Because whatever procedures are adopted for distinguishing ESTs will necessarily be imperfect and controversial, is the effort doomed to failure?

How one answers these questions follows largely from the value one places on quantitative outcome research in the development, evaluation, and dissemination of psychotherapies (Craighead & Craighead, 1998; Nathan, 2000). Those who see little value in controlled clinical trials are unlikely to be persuaded that any listing of ESTs is a good idea. On the other hand, those who view research as critical to the evaluation of psychotherapy (this would include various stakeholders in addition to scientifically minded psychologists, including many clients and third-party payers) are likely to be more favorably disposed to lists of ESTs, provided that the standards on which they are based are scientifically sound.

Perhaps the strongest case for the effort to identify ESTs lies in the failure of many practitioners to use evidenced-based treatments in their work, despite dramatic advances in psychotherapy over the past three decades (Barlow, Levitt, & Bufka, 1999; Hayes, 1996; Sanderson, 2002; Sechrest, 1992). For example, studies have found that only a minority of patients treated for an anxiety disorder receive an empirically supported treatment, despite the well-documented effectiveness of such treatments (Goisman et al., 1993; Goisman, Warshaw, & Keller, 1999; Taylor, King, & Margraf, 1989). For a variety of reasons,

many clients do not routinely receive empirically supported, state-of-the-art psychotherapies. Several factors contribute to this science-practice gap, including cognitive biases inherent in common methods of clinical decision making, widespread theoretical perspectives that devalue empiricism, and the sheer volume of the research literature. Human judgment and decision making are subject to a variety of cognitive heuristics that often lead to erroneous conclusions (Kahneman, Slovic, & Tversky, 1982; Mumma, 2001; Nisbett & Ross, 1980). The effects of these cognitive biases on clinical prediction are well known, with a large body of literature demonstrating consistently that the predictions of expert clinical judges are less accurate than those generated by relatively simple statistical models (Dawes, 1994; Dawes, Faust, & Meehl, 1989; Meehl, 1986). These biases can also readily lead clinicians to overestimate the effectiveness of their interventions. For example, positive clinical outcomes may be more likely to be recalled than treatment failures, leading to an exaggerated perception of effectiveness. The typical clinical setting does not permit the distinction of true intervention effects from those due to a variety of other sources, including developmental maturation, statistical regression to the mean, demand characteristics inherent in the setting, and various placebo effects. Not surprisingly, clinicians tend to attribute positive changes to their interventions—especially to the unique, distinctive, and highly vivid aspects of their interventions—rather than these other factors.

A closely related factor contributing to the science-practice gap is the popularity of theoretical paradigms that embrace epistemologies based on personal experience rather than controlled data. Psychoanalytic theory has traditionally been based on the idea that valid theoretical generalizations can only be gleaned from the careful idiographic study of individual cases in uncontrolled clinical settings (Safran, 2001). Likewise, existential and humanistic therapists tend to view controlled research as dehumanizing and unable to capture the critical therapist-client relationship features responsible for change (Bohart, O'Hara, & Leitner, 1998). The devaluation of empirical research has been reinforced by the recent growth of postmodernism in both academic and clinical settings, which eschews the methods of natural science in favor of subjective "narrative truths." The increasing popular-

ity of "holism," the idea that one must consider the complex whole rather than individual mechanisms of change, has further reinforced the devaluation of scientific data in clinical decision making (Ruscio, 2002).

Finally, even among clinicians favorably disposed to empiricism, the rapid pace of developments and the enormous volume of the scientific psychological literature can make it difficult to stay abreast of the latest research-based findings. Expecting clinicians to read and digest the primary research literature in psychotherapy, especially for the full range of disorders and problems that they encounter in their work, is simply unrealistic. Indeed, even psychotherapy researchers struggle to keep up with new developments outside their own areas of specialization.

The development of empirically based practice guidelines may be one way to begin bridging the science-practice gap.[1] Identifying empirically supported treatments can be a useful first step in the development of such guidelines. Although lists of ESTs may be useful, they are in and of themselves insufficient guides for the clinician because they do not address several important issues that must be considered in clinical decision making. These include the relative effectiveness of alternative interventions, the average effect sizes and numbers of individuals helped by alternative interventions, and analyses of cost-effectiveness. These issues are discussed further below.

## PROBLEMS WITH THE CURRENT GUIDELINES FOR DEFINING ESTs

To be ultimately useful, practice guidelines and the lists of treatment procedures on which they are based must be grounded in science. As I have argued elsewhere (see Herbert, 2000), a serious problem arises when the sociopolitical goals of developing lists of ESTs trump the scientific goals of ensuring that such lists are scientifically defensible. Just as the internal validity of a study is a prerequisite to considerations of external validity, questions of how best to disseminate an effective psychotherapy to clinicians and the public are only meaningful after the psychotherapy has in fact been demonstrated to

be effective relative to appropriate control conditions. The sociopolitical purposes of developing lists of ESTs include raising awareness of the value of psychotherapy among various stakeholders and promoting its use among practitioners. These goals, however laudable, can only be sustained if the criteria used to define ESTs are scientifically sound. In the short term, efforts to promote psychotherapy may appear to be best served by lists composed of a relatively large number of treatments. Hence, there may be pressure to establish criteria that define ESTs more liberally than may be scientifically warranted. To be useful over the long term, however, the emphasis must be on the scientific integrity of the procedures used to identify ESTs, regardless of how many treatments meet these criteria today. What follows is a discussion of the most important scientific concerns with the current procedures established by the CSP for identifying ESTs.

## THE CRITERIA

The most fundamental problem with the CSP's approach to defining ESTs centers on the criteria established to define ESTs. According to the current criteria, to be listed as "probably efficacious," a treatment need only demonstrate that it is statistically superior to no treatment (e.g., a wait-list control condition). Several authors have raised serious objections to this criterion (Herbert, 2000; Klein, 2002; Lohr, DeMaio, & McGlynn, this issue). We now know that virtually any intervention is superior to no treatment, especially for mood and anxiety disorders, the most common problems for which clients seek psychotherapy. Using no treatment as a baseline for establishing ESTs therefore imparts little new information about the effects of a psychotherapy. The absurdity of the no-treatment standard is illustrated by the fact that both prayer (Benor, 1990; Brown, 1995; Hodges & Scofield, 1995) and even placebo qualify as ESTs under this standard because both have been shown to be superior to no treatment.

Chambless and Hollon (1998) defend the no-treatment baseline on the grounds that it addresses the fundamental question that treatment-seeking clients ask: "Is this treatment likely to benefit me?" This solution is problematic for two reasons. First, implicit in this question is the presumption that treatment effects involve something more than

simply talking with another party. Like medical professionals, the professional psychotherapist is assumed to possess certain specialized skills relevant to the problem at hand. Otherwise, why would clients pay to see a professional rather than talking to a family member, friend, bartender, or hairdresser? Second, the most common alternative class of interventions to psychotherapy—psychotropic medications—are held to a higher standard (i.e., superiority to pill placebo), and it is considered unethical to deceive clients about whether they are taking an active medication. The no-treatment standard, in effect, suggests that such deception is somehow permitted with respect to psychotherapy, a clear double standard.

### COUNTING ONLY HITS

The no-treatment baseline is not the only problem with the current criteria for identifying ESTs. The "box score" method of summarizing the literature adopted by the CSP counts only "hits" (i.e., the number of studies yielding positive results), ignoring "misses." Thus, in principle, a treatment could be listed as empirically supported based on 2 supportive studies, even if 50 other studies failed to find effects. This failure to consider explicitly the entire literature is compounded by the file-drawer problem, such as the bias against publication of null results (Rosenthal, 1979). Conversely, consider a treatment that resulted in consistent but statistically insignificant trends relative to a strong control condition (e.g., a pill or psychological placebo) across several individual studies. The current criteria make no provision for meta-analytic strategies that might detect such findings.

### METHODOLOGICAL STANDARDS

The CSP's criteria are not explicit with respect to the minimal standards that a study must satisfy to ensure methodological rigor. Although the CSP discusses the importance of methodological features in general terms, they do not explicate minimal standards. Psychotherapy research varies widely in methodological rigor. The absence of explicit standards raises the possibility that seriously flawed studies could be counted as evidence in favor of a treatment,

and there is in fact reason to believe that this has already occurred (see Herbert, 2000, for a discussion of the use of a seriously flawed study to support the inclusion of a controversial treatment on the list of ESTs).

## DISTINCTIONS BETWEEN TREATMENTS

One of the most vexing problems with the current EST criteria is the problem of distinguishing apparently similar treatments. If the boundaries around different treatments are established solely on the basis of superficial appearance, then a potentially infinite number of variations of established treatments could be added to the list. Rosen and Davison (this issue) use "purple hat therapy" to illustrate the point. Imagine an intervention consisting of in vivo exposure for a phobia (an established procedure supported by a large research literature), but in which the therapist is instructed to wear a large purple hat with magnets arranged in a specific manner in the headband. The therapist might develop a theory about how the hat produces its effects. Moreover, imagine that the treatment is shown to be effective in a couple of studies, and the effects are attributed by proponents to the highly vivid hat, rather than systematic exposure to the feared stimulus. There is nothing in the current EST criteria that preclude purple hat therapy from taking its rightful place on the list.

If this example seems unrealistic, consider the case of eye movement desensitization and reprocessing (EMDR) (Shapiro, 1995). The various issues surrounding this highly controversial treatment are beyond the scope of this article; the interested reader is referred to reviews by Davidson and Parker (2001), Herbert et al. (2000), and Lohr, Tolin and Lilienfeld (1998). Suffice it to say that EMDR is essentially imaginal exposure (a well-established behavioral procedure) with the added feature of therapist-guided bilateral eye movements (or similar bilateral stimulation like alternating knee taps). Several studies have now established that the very feature that distinguishes EMDR from imaginal exposure (e.g., the eye movements) are superfluous (see Davidson & Parker, 2001, and Herbert et al., 2000). In other words, EMDR is no more effective than the EMDR procedure minus the eye movements (i.e., imaginal exposure).

Furthermore, there is no evidence that the mechanisms of change in EMDR are in any way unique. Nevertheless, the CSP includes EMDR on the list of ESTs. Other forms of exposure procedures (e.g., systematic desensitization, flooding), however, are not listed as distinct treatments but are instead listed under the general rubric of "exposure." No rationale is offered for this glaring inconsistency.

O'Donohue and Yeater (this issue) argue that for treatments to be considered distinct, they must differ either in the causal mechanisms by which they produce their effects and/or the procedures through which those mechanisms are instantiated. These authors recognize the difficulties inherent in making such judgments. Nevertheless, adoption of this principle would provide at least some rational basis for distinguishing treatments, in contrast to the current CSP procedures. I would only add that the burden of proof should rest clearly on the proponent of novel therapies to demonstrate the distinctiveness of new procedures relative to established ones and that the uniqueness of novel procedures should not be assumed (the principle of parsimony).

## OVERRELIANCE ON THE DSM FOR DEFINING THE TARGETS OF INTERVENTION

Several authors have decried the influence of the psychiatric *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-IV*) (American Psychiatric Association, 1994) on the process of identifying ESTs (Follette, 1996). The *DSM* is largely rooted in a medical model of psychopathology. It has been criticized on many grounds, such as for failing to capture natural boundaries between different forms of psychopathology, for discouraging idiographic assessment of the individual, and for failing to address sufficiently contextual factors.

In all fairness, however, the CSP has not required that treatments be linked to disorders as defined by the *DSM* to qualify as empirically supported. The criteria only require that the target of treatment be reliably identified. The reason that so many of the treatments on the current list of ESTs are defined by *DSM* categories is not because the CSP required it but rather because this is how most psychotherapy researchers have chosen to define the intervention targets of their studies, for better or for worse. There is no reason that ESTs must be iden-

tified in terms of *DSM* categories. Nevertheless, a reliable and comprehensive alternative to the *DSM* has yet to emerge, and until such an alternative becomes widely accepted, the *DSM* will continue to dominate as a system for defining psychopathology.

#### REMOVING TREATMENTS FROM THE LIST

The current criteria make no provisions for removing treatments from the list. The need to remove treatments might occur, for example, when alternative procedures are found to be significantly more effective or efficient or, as in the case of purple hat therapy or EMDR, when the unique features of the treatment are found to be superfluous.

#### HARMFUL TREATMENTS

Although most psychological interventions are at least minimally helpful, there is growing evidence that some procedures are actually harmful. Certain forms of psychological debriefing for trauma, for example, have been shown in several studies to delay the natural course of recovery following a traumatic event (for reviews, see Gist & Woodall, 1999; Gist, Woodall, & Magenheimer, 1999). Facilitated communication for autism resulted not only in false hopes for families of autistic individuals but also in false accusations and even civil and criminal charges of physical and sexual abuse against parents of autistic persons (Herbert, Sharp, & Gaudiano, 2002). Relapse prevention for sexual offenders may actually lead to higher rates of recidivism (McConaghy, 1999). Similarly, certain adolescent peer-group interventions may increase delinquent behavior (Dishion, McCord, & Poulin, 1999). Rebirthing therapy led to the death of a 10-year-old girl in Colorado in 2000 (Radford, 2001). Consistent with the principle of *primum non nocere* (first do no harm), such treatments have led some scholars to suggest that a list of treatments empirically demonstrated to be harmful would be an even more useful endeavor than a list of ESTs (Follette & Beitz, this issue).

#### TREATMENTS, PROCEDURES, OR MECHANISMS OF CHANGE?

Rosen and Davison (this issue) make a strong case for identifying empirically supported principles of change rather than trademarked

therapies. They echo some of the concerns of Lohr et al. (this issue), who advocate for experimental conditions (e.g., component analysis studies) aimed at evaluating theoretically derived mechanisms of change. A focus on trademarked therapies presents many problems, and shifting attention to principles of change addresses some of these. The link between theory and technology is emphasized, which is hoped to be the most effective approach to developing better interventions (Kazdin, 2001). The political and financial pressures to have treatments listed may be attenuated. Ultimately, it is difficult to imagine anyone arguing against the desirability of elucidating causal mechanisms underlying effective treatments.

Shifting the focus from treatments to principles of change does not, however, solve several key problems and in fact raises problems of its own. First, the issue of how to distinguish two putatively distinct treatments (or procedures) is not resolved. Second, it is not clear whether principles of change refer to causal mechanisms or to effective treatment procedures (regardless of the mechanisms by which they produce their effects). This is an important distinction that is often overlooked in discussions of this issue (see, for example, Haynes, Kaholokula, & Nelson, 1999). The problem with causal mechanisms is that they are largely unknown even in cases in which treatment effects have been reliably demonstrated. The causal mechanisms of even well-established treatment procedures (e.g., systematic exposure for phobias) remain controversial (e.g., the learning process of habituation vs. the modification of cognitive fear structures). Nevertheless, one may isolate the exposure procedure as the effective component of a larger treatment protocol without necessarily understanding the specific causal mechanisms responsible for the effect. In other words, it is quite possible to demonstrate that a treatment procedure works without knowing why it works. Indeed, the history of medicine is replete with examples of treatments that were shown to be effective relative to appropriate control conditions long before the causal mechanisms were understood (e.g., the analgesic effects of aspirin).

A related problem that a focus on empirically supported principles of change does not resolve concerns multicomponent treatment packages. What are we to make of programs that have been shown to be efficacious in well-controlled studies, but for which the specific com-

ponents responsible for the package's effects (much less the precise causal mechanisms of those active components) remain unknown or controversial (e.g., cognitive behavior therapy for depression) (see Jacobson et al., 1996). If we wait until causal debates are resolved before listing a treatment or procedure as empirically supported, we may be waiting quite a long time, during which time psychotherapy stakeholders are deprived of the potential benefits that lists of ESTs and corresponding practice guidelines might provide.

## RECOMMENDATIONS

A list of intervention procedures supported by controlled empirical research can play a vital role in bridging the scientist-practitioner gap but only if the list is based on sound science. The work of the CSP has been an important first step in the effort to identify ESTs. The project has been especially valuable in bringing into focus many of the difficult issues inherent in distinguishing treatment procedures that are supported by scientific data from those that are not. To achieve the goals for which they were designed, the procedures currently employed by the CSP for identifying ESTs will require significant modification. The following recommendations are not meant to be exhaustive but are instead offered as possible directions for resolving some of the problems with the current procedures for identifying ESTs.

1. First and foremost, the no-treatment baseline provides an insufficient comparison group for making meaningful statements about the specific effects of an intervention program and should therefore be dropped. Treatments should instead be evaluated in reference to a pill or psychological placebo. The concept of psychological placebo is itself fraught with controversy. Nevertheless, at a minimum, comparison conditions that control for widely recognized "nonspecific" factors common to all forms of psychotherapy should be required. An understanding of mechanisms of change and the link between theory and technology are both highly desirable goals and should be strongly encouraged in psychotherapy research programs. However, neither should be required for a treatment procedure to be listed as empirically supported.

2. Clear guidelines need to be put in place for distinguishing treatment procedures, with the principle of parsimony as a guiding principle. To be considered a distinct treatment, a procedure must differ from existing procedures in at least one of three ways: (a) effectiveness (broadly defined, to include issues of average effect size, clinical significance, the number of individuals who achieve responder status, etc.); (b) underlying causal mechanisms; or (c) substantial differences in treatment procedures. Regarding the third criterion, a putatively new treatment must differ from established procedures in significant, nontrivial procedural details to be considered distinct. It is not enough to add some twist—no matter how apparently distinctive or vivid—to an already established treatment, unless the incremental effects of that addition are empirically demonstrated.
3. Specific procedures for consideration of the full range of research literature should be incorporated into criteria for identifying ESTs. At least as much effort must be made to identify studies that failed to support an intervention (e.g., through examination of unpublished dissertations) as studies with supportive findings. The proportion of supportive to unsupportive studies must be considered rather than simply a count of the number of supportive studies. Meta-analytic procedures should be employed when appropriate.
4. Clear methodological standards for studies to be counted as probative must be articulated. Although no single study is flawless, in some cases a number of methodological problems render a study's results uninterpretable, and in other cases a single serious flaw has the same effect. Both "gold standard" (e.g., Foa & Meadows, 1997) criteria and the minimal standards for considering a study probative should be clearly explicated.
5. Procedures for removing treatments from the list must be established, including an ongoing review process aimed at critically reviewing each entry's continued presence on the list. This is important in several cases. First, a treatment might initially appear effective, but subsequent research reveals that it is ineffective or even harmful. Second, a multicomponent treatment package might be initially listed as effective, but subsequent component analyses reveal that certain components of the package are superfluous. It makes no sense to continue listing the entire package as effective; instead, ineffective components should be dropped from the list. The list should be viewed as a continuously evolving document rather than one that only grows ever larger.
6. A list of harmful treatments should be established. There is a growing empirical literature demonstrating that certain treatments produce iatragenic effects under certain conditions. The specific population and conditions in which the treatment has been shown to be harmful should be specified. In fact, it is conceivable that a treatment could be listed

simultaneously on a list of ESTs and a list of harmful procedures if it is shown to be helpful under certain conditions and harmful under others.

7. The link between a list of ESTs and practice guidelines should be made explicit, with the list representing only the first step in the development of such guidelines. Any list of ESTs, regardless of the specific criteria employed, will necessarily ignore many of the relevant factors that a clinician must consider in deciding on an intervention strategy. For example, such lists do not address the relative effectiveness of alternative procedures for a given target problem and ignore many of the contextual factors in which that target is embedded. Likewise, they do not address the magnitude of treatment effects (beyond statistical significance), the likelihood that an individual client will receive at least some benefit, the availability of alternative treatments, or the cost effectiveness of delivering treatments. It is difficult to see how a list of ESTs could possibly address each of these issues without becoming overly cumbersome. This is where treatment guidelines come into play. Once a list of ESTs is identified for a particular problem, a panel of experts can develop best practice guidelines that address these issues. Obviously, more effective treatments (as assessed either through average effect sizes and/or average number of treated patients who benefit from treatment) and more cost-effective treatments would be given priority over less effective, more costly treatments. It is imperative that practice guidelines stick as closely as possible to empirical findings. A frequent criticism of prior attempts to develop practice guidelines is "clinical judgment" overriding empirical findings (Persons, Thase, & Crits-Christoph, 1996).
8. Finally, steps must be taken to identify and address the various factors motivating the resistance of many practicing clinicians to ESTs. These factors range from the pragmatic (e.g., the difficulties for the established professional to obtain training and supervision in newer ESTs) to the economic (e.g., the association of the EST movement with the much-maligned managed care industry in the United States) to the philosophical (e.g., the belief that psychotherapy outcome is inherently not amenable to empirical evaluation). Addis et al. (1999) have offered several useful suggestions in this regard (see also commentaries by Abrahamson, 1999; Goldfried, 1999; Norcross, 1999). Mueser, Torrey, Lynde, Singer, and Drake (this issue) describe an interesting community-based collaborative program to increase the use of ESTs for persons with chronic mental illnesses and highlight several steps to increase the use of such interventions among front-line clinicians. In an effort to be sensitive to the concerns of practitioners, however, it is imperative that the scientific integrity of the process not be sacrificed in the process. Moreover, it must be kept in mind that although practition-

ers are an important constituency with respect to psychotherapy, they are not the only stakeholder group.

## CONCLUSION

There can be little doubt that the work of the CSP (and its predecessor task forces) has had a substantial impact on the field. Perhaps most significant is the healthy debate it has generated about the many complex issues involved in evaluating the empirical support for psychological interventions. The CSP's deliberations have served to improve the quality of psychotherapy research by highlighting important methodological issues. To its credit, the CSP has operated from the beginning in a spirit of relative transparency and openness, thereby encouraging healthy skepticism and debate. Despite criticisms from both practicing clinicians and academic scientists, the CSP has continued its work. The committee has temporarily suspended the review process to reevaluate the procedures used for identifying ESTs (Chambless & Ollendick, 2001). Issues currently under consideration include the importance of standardization in literature review procedures, assessing the interrater reliability of reviewer decisions, the merits of multiple lists versus a single list, the possibility of ranking ESTs according to relative efficacy, and the development of vehicles for dissemination of ESTs directly to the public (Weisz et al., 2000).

A revolution of sorts appears to be underway within the field of mental health. A growing scientific literature suggests that certain psychotherapies are effective for certain problems, and that those effects go beyond those attributable to so-called nonspecific factors. Nevertheless, these developments have had limited impact in frontline clinical settings, and the gap between science and practice continues to grow. The effort to identify ESTs and to promote their use represents an important step in addressing this gap. The current procedures, however, are seriously flawed, and unless corrected, the effort will likely fail. In addition, to be maximally beneficial, a list of ESTs, rather than being a goal in and of itself, must be developed in the context of an effort to develop best practice guidelines.

## NOTE

1. Many of the issues involved in the development of practice guidelines and their acceptance by clinicians parallel the controversy surrounding the effort to list ESTs. In addition, practice guidelines present some unique problems as well; see Abrahamson (1999); Hayes, V. M. Follette, Dawes, and Grady (1995); and Nathan (1998) for interesting discussions.

## REFERENCES

Abrahamson, D. J. (1999). Outcomes, guidelines, and manuals: On leading horses to water. *Clinical Psychology: Science and Practice*, *6*, 467-471.

Addis, M. E., Wade, W. A., & Hatgis, C. (1999). Barriers to dissemination of evidence-based practices: Addressing practitioners' concerns about manual-based psychotherapies. *Clinical Psychology: Science and Practice*, *6*, 403-407.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Barlow, D. H., Levitt, J. T., & Bufka, L. F. (1999). The dissemination of empirically supported treatments: A view to the future. *Behaviour Research and Therapy*, *37*, 147-162.

Benor, D. K. (1990). Survey of spiritual healing research. *Complementary Medicine Research*, *4*, 9-33.

Bohart, A. C., O'Hara, M., & Leitner, L. M. (1998). Empirically violated treatments: Disenfranchisement of humanistic and other psychotherapies. *Psychotherapy Research*, *8*, 141-157.

Brown, C. K. (1995). Spiritual healing in a general practice: Using a quality-of-life questionnaire to measure outcome. *Complementary Therapies in Medicine*, *3*, 230-233.

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. *The Clinical Psychologist*, *51*, 3-15.

Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, *66*, 7-18.

Chambless, D. L., & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology*, *52*, 685-716.

Chambless, D. L., Sanderson, W. C., Shoham, V., Johnson, S. B., Pope, K., Crits-Christoph, P., et al. (1996). An update on empirically validated therapies. *The Clinical Psychologist*, *49*, 5-18.

Craighead, W. E., & Craighead, L. W. (1998). Manual-based treatments: Suggestions for improving their clinical utility and acceptability. *Clinical Psychology: Science and Practice*, *5*, 403-407.

Davidson, P. R., & Parker, K. C. H. (2001). Eye movement desensitization and reprocessing (EMDR): A meta-analysis. *Journal of Consulting and Clinical Psychology*, *69*, 305-316.

Dawes, R. M. (1994). *House of cards*. New York: Free Press.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, *243*, 1668-1674.

Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and problem behavior. *American Psychologist*, *54*, 755-764.

Fensterheim, H., & Raw, S. D. (1996). Psychotherapy research is not psychotherapy practice. *Clinical Psychology: Science and Practice*, *3*, 168-171.

Foa, E. B., & Meadows, E. A. (1997). Psychosocial treatments for posttraumatic stress disorder: A critical review. *Annual Review of Psychology*, *48*, 449-480.

Follette, W. C. (1996). Introduction to the special section on the development of theoretically coherent alternatives to the *DSM* system. *Journal of Consulting and Clinical Psychology*, *64*, 1117-1119.

Gambrill, E. (1999). Evidence-based clinical behavior analysis, evidence-based medicine, and the Cochrane collaboration. *Journal of Behavior Therapy and Experimental Psychiatry*, *30*, 1-14.

Garfield, S. L. (1996). Some problems associated with "validated" forms of psychotherapy. *Clinical Psychology: Science and Practice*, *3*, 218-229.

Gist, R., & Woodall, S. J. (1999). There are no simple solutions to complex problems: The rise and fall of critical incident stress debriefing as a response to occupational stress in the fire service. In R. Gist & B. Lubin (Eds.), *Response to disaster: Psychosocial, community, and ecological approaches* (pp. 211-235). Philadelphia: Brunner/Mazel.

Gist, R., Woodall, S. J., & Magenheimer, L. K. (1999). And then you do the hokey-pokey and you turn yourself about. . . . In R. Gist & B. Lubin (Eds.), *Response to disaster: Psychosocial, community, and ecological approaches* (pp. 269-290). Philadelphia: Brunner/Mazel.

Goisman, R. M., Rogers, M. P., Stekettee, G. S., Warshaw, M. G., Cuneo, P., & Keller, M. B. (1993). Utilization of behavioral methods in a multicenter anxiety disorders study. *Journal of Clinical Psychiatry*, *54*, 213-218.

Goisman, R. M., Warshaw, M. G., & Keller, M. B. (1999). Psychosocial treatment prescriptions for generalized anxiety disorder, panic disorder, and social phobia, 1991-1996. *American Journal of Psychiatry*, *156*, 1819-1821.

Goldfried, M. R. (1999). The pursuit of consensus in psychotherapy research and practice. *Clinical Psychology: Science and Practice*, *6*, 462-466.

Hayes, S. C. (1996). Creating the empirical clinician. *Clinical Psychology: Science and Practice*, *3*, 179-181.

Hayes, S. C., Follette, V. M., Dawes, R. B., & Grady, K. E. (1995). *Scientific standards of psychological practice: Issues and recommendations*. Reno, NV: Context Press.

Haynes, S. N., Kaholokula, J. K., & Nelson, K. (1999). The idiographic application of nomothetic, empirically based treatments. *Clinical Psychology: Science and Practice*, *5*, 456-461.

Herbert, J. D. (2000). Defining empirically supported treatments: Pitfalls and possible solutions. *The Behavior Therapist*, *23*, 113-134.

Herbert, J. D., Lilienfeld, S. O., Lohr, J. M., Montgomery, R. W., O'Donohue, W. T., Rosen, G. M., et al. (2000). Science and pseudoscience in the development of eye movement desensitization and reprocessing: Implications for clinical psychology. *Clinical Psychology Review*, *20*, 945-971.

Herbert, J. D., Sharp, I. R., & Gaudiano, B. A. (2002). Separating fact from fiction in the etiology and treatment of autism: A scientific review of the evidence. *The Scientific Review of Mental Health Practice*, *1*, 25-45.

Hodges, R. D., & Scofield, A. M. (1995). Is spiritual healing a valid and effective therapy? *Journal of the Royal Society of Medicine*, *88*, 203-207.

Jacobson, N. S., Dobson, K. S., Truax, P. A., Addis, M. E., Koerner, K., Gollan, J. K., et al. (1996). A component analysis of cognitive-behavioral treatment for depression. *Journal of Consulting and Clinical Psychology*, *64*, 295-304.

Kahnemann, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.

Kazdin, A. E. (2001). Progression of therapy research and clinical application of treatment require better understanding of the change process. *Clinical Psychology: Science and Practice*, *8*, 143-151.

Klein, D. F. (2002). Dodo deliberations. *Clinical Psychology: Science and Practice*, *9*, 28-29.

Lambert, M. J. (1998). Manual-based treatment and clinical practice: Hangman of life or promising development? *Clinical Psychology: Science and Practice*, *5*, 391-395.

Lohr, J. M., Tolin, D. F., & Lilienfeld, S. O. (1998). Efficacy of eye movement desensitization and reprocessing: Implications for behavior therapy. *Behavior Therapy*, *29*, 123-156.

McConaghy (1999). Methodological issues concerning evaluation of treatment for sexual offenders: Randomization, treatment dropouts, untreated controls, and within-treatment studies. *Sexual Abuse: A Journal of Research and Treatment*, *11*.

Meehl, P. E. (1986). Causes and effects of my disturbing little book. *Journal of Personality Assessment*, *50*, 370-375.

Mumma, G. H. (2001). Increasing accuracy in clinical decision making: Toward an integration of nomothetic-aggregate and intraindividual-idiographic approaches. *The Behavior Therapist*, *24*, 77-94.

Nathan, P. E. (1998). Practice guidelines: Not yet ideal. *American Psychologist*, *53*, 290-299.

Nathan, P. E. (2000). The Boulder model: A dream deferred—or lost? *American Psychologist*, *55*, 250-251.

Nezu, A. M. (1996). What are we doing to our patients and should we care if anyone else knows? *Clinical Psychology: Science and Practice*, *3*, 160-163.

Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice Hall.

Norcross, J. D. (1999). Collegially validated limitations of empirically validated treatments. *Clinical Psychology: Science and Practice*, *6*, 472-476.

Parloff, M. B. (1998). Is psychotherapy more than manual labor? *Clinical Psychology: Science and Practice*, *5*, 376-381.

Persons, J. B., Thase, M. E., & Crits-Christoph, P. (1996). The role of psychotherapy in the treatment of depression. *Archives of General Psychiatry*, *53*, 283-290.

Radford, B. (2001). Rebirthing update: Therapists convicted, therapy outlawed in Colorado. *Skeptical Inquirer, 25*(4), 7.

Raw, S. D. (1993). Does psychotherapy research teach us anything about psychotherapy? *The Behavior Therapist*, *16*, 75-76.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin*, *86*, 638-641.

Ruscio, J. (2002, March/April). The emptiness of holism. *Skeptical Inquirer*, *26*, 46-50.

Sackett, D. L., Richardson, W. S., Rosenberg, W., & Haynes, R. B. (1997). *Evidence-based medicine*. New York: Churchill Livingstone.

Safran, J. D. (2001). When worlds collide: Psychoanalysis and the empirically supported treatment movement. *Psychoanalytic Dialogues*, *11*, 659-681.

Sanderson, W. C. (2002). Are evidenced-based psychological interventions practiced by clinicians in the field? *Medscape Mental Health*, *7*, 1-3.

Sechrest, L. (1992). The past and future of clinical psychology: A reflection on Woodward. *Journal of Consulting and Clinical Psychology*, *60*, 18-23.

Shapiro, F. (1995). *Eye movement desensitization and reprocessing: Basic principles, protocols, and procedures*. New York: Guilford.

Silverman, W. H. (1996). Cookbooks, manuals, and paint-by-numbers: Psychotherapy in the '90s. *Psychotherapy*, *33*, 207-215.

Task Force on Promotion and Dissemination of Psychological Procedures. (1995). Training in and dissemination of empirically validated treatments: Report and recommendations. *The Clinical Psychologist*, *48*, 3-23.

Taylor, C. B., King, R., & Margraf, J. (1989). Use of medication and in vivo exposure in volunteers for panic disorder research. *American Journal of Psychiatry*, *146*, 1423-1426.

Weisz, J. R., Hawley, K. M., Pilkonis, P. A., Woody, S. R., & Follette, W. C. (2000). Stressing the (other) three Rs in the search for empirically supported treatments: Review procedures, research quality, relevance to practice and the public interest. *Clinical Psychology: Research and Practice*, *7*, 243-258.

*James D. Herbert is currently associate professor of psychology at Drexel University in Philadelphia. He has a bachelor's degree from the University of Texas at Austin and master's and doctoral degrees in clinical psychology from the University of North Carolina at Greensboro. His research interests include assessment and treatment of anxiety disorders, psychotherapy outcome, and the role of science and pseudoscience in clinical psychology.*