

Chapter 21

Empirically Supported Treatments, Evidence-Based Treatments, and Evidence-Based Practice

Barry L. Duncan and Robert J. Reese

The great tragedy of science—the slaying of a beautiful hypothesis by an ugly fact.

Thomas Henry Huxley

Accountability via the application of research to practice is the *raison d'être* of the empirically supported treatment (EST), evidenced-based treatment (EBT), and evidence-based practice (EBP) movements, and truly the zeitgeist of our time. The idea that clinical practice can be informed by empirical research, however, is not new and has been integral to psychology since the late 19th century, marked by Lightner Witmer's first psychology clinic in 1896 (see McReynolds, 1997). The Boulder Conference in 1949 formalized clinical psychology's commitment to an empirical base with the scientist-practitioner paradigm of training and practice. Since that time, *EST*, *EBT*, and *EBP* have all become commonplace acronyms within clinical psychology and across the mental health and substance abuse fields. Although basing practice on empirical findings and using treatments with demonstrated efficacy seems the only reasonable course of action, such a straightforward idea becomes increasingly complex when unfurled in the various social, political, economic, and other ideological contexts that influence the delivery of mental health services (Norcross, Beutler, & Levant, 2006). In truth, what constitutes evidence and how it should influence practice is, perhaps, the fiercest debate of our times.

This chapter examines ESTs, EBTs, and EBPs and describes two fundamentally different approaches to defining and disseminating evidence (Littell, 2010)—one that seeks to improve clinical practice via the dissemination of treatments meeting a minimum standard of empirical support (EBT) and another that describes a process of research application to practice that includes clinical judgment and client preferences (EBP). We unfold the differences

between the approaches by addressing the nature of evidence itself, how it is transported to real-world settings, and ultimately, whether such evidence improves client outcomes. To further inform the controversy, this chapter also discusses the advantages and disadvantages of the randomized clinical trial (RCT), its specificity assumption, and the connection of the RCT to a medical-model way of understanding psychotherapy. Finally, we strike at the heart of the EBT-versus-EBP debate by tackling the thorny question of whether evidence-based treatments should be mandated.

From EBP to ESTs and EBTs, and Back Again: The Evolution of Evidenced-Based Practice

There is no new thing under the sun.

Ecclesiastes 1:9

Evidence-based practice in psychology evolved from evidence-based medicine (EBM). Leff (2002) posited three important events that shaped the evolution of EBM. First, in 1910, Abraham Flexner wrote of the conditions in medical schools that led to sweeping reforms in physician training with an increased emphasis on a curriculum undergirded by science. Second was the publication of the first RCT in 1948 in the *British Medical Journal*. The third major influence was the creation of the Food and Drug Administration (FDA), and, in the early 1960s, the establishment of the double-blinded RCT as the gold standard for demonstrating efficacy and safety.

Inspired by Archi Cochrane and David Sackett (Claridge & Fabian, 2005), EBM took shape in the early 1990s. Cochrane, a British epidemiologist, wrote *Effectiveness and Efficiency: Random Reflections on Health Services* (1972), illuminating the lack of routine empirical application to medical practice. He recommended a reliance on the RCT and called for a compilation of research by discipline to guide medical treatment. His vision resulted in the formation of the Cochrane Collaboration in 1993, which reviews, critiques, and synthesizes RCTs in medicine and mental health for the purpose of dissemination to the public (www.cochrane.org).

Sackett and colleagues are typically credited with defining EBM as the integration of the best research evidence with clinical expertise, including patient values, to make informed decisions about individual cases (Sackett, Rosenberg, Muir-Gray, Haynes, & Richardson, 1996). Importantly, EBM was viewed as a

process driven by clinicians asking specific questions having practical value for the patient at hand (Littell, 2010). This process perspective of EBM was adopted in definitions by the Institute of Medicine (2001), in models of evidence-based practice for human services (Gibbs, 2003), public policy (Davies, 2004), and importantly, by the American Psychological Association (APA) Presidential Task Force on EBP (hereafter APA Task Force, 2006).

Empirically Supported Treatments and Evidenced-Based Treatments

Simultaneous with Sackett's influence in medicine, a completely different approach to the application of evidence to practice occurred in psychology. It started with the American Psychiatric Association's development of practice guidelines. Beginning in 1993, psychiatrists produced guidelines for disorders ranging from major depression to nicotine dependence. Psychiatry's imprimatur gave an aura of scientific legitimacy to what was primarily an agreement among psychiatrists about their preferred practices, with an emphasis on biological treatment.

Concurrently, psychologists rushed to offer magic bullets to counter psychiatry's magic pills—to establish ESTs (originally empirically validated treatments, or EVT's). Perhaps fearing psychiatry's historical hegemony in health care, ESTs were promoted as the rallying point, a "common cause" for a clinical profession fighting exclusion (Nathan, 1997, p. 10). Arguing that clients have a right to "proven," not consensus, treatments, a special task force (Task Force on Promotion and Dissemination of Psychological Procedures, 1995) acting under the auspices of APA Division 12 (Society of Clinical Psychology) derided psychiatry's approved treatment list as medically biased and unrepresentative of the clinical literature and set forth its conclusions about what constituted scientifically valid treatments. Instead of clinical consensus and comprehensive guidelines, the Task Force concentrated its efforts on research demonstrations that a particular treatment has proven to be beneficial for clients in RCTs. The Task Force reviewed available research and catalogued treatments of choice for specific diagnoses based on their efficacy criteria, similar to the standards of the FDA.

Outlining criteria for "empirically validated treatments," the Task Force identified 18 treatments that were "well-established." The terms *empirically supported treatment* or *evidence-based treatment* later replaced *empirically validated treatments* due to the recognition that completely establishing validity in the social sciences is difficult, at best (Ollendick & King, 2004). The criteria for a *well-established* treatment could be met in two ways: First was to have at least two experimental, between-group studies that demonstrate statistically significant gains when compared to another treatment, pill or psychological placebo, or show

equivalence with another established treatment. Second was for a treatment to have had more than nine single-subject design studies considered to be of high quality, and included a comparison of the treatment to another treatment. For both, the studies had to use treatment manuals with a well-defined population, and include studies from two independent researchers (Chambless et al., 1998).

According to the Division 12 criteria (and others—see ahead), the RCT represents the highest form of research evidence. The RCT is unequivocally the best research design to isolate treatment effects and control for threats to internal validity; it has served psychology well, contributing to its reputation as a discipline rooted in rigorous science. RCT efficacy research has documented that psychological interventions are better than no treatment or nonspecific intervention (Lipsey & Wilson, 1993; Smith & Glass, 1977) and are at least as efficacious as medication for many psychological disorders, and more so in the long run (DeRubeis et al., 2005).

Concurrent with Division 12 efforts, criteria for evaluating the efficacy of therapeutic interventions have been developed by many professional and government organizations. Examples include the Blueprints for Violence Prevention series (Mihalic, Fagan, Irwin, Ballard, & Elliott, 2004) and the National Registry of Evidence-based Programs and Practices (NREPP) of the Substance Abuse and Mental Health Services Administration (SAMHSA; 2007). These efforts have resulted in lists of effective or model programs. This approach to evidence, which seeks to identify interventions or programs that meet certain evidentiary criteria, will hereafter be called *EBT*, distinguishing it from EBP (Littell, 2010).

Evidenced-Based Practice in Psychology

The criteria and list of EBTs by Division 12 and others touched off a firestorm of controversy resulting in sometimes-contentious camps regarding what constitutes evidence and how such evidence should be gathered, disseminated, and implemented (Norcross et al., 2006). In the face of growing criticism of EBTs, 2005 APA President Ronald Levant appointed the Presidential Task Force on Evidence-Based Practice in Psychology. The APA Task Force defined *EBP* as “the integration of the best available research with clinical expertise in the context of patient [*sic*] characteristics, culture, and preferences” (APA Task Force 2006, p. 273).

The Best Available Research

Consisting of both researchers and practitioners, the APA Task Force defined “the best available research” as “results related to intervention strategies,

assessment, clinical problems, and patient populations in laboratory and field settings as well as to clinically relevant results of basic research in psychology and related fields” (2006, p. 274). In contrast to the efforts of Division 12 and others that delineate the RCT as the gold standard of research, the Task Force did not identify one research methodology to be superior, maintaining that different methodologies are required to answer different research questions, including effectiveness studies, process research, single-subject designs, case studies, and qualitative methodologies.

Effectiveness studies evaluate treatment outcomes in naturalistic settings, therefore offering improved generalizability. For example, clients diagnosed with more than one disorder¹ could participate, in contrast to their likely exclusion in RCTs. The downside is the loss of control of internal validity given the lack of random assignment and placebo conditions. An often-discussed example of an effectiveness study is the *Consumer Reports* (1995) survey of readers about their experience with psychotherapy. Survey respondents indicated that psychotherapy was helpful, that longer treatment duration was better than shorter, and the inclusion of medication with psychotherapy did not lead to better outcomes. This study, and effectiveness research in general, has stirred considerable debate (VandenBos, 1996) regarding whether effectiveness studies are a viable methodology given their inherent limitations. Some have advocated (Clarke, 1995) that rather than being an either-or issue, effectiveness studies can be used in unison with efficacy research. This is the stance of the APA Task Force.

Process research plays an integral role in addressing the *why* and *what* of an effective treatment. Understanding why a treatment works can facilitate a deeper understanding of the processes of therapy as conducted in actual clinical settings. For example, the finding that different treatment approaches yield similar results (see ahead) has led to research addressing the *common factors* across all treatments that are integral to outcome. Common factors research has made a substantial contribution to the psychotherapy literature regarding the key processes that promote change. A specific example is the *therapeutic alliance*, one of the most consistent predictors of psychotherapy outcome (e.g., Anker, Owen, Duncan, & Sparks, 2010; Horvath, 2001) regardless of the theoretical approach or orientation used by the clinician. Process research offers the advantage of being responsive to contextual factors, allowing for evaluation of the complex and nuanced exchanges that comprise psychotherapy.

Single-subject research is considered an acceptable treatment methodology for identifying a treatment as an EBT because it addresses threats to internal and

¹The use of the word “*disorder*” or reference to any specific diagnosis is done only as a matter of convenience to note the related research and in no way reflects any endorsement of the science or ethics of diagnosis.

external validity (Chambless et al., 1996). Relying on a small number of participants, or even one participant, behavior is measured over time using the presentation and withdrawal of an independent variable. Instead of inferential statistics, data are graphically presented. Examples of single-subject designs include: the simple baseline (AB) design, the reversal design (ABA or ABAB), the alternating treatments design (comparing multiple treatments), and the multiple baseline design (see Kazdin, 2010). Single-subject design offers the advantage of implementation in settings where an RCT is not feasible or with populations of interest that are smaller in number (e.g., persons diagnosed with trichotillomania). Single-subject designs are also useful for exploratory analysis to determine if large-scale group comparisons are warranted. Given its feasibility, it can also be easily integrated with other methodologies (Stiles, Hurst, et al., 2006). The biggest disadvantage is the small sample size and resulting limitation on generalization.

Case studies and qualitative research share the same limitation of generalizability as the single-subject design. However, both case studies and qualitative designs obtain information that adds richness and color to data from other methodologies. For example, aggregated data can yield a mean score on a construct. Case studies and qualitative research, however, provide information regarding what a score *means* to a certain individual. These methodologies evaluate and deepen understanding of a theory or an approach (Stiles, Hurst, et al., 2006). In addition to serving as a form of evidence in their own right, case studies and qualitative research identify potential variables or constructs that could be considered in larger scale studies. Finally, such methodologies can also provide further, more specific evidence for treatments that have been identified as efficacious in RCT studies. For example, a case study or qualitative design could evaluate the effectiveness of cognitive-behavioral treatment conducted in a real-world setting with a population of clients difficult to obtain (e.g., adolescents diagnosed with both an intellectual disability and major depression).

What qualifies as evidence differentiates the EBP from the EBT approach. With EBT, the emphasis is on the RCT while the EBP views evidence from a more broad-based perspective. There is an inevitable trade-off between internal and external validity—EBT falls on the side of internal at the expense of external validity. In addition, the EBT approach of Division 12 and others has focused on the treatment model that is administered with less emphasis on who is providing or receiving the treatment. More recently, researchers in the social sciences and education have acknowledged that *who* provides an intervention is an important, if not critical, variable to consider. Enter the next component of the definition of EBP: clinical expertise.

Clinical Expertise

APA's definition of evidence-based practice includes the clinician, or more precisely the role of "clinical expertise." Clinical expertise encompasses the assessment of clients and the provision of appropriate services. A therapist must ultimately use a decision-making process (i.e., clinical judgment) to determine if an intervention, based on the latest research, is likely to be effective for a particular client given his or her unique circumstance. This component of the definition acknowledges the inherent limitation of research findings—that the individual application of research is constrained by myriad client and environmental factors that could potentially influence the effectiveness of a type of treatment. Practitioners must use their clinical judgment and expertise to determine how to implement, and if necessary, modify a given approach for a particular client, in a particular circumstance, at a particular time.

The controversy here lies in the extent of the role that expertise and judgment should play in clinical decision making. Should clinical decision making be kept to a minimum (the EBT approach) or should research be necessarily contextualized by the therapist's expertise and experience (the EBP approach)? Psychologists are trained to be scientist-practitioners or practitioner-scholars. The foundational knowledge of research design and scientific inquiry prepares psychologists to make idiographic treatment decisions based on a nomothetic research literature. This encompasses the ability to critically evaluate the research literature, synthesize an area of research, and make general conclusions about the appropriateness of a set of potential bona fide treatment options for a particular person.

Although psychologists may possess specific evaluative skills, it is questionable whether they have the time and/or inclination to review the latest research. This is what EBTs bring to the table regarding interventions and programs: SAMSHA's National Registry, APA Division 12's ESTs, and the international Cochrane Collaboration summarize treatments through reviews and meta-analyses that are accessible and easy to digest for clinician consumption. Clinicians also rely on articles that summarize a group of studies focused on a specific presenting problem or treatment approach. Such reviews provide an attractive option given that data have already been culled, synthesized, and summarized.

Although reviews and summaries can make a practitioner's life easier, Littell (2010) warns that they are at risk for promoting misinformation and ultimately influencing policy in ways that the data do not suggest. Gambrill and Littell (2010), reprinted here in the sidebar, provides

Gambrill, E., & Littell, J. H. (2010). Do haphazard reviews provide sound directions for dissemination efforts? *American Psychologist*, 65, 927.

The lead article in the February-March issue by McHugh and Barlow (2010) emphasizes the need for “dissemination and implementation of evidence-based psychological treatments.” The authors identify a number of intervention programs as evidence-based and in need of dissemination. One is multisystemic therapy (MST). They claimed that this program is among “the most successful dissemination efforts . . . pursued by treatment developers” (p. 79). One randomized-controlled trial (RCT) was cited in support of the effectiveness of MST (Henggeler, Melton, Brondino, Scherer, & Hanley, 1997). The remaining citations were to nonexperimental or weak quasi-experimental studies and non-systematic reviews, including a 1998 review by Kazdin and Weiss. The systematic review of RCTs on the effects of MST by Littell, Popa, and Forsythe (2005) was not mentioned. This review, which is in the *Cochrane Library of Systematic Reviews* and the *Campbell Library of Systematic Reviews*, reported that MST is no more effective than other treatments. This review found a number of concerning lapses in methodology and data analysis in related RCTs, including failure in all but a few studies to conduct an intention-to-treat analysis.

An analysis of previously published reviews of MST trials showed that, like the McHugh and Barlow (2010) article, most published reviews provided information that was incomplete and potentially misleading (Littell, 2008).

McHugh and Barlow’s (2010) discussion of the implementation of MST in Hawaii is troubling, because it neglected to mention concerns about the perceived lack of cultural sensitivity of the MST program in that state. “Clearly, the use of MST in Hawaii has been controversial and resulted in reports that strongly questioned the appropriateness of using MST in the state” (Rosenblatt et al., 2001, p. 2). McHugh and Barlow did not mention the fact that a controlled trial of the MST-based Continuum of Care was stopped early in Hawaii in the wake of “bad press” (Rosenblatt et al., 2001). The “open trial” (p. 78) cited by McHugh and Barlow (2010) had no parallel comparison or control groups.

Also troubling are repeated claims that fidelity to MST predicts better outcomes. The MST Therapist Adherence Measure (TAM) and related instruments tap common factors, client satisfaction, and early outcomes (Littell, 2006). It is not surprising that such measures predict outcomes, but that does not make them valid measures of fidelity.

The Cochrane Collaboration and the Campbell Collaboration provide syntheses of evidence related to specific practice and policy questions. Cochrane and Campbell reviews are based on an exhaustive search for and rigorous appraisal of all research related to a question. Why would we base recommendations for dissemination on a haphazard review of research, such as the one provided by

McHugh and Barlow, when such haphazard reviews provide misleading information?

References

- Henggeler, S. W., Melton, G. B., Brondino, M. J., Scherer, D. G., & Hanley, J. H. (1997). Multisystemic therapy with violent and chronic juvenile offenders and their families: The role of treatment fidelity in successful dissemination. *Journal of Consulting and Clinical Psychology, 65*, 821–833.
- Kasdin, A. E., & Weisz, J. R. (1998). Identifying and developing empirically supported child and adolescent treatments. *Journal of Consulting and Clinical Psychology, 66*, 19–36.
- Littell, J. H. (2006). The case for multisystemic therapy: Evidence or orthodoxy? *Children and Youth Services Review, 28*, 458–472.
- Littell, J. H. (2008). Evidence-based or biased? The quality of published reviews of evidence-based practices. *Children and Youth Services Review, 30*, 1299–1317.
- Littell, J. H., Popa, M., & Forsythe, B. (2005). Multisystemic therapy for social, emotional, and behavioral problems in youth aged 10–17 (Cochrane Review). *Cochrane Database of Systematic Reviews, 4*. doi:10.1002/14651858.CD004797.pub4
- McHugh, R. K., & Barlow, D. H. (2010). The dissemination and implementation of evidence-based psychological treatments. *American Psychologist, 65*(2), 73–84. doi:10.1037/a0018121
- Rosenblatt, A., Deuel, L., Mak, W., Thornton, P., Baize, H., Morea, J., & Smucker, S. (2001). Evaluation of two therapeutic programs for children with serious mental health problems and their families: Home-based multisystemic therapy (MST) and the MST continuum of care. San Francisco: University of California, San Francisco, Child Services Research Group.

a recent example. Meta-analytic strategies were developed to help minimize these biases, and have generally been more successful in doing so, but such studies also fall prey to the same study selection strategies and potential for being influenced by publication bias (Reese, Prout, Zirkelback, & Anderson, 2010).

The clinical expertise component of EBP also distinguishes the two different approaches to evidence. EBP tends to appeal to those who value clinicians' autonomy and individualized treatment decisions while EBT tends to appeal to those who believe that more structure and consistency is needed to ensure positive outcomes (Littell, 2010). As noted, EBTs focus on the treatment, intervention, or program itself and not on who is delivering or receiving it. This is perhaps the biggest difference between EBT and EBP.

In the Context of Patient Characteristics, Culture, and Preferences

The last portion of the EBP definition includes the client's preferences, cultural context, and idiographic needs (Messer, 2006), suggesting that the clinical decision-making process is a collaborative one. Goodheart, Kazdin, and Sternberg (2006) articulate that a DSM-based diagnosis and a corresponding EBT, complete with treatment manual, cannot be implemented with a nomothetic understanding of the client. Importantly, diversity is part of the idiographic mix that requires consideration (see Sue et al., 2006). For example, EBT research of racial/ethnic minority, sexual minority, or economically disadvantaged populations is limited, and therefore it is unknown if the efficacy of EBTs extend to such groups. Two

recent meta-analyses demonstrate that attention to both culture and client preferences make empirical as well as clinical sense. Smith, Domenech Rodríguez, and Bernal (2011) examined 65 experimental and quasi-experimental studies involving 8,620 participants and found that treatments specifically adapted for clients of color were moderately more effective with that clientele than traditional treatments ($d = .46$). A meta-analysis of 35 studies found that clients who were matched to their preferred therapy conditions were less likely to drop out and showed greater improvements ($d = .31$; Swift, Callahan, & Vollmer, 2011).

From an EBP vantage point, the unique preferences, values, needs, strengths, weaknesses, and other characteristics unique to the client may not be served adequately by sole reliance on an EBT. For example, a client may meet criteria for major depressive disorder, but an assessment reveals an unfulfilling career is underlying the feelings of dissatisfaction. Given the career problem, a manualized treatment for depression may not be helpful or even appropriate. Career counseling may be warranted, rendering the DSM diagnosis (and accompanying intervention) secondary and not the primary focus that will yield treatment benefit.

Both EBT and EBP approaches strive to make the delivery of psychosocial services a scientific endeavor, although it is debatable (Norcross et al., 2006) how scientific such services can ultimately be. Regardless of the stance on the role of science and clinical judgment in providing efficacious treatments, the value system, cultural expectations, and preferences of the client cannot be removed—even in medicine. To illustrate, does a woman who needs cataract surgery select a monofocal intraocular lens that typically restores excellent long distance vision but will require reading glasses? Or does she select a newer, more expensive multifocal intraocular lens that will likely eliminate the need for corrective lenses but increase the potential difficulty of night driving? An ophthalmologist can readily go to the Cochrane Collaborative web site and pull up a summary of 10 RCTs that compared the two lenses. The review concluded that both lenses were excellent and provided trade-offs depending on the patient's lifestyle and preference. Client preference matters, even in medicine, for an issue where several RCTs exist. Did the research help inform the decision? Absolutely. Did the research dictate the ultimate treatment? No. A good practitioner would share the results of the research and help the patient prioritize what her preferences are in relation to the lens replacement options, as well as share his or her experiences with patients who have had both lenses.

Two fundamentally different perspectives to the dissemination and implementation of the empirical research have emerged, the EBT and the EBP approaches (Littell, 2010). EBP emphasizes a process through which clinicians can integrate empirical evidence with clinical expertise and client preferences, to make informed judgments in individual cases. EBT seeks to identify treatments

that are effective for specific conditions and insure the widespread availability of these treatments via lists of interventions and programs meeting specified criteria.

Can't We All Just Get Along?

To follow knowledge like a sinking star,
Beyond the upmost bound of human thought . . .
To strive, to seek, to find, and not to yield.

Tennyson, *Ulysses*

Without question, it makes good clinical sense to be “evidence based.” In truth, no one says, “Evidence, shmevidence! It means nothing to my work—I fly by the seat of my pants, meander willy nilly through sessions, and rely totally on the wisdom of the stars to show the way.” Saying you don’t believe in the almighty evidence, especially if you are a psychologist, is tantamount to not believing in Mom or apple pie, or whatever your sacrosanct cultural icons happen to be. So, what is the controversy about?

On one hand, the Division 12 Task Force effectively increased recognition of the efficacy of psychological interventions among the public, policymakers, and training programs; on the other hand, it simultaneously promulgated gross misinterpretations—that EBTs have proven superiority over other approaches, and therefore, should be mandated and exclusively reimbursed. Taking it a hyperbolic step further, some even have suggested that not administering EBTs is unethical (Chambless & Crits-Christoph, 2006), and perhaps even “prosecutable”! A *New York Times* article reported:

Using vague, unstandardized methods to assist troubled clients
“should be prosecutable” in some cases, said Dr. Marsha Linehan. .
. . (Carey, 2005, p. 2)

Unfortunately, because of such statements, many believe, to paraphrase Orwell, that some therapies are more equal than others.

For example, The President's New Freedom Commission on Mental Health (PNFC) called for incentives to implement EBTs (PNFC, 2005). The National Institutes of Health (NIH) and the Department of Health and Human Services funded state implementation of EBTs as well as research on their transportability. The Division 12 list not only has been referenced by local, state, and federal funding agencies, but also has been used to restrict reimbursement. For example, in 2003, the state of Oregon mandated use of EBTs in their mental health and addiction service systems; and in 2004, Iowa mandated that of the 70% of block grant funds to go to community mental health centers, all must be tied to EBTs (Littell, 2010). Other states are following suit. Given such funding and regulatory

mandates for EBTs, they are now inextricably woven into the fabric of mental health and substance abuse policy and practice.

Clinicians, however, are suspicious of the whole idea. Woody, Weisz, and McLean (2005) described barriers to training in EBTs:

Some of this opposition was based on the idea that lists of ESTs reflect a political or theoretical bias more than they reflect treatments that work. Others opposed what they see as an erosion of their autonomy as professionals due to pressure to conduct ESTs. In this view, the manualized approach is seen as too rigid and objectifying rather than humanizing clients. Some Training Directors also expressed a lack of trust in researchers, pointing to stories of misleading reporting of clinical trials from the drug industry in support of this view. (p. 11)

EBT has been described as the cause of “psychological warfare between therapists and scientists” (Tavris, 2003, p. xii), and is perhaps the current manifestation of a long history of mutual antagonism between practitioners and researchers. Clinicians often see researchers as pencil-headed geeks, detached from the real world, who contemplate the obvious and irrelevant, and write only for other researchers in incomprehensible gobbledygook. Researchers often regard clinicians as flying by the seat of their pants, fueled by soft-headed intuition and new-age fads—intellectually lazy mercenaries who ignore empirical findings. But the real issue is not the divide between researchers and clinicians or even between science and practice, but whether the call for accountability via mandating EBTs is empirically justified. Three research areas address this thorny question.

Psychotherapy and Specific (Unique) Ingredients

Seek facts and classify them and you will be the
workmen of science. Conceive or accept theories
and you will be their politicians.

Nicholas Maurice Arthus, *De l'Anaphylaxie a l'immunité*

First, the good news is that the efficacy of psychotherapy is very good—the average treated person is better off than about 80% of the untreated sample (Duncan, Miller, Wampold, & Hubble, 2010), translating to an effect size (ES) of

about 0.8.² Moreover, these substantial benefits apparently extend from the laboratory to everyday practice. For example, a real-world study in the UK (Stiles, Barkham, Twjgg, Mellor-Clark, & Cooper, 2006) comparing cognitive-behavioral therapy (CBT), psychodynamic therapy (PDT), and person-centered therapy (PCT) as routinely practiced reported a pre-post ES of around 1.30 across treatment approaches. In short, psychotherapy works.

But how it works gets to the controversy at hand. Consider the RCT. It was designed to compare the effects of a drug (an active compound) to a placebo (a therapeutically inert or inactive substance) for a specific illness. The basic assumption of the RCT is that the specific (unique) ingredients of different drugs (or psychotherapies) will produce different effects, superior over placebo, with different disorders. In effect, this assumption likens psychotherapy to a pill, with discernable unique ingredients that can be shown to have more potency than other active ingredients of other drugs. The specific- or unique-ingredients assumption or drug metaphor of the RCT, however, does not seem to fit psychotherapy (Stiles & Shapiro, 1989; Wampold et al., 1997).

Drug trials include, as a critical component, the employment of double-blind methodology. Neither the participant in the study nor the person administering the medication presumably know if any particular person is receiving the medication under investigation or an identical-in-appearance placebo. When the side effects of the active medication are discerned by either party, the blind is penetrated and the internal validity of the study is seriously compromised. This is a major confound and criticism of drug trials (Sparks, Duncan, Cohen, & Antonuccio, 2010). In studies of psychotherapy, everyone knows which treatment is being offered or received—there is no true placebo. The ongoing penetration of the double-blind in psychotherapy research introduces other, so-called nonspecific variables of influence (e.g., allegiance effects; Luborsky et al., 1999), calling into question the assumption of specific effects of treatment models.

Moreover, there are three empirical arguments that cast doubt upon the specific effects assumption. First is the *dodo bird verdict*, which colorfully summarizes the robust finding that therapy approaches do not show specific effects or relative efficacy. In 1936, Saul Rosenzweig first invoked the dodo's words from *Alice's Adventures in Wonderland*, "Everybody has won and all must have prizes," to illustrate his observation of the equivalent success of diverse

²*Effect size* (ES) refers to the magnitude of change attributable to treatment, compared to an untreated group. The ES most associated with psychotherapy is 0.8 standard deviations above the mean of the untreated group. An ES of 1.0 indicates that the mean of the treated group falls at approximately the 84th percentile of the untreated one. Consequently, the average treated person is better off than approximately 80% of those without the benefit of treatment.

psychotherapies. Almost 40 years later, Luborsky, Singer, and Luborsky (1975) empirically substantiated Rozenzweig's conclusion in their now-classic review of comparative clinical trials. The dodo bird verdict has since become the most replicated finding in the psychological literature, encompassing a broad array of research designs, problems, and clinical settings.

Three methodologically sophisticated comparative clinical trials illustrate the dodo verdict. Ushering in the RCT in psychotherapy research was the Treatment of Depression Collaborative Research Program (TDCRP; Elkin et al., 1989). The TDCRP randomly assigned 250 depressed participants to four different conditions: CBT, interpersonal therapy (IPT), antidepressants plus clinical management (IMI), and a pill placebo plus clinical management. The four conditions—including placebo—achieved comparable results, although both IPT and IMI surpassed placebo (but not the other treatments) on the recovery criterion.

Project MATCH, considered the largest and most statistically powerful clinical trial in the history of alcohol and drug treatment (Project MATCH Research Group, 1997), was designed to examine differential efficacy and treatment matching. Three widely divergent approaches were included: motivational enhancement therapy (MET), 12-step facilitation (TSF), and CBT. The results revealed considerable improvement, but no differences in outcome emerged among the three approaches. Follow-up 10 years later (Tonigan et al., 2003) found no support for differential outcomes among the three therapies on percent days abstinent, drinks per drinking day, or total standard drink measures.

In the Cannabis Youth Treatment (CYT) Study (Dennis et al., 2004), considered by many to be the largest and most methodologically sound investigation of adolescents to date, 600 adolescents were assigned either to treatment with MET plus CBT (5 or 12 sessions), family education and therapy, Adolescent Community Reinforcement Approach, or Multidimensional Family Therapy (MDFT). Comparisons between conditions found roughly equivalent statistically significant pre-post treatment effects that were stable in terms of days of abstinence and percent in recovery by the end of the study.

Meta-analyses have yielded similar results. A meta-analysis designed specifically to test the dodo bird verdict (Wampold et al., 1997) included some 277 studies conducted from 1970 to 1995. This analysis verified that no approach has reliably demonstrated superiority over any other. At most, the ES of treatment differences was a weak 0.2. "Why," Wampold et al. ask, "[do] researchers persist in attempts to find treatment differences, when they know that these effects are small?" (p. 211).

Perhaps a more controversial illustration is provided by the treatments for the diagnosis du jour, *posttraumatic stress disorder* (PTSD). CBT has been demonstrated to be effective and is widely believed to be the treatment of choice, but several approaches with diverse rationales and methods have also been shown

to be effective: eye-movement desensitization and reprocessing, cognitive therapy without exposure, hypnotherapy, psychodynamic therapy, and present-centered therapy.

A recent meta-analysis comparing these treatments found all of them about equally effective (Benish, Imel, & Wampold, 2007). What is remarkable is the diversity of methods that achieve similar results. Two of the treatments, cognitive therapy without exposure and present-centered therapy, were designed to exclude any therapeutic actions that might involve exposure (clients were not allowed to discuss their traumas because it invoked imaginal exposure). Despite the presumed extraordinary benefits of exposure for PTSD, the two treatments without it, or in which it was incidental (psychodynamic), were just as effective (Benish et al., 2007).

The dodo bird verdict has been replicated in real-world studies as well. For example, the study mentioned above (Stiles, Barkham, et al., 2006) comparing CBT, PDT, and PCT as routinely practiced found no differences among the approaches. The preponderance of the data, therefore, indicate a lack of specific effects and refute any claim of superiority when two or more bona fide treatments fully intended to be therapeutic are compared. If there are no specific technical operations that can be reliably shown to produce a specific effect, then mandating specific models and techniques for particular disorders seems to make little sense.

It is important to keep in mind that despite attaining the status of EBT, attaining the list merely means that a model of treatment has shown itself only to be better than placebo, sham treatments, or no treatment at all, which is not really newsworthy given that it has been known for five decades that therapy is superior to placebo or no treatment. Think about it: What if one of your friends went on a date with a new person, and when you asked about the guy, your friend replied, “He was better than nothing—he was unequivocally better than watching TV or washing my hair.” (Or, if your friend was a researcher: “He was *significantly* better, at a 95% confidence level, than watching TV or washing my hair.) How impressed would you be?

The second argument shining a light on the specific-ingredients assumption comes from component studies. Component studies, which dismantle approaches to tease out unique ingredients, have similarly found little evidence to support any specific effects of therapy. A prototypic component study can be found in an investigation by Jacobson et al. (1996) of CBT and depression. Clients were randomly assigned to (1) behavioral activation treatment, (2) behavioral activation treatment plus coping skills related to automatic thoughts, or (3) the complete cognitive treatment (the above two conditions plus identification and modification of core dysfunctional schemas). Results generally indicated no differences at termination and follow-up. Perhaps putting this issue to rest, a meta-analytic investigation of component studies (Ahn & Wampold, 2001) located 27

comparisons in the literature between 1970 and 1998 that tested an approach against that same approach without a specific component. The results revealed no differences. These studies have shown that it doesn't matter what component you leave out—the approach still works as well as the treatment containing all of its parts.

A final empirical argument challenging the assumption comes from estimates regarding the impact of specific technique on outcome. After an extensive but nonstatistical analysis of decades of outcome research, Lambert (1986, 1992) suggests that model/technique factors account for about 15% of outcome variance. An even smaller role for specific technical operations of various psychotherapy approaches is proposed by Wampold (2001). His meta-analysis assigns a 13% (derived from a 0.8 ES) contribution to the impact of therapy, both general and specific factors combined. Of that 13%, a mere 8% is portioned to the contribution of model effects. Of the total variance of change, only 1% can be assigned to specific technique. This surprisingly low number is derived from the 1997 meta-analytic study, in which the most liberally defined ES for treatment differences was 0.2—indicating that only 1% of the variance in outcomes can be attributed to specific treatment factors.

When taken in total—the equivalent results of comparative clinical trials and meta-analytic investigations, component studies, and analyses of the amount of variance attributed to specific effects—the evidence points in the same direction. There are no significant unique ingredients to therapy approaches, offering no justification for mandating EBTs.

EBTs and the Known Sources of Variance: The Common factors

Whoever acquires knowledge and does not practice it resembles him who ploughs his land and leaves it unsown.

Sa'di, *Gulistan*

There is a certain seductive appeal to the idea of making psychological interventions dummy-proof, where the users—the client and the therapist—are basically irrelevant. This product view of therapy is perhaps the most empirically vacuous aspect of EBTs because the treatment itself accounts for so little of outcome variance, while the client and the therapist—and their relationship—account for so much. These are the common factors of psychotherapy.

The common factors—what works in therapy—have a storied history that started with Rosenzweig’s (1936) classic article “Implicit Common Factors in Diverse Forms of Psychotherapy.” In addition to the original invocation of the dodo bird and seminal explication of the common factors of change, Rosenzweig also provided the best explanation for the common factors, still used today. Namely, given that all approaches achieve roughly similar results, there must be pantheoretical factors accounting for the observed changes beyond the presumed differences among schools (Duncan, 2010b).

The factors are interdependent, fluid, dynamic, and dependent on who the players are and what their interactions are like. The common factors provide a big picture view of what really works, suggesting that efforts in therapy be commensurate to each element’s differential impact on outcome. Five factors comprise this perspective: client, therapist, alliance, the model/technique delivered, and feedback.

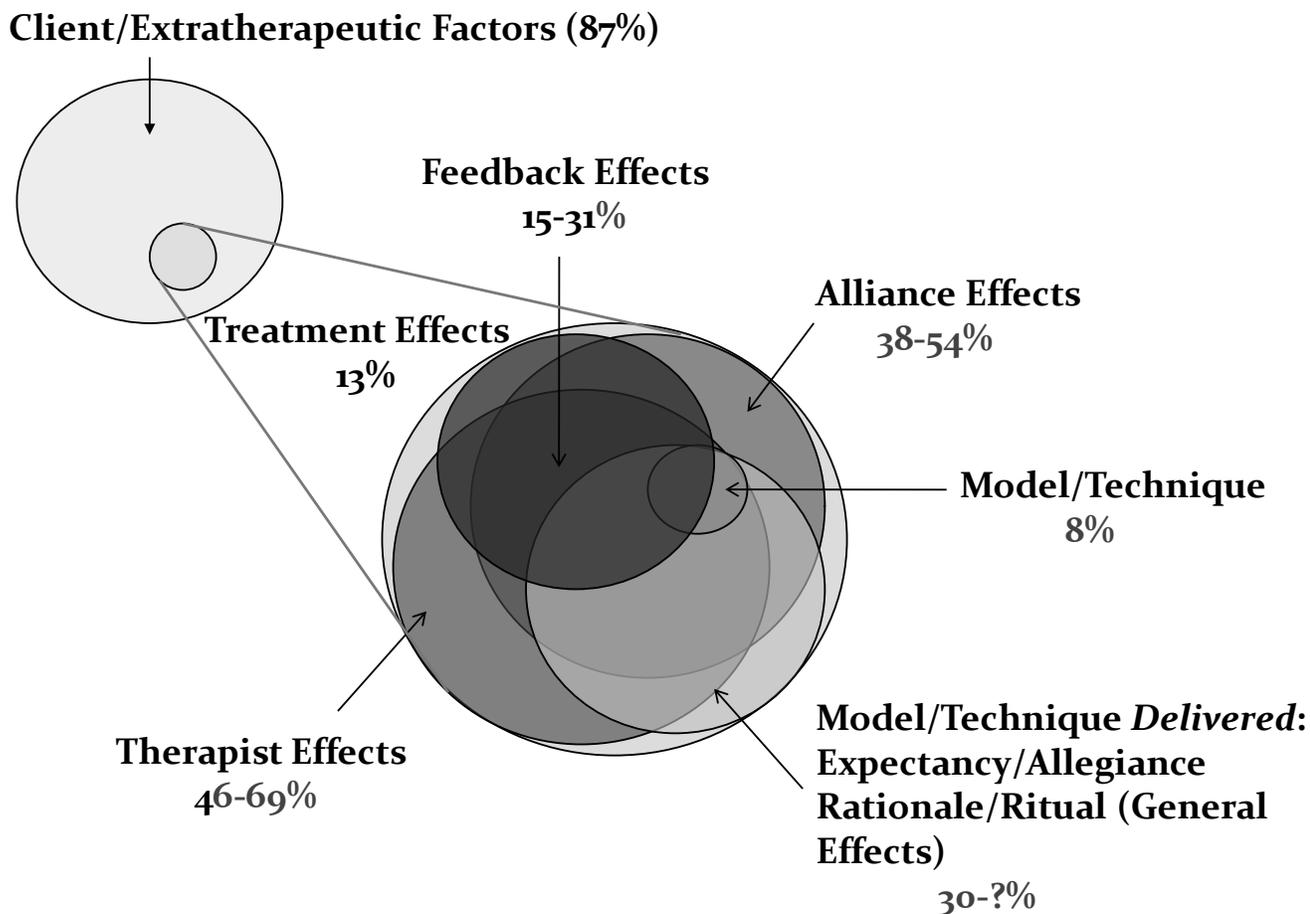
Extratherapeutic/Client Factors

To understand the common factors, it is first necessary to separate the variance due to psychotherapy from that attributed to *extratherapeutic factors*, those variables incidental to the treatment model, idiosyncratic to the specific client, and part of the client’s life circumstances that aid in recovery despite participation in therapy (Asay & Lambert, 1999)—everything that the client brings to therapy. (See Figure 21.1.)

The proportion of outcome variance attributable to client factors is represented by the circle on the left. The variance accounted for by treatment is depicted by the small circle nested within client factors (at the lower-right side). Even a casual inspection reveals the disproportionate influence of what the client brings to therapy. Client factors, including unexplained and error variance, account for 87% of the variance of change, leaving 13% of the variance accounted for by psychotherapy (Wampold, 2001). These extratherapeutic aspects consist of client strengths, struggles, motivations, distress, supportive elements in the environment, and even chance events. These elements are the most powerful of the common factors in therapy—the client is the engine of change (Bohart & Tallman, 2010).

Figure 21.1 also illustrates the second step in understanding the common factors. The second, larger circle in the center depicts the overlapping elements that form the 13% of variance attributable to treatment. Visually, the relationship among the common factors, as opposed to a static pie-chart depicting discrete elements adding to a total of 100%, is more accurately represented with a Venn

Figure 21.1 Common Factors



Reprinted from *On becoming a better therapist*, by B. Duncan, 2010, Washington, DC, American Psychological Association

diagram, using overlapping circles and shading to demonstrate mutual and interdependent action. The factors, in effect, act in concert and cannot be separated into disembodied parts (Duncan, Solovey, & Rusk, 1992). To exemplify the various factors and their attending portions of the variance, the tried and true TDCRP (Elkin et al., 1989) will be enlisted.

Therapist Effects

Therapist effects represent the amount of variance attributable not to the model wielded, but rather to *who* the therapist is. Indeed, therapist factors have emerged as potent and predictive aspects of therapeutic services, accounting for more of the variance of outcome than any treatment provided, second only to what the client brings (Wampold & Brown, 2005). The explosion of EBTs has not

eliminated the influence of the individual therapist on outcomes. Conservative estimates indicate that between 6% (Crits-Christoph et al., 1991) and 9% (Project MATCH Research Group, 1998) of the overall variance in outcomes is attributable to therapist effects or 46 to 69% of the variance attributed to treatment.³ Putting this into perspective, the amount of variance attributed to therapist factors is about six to nine times more than that of model differences. In the TDCRP, 8% of the variance in the outcomes within each treatment was due to therapists (Kim, Wampold & Bolt, 2006). The psychiatrists in the study highlight this finding—the clients receiving sugar pills from the top-third-most-effective psychiatrists did better than the clients taking antidepressants from the bottom-third-least-effective psychiatrists.

What accounts for the variability? The alliance accounts for the lion's share of therapist variability. Baldwin, Wampold, and Imel (2007), for example, found that therapists who generally form better alliances also had better outcomes.

The Alliance

Researchers repeatedly find that a positive alliance—an interpersonal partnership between the client and therapist to achieve the client's goals (Bordin, 1979)—is one of the best predictors of outcome (Horvath, & Bedi, 2002; Horvath & Symonds, 1991; Martin, Garske, & Davis, 2000). The amount of variance attributed to the alliance ranges from 5 to 7% of overall variance or 38 to 54% of the variance accounted for by treatment. Putting this into perspective, the amount of change attributable to the alliance is about five to seven times that of specific model or technique. Krupnick et al. (1996) analyzed data from the TDCRP and found that the alliance, from the client's perspective, was predictive of success for all conditions—the treatment model was not. Mean alliance scores explained 21% of the variance, whereas treatment differences accounted for 0% to at most 2% of outcome variance (Wampold, 2001). Keep in mind that treatment accounts for, on average, 13% of the variance. The alliance in the TDCRP explained more of the variance by itself, illustrating how the percentages are not fixed and depend on the particular context of client, therapist, alliance, and treatment model.

Research on the power of the alliance reflects over 1,000 findings, and counting (Orlinsky, Rønnestad, & Willutzki, 2004). In Project MATCH, the alliance, regardless of the treatment employed, was a significant predictor of participation, drinking behavior during treatment, and drinking at 12-month

³The percentages are best viewed as a defensible way to understand outcome variance but not as representing any ultimate truths. Because of the overlap among the common factors, the percentages for the separate factors will not add to 100%.

follow-up (Connors, Carroll, DiClemente, Longabaugh, & Donovan, 1997). In the CYT, Shelef, Diamond, Diamond, and Liddle (2005) examined adolescent–therapist and parent–therapist alliances, dropout, and outcome in the MDFT condition of the CYT. Positive parent–counselor alliance scores predicted retention, and adolescent alliance predicted fewer substance abuse symptoms, accounting for 7% of the variance; the adolescent–parent alliance interaction accounted for an additional 6% of the variance. Finally, adolescent ratings of the alliance predicted substance-related problems at 3- and 6-month follow-up (Tetzlaff et al., 2005).

Model/Technique Delivered:⁴ Allegiance and Placebo (Expectancy) Factors

Model/technique factors are the beliefs and procedures unique to any given treatment. But these specific effects, the impact of the differences among treatments, are very small, only about 1% of the overall variance or 8% of that attributable to treatment. But the *general effects* of delivering a treatment are far more potent. As Jerome Frank (1973) seminally noted, all models include a rationale or myth, an explanation for the client’s difficulties, and a procedure or ritual, strategies to follow for resolving them. Models achieve their effects, in large part, if not completely, *through* the activation of placebo, hope, and expectancy, combined with the therapist’s belief in (allegiance to) the treatment administered. As long as a treatment makes sense to, is accepted by, and fosters the active engagement of the client, the particular approach used is unimportant. Said another way, therapeutic techniques are placebo-delivery devices (Kirsch, 2005).

Allegiance and expectancy are two sides of the same coin—the belief by both the therapist and the client in the restorative power and credibility of the therapy’s rationale and related rituals. When a placebo or technically “inert” condition is offered in a manner that fosters positive expectations for improvement, it reliably produces effects almost as large as a bona fide treatment (Baskin, Tierney, Minami, & Wampold, 2003). The TDCRP is again instructive. First, across all conditions, client expectation of improvement predicted outcome (Sotsky et al., 1991). And second, an inspection of the Beck Depression Inventory scores of those who completed the study (see Elkin et al., 1989) reveals that the placebo plus clinical management condition accounted for nearly 93% of the average response to the active treatments (Duncan, 2010a).

⁴This term was coined by Bruce Wampold, and the idea grew out of a discussion during the preparation of the introductory chapter in *The Heart and Soul of Change*, but was not included or developed in that chapter.

To punctuate the point about the more powerful general effects, consider *present-centered therapy*, mentioned earlier as a treatment that works for PTSD (see Wampold, 2007 for a full description). Researchers testing the efficacy of CBT for PTSD wanted a comparison group that contained curative factors shared by all treatments (warm empathic relationship) while excluding those believed unique to CBT (exposure). This control treatment, present-centered therapy (PRCT), contained no treatment rationale and no therapeutic actions. Moreover, to rule out any possibility of exposure, even covert in nature, clients were not allowed to talk about the traumatic events that had precipitated therapy. PRCT was, of course, found to be less effective than CBT—it wasn't really a treatment with professed "active" ingredients. However, when later a manual containing a rationale and condition-specific treatment actions was added to facilitate standardization in training and delivery, few differences in efficacy were found between PRCT and CBT in the treatment of PTSD (McDonagh et al., 2005). In fact, significantly fewer clients dropped out of PRCT than CBT. Thus, when PRCT was made to resemble a bona fide treatment, that is, it added placebo, expectancy, and allegiance variables, it was not only as effective but also more acceptable than CBT.

The act of administering treatment—the *model/technique delivered*—is the vehicle that carries allegiance and placebo effects in addition to the specific effects of the given approach. Placebo factors are also fueled by a therapist belief that change occurs naturally and almost universally—the human organism, shaped by millennia of evolution and survival, tends to heal and to find a way, even out of the heart of darkness (Sparks & Duncan, 2010).

Finally, it is important to note that suggesting specific effects are small in comparison to general effects, and that psychotherapy approaches achieve about the same results, does not mean that models and techniques are not important. On the contrary, a particular orientation or method may be just the ticket for a given client—what Beutler (see Beutler, Harwood, Michelson, Song, & Holman, 2011) and others refer to as *treatment by client interaction*. While there is no differential efficacy on aggregate, there are approaches that are likely better or worse for individual clients. Moreover, model/techniques are essential components of a common factors perspective. The alliance, expectancy, and model/technique are interdependent and overlapping. Technique is the alliance in action, carrying an explanation for the client's difficulties and a remedy for them—an expression of the therapist's belief that it could be helpful in hopes of engendering the same response in the client. Indeed, you cannot have an alliance without a treatment, an agreement between the client and therapist about how therapy will address the client's goals. Similarly, you cannot have a positive expectation for change without a credible way for both the client and therapist to understand how change can happen.

EBTs neither explain nor capitalize on these sources of variance known to effect treatment outcome. A simpler path to effective, efficient, and accountable intervention exists. Rather than attempting to fit clients into manualized treatments via “evidence-based treatments,” we recommend that therapists and systems of care tailor their work to individual clients through “practice-based evidence.”

Feedback Effects

Practice-based evidence or consumer-based outcome feedback will likely become the rage of the next decade—and for good reason: Monitoring client-based outcome, when combined with feedback to the clinician, significantly increases the effectiveness of services. Lambert (2010) reports that ESs for the difference between feedback and treatment-as-usual (TAU) ranges from 0.34 to 0.92, unusually large considering that the estimates of the ES of the difference between empirically supported and comparison treatments are about 0.20. Putting this in perspective, feedback has two to four times the impact of model differences. Given its broad applicability, lack of theoretical baggage, and independence of a specific instrument or defined practice, feedback can be argued to be a factor that demonstrably contributes to outcome regardless of the theoretical predilection of the clinician.

The APA Task Force (2006) commented that client feedback was an important area of research that needed to be considered as a means to improve treatment by “providing clinicians with real-time patient feedback to benchmark progress in treatment and clinical support tools to adjust treatment as needed” (p. 278). APA’s Division 29 Task Force on Empirically Supported Relationships also supported the use of feedback by advising practitioners to “routinely monitor patients’ responses to the therapy relationship and ongoing treatment. Such monitoring leads to increased opportunities to repair alliance ruptures, to improve the relationship, and to avoid premature termination” (Ackerman et al., 2001, p. 496).

Although there are several feedback systems available (see Lambert, 2010), only two have extensive empirical support. Lambert, the pioneer of outcome feedback, has conducted five RCTs using the Outcome Questionnaire 45.2, and all five demonstrated statistically significant gains for feedback groups over treatment as usual (TAU) for clients at risk for a negative outcome. Twenty-two percent of TAU at-risk cases reached reliable improvement and clinically significant change compared with 33% for feedback to therapist groups, 39% for feedback to therapists and clients, and 45% when feedback was supplemented with support tools such as measures of the alliance (Lambert, 2010). The addition of client feedback alone, without new techniques or models of treatment and leaving therapists to practice as they saw fit, enabled more than twice the amount of at-risk clients to benefit from psychotherapy.

The Partners for Change Outcome Management System (PCOMS; Duncan, 2010a, in press; Duncan, Miller, & Sparks, 2004; Miller, Duncan, Sorrell, & Brown, 2005) appeal rests on the brevity of the measures and therefore its feasibility for everyday use in the demanding schedules of frontline clinicians. The Outcome Rating Scale (ORS) and the Session Rating Scale (SRS) (available free for individual clinician use at www.heartandsoulofchange.com) are both four-item measures designed to track outcome and the therapeutic alliance, respectively. PCOMS was based on Lambert's continuous-assessment model using the Outcome Questionnaire 45.2 (Lambert et al., 1996), but there are differences beyond the measures. First, PCOMS is integrated into the ongoing psychotherapy process and routinely includes a transparent discussion of the feedback with the client (Duncan, 2010a; Duncan et al., 2004). Session-by-session interaction is focused by client feedback about the benefits or lack thereof of psychotherapy. Second, PCOMS assesses the therapeutic alliance every session and includes a discussion of any potential problems. Lambert's system includes alliance assessment only when there is a lack of progress. Finally, the ORS, rather than a list of symptoms rated on a Likert Scale, is a clinical tool as well as an outcome instrument that evolves in collaboration with clients from a general framework of client distress to a specific representation of the client's idiosyncratic experience and reasons for service (Duncan, in press).

Four studies have demonstrated the benefits of client feedback with the ORS and SRS. In an effectiveness study, Miller, Duncan, Brown, Sorrell, and Chalk (2006) explored the impact of feedback in a large culturally diverse sample utilizing a telephonic employee assistance program (EAP). Although the study's quasi-experimental design qualifies the results, the use of outcome feedback doubled overall effectiveness and significantly increased retention. Three recent RCTs used PCOMS to investigate the effects of feedback versus TAU. First, in an independent investigation, Reese, Norsworthy, and Rowlands (2009) found that clients who attended therapy at a university counseling center or a graduate training clinic demonstrated significant treatment gains for feedback when compared to TAU ($ES = 0.49$). Second, a recent study in Norway (Anker, Duncan, & Sparks, 2009), the largest RCT of couple therapy ever done, found that feedback clients reached clinically significant change nearly four times more than non-feedback couples ($ES = 0.50$). The feedback condition maintained its advantage at 6-month follow-up and achieved a 46% lower separation/divorce rate. Feedback improved the outcomes of 9 out of 10 therapists in this study. Third, Reese, Toland, Slone, and Norsworthy (2010) replicated the Norway feedback with couples, finding nearly identical results ($ES = 0.48$). Finally, a recent meta-analysis of PCOMS studies (Lambert & Shimokawa, 2011) found that those in the feedback group had 3.5 higher odds of experiencing reliable change and less than half the odds of experiencing deterioration.

An inspection of Figure 21.1 shows that feedback overlaps and affects all the factors—it is the tie that binds them together—allowing the other common factors to be delivered one client at a time. Soliciting systematic feedback is a living, ongoing process that engages clients in the collaborative monitoring of outcome, heightens hope for improvement, fits client preferences, maximizes therapist–client fit and client participation, and is itself a core feature of therapeutic change. Securing client feedback also exemplifies what Stricker and Trierweiler (1995) called the “local clinical scientist.” Positing that the clinical setting is analogous to the laboratory, Stricker and Trierweiler suggested that the inadequacy of any one model is reduced by embracing local observations and solutions to problems that are then subjected to the same need for verifiability that greets all scientific enterprises.

Common factors research provides general guidance for enhancing those elements shown to be most influential in positive outcomes. The specifics, however, can be derived only from the client’s response to any treatment delivered—the client’s feedback regarding progress in therapy and the quality of the alliance. Therapists need not know what approach should be used with each disorder as suggested by the mandate for EBTs, but rather whether the delivered approach is a good fit for and benefits the client as suggested by practice-based evidence. The empirical justification for mandating EBTs can be further examined by taking a closer look at the RCT and its close association with the medical-model perspective of psychotherapy.

Randomized Clinical Trials, Evidence-Based Treatments, and the Medical Model

To exchange one orthodoxy for another is not necessarily an advance. The enemy is the gramophone mind, whether or not one agrees with the record that is being played at the moment.
George Orwell

The RCT has often been criticized for its limited ecological validity, manualized treatments, fixed number of sessions, and homogeneity of participants that typically focus on diagnosis (Duncan, 2002; Duncan et al., 2004; Seligman, 1995). In other words, the RCT is not reflective of the complexity and nuances of how treatment is delivered in the real world. EBTs’ alignment with the RCT, and the accompanying reliance on diagnosis and treatment manuals (to mimic drug

protocols), inextricably aligns EBTs with the medical model. The trend, however, toward describing, researching, teaching, practicing, and regulating psychotherapy in the terms of the medical model began long before the push for EBTs. George Albee (2000) suggested that psychology made a Faustian deal with the medical model over 50 years ago. The deal was sealed, he asserted, at the famed Boulder Conference in 1949, where psychology's bible of training was developed with a fatal flaw:

[The fatal flaw] . . . was the uncritical acceptance of the medical model, the organic explanation of mental disorders, with psychiatric hegemony, medical concepts, and language. (Albee, 2000, p. 247)

Later, in the 1970s, with the passing of freedom-of-choice legislation guaranteeing parity with psychiatrists, psychologists (and, later, others) learned to collect from third-party payers using only a psychiatric diagnosis for reimbursement. Thereafter, drowning any possibilities for other psychosocial systems of understanding human challenges, the NIMH, the leading source of research funding for psychotherapy, decided to apply the same methodology used in drug research to evaluate psychotherapy (Goldfried & Wolfe, 1996)—the randomized clinical trial requiring both diagnosis and manualized treatments. Diagnosis reached its pinnacle. Now both reimbursement and research funding depended on it. Funding for studies not related to specific treatments for specific disorders precipitously dropped as both research and psychotherapy itself became more and more medicalized and dependent on diagnosis and manualization for credibility.

Diagnosis

Diagnosis is the beginning point, the foundation of both the medical model as well as the RCT. Unlike with medical treatments, diagnosis is an ill-advised starting point for psychotherapy. The *Diagnostic and Statistical Manual (DSM)* of the American Psychiatric Association, “the professional digest of human disasters” (Duncan, Miller & Sparks, 2004, p. 23), dates back to 1952. The original DSM contained 66, while the latest version (*DSM-IV-TR*) totals 397 disorders/diagnoses (American Psychiatric Association, 1952, 2000)—the volume itself has swelled from 100 to 943 pages. Still ringing true today, in 1961, in the first edition of his classic book, Jerome Frank wrote, “Psychotherapy is the only form of treatment which, at least to some extent, appears to create the illness it treats” (p. 7). *DSM-5*, coming soon, promises even more disorders and more pages.

It simply lacks reliability. The last major study of the DSM, using highly trained clinicians at multiple sites under supervision of the very writers of the DSM (Williams et al., 1992), found reliability coefficients not much different from studies in the 1950s and 1960s.

Robert Spitzer, the architect of the *DSM-III*, admitted:

To say that we've solved the reliability problem is just not true. . . . It's been improved. But if you're in a situation with a general clinician it's certainly not very good. There's still a real problem, and it's not clear how to solve the problem. (Spiegel, 2005, p. 63)

In addition to underwhelming reliability, psychiatric diagnosis lacks validity.

Allen Frances, lead editor of the fourth edition of the DSM, recently confessed that “there is no definition of a mental disorder. It's bullshit. I mean, you just can't define it” (Greenberg, 2010, p. 1). This candid admission merely confirms what has been known for many years, in fact since its inception (e.g., Frank, 1961). For example, Kendell and Zablansky (2003, p. 7), writing in the *American Journal of Psychiatry*, conclude that “[a]t present there is little evidence that most contemporary psychiatric diagnoses are valid, because they are still defined by syndromes that have not been demonstrated to have natural boundaries.” Psychiatric diagnoses fail the most basic definition of validity—they lack empirical standards to distinguish the hypothesized pathological states from normal human variation or other disorders. Medicine is able to define the conditions of a disease-free organism as a basis to understand illness. Physicians know, for example, the normal range for glucose levels in the blood. They are able to discern deviations beyond established parameters and can reliably diagnose diabetes.

In mental health, no such “normal” parameters exist for the wide variations of human behavior. Consequently, diagnosis always begs numerous, unanswered questions concerning cultural expectations and the role that power, privilege, gender, and race play in identifying, cataloguing, and addressing client distress. The result is a set of murky, over-inclusive criteria, often disadvantaging those who are racially or ethnically different, for an ever-growing list of disorders (Duncan et al., 2004). Diagnosis locates problems inside the individual, giving a free pass to social conditions like poverty and racism that breed fear and despair (Albee, 2000).

Finally and particularly germane to practitioners, diagnosis tells little about a person that is relevant to therapeutic change. Diagnosis in mental health is not correlated with outcome or length of stay (Brown et al., 1999; Wampold & Brown, 2005), and given the dodo bird verdict cannot provide reliable guidance to clinicians or clients regarding the best approach to resolving a problem. And diagnosis does not address what is most relevant to the helping process, namely the impact of the “disorder” in the client's life and what can be done about it. Diagnosis also does not cover the range of reasons for which people seek therapy—relational, situational, and quality-of-life related, not symptom oriented. Nevertheless, the DSM, in spite of a long history of detailed critique (Carson, 1997; Duncan et al., 2004; Kirk & Kutchins, 1992), poor reliability and validity,

and limited power to predict treatment outcome, lives on. It remains a fixed part of graduate training programs, a prominent feature of EBTs, and a prerequisite for funding in most mental health and substance abuse delivery systems—all engendering an illusion of scientific aura and clinical utility that far overreaches the DSM’s deeply flawed infrastructure. Change, however, is afoot and a substantial protest to the upcoming DSM V has mounted. The Society for Humanistic Psychology (Division 32 of APA) in alliance with several other APA Divisions as well as professional organizations from around the world has circulated a petition entitled “An Open Letter to the DSM-5” (visit: http://www.ipetitions.com/petition/dsm5/?utm_medium=email&utm_source=system&utm_campaign=Send%2Bto%2BFriend).

Manualization

Manuals date back to the 1960s (Lang & Lasovik, 1963), but bringing the RCT to psychotherapy, and ultimately the EBT movement, brought them to life. Drawing on 8 of the 12 overlapping lists of empirically supported therapies, Chambless and Ollendick (2001) noted that 108 different manualized treatments have met the specific criteria of empirical support—a daunting number for any clinician to consider. Although the move to manualize psychotherapy emerged from its increasing medicalization of psychotherapy, manuals have a positive role to play. They enhance the internal validity of comparative outcome studies, facilitate treatment integrity and therapists’ technical competence, ensure the possibility of replication, and provide a systematic way of training and supervising therapists in specific models (Lambert & Ogles, 2004).

Manuals, however, bring two critical disadvantages. Manuals emphasize specific technical operations in the face of evidence, as discussed above, that psychotherapies demonstrate few, if any, specific effects. Moreover, in direct contrast to the move to transfer manualized therapies to clinical settings, manuals have demonstrated little relationship to outcome, and perhaps even detract from positive results.

For example, Henry and colleagues (Henry, Schacht et al., 1993; Henry, Strupp, Butler, Schacht, & Binder, 1993) found that therapist interpersonal skills were negatively correlated with the ability to learn a manual in the Vanderbilt II project, which examined the effects of training in *time-limited dynamic psychotherapy* (TLDP) for 16 therapists. During the year of training, therapists participated in weekly group supervision and attended workshops teaching the manualized approach. Evaluation of the training revealed that the therapists learned the manualized protocol (Henry, Schacht et al., 1993; Henry, Strupp et al., 1993). The extensive training, however, did not result in improved treatment outcomes. Clients prior to their therapists’ manualized training were as likely to improve as those seen after training (Bein et al., 2000).

This study and others indicate that manuals can effectively train therapists in a given psychotherapy approach. The same research shows no resulting improvement in outcome and the strong possibility of untoward negative consequences (Beutler et al., 2004; Lambert & Ogles, 2004). With regard to the former, researchers Shadish, Matt, Navarro, and Phillips (2000) found non-manualized psychotherapy as effective as manualized in a meta-analysis of 90 studies. Comparing an individualized cognitive therapy to a manualized cognitive therapy, Emmelkamp, Bouman, and Blaauw (1994) found a modest, mean negative effect of manualization at treatment end and follow-up. On the other hand, Schulte, Kunzel, Pepping, and Schulte-Bahrenberg (1992) found small positive effects of manualization. Finally, a mega-analysis of 302 meta-analyses of various forms of psychotherapy and psychoeducation (Lipsey & Wilson, 1993) also revealed very similar outcomes between highly structured research treatments and those applied in naturalistic settings. The consistency of these results suggests few differences in outcome following the use of manuals in clinical settings. Finally, in a recent meta-analysis examining the relationship between adherence to, and competence in delivering, a particular manualized approach and outcome, Webb, DeRubeis, and Barber (2010) concluded that “neither adherence nor competence was . . . related to patient outcome and indeed that the aggregate estimates of their effects were very close to zero” (p. 207).

Moreover, high levels of adherence to specific technical procedures can interfere with the development of a good relationship (Henry, Strupp et al., 1993), and with positive outcomes (Castonguay, Goldfried, Wiser, Raue, & Hayes, 1996). In a study of 30 depressed clients, Castonguay and colleagues (1996) compared the impact of a technique specific to cognitive therapy—the focus on correcting distorted cognitions—with two other nonspecific factors: the alliance and the client’s emotional involvement with the therapist. Results revealed that while the two common factors were highly related to progress, the technique unique to cognitive-behavioral therapy—eliminating negative emotions by changing distorted cognitions—was negatively related to successful outcome. In effect, therapists who do therapy by the book develop better relationships with their manuals than with clients and seem to lose the ability to respond creatively. Little evidence, therefore, exists that manualized treatments have any impact on outcome, although there is some indication of negative effects.

Thinking more clinically, the use of manuals contains a fatal flaw. RCTs require that the treatments being assessed not contain the inevitable improvisations of therapy as practiced in the real world. Instead, the approaches studied are all required to follow a script, a manual, so that the variable presumably being examined—a precisely defined and structured form of treatment—can be strictly controlled. From a clinical perspective, manuals fall flat and seem like incontrovertible proof that researchers are card-carrying nerds. Experienced therapists know that the work requires the unique tailoring of any approach to a

particular client and circumstance. The nuances and creativity of an actual encounter flows from the moment-to-moment interaction of the participants—from the client, relational, and therapist idiographic mix—not from step a to step b on page 39.

Although certain kinds of therapy can be scripted—CBT being the most prominent—most cannot. So it should come as no surprise that CBT and other behavioral approaches dominate, amounting to about 80 percent of the list of EBTs. Consider also that very few approaches ever have the privilege of being researched. So the EBT list really is just those approaches that are practiced in settings that support research endeavors and that are able to attain funding. This privilege does not extend to some 250 other approaches around today. On this point, Wachtel (2010) concludes:

. . . to make manualization a requirement for regarding a treatment approach as evidence-based is not a reflection of commitment to scientific rigor, but a political ploy that effectively excludes from the lists of evidence-based treatments a variety of treatments for which there is in fact a very substantial body of evidence . . . , but which do not happen to have approached the task of empirical validation via the particular investigative strategies that the “EST” movement advocates (p. 261).

Diagnosis and manualization have little empirical support, and both have potential downsides. Although there is obvious value in RCTs, in actual clinical practice, manuals are not commonly used; therapies are never purely practiced, but rather are predominantly eclectic or integrative (Stricker, 2006); clients are not randomly assigned to treatments; and clients rarely enter therapy for singular DSM-defined disorders (Duncan et al., 2004).

Debunking Claims of Superiority: The Truth Is in the Tables

Believe those who are seeking the truth; doubt those who find it.

Andre Gide

Although the preponderance of the evidence suggests the dodo bird verdict to be true, the view that some approaches are better than others persists and this singular issue likely represents the crux of the controversy regarding EBTs, striking at the heart of any mandate to exclusively fund or implement any approach. Regardless of the overwhelming evidence that supports the dodo bird verdict, it is quite easy for a detractor to cite this or that study as an exception. The

number of exceptions is less than would be attributed to chance alone (Wampold, 2001). To resolve this apparent conundrum, however, an examination of what constitutes claims of superiority—studies that report differential efficacy—deserve a closer look. Two factors must always be kept in mind when a report of differential efficacy is advanced: allegiance factors and unfair comparisons.

Allegiance

Allegiance refers to the researchers' belief in and commitment to a particular approach. Allegiance can exert a large influence on outcome in comparative studies. For example, Luborsky et al. (1999) used three types of allegiance measures (reprint method, ratings by colleagues, and researcher self-ratings) and found that allegiance explained 69% of the variance in outcomes. Allegiance effects can trickle down to the therapist level as well. Often, allegiance-bound therapists are compared to colleagues without similar ties to models. As a point of comparison, in the TDCRP and CYT mentioned earlier, the principal investigators had no particular allegiance to the models compared, and the therapists in each condition believed their approach superior and were equally committed to their models. As a result, allegiance was controlled for and no differences were found.

One step further, when therapists in trials are trained and supervised by the model advocate, at a site where the model is taught, and in a study designed by a model proponent, they most likely will have allegiance to the researcher/trainer's model (Wampold, 2001). Consider the role of allegiance in findings for the efficacy of emotionally focused couple therapy (EFT). Johnson (2003) refers to a meta-analysis of four EFT studies (Johnson, Hunsley, Greenberg, & Schindler, 1999) indicating an ES of 1.3. This estimate significantly outstrips the 0.84 reported by Shadish and Baldwin (2002) for couple therapy. Calling the dodo bird verdict the "dodo cliché," Johnson explains, "Some researchers . . . believe that, like the Dodo bird, the idea of some models of intervention being more effective than others is extinct . . ." (2003, p. 367). Setting aside this erroneous interpretation of the dodo bird verdict, an examination of allegiance in the meta-analyzed studies addresses the assertion that "EFT appears to demonstrate the best outcomes at present" (p. 365).

First, two trials of the four compared EFT with a waitlist control and predictably found superior outcomes—demonstrations of efficacy over placebo or no treatment are not comparisons with other approaches and therefore have no bearing on the dodo bird verdict. Two studies investigated differential effects. In Johnson and Greenberg (1985), EFT was superior to problem-solving treatment (PS) on 6 of 13 outcome indices at termination and 2 of the 5 reported at 8-week follow-up. Both EFT and PS achieved statistically significant differences over the waitlist and clinically significant change (recovery into a non-distressed range),

with equivalent maintenance of that change. This article acknowledged that the first author had served as a therapist in the study and that the authors developed EFT, raising concerns about therapist allegiance to the contrasted approach conducted in an EFT hotbed. In the second trial addressing differential efficacy, Goldman & Greenberg (1992), researchers had comparable allegiance to the treatments delivered—EFT and integrated systemic therapy (IST)—and no significant differences were found.

In all four EFT studies cited by Johnson et al. (1999), authors are model developers or developers' student/trainees, and study sites are locations where model creators trained, facts acknowledged by the authors. It is worthy of note that, in the only direct comparison of EFT with another couple approach in which the comparative model was delivered by therapists with equal allegiance, no differences in outcomes were reported. Magnitudes of effect sizes and claims of superiority in the EFT meta-analysis clearly must be interpreted with allegiance as a point of reference. The robust impact of allegiance factors illustrated in these instances suggests that the portion of outcome variance attributable to allegiance factors in the literature warrants close scrutiny in evaluating claims of differential efficacy.

Unfair Comparisons

Inequality in important attributes between treatments constitutes a significant confound in evaluating comparative trial findings (Duncan et al., 2004; Sparks & Duncan, 2010; Wampold, 2001). Looking for unfair comparisons speaks to the old but relevant question: “As compared to what?” Unequal comparisons significantly inflate the meanings often attributed to results. Consider that, on average, any systematically applied treatment is four times more effective than no treatment (Lambert & Ogles, 2004). So when *functional family therapy* (FFT), for example, reports that the no-treatment group had a 41% recidivism rate while FFT achieved 9% (Gordon, Arbuthnot, Gustafson, & McGreen, 1988), the findings are laudable but nothing more than would be expected. Moreover, comparisons to no treatment have no relevance to differential efficacy.

In the minority of studies that claim superiority over TAU or another approach, you need only ask this question of the investigation (Duncan, 2010; Duncan et al., 2004; Sparks & Duncan, 2010): Is it a fair contest? In other words, is the study a comparison of two valid approaches intended to be therapeutic administered in equal amounts by therapists who equally believe in what they are doing and who are equally supported to do it—are the therapists from the same pool with equal caseloads or is the experimental group specially selected, trained, and supervised by the researcher/founder of the approach, who has reduced caseloads or other advantages? As a point of comparison, consider the Norway feedback trial (Anker et al., 2009) conducted in a real-world setting. Therapists

served as their own controls, so there was no special group pitted against another disadvantaged group. And regarding allegiance, the therapists were selected based on the naiveté regarding feedback and therefore had no special affinity to the feedback condition. In contrast, a pre-study survey revealed that they believed feedback would not improve their outcomes.

A recent investigation of Parent Management Training, the Oregon Model (PMTO), illustrates. After an uncritical account of reviews claiming PMTO efficacy (confirming Littell's fears highlighted earlier), Ogden and Hagen (2008) reported that PMTO was more effective than TAU, concluding:

The findings thus indicate that PMTO is an effective treatment program . . . with children exhibiting serious behavioral problems and moreover that an evidence-based treatment program can be transported successfully to a new participant group. (p. 617)

An examination of the tables revealed that the initial analysis included 16 outcome measures. Only 4 found a difference favoring PMTO. On one of the four measures reporting a significant effect for PMTO (the Child Behavior Check List [CBCL] Total), the difference between the means at the end of treatment of PMTO versus TAU was 1.91 points (T score means of 62.48 vs. 60.57). On another (CBCL Externalizing Total), the difference between posttreatment means was 1.53 points (T score means of 61.22 vs. 59.69). The clinical significance of these differences is questionable at best. The secondary analysis looked at treatment differences by age of the child. Once again, they found a superior finding for PMTO on 4 of 16 measures for children 7 and younger *only*, and no differences between TAU and PMTO on 15 of 16 measures for children 8 and older; 1 measure favored TAU over PMTO.

In addition to these underwhelming results, the PMTO therapists received 18 months of training and ongoing support/supervision during the study, while the TAU therapists received no additional training, support, or supervision. Finally, the dose of treatment favored PMTO (work with parents), 40 versus 21 hours. The meager results, no findings on 12 of 16 measures, and no effects favoring PMTO for children 8 and over, combined with the confounds of the differential training and support of the two therapist groups, and unequal doses of treatment, cast significant doubt on this study's conclusions. The cost effectiveness of implementing an approach that requires 18 months of training while yielding minimal results is dubious.

Another example is provided by dialectical behavioral therapy (DBT) for borderline personality disorder. Perhaps the most publicized study (Linehan et al., 2006) compared DBT with community experts, examining suicidal behavior, emergency room and hospital admissions, as well as other variables. Results indicated the DBT led to significantly fewer suicide attempts and emergency room and hospital admissions, as well as reduced medical risk, but no differences were

found for TAU with community experts on the rest of the outcome measures: suicidal ideation, the Reasons for Living Inventory, and the Hamilton Rating Scale for Depression. DBT therapists received 45 hours of specialized training as well as pre- and during-study weekly supervision and support. The TAU community expert therapists received no training, supervision, or consultation. Moreover, in addition to the individual treatment component of DBT, the DBT therapists administered 38 group therapy sessions of 2.5 hours' duration largely focused on keeping people out of the hospital, perhaps accounting for the reduced ER and hospital admissions. Although the study reports that the dose of treatment was comparable when considering all the treatments together (day treatment, vocation counseling, hospitalization, etc.), an examination of the tables revealed that the 2.5-hour group sessions were counted only as 20 minutes of therapy, a somewhat curious way to record 95 hours of additional treatment. Given the unequal doses of treatment as well as the differential training and attention that the DBT therapists received, it is surprising that DBT didn't outperform TAU on all measures.

A final RCT example is provided by trauma-focused (TF)-CBT, an approach to child sexual abuse compared to child-centered treatment (CCT: Cohen, Deblinger, Mannarino, & Steer, 2004). In the CCT condition, therapists did not see the children and parents together at all, whereas the TF-CBT therapists saw children with their parents 3 times out of the 12 possible sessions. CCT is not a therapy practiced in the real world. It is not reasonable to treat children who have been sexually abused without meeting with both the child and parent (or caring adult) together to make sense of what has happened. Moreover, therapists in the CCT condition did not provide advice or suggestions to the children or parents. Again, this is not a real treatment. In the face of such serious concerns, even the most dyed-in-the-wool "client-centered" therapist would address client requests for suggestions and guidance.

Given this mock therapy, one might also suspect that the therapists likely believed that the TF-CBT offered some advantages over CCT given there was at least some structure and ideas offered to these struggling families. Enter allegiance factors. Therapists served as their own controls (performed both TF-CBT and CCT) and were monitored for fidelity to ensure they did not offer guidance (beyond processing feelings and finding client solutions) in the CCT condition.

So, given that it was an unfair comparison of an active treatment model to one unlikely to ever happen in the real world, and given the therapists in the study could hardly help but like to offer some guidance to clients when asked and therefore likely were more committed to TF-CBT, the results are particularly underwhelming. First, there was a main effect for both conditions. There were 16 measures for the children and 4 for the caregivers. Of the 16 outcome measures, 8 found a significant advantage for TF-CBT but 3 of those were from the clinician's point of view. Only 5 of 13 client-rated measures found an advantage for TF-CBT.

All 4 of the adult measures found an advantage for TF-CBT, which is not surprising given that caregivers were not involved in the CCT condition.

Meta-analytic studies comparing EBTs to TAU also illustrate unfair comparisons. For example, in a meta-analysis of 32 studies comparing EBT to TAU for child problems, Weisz, Jensen-Doss, and Hawley (2006) report an ES of 0.30 in favor of EBT. This meager difference becomes even more so when considering: (1) When the EBT was not added to the TAU, which is a fairer comparison than comparing the combination to TAU, the effect was smaller; (2) if the dose of EBT was not greater than the dose of TAU, the difference was not statistically significant; and (3) several of the comparisons were between EBT and a TAU that was not a psychotherapy (e.g., case management or minimal contact—comparisons must be a legitimate psychotherapy approach). When the TAU was a psychotherapy approach, the effect was not statistically significant. Furthermore, many comparisons did not draw the therapists for EBT and TAU from the same pool. When therapists were drawn from the same pool, the superiority of EBT was nonsignificant.

When you scratch below the surface of superiority claims (and examine the tables), they do not hold up to critical scrutiny. There are always allegiance effects and unfair comparisons to temper the findings. The critical question: Does the study show genuine superiority or is the favored approach of the experimenter pitted against a far-less-equal opponent? Of course, there is nothing wrong with PMTO, DBT, or TF-CBT. They all offer good ideas and possibilities, but the data do not support any mandates for practice.

Empirically Supported Treatments, Evidence-Based Treatments, and Evidence-Based Practice: A Rose by Any Other Name?

At bottom every man [*sic*] knows well enough that he is a unique being, only once on this earth; and by no extraordinary chance will such a marvelously picturesque piece of diversity in unity as he is, ever be put together a second time.

Friedrich Nietzsche

That psychotherapists might possess the psychological equivalent of a “pill” for emotional distress resonates strongly with many, and is nothing if not seductive as it teases the desire to be as helpful as possible to clients. A treatment for a specific “disorder,” from this perspective, is like a silver bullet, potent and transferable from research setting to clinical practice. Any therapist need only load the silver bullet into any psychotherapy revolver and shoot the psychic werewolf stalking the client. This is the essence of an EBT approach, characterized by Division 12, depicting confidence in the available evidence and appealing to those who believe that more structure and consistency and less clinician judgment is needed to bring about positive outcomes in mental health and substance abuse services.

On the other hand, EBP reflects the understanding that scientific evidence is tentative and that outcome is dependent not only on applying the various types of empirical research but also on the participants. EBP appeals to those who value clinician autonomy and individualized treatment decisions based on unique presentations of clients. The APA Task Force on EBP exemplifies this approach to the evidence.

Which approach is right? Although it is tempting to say that they both are, and clearly they both have pros and cons, there are far reaching implications that make a noncommittal answer unsatisfactory. EBTs influence the research priorities of funding sources, the editorial policies of journals, training priorities, and the programs of scholarly conferences. Moreover, the linkage of federal, state, and managed care resources to the provision of EBTs significantly amps up the importance of the answer and brings forth another question: Is mandating EBTs justified by the evidence?

This chapter provided our answer: an unequivocal “no.” We reviewed evidence that challenged any mandate of EBTs, including the dodo bird verdict, component studies, and the common factors. Similarly we challenged the medical model of understanding psychotherapy, showing how diagnosis and the use of manuals, ironically, are not empirically supported. Finally, we offered a look at studies often cited as demonstrating the superiority of EBTs, and showed that a closer look reveals that claims of superiority are often exaggerated and plagued by allegiance effects and unfair comparisons.

In addition, we believe the EBP approach provides a better understanding of the inherent complexities of the beautifully human, interpersonal endeavor of psychotherapy. The APA Task Force definition illustrates the critiques outlined in this chapter: The first part, “the integration of the best available research,” includes the consideration of EBTs without privileging them, as well as the wide range of findings regarding the alliance and other common factors. Next, “with clinical expertise,” in contrast to the EBT mentality of the therapist as an interchangeable part, brings the therapist into the equation—highlighting what

therapists bring is consistent with emerging research about the importance of clinician variability to outcome. Moreover, the Task Force submitted: “Clinical expertise also entails the monitoring of patient progress (and of changes in the patient's circumstances—e.g., job loss, major illness) that may suggest the need to adjust the treatment” (Lambert, Bergin, & Garfield, 2004).

If progress is not proceeding adequately, the psychologist alters or addresses problematic aspects of the treatment (e.g., problems in the therapeutic relationship or in the implementation of the goals of the treatment) as appropriate. (APA, 2006, p. 276–277)

So, attaining feedback as described earlier is an evidence-based practice.

Next, “in the context of patient characteristics, culture, and preferences” rightfully emphasizes what the client brings to the therapeutic stage as well as the acceptability of any intervention to the client’s expectations, how well any model or technique resonates. In short, EBP accommodates the common factors, reinforces the importance of the therapist and client, and includes client feedback as a necessary component.

Finally, the Task Force said:

The application of research evidence to a given patient always involves probabilistic inferences. Therefore, ongoing monitoring of patient progress and adjustment of treatment as needed are essential. (Task Force, 2006, p. 280)

Proponents from both sides of the EBT-versus-EBP aisle recognize that outcome is not guaranteed regardless of evidentiary support of a given technique or the expertise of the therapist. Practice-based evidence, in other words, must become routine. EBP supports an identity of plurality, essential attention to client preferences, a focus on therapist expertise, and the importance of feedback.

The history of psychotherapy can be characterized as the search for the specific mechanisms or processes that reliably produce change. Few would debate the success of this perspective in medicine, where an organized knowledge base, coupled with improvements in diagnosis and pathology, and the development of treatments containing *specific* therapeutic ingredients, have led to the near-extinction of a number of once-fatal diseases. Unfortunately, for all the claims and counterclaims, psychotherapy, in spite of numerous years of research and development, can boast of no similar accomplishments. The evidence is difficult to ignore: Psychotherapy does not work in the same way as medicine. Psychotherapy is a *relational* endeavor, not a medical one.

EBP does a better job of capturing what empirical research can offer therapists. It calls for a more sophisticated and empirically informed clinician who chooses from a variety of orientations and methods to best fit client preferences and cultural values. Although there has not been convincing evidence for differential efficacy among approaches, whether they are called “evidence based”

or not, there is indeed differential effectiveness for the client in the room now: Therapists need expertise in a broad range of intervention options, including EBTs, but must remember that the essence of the rose is how it smells—the sweet aroma of a successful outcome gleaned from client-based outcome feedback.

References

- Ackerman, S. J., Benjamin, L. S., Beutler, L. E., Gelso, C. J., Goldfried, M. R., Hill, C., . . . Rainer, J. (2001). Empirically supported therapy relationships: Conclusions and recommendations of the Division 29 Task Force. *Psychotherapy: Theory, Research, Practice, Training*, *38*, 495–497. doi:10.1037/0033-3204.38.4.495
- Ahn, H., & Wampold, B. E. (2001). Where oh where are the specific ingredients? A meta-analysis of component studies in counseling and psychotherapy. *Journal of Counseling Psychology*, *48*, 251–257. doi:10.1037/0022-0167.48.3.251
- Albee, G. (2000). The Boulder Model's fatal flaw. *American Psychologist*, *55*, 247–248. doi:10.1037/0003-066X.55.2.247
- American Psychiatric Association. (1952). *Diagnostic and statistical manual of mental disorders*. Washington, DC: Author.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.; Text rev.). Washington, DC: Author.
- American Psychological Association Presidential Task Force on Evidence-Based Practice (2006). Evidence-based practice in psychology. *American Psychologist*, *61*, 271–285. doi:10.1037/0003-066X.61.4.271
- Anker, M. G., Duncan, B. L., & Sparks, J. A. (2009). Using client feedback to improve couple therapy outcomes: A randomized clinical trial in a naturalistic setting. *Journal of Consulting and Clinical Psychology*, *77*, 693–704. doi:10.1037/a0016062
- Anker, M. G., Owen, J., Duncan, B. L., & Sparks, J. A. (2010). The alliance in couple therapy. *Journal of Consulting and Clinical Psychology*, *78*, 635–645.
- Asay, T. P., & Lambert, M. J. (1999). The empirical case for the common factors in therapy: Quantitative findings. In M. A. Hubble, B. L. Duncan, and S. D. Miller (Eds.), *The heart and soul of change: What works in therapy* (pp. 33–56). Washington, DC: American Psychological Association.
- Baldwin, S. A., Wampold, B. E., & Imel, Z. E. (2007). Untangling the alliance-outcome correlation: Exploring the relative importance of therapist and patient variability in the alliance. *Journal of Consulting and Clinical Psychology*, *75*, 842–852. doi:10.1037/0022-006X.75.6.842
- Baskin, T. W., Tierney, S. C., Minami, T., & Wampold, B. E. (2003). Establishing specificity in psychotherapy: A meta-analysis of structural equivalence of placebo controls. *Journal of Consulting and Clinical Psychology*, *71*, 973–979. doi:10.1037/0022-006X.71.6.973

- Bein, E., Anderson, T., Strupp, H. H., Henry, W. P., Schacht, T. E., Binder, J. L., & Butler, S. F. (2000). The effects of training in Time-Limited Dynamic Psychotherapy: Changes in therapeutic outcome. *Psychotherapy Research, 10*, 119–132. doi:10.1093/ptr/10.2.
- Benish, S., Imel, Z. E., & Wampold, B. E. (2007). The relative efficacy of bona fide psychotherapies of post-traumatic stress disorder: A meta-analysis of direct comparisons. *Clinical Psychology Review, 28*, 746–759.
- Beutler, L. E., Harwood, T. M., Michelson, A., Song, X., & Holman, J. (2011). Resistance/reactance level. *Journal of Clinical Psychology: In Session, 67*, 133–142.
- Beutler, L. E., Malik, M., Alimohamed, S., Harwood, T. M., Talebi, H., Noble, S., & Wong, E. (2004). Therapist variables. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 227–306). Hoboken, NJ: Wiley.
- Blatt, S. J., Sanislow, C. A., Zuroff, D. C., & Pilkonis, P. A. (1996). Characteristics of effective therapists: Further analyses of data from the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Journal of Consulting and Clinical Psychology, 64*, 1276–1284. doi:10.1037/0022-006X.64.6.
- Bohart, A. C., & Tallman, K. (2010). Clients: The neglected common factor in psychotherapy. In B. L. Duncan, S. C. Miller, B. E. Wampold, & M. A. Hubble (Eds.), *Heart and soul of change: Delivering what works in therapy* (2nd ed., pp. 83–112). Washington, DC: American Psychological Association.
- Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy, 16*, 252–260. doi:10.1037/h0085885
- Brown, J., Dreis, S., & Nace, D. (1999). What really makes a difference in psychotherapy outcomes? Why does managed care want to know? In M. Hubble, B. Duncan, & S. Miller (Eds.), *The heart and soul of change* (pp. 389–406). Washington, DC: American Psychological Association.
- Carey, B. (2005, December 27). Psychotherapy on the road . . . to where? *New York Times*. Retrieved from <http://www.nytimes.com/2005/12/27/science/27ther.html>
- Carson, R. C. (1997). Costly compromises: A critique of *The Diagnostic and Statistical Manual of Mental Disorders*. In S. Fisher & R. P. Greenberg (Eds.), *From placebo to panacea: Putting psychiatric drugs to the test* (pp. 98–114). New York, NY: Wiley.
- Castonguay, L. G., Goldfried, M. R., Wiser, S., Raue, P. J., & Hayes, A. M. (1996). Predicting the effect of cognitive therapy for depression: A study of unique and common factors. *Journal of Consulting and Clinical Psychology, 64*, 497–504. doi:10.1037/0022-006X.64.3.497
- Chambless, D. L., Baker, M. J., Baucom, D. H., Beutler, L. E., Calhoun, K. S., Crits-Christoph, P., . . . Woody, S. R. (1998). Update on empirically validated therapies, Vol. II. *The Clinical Psychologist, 51*(1), 3–16.
- Chambless, D. L., & Crits-Christoph, P. (2006). The treatment method. In J. C. Norcross, L. E. Beutler, & R. F. Levant (Eds.) *Evidence-based practices in mental health* (pp. 191–199). Washington, DC: American Psychological Association.

- Chambless, D. L., & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology*, *52*, 685–716. doi:10.1146/annurev.psych.52.1.685
- Chambless, D., Sanderson, W. C., Shoham, V., Johnson, S. B., Pope, K. S., Crits-Christoph, P., . . . McCurry, S. (1996). An update on empirically validated therapies. *The Clinical Psychologist*, *49*, 5–14.
- Claridge, J. A., & Fabian, T. C. (2005). History and development of evidence-based medicine. *World Journal of Surgery*, *29*, 547–553. doi:10.1007/s00268-005-7910-1
- Clarke, G. N. (1995). Improving the transition from basic efficacy research to effectiveness studies: Methodological issues and procedures. *Journal of Consulting and Clinical Psychology*, *63*, 718–725. doi:10.1037/0022-006X.63.5.718
- Cochrane, A. L. (1972). *Effectiveness and efficiency: Random reflections on health services*. London, UK: Royal Society of Medicine Press.
- Cohen, J. A., Deblinger, E., Mannarino, A. P., & Steer, R. A. (2004). A multisite, randomized controlled trial for children with sexual abuse–related PTSD symptoms. *Journal of the American Academy of Child and Adolescent Psychiatry*, *43*, 393–402. doi:10.1097/01.chi.0000135645.81066.24
- Connors, G. J., Carroll, K. M., DiClemente, C. C., Longabaugh, R., & Donovan, D. M. (1997). The therapeutic alliance and its relationship to alcoholism treatment participation and outcome. *Journal of Consulting and Clinical Psychology*, *65*, 588–598. doi:10.1037/0022-006X.65.4.588
- Consumer Reports*. (1995, November). Mental health: Does therapy help? *60*, 734–739.
- Crits-Christoph, P., Barancackie, K., Kurcias, J. S., Beck, A. T., Carroll, K., Perry, K., . . . Zitrin, C. (1991). Meta-analysis of therapist effects in psychotherapy outcome studies. *Psychotherapy Research*, *1*, 81–91. doi:10.1080/10503309112331335511
- Davies, P. (2004, February). *Evidence-based government: Is it possible?* Paper presented at the 4th Annual Campbell Collaboration Colloquium, Washington, DC.
- Dennis, M., Godley, S. H., Diamond, G., Tims, F. M., Babor, T., Donaldson, J., . . . Funk, R. (2004). The Cannabis Youth Treatment (CYT) Study: Main findings from two randomized trials. *Journal of Substance Abuse Treatment*, *27*, 197–213. doi:10.1016/j.jsat.2003.09.005
- DeRubeis, R. J., Hollon, S. D., Amsterdam, J. D., Shelton, R. C., Young, P. R., Salomon, R. M., . . . Gallop, R. (2005). Cognitive therapy vs. medications in the treatment of moderate to severe depression. *Archives of General Psychiatry*, *62*, 409–416. doi:10.1001/archpsyc.62.4.409
- Duncan, B. L. (2002). The legacy of Saul Rosenzweig: The profundity of the dodo bird. *Journal Of Psychotherapy Integration*, *12*, 32-57.
- Duncan, B. L. (2010a). *On becoming a better therapist*. Washington, DC: American Psychological Association.

- Duncan, B. L. (2010b). Saul Rosenzweig: The founder of the common factors. In B. L. Duncan, S. C. Miller, B. E. Wampold, & M. A. Hubble (Eds.), *The heart and soul of change: Delivering what works* (2nd ed., pp. 3–22). Washington, DC: American Psychological Association.
- Duncan, B. L. (in press). The Partners for Change Outcome Management System (PCOMS): The Heart and Soul of Change Project. *Canadian Psychology*.
- Duncan, B. L., Miller, S. D., & Sparks, J. (2004). The heroic client: A revolutionary way to improve effectiveness through client directed outcome informed therapy (2nd ed.). San Francisco, CA: Jossey-Bass.
- Duncan, B. L., Miller, S. D., Wampold, B. E. & Hubble, M. A. (2010). *The heart and soul of change: Delivering what works* (2nd ed.). Washington, DC: American Psychological Association.
- Duncan, B. L., & Moynihan, D. (1994). Applying outcome research: Intentional utilization of the client's frame of reference. *Psychotherapy, 31*, 294–301.
- Duncan, B. L., Solovey, A. D., & Rusk, G. S. (1992). *Changing the rules: A client-directed approach*. New York: Guilford.
- Duncan, B. L. , & Sparks, J. A. (2010). *Heroic clients, heroic agencies: Partners for change* (2nd ed.). Jensen Beach, FL: Author
- Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., . . . Parloff, M. B. (1989). National Institute of Mental Health Treatment of Depression Collaborative Research Program: General effectiveness of treatments. *Archives of General Psychiatry, 46*, 971–982.
- Emmelkamp, P.M.G., Bouman, T. K., & Blaauw, E. (1994). Individualized versus standardized therapy: A comparative evaluation with obsessive-compulsive patients. *Clinical Psychology & Psychotherapy, 1*, 95–100.
- Frank, J. (1961). *Persuasion and Healing: A comparative study of psychotherapy* Baltimore, MD: Johns Hopkins.
- Frank, J. (1973). *Persuasion and Healing: A comparative study of psychotherapy* (2nd ed.) Baltimore, MD: Johns Hopkins.
- Frank, J. D., Frank, J. B. (1991). *Persuasion and healing* (3rd ed.). Baltimore, MD: John Hopkins.
- Gambrill, E., & Littell, J. H. (2010). Do haphazard reviews provide sound directions for dissemination efforts? *American Psychologist, 65*, 927.
- Gibbs, L. E. (2003). *Evidence-based practice for the helping professions: A practical guide with integrated multimedia*. Pacific Grove, CA: Brooks/Cole–Thompson Learning.
- Goldfried, M. R., & Wolfe, B. E. (1996). Psychotherapy practice and research: Repairing a strained alliance. *American Psychologist, 51*, 1007–1016. doi:10.1037/0003-066X.51.10.1007

- Goldman, A., & Greenberg, L. (1992). Comparison of integrated systemic and emotionally focused approaches to couple therapy. *Journal of Consulting and Clinical Psychology, 61*, 6–15. doi:10.1037/0022-006X.60.6.962
- Goodheart, C. D., Kazdin, A. E., & Sternberg, R. J. (2006). *Evidence-based psychotherapy: Where practice and research meet*. Washington, DC: American Psychological Association.
- Gordon, D. A., Arbuthnot, J., Gustafson, K. E., & McGreen, P. (1988). Home-based behavioral-systems family therapy with disadvantaged juvenile delinquents. *American Journal of Family Therapy, 16*, 243–255. doi:10.1080/01926188808250729
- Greenberg, G. (2011, January). Inside the battle to define mental illness. *Wired, 19*(1). Retrieved from <http://www.wired.com/magazine/19-01/>
- Henry, W. P., Schacht, T. E., Strupp, H. H., Butler, S. F., & Binder, J. L. (1993). Effects of training in time-limited dynamic psychotherapy: Mediators of therapists' responses to training. *Journal of Consulting and Clinical Psychology, 61*, 441–447. doi:10.1037/0022-006X.61.3.441
- Henry, W. P., Strupp, H. H., Butler, S. F., Schacht, T. E., & Binder, J. L. (1993). Effects of training in time-limited dynamic psychotherapy: Changes in therapist behavior. *Journal of Consulting and Clinical Psychology, 61*, 434–440. doi:10.1037/0022-006X.61.3.434
- Horvath, A. O. (2001). The alliance. *Psychotherapy: Theory, Research, Practice, Training, 38*, 365–372. doi:10.1037/0033-3204.38.4.365
- Horvath, A. O., & Bedi, R. P. (2002). The alliance. In J. C. Norcross (Ed.), *Psychotherapy relationships that work* (pp. 37–69). New York, NY: Oxford University Press.
- Horvath, A. O., & Symonds, B. D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology, 38*, 139–149. doi:10.1037/0022-0167.38.2.139
- Institute of Medicine. (2001). *Crossing the quality chasm: A new health system for the 21st century*. Washington, DC: National Academy Press.
- Jacobson, N. S., Dobson, K. S., Truax, P. A., Addis, M. E., Koerner, K., Gollan, J., . . . Prince, S. E. (1996). A component analysis of cognitive-behavioral treatment for depression. *Journal of Consulting and Clinical Psychology, 64*, 295–304. doi:10.1037/0022-006X.64.2.295
- Johnson, S. M. (2003). The revolution in couple therapy: A practitioner-scientist perspective. *Journal of Marital and Family Therapy, 29*, 365–384.
- Johnson, S. M., & Greenberg, L. S. (1985). Differential effects of experiential and problem-solving interventions in resolving marital conflict. *Journal of Consulting and Clinical Psychology, 53*, 175–184. doi:10.1037/0022-006X.53.2.175
- Johnson, S. M., Hunsley, J., Greenberg, L., & Schindler, D. (1999). Emotionally focused couples therapy: Status and challenges. *Clinical Psychology: Science and Practice, 6*, 67–79. doi:10.1093/clipsy/6.1.67

- Kazdin, A. E. (2010). *Single-case research designs: Methods for clinical and applied settings* (2nd ed.). New York, NY: Oxford University Press.
- Kendell, R., & Zabransky, A. (2003). Distinguishing between the validity and utility of psychiatric diagnoses. *American Journal of Psychiatry*, *160*, 4–12. Retrieved from <http://ajp.psychiatryonline.org/cgi/content/full/160/1/4>
- Kim, D. M., Wampold, B. E., & Bolt, D. M. (2006). Therapist effects in psychotherapy: A random effects modeling of the NIMH TDCRP data. *Psychotherapy Research*, *16*, 161–172. doi:10.1080/10503300500264911
- Kirk, S. A., & Kutchins, H. (1992). *The selling of DSM: The rhetoric of science in psychiatry*. New York, NY: Aldine.
- Kirsch, I. (2005). Placebo psychotherapy: Synonym or oxymoron? *Journal of Clinical Psychology*, *61*, 791–803. doi:10.1002/jclp.20126
- Krupnick, J. L., Sotsky, S. M., Simmens, S., Moyher, J., Elkin, I., Watkins, J., & Pilkonis, P. A. (1996). The role of the therapeutic alliance in psychotherapy and pharmacotherapy outcome: Findings in the National Institute of Mental Health Treatment of Depression Collaborative Research Project. *Journal of Consulting and Clinical Psychology*, *64*, 532–539. doi:10.1037/0022-006X.64.3.532
- Lambert, M. J. (1986). Implications of psychotherapy outcome research for eclectic psychotherapy. In J. C. Norcross (Ed.), *Handbook of eclectic psychotherapy* (pp. 436–462). New York, NY: Brunner/Mazel.
- Lambert, M. J. (1992). Psychotherapy outcome research: Implications for integrative and eclectic therapists. In J. C. Norcross & M. R. Goldfried (Eds.), *Handbook of psychotherapy integration* (pp. 94–129). New York, NY: Basic Books.
- Lambert, M. J. (2010). Yes, it is time for clinicians to monitor treatment outcome. In B. L. Duncan, S. C. Miller, B. E. Wampold, and M. A. Hubble (Eds.), *Heart and soul of change: Delivering what works in therapy* (2nd ed., pp. 239–266). Washington, DC: American Psychological Association.
- Lambert, M. J., Bergin, A. E., & Garfield, S. L. (2004). Introduction and overview. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy & behavior change* (5th ed., pp. 3–15). Hoboken, NJ: Wiley.
- Lambert, M. J., Hansen, N. B., Umphress, V., Lunnen, K., Okiishi, J., Burlingame, G., . . . Reisinger, C. (1996). *Administration and scoring manual for the OQ 45.2*. Stevenson, MD: American Professional Credentialing Services.
- Lambert, M. J., & Ogles, B. (2004). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 139–193). Hoboken, NJ: Wiley.
- Lambert, M. J., & Shimokawa, K. (2011). Collecting client feedback. In J. C. Norcross (Ed.), *Psychotherapy relationships that work* (2nd ed., pp. 203–223). New York, NY: Oxford University Press.
- Lang, P. J., & Lazovik, D. A. (1963). Experimental desensitization of phobia. *Journal of Abnormal and Social Psychology*, *66*, 519–525. doi:10.1037/h0039828

- Leff, H. S. (2002). A brief history of evidence-based practice and a vision for the future. In R. W. Manderscheid & M. J. Henderson (Eds.), *Mental health, 2003* (pp. 224–241). Rockville, MD: U.S. Department of Health and Human Services.
- Linehan, M. M., Comtois, K. A., Murray, A. M., Brown, M. Z., Gallop, R. J., Heard, H. L., . . . Lindenboim, N. (2006). Two-year randomized controlled trial and follow-up of dialectical behavior therapy vs. therapy by experts for suicidal behaviors and borderline personality disorder. *Archives of General Psychiatry, 63*, 757–766. doi:10.1001/archpsyc.63.7.757
- Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist, 48*, 1181–1209. doi:10.1037/0003-066X.48.12.1181
- Littell, J. H. (2010). Evidence-based practice: Evidence or orthodoxy? In B. L. Duncan, S.C. Miller, B. E. Wampold, and M. A. Hubble (Eds.), *Heart and soul of change: Delivering what works in therapy* (2nd ed., pp. 167–198). Washington, DC: American Psychological Association.
- Luborsky, L., Barber, J., Siqueland, L., Johnson, S., Najavits, L., Frank, A., & Daley, D (1996). The revised Helping Alliance Questionnaire (HAQ-II). *Journal of Psychotherapy Practice and Research, 5*, 260–271.
- Luborsky, L., Diguier, L., Seligman, D. A., Rosenthal, R., Krause, E. D., Johnson, S., . . . Schweizer, E. (1999). The researcher's own therapy allegiances: A “wild card” in comparisons of treatment efficacy. *Clinical Psychology: Science and Practice, 6*, 95–106. doi:10.1093/clipsy/6.1.95
- Luborsky, L., Singer, B., & Luborsky, L. (1975). Comparative studies of psychotherapies: Is it true that “everyone has won and all must have prizes”? *Archives of General Psychiatry, 32*, 995–1008.
- 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. doi:10.1037//0022-006X.68.3.438
- McDonagh, A., Friedman, M., McHugo, G., Ford, J., Sengupta, A., Mueser, K., . . . Descamps, M. (2005). Randomized trial of cognitive-behavioral therapy for chronic posttraumatic stress disorder in adult female survivors of childhood sexual abuse. *Journal of Consulting and Clinical Psychology, 73*, 515–524. doi:10.1037/0022-006X.73.3.515
- McReynolds, P. (1997). *Lightner Witmer: His life and times*. Washington, DC: American Psychological Association.
- Messer, S. B. (2001). Introduction to the Special Issue on assimilative integration. *Journal of Psychotherapy, 11*, 1–4. Retrieved from <http://www.springerlink.com/content/1053-0479/11/1/>
- Mihalic, S., Fagan, A., Irwin, K., Ballard, D., & Elliott, D. (2004). *Blueprints for violence prevention* (No. NCJ 204274). Washington, DC: U.S. Department of Justice Office of Juvenile Justice and Delinquency Prevention.

- Miller, S. D., Duncan, B. L., Brown, J., Sorrell, R., & Chalk, B. (2006). Using outcome to inform and improve treatment outcomes. *Journal of Brief Therapy, 5*, 5–22
- Miller, S. L., Duncan, B. L., Sorrell, R., & Brown, G. S. (2005). The partners for change outcome management system. *Journal of Clinical Psychology, 61*, 199–208. doi:10.1002/jclp.20111
- Nathan, P. E. (1997). Fiddling while psychology burns? *Register Report, 23*(2), 1,4-5, 10.
- Norcross, J. C. (Ed.). (2001). Empirically supported therapy relationships: Summary report of the Division 29 Task Force. *Psychotherapy, 38*, 345–356.
- Norcross, J. C., Beutler, L. E., & Levant, R. F. (2006). *Evidence-based practices in mental health: Debate and dialogue on the fundamental questions*. Washington, DC: American Psychological Association.
- Ogden, T., & Hagen, K. A. (2008). Treatment effectiveness of Parent Management Training in Norway: A randomized controlled trial of children with conduct problems. *Journal of Consulting and Clinical Psychology, 76*, 607–621. doi:10.1037/0022-006X.76.4.607
- Ollendick, T. H., & King, N. J. (2004). Empirically supported treatments for children and adolescents: Advances toward evidence-based practice. In P. M. Barrett & T. H. Ollendick (Eds.), *Handbook of interventions that work with children and adolescents: Prevention and treatment* (pp. 3–25). Hoboken, NJ: Wiley.
- Orlinsky, D. E., Rønnestad, M. H., & Willutzki, U. (2004). Fifty years of process-outcome research: Continuity and change. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 307–390). Hoboken, NJ: Wiley.
- President's New Freedom Commission on Mental Health. (2005). *Subcommittee on evidence-based practices: Background paper*. (DHHS Pub. No. SMA-05-4007). Rockville, MD: Author. Available at http://www.mentalhealthcommission.gov/reports/EBP_Final_040605.pdf
- Project MATCH Research Group. (1997). Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcomes. *Journal of Studies on Alcohol, 58*, 7–29.
- Project MATCH Research Group. (1998). Therapist effects in three treatments for alcohol problems. *Psychotherapy Research, 8*, 455–474. doi:10.1093/ptr/8.4.455
- Reese, R. J., Norsworthy, L. A., & Rowlands, S. R. (2009). Does a continuous feedback system improve psychotherapy outcome? *Psychotherapy: Theory, Research, Practice, Training, 46*, 418–431. doi:10.1037/a0017901
- Reese, R. J., Prout, H. T., Zirkelback, E. A., & Anderson, C. R. (2010). Effectiveness of school-based psychotherapy: A meta-analysis of dissertation research. *Psychology in the Schools, 47*, 1035–1045. doi:10.1002/pits.20522
- Reese, R. J., Toland, M. D., Slone, N. C., & Norsworthy, L. A. (2010). Effect of client feedback on couple psychotherapy outcomes. *Psychotherapy: Theory, Research, Practice, Training, 47*, 616–630. doi:10.1037/a0021182

- Rosenzweig, S. (1936). Some implicit common factors in diverse methods of psychotherapy. *American Journal of Orthopsychiatry*, 6, 412–415.
- Sackett, D. L., Rosenberg, W.M.C., Muir-Gray, J. A., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't. *British Medical Journal*, 312, 71–72. Retrieved from <http://www.bmj.com/content/312/7023/71.long>
- Sackett, D. L., Straus, S. E., Richardson, W. S., Rosenberg, W., & Haynes, R. B. (2000). *Evidence based medicine: How to practice and teach EBM* (2nd ed.). London: Churchill Livingstone.
- Schulte, D., Künzel, R., Pepping, G., & Schulte-Bahrenberg, T. (1992). Tailor-made versus standardized therapy of phobic patients. *Advances in Behaviour Research & Therapy*, 14, 67–92. doi:10.1016/0146-6402(92)90001-5
- Seligman, M.E.P. (1995). The effectiveness of psychotherapy: The *Consumer Reports* study. *American Psychologist*, 50, 965–974. doi:10.1037/0003-066X.50.12.965
- Shadish, W. R., & Baldwin, S. A. (2002). Meta-analysis of MFT interventions. In D. H. Sprenkle (Ed.), *Effectiveness research in marriage and family therapy* (pp. 339–370). Alexandria, VA: American Association for Marriage and Family Therapy.
- Shadish, W. R., Navarro, A. M., Matt, G. E., & Phillips, G. (2000). The effects of psychological therapies under clinically representative conditions: A meta-analysis. *Psychological Bulletin*, 126, 512–529. doi:10.1037/0033-2909.126.4.512
- Shelef, K., Diamond, G. M., Diamond, G. S.; Liddle, H. A. (2005). Adolescent and parent alliance and treatment outcome in multidimensional family therapy. *Journal of Consulting and Clinical Psychology*, 73, 689–698. doi:10.1037/0022-006X.73.4.689
- Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 32, 752–760. doi:10.1037/0003-066X.32.9.752
- Smith, T. B., Domenech Rodríguez, M., & Bernal, G. (2011). Culture. *Journal of Clinical Psychology: In Session*, 67, 166–175.
- Sotsky, S.M., Glass, D.R., Shea, M. T., Pilkonis, P.A. Collins, J.F., Elkin, I., et al. (1991). Patient predictors of response to psychotherapy and pharmacotherapy: Findings in the NIMH Treatment of Depression Collaborative Research Program. *American Journal of Psychiatry*, 148, 997-1008.
- Sparks, J. A., & Duncan, B. L. (2010). Common factors in couple and family therapy: Must all have prizes? In B. L. Duncan, S. C., Miller, B. E. Wampold, and M. A. Hubble (Eds.), *Heart and soul of change: Delivering what works in therapy* (2nd ed., pp. 357–392). Washington, DC: American Psychological Association.
- Sparks, J. A., Duncan, B. L., Cohen, D., Antonuccio, D. O. (2010). In B. L. Duncan, S.C., Miller, B. E. Wampold, and M. A. Hubble (Eds.), *Heart and soul of change: Delivering what works in therapy* (2nd ed., pp. 199–236). Washington, DC: American Psychological Association.
- Spiegel, A. (2005). The dictionary of disorder: How one man redefined psychiatric care. *The New Yorker*, January 3, 56-63.

- Stiles, W. B., Barkham, M., Twigg, E., Mellor-Clark, J., & Cooper, M. (2006). Effectiveness of cognitive-behavioural, person-centred and psychodynamic therapies as practised in UK National Health Service settings. *Psychological Medicine*, *36*, 555–566. doi:10.1017/S0033291706007136
- Stiles, W. B., Hurst, R. M., Nelson-Gray, R., Hill, C. E., Greenberg, L. S.; Watson, J. C., . . . Hollon, S. D. (2006). What qualifies as research on which to judge effective practice? In J. C. Norcross, L. E. Beutler, and R. F. Levant (Eds.), *Evidence-based practices in mental health: Debate and dialogue on the fundamental questions* (pp. 56–130). Washington, DC: American Psychological Association.
- Stiles, W. B., & Shapiro, D. A. (1989). Abuse of the drug metaphor in psychotherapy process-outcome research. *Clinical Psychology Review*, *9*, 521–543. doi:10.1016/0272-7358(89)90007-X
- Stricker, G. (2006). A poor fit between empirically supported treatments and psychotherapy integration. In J. C. Norcross, L. E. Beutler, & R. F. Levant (Eds.), *Evidence-based practices in mental health* (pp. 275–281). Washington, DC: American Psychological Association.
- Stricker, G., & Trierweiler, S. J. (1995). The local clinical scientist: A bridge between science and practice. *American Psychologist*, *50*, 995–1002. doi:10.1037/0003-066X.50.12.995
- Substance Abuse and Mental Health Services Administration. (2007, October 18). *National registry of evidence-based programs and practices*. Retrieved from <http://www.nrepp.samhsa.gov/>
- Sue, S., Zane, N., Levant, R. F., Silverstein, L. B., Brown, L. S., Olkin, R., & Taliaferro, G. (2006). How well do both evidence-based practices and treatment as usual satisfactorily address the various dimensions of diversity? In J. C. Norcross, L. E. Beutler, & R. F. Levant (Eds.), *Evidence-based practices in mental health: Debate and dialogue on the fundamental questions* (pp. 329–374). Washington, DC: American Psychological Association.
- Swift, J. K., Callahan, J. L., & Vollmer, B. M. (2011). Preferences. *Journal of Clinical Psychology: In Session*, *67*, 155–165.
- Task Force on Promotion and Dissemination of Psychological Procedures. (1995). Training in and dissemination of empirically validated treatments: Report and recommendations of the Task Force on Promotion and Dissemination of Psychological Procedures of Division 12 (Clinical Psychology) of the American Psychological Association. *The Clinical Psychologist*, *48*(1), 3–23.
- Tavris, C. (2003). Forward: The widening scientist-practitioner gap: A view from the bridge. In S. O. Lilienfeld, S. J. Lynn & J. M. Lohr (Eds.), *Science and pseudoscience in clinical psychology* (pp. ix–xviii). New York, NY: Guilford Press.
- Tetzlaff, B. T., Kahn, J. H., Godley, S. H., Godley, M. D., Diamond, G. S., & Funk, R. R. (2005). Working alliance, treatment satisfaction, and patterns of posttreatment use among adolescent substance users. *Psychology of Addictive Behaviors*, *19*, 199–207. doi:10.1037/0893-164X.19.2.199

- Tonigan, J.S., Miller, W.R., Chavez, R., Porter, N., Worth, L., Westphal, V., Carroll, L., Repa, K., Martin, A., & Tracy, L.A. (2003). Project Match 10-Year Treatment Outcome: Preliminary Findings Based On The Albuquerque Clinical Research Unit. <http://casaa.unm.edu/posters/project%20match%2010-year%20treatment%20outcome.pdf>
- VandenBos, G. R. (1996). Outcome assessment of psychotherapy [Special Issue]. *American Psychologist, 51*(10).
- Wachtel, P. (2010). Beyond ESTs: Problematic assumptions in the pursuit of evidence based practice. *Psychoanalytic Psychology, 27*, 251–272.
- Wampold, B. E. (2001). *The great psychotherapy debate: Model, methods, and findings*. Mahwah, NJ: Erlbaum.
- Wampold, B. E. (2007). Psychotherapy: The humanistic (and effective) treatment. *American Psychologist, 62*, 857–873. doi:10.1037/0003-066X.62.8.857
- Wampold, B. E., & Brown, G. (2005). Estimating therapist variability in outcomes attributable to therapists: A naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology, 73*, 914–923. doi:10.1037/0022-006X.73.5.914
- Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, “all must have prizes.” *Psychological Bulletin, 122*, 203–215. doi:10.1037/0033-2909.122.3.203
- Webb, C. A., DeRubeis, R. J., & Barber, J. P. (2010). Therapist adherence/competence and treatment outcome: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 78*, 200–211. doi:10.1037/a0018912
- Weisz, J. R., Jensen-Doss, A., & Hawley, K. M. (2006). Evidence-based youth psychotherapies versus usual clinical care: A meta-analysis of direct comparisons. *American Psychologist, 61*, 671–689. doi:10.1037/0003-066X.61.7.671
- Westen, D., Novotny, C., & Thompson-Brenner, H. (2004). The empirical status of empirical supported therapies: Assumptions, methods, and findings. *Psychological Bulletin, 130*, 631–663. doi:10.1037/0033-2909.130.4.631
- Williams, J., Gibbon, M., First, M., Spitzer, R., Davies, M., Borus, J., Howes, M., Kane, J., Pope, H., Rounsaville, B., & Wittchen, H. (1992). The structured clinical interview for DSM-III-R (SCID)II: Multi-site test-retest reliability. *Archives of General Psychiatry, 49*, 630–636.
- Woody, S. R., Weisz, J., & McLean, C. (2005). Empirically supported treatments: Ten years later. *The Clinical Psychologist, 58* (4), 5–11. Retrieved from http://www.apa.org/divisions/div12/tcp_journals/TCP_fa05.pdf