C P S P		1 1 6 8	B	Dispatch: 1.7.09	Journal: CPSP	CE: R.Balaji
Journal Name	Manuscript No.			Author Received:	No. of pages: 17	PE: HariPrasad

Making Science Matter in Clinical Practice: Redefining Psychotherapy

Larry E. Beutler, Pacific Graduate School of Psychology, Stanford University College of Medicine

Evidence suggests that the well-known chasm that exists between science and practice may be maintained less by the intransigence of practitioners than by the failure of scientists to: (a) offer a workable model of how to integrate science to clinicians and (b) to recognize how weak the evidence is for certain widely held beliefs about the nature of empirically supported treatments (ESTs). A rational weighing of the status of current evidence behooves scientists to take another, more careful look at why ESTs have failed to distinguish themselves from other treatments and to use this information in framing a broader approach to psychotherapy research.

Key words: **empirically supported treatment, methodology, principles of change, psychotherapy, research-informed practice, systematic treatment selection.** ***[Clin Psychol Sci Prac 16: 301–317, 2009]***

Originally, the preponderance of this article was prepared and presented as an invited address for the 2008 American Psychological Association meeting in Boston. It was originally to be addressed to clinicians under the title, ''Why Science Matters to Clinicians, Even if They Don't Know It.'' As I was reviewing the research literature in order to develop this latter theme, however, it rapidly became apparent that there was considerably less evidence than I had supposed to suggest that science mattered or should matter to clinicians. As I tried to summarize an arsenal of research findings, I concluded that under the original title, I could make only three relatively weak points: (a) Psychotherapy works better than no treatment at all and about as well as or better than most other treatments for most problems; (b) Many things that are done in the name of psychotherapy do not work and in fact can be harmful; and (c) both the therapist and the patient are important to the outcome of psychotherapy. There was not a lot of earthshaking news there.

As I struggled unsuccessfully with how to make these obvious points meaningful and interesting, it occurred to me that the problem was that I was speaking to the wrong audience. My intended audience of clinicians already knew these things. Contemporary scientific findings, I realized, had much more to say to scientists than to clinicians about what would advance our understanding and optimization of clinical effects. More specifically, I became convinced that scientists were intentionally obscuring many important results because of an unwarranted devotion to a limited number of scientific methods. In fact, I came to believe that they may be using methods and defining psychotherapy and research-informed practice in ways that hindered clinicians from being optimally effective. As I believed (and still do) that scientific methods offer the best hope of finding optimal and effective ways to intervene with behavioral health problems, I was forced to re-think what science has given us and then to identify the disconnects between scientific assertions and scientific evidence.

I began by revising my title—''Making Science Matter in Clinical Practice: Redefining Psychotherapy''—

Address correspondence to Larry E. Beutler, Pacific Graduate School of Psychology, 405 Broadway St., Redwood City, CA 94063. E-mail: larrybeutler@yahoo.com.

Portions of this article were originally presented under a different title as an invited address at the 2008 American Psychological Association meetings in Boston, Massachusetts.

in order to capture the new direction I was compelled to take. Under this revised title, the primary objective of this article became that of stimulating a dialogue about the nature of ''Research-Informed'' practices that (I hope) will result in a more practice-friendly role of research than is currently used in the contemporary empirically supported treatment (EST) movement.

To stimulate the desired discussion and, at the same time, to dispel the notion that I have reached this point in developing and then reconstructing this article while maintaining a dispassionate objectivity, I will begin with my conclusions. I believe the evidence and argument that I will make confirm that there is a pressing need to: (a) establish a broad research agenda and associated armamentarium of procedures to replace the limited one that seems to have shackled clinical science in the past 30 years; and (b) revise our definition of ''Research-informed psychotherapy practice'' so that it addresses those factors that actually comprise psychotherapy.

At the risk of offending some, but within the guidelines for which I was originally invited to present this article at the American Psychological Association convention, I will illustrate the advantages of the conclusions and definitions that I will propose by reviewing a coordinated series of studies from my own laboratory. I hasten to emphasize that this use of my own research rather than that of others reflects my familiarity with the subject matter; it is not an assertion of some special calling that only I can hear. It is offered not as proof of the validity of the findings which were obtained in these studies, but to illustrate some compelling findings that have arisen from employing a focused sequence of studies that arise directly from an explicit decision to broaden the definitions that are applied both to psychotherapy and evidence–based practice. The findings, of course, are also interesting, I believe, and will be used to illustrate the kind of conclusions that may arise from incorporating the broad definitions of practice and science for which I will argue as a backdrop to these examples.

MYTHS ABOUT EMPIRICALLY SUPPORTED PSYCHOTHERAPY

Humankind has always been concerned with what identifies ''truth'' and how it is distinguished from strongly held but inaccurate beliefs. In the earliest history of humans, ''truth'' was considered to be the province of religious authorities. Beliefs that were at odds with those ''truths'' that emanated from learned prophets and religious leaders were judged to be myths or heresy. Early scientists offered an alternative to the pronouncements of religious authorities based on discovery and objectively observed evidence. The contemporary struggle between the use of theory-driven interventions or research-derived ones represent a modern remnant of this ancient conflict between the value of authoritative versus discovered ''truth.'' The former perspective places relatively more emphasis on the views of charismatic scholars and relatively less on quantifiable measurement. Conversely, and perhaps with equally poor insight into its dangers, the latter often places relatively more emphasis on the method of ''knowing'' than on what specifically is known.

A more obscure parallel with this historical conflict of world views is I believe found within the scientific community itself. Like the devotion to oracles of knowledge that is seen in times past, one perspective within the field of research-informed practice (RIP) places its faith in a particular scientific method of investigation that is used rather than in a variety of scientific methods whose appropriateness to the questions asked may vary with the nature of the variables studied. In this latter view of science, one has no devotion to a particular research design and constructs, develops, or employs the method that is the most reasonable fit to the types of questions asked.

To the degree that the effort to identify EST or research-informed psychotherapies is viewing evidence through the lense of a single or preferred research methodology, when there are several competent methods available, is the degree to which the scientist has fallen prey to worshiping the method rather than ''truth.'' A cardinal value of science is openness within a discovery–based system that is guided by the questions asked rather than by the method used to answer them. By automatically excluding certain scientifically respectable methods in favor of a single ''gold standard,'' scientists of this ilk have merely transferred the mantel of authority from a person to a method.

Let me illustrate the implications and perhaps, even the value of this assertion with a thought experiment. Regardless of whether you consider yourself more of a

scientist than a practitioner or vice versa, answer the following three questions *as you believe the preponderance of scientific studies has found*.

1. (True or False) Psychotherapy would be more effective if everyone practiced an ''Empirically-Supported Treatment.''
2. (True or False) Cognitive and cognitive-behavioral therapy are more effective than relational and insight-oriented forms of psychotherapy.
3. (True or False) The relationship between the patient and therapist determines most of the meaningful outcomes that can be attributed to psychotherapy.

How certain are you that these conclusion are ''true.'' And, how have you derived this level of certainty? Does your answer depend on what you have heard some ''expert'' say at a meeting? Or, does it just fit with your own theory of what ''should'' be true, rationally? Or, do you remember a research study that addressed these questions?

The following paragraphs will inspect the evidence for these beliefs as derived from a series of meta-analyses which have identified and assessed the relative importance of the factors that contribute to psychotherapy outcome. I have chosen to limit my discourse to the results of meta-analysis because this method is specifically adapted to speak to the issues of replicability and consistency of research findings. In particular, as the literature is far too vast to review exhaustively in the current article, meta-analysis avoids the problems of inadvertently cherry picking results in order to support a preconceived opinion.

Meta-analysis offers a way of combining results across studies and as such, allows us to determine what trends and effects are present when the potential of errors that exist in individual studies are averaged out. Such analyses average the effects found in all the available studies that meet defined criteria and report them as that proportion of a normal curve that separates the compared treatments. This comparison, or *effect size* (ES), is reported as the statistic, *d*.

It should be said that meta-analyses are not without their problems. For example, they are not sensitive to the variations of outcome that may accrue from slight changes in the methods of selecting patients and implementing treatments, and they assume that all treatments within a class or brand name are very similar. Nonetheless, these disadvantages are outweighed by the advantages of being able to objectively identify patterns across a large number of studies.

ASSERTION 1: PSYCHOTHERAPY WOULD BE MORE EFFECTIVE IF EVERYONE PRACTICED AN ''EMPIRICALLY-SUPPORTED TREATMENT''

There are many scientists and scholars who would probably accept this assertion as being true, on its face. For evidence, they would probably cite two bodies of research. One of these would be drawn from the many studies and meta-analyses that have demonstrated that psychotherapy is an effective means of treatment when it is compared to the outcomes of those who have received no treatment, a delayed treatment, or a placebo treatment (e.g., Beutler et al., 2003a; Shapiro & Shapiro, 1982; Smith, Glass, & Miller, 1980; Wampold, 2001). Typically, such comparisons earn ES estimates ranging from $d = .60$ to $d > 1.0$ and average around $d = .80$), testifying to psychotherapy's effectiveness when treated patients are contrasted with those who receive no treatment.

The other body of literature cited would be that derived from the many task force reports that identify the treatments that have met the various criteria that have been used to identify ESTs. Almost universally, such criteria require treatments to have been shown to be effective in two or more randomized clinical trial (RCT) studies (e.g., see reviews of these task force findings in Chambless & Ollendick, 2001 and in Beutler et al., 2003a). Over 150 models and brands of psychotherapy have met criteria established by one or more of a dozen task forces and have done so with sufficient strength as to be identified as being empirically supported.

Unfortunately, both of these bodies of literature are frequently but incorrectly interpreted as support for the value of ESTs over the usual treatments that are used. Neither of these bodies of literature, however, has directly compared manualized and structured psychotherapies that may be identified as an EST, with therapy as usual (TAU) in clinical settings. Fortunately,

direct comparisons of ESTs and TAUs have been done and an answer to the first question is possible.

In a meta-analytic comparison of 90 studies (Shadish, Matt, Navaro, & Phillips, 2000), all of which compared a research–based, EST treatment and a ''treatment as usual'' condition, revealed few differences in the benefits achieved. The sum of the various ESs averaged about zero indicating that ESTs and TAUs obtained equivalent results across patient populations, including both clinically representative and nonrepresentative samples.

In a large mega-analysis, which combined the results of 302 meta-analyses, Lipsey and Wilson (1993) found a similar result. Each of the separate meta-analyses were restricted to studies that compared a research–based form of psychotherapy with various naturalistic and typically offered interventions that were largely based on psychoeducational models of change. The authors found that highly structured, research treatments and the usual naturalistic treatments applied in uncontrolled, clinical and educational settings achieved equivalent results. As before, the ESs were near or at zero.

A concern with the foregoing analysis, however, was that it did not systematically differentiate between comparisons that used real clinical populations and those that used analog or minimally distressed populations. Nor did this latter analysis distinguish among true psychotherapy conducted by psychotherapists and either educational interventions or those that used student therapists.

A third meta-analysis corrected for the foregoing problems (Wampold, 2001). This latter analysis included only those studies that contrasted various treatments within actual clinical settings, using treatments that were applied to actual patients and that were conducted by experienced psychotherapists. All identified treatment types were compared with one another. The obtained mean ES associated with these comparisons was again zero—the structured and ''proven'' treatments did not differ from the usual treatments or with less structured ones.

This latter study also provides a good illustration of the advantages of using meta-analysis rather than simply a subjective review of studies in addressing these important questions. While the mean ES was zero, the results of individual studies varied from one to another. Depending on one's proclivities, therefore, a subjective review could emphasize either those studies whose results favored an EST over a TAU (e.g., Schulte, Kunzel, Pepping, & Schulte-Bahrenbert, 1992) or those that favored a TAU over an EST (e.g., Emmelkamp, Bouman, & Blaauw, 1994). It should be no surprise, therefore, that impressionistic, nonmeta-analytic reviews of literature often result in diametrically opposed conclusions in including very different bodies of research (e.g., Addis & Cardemil, 2007a,b vs Ollendick & King, 2007).

ASSERTION 2: COGNITIVE AND COGNITIVE-BEHAVIORAL THERAPY ARE MORE EFFECTIVE THAN MOST RELATIONAL OR INSIGHT-ORIENTED FORMS OF PSYCHOTHERAPY

Among both research-minded practitioners and research-oriented academic psychologists, it is commonly thought that cognitive and cognitive-behavioral therapies are more effective than other procedures, particularly psychodynamic and relationship–based ones (Beutler, Williams, Wakefield, & Entwistle, 1995). This impression probably arises from three sources: (a) a few highly cited studies that have found differences favoring these latter therapies, (b) early meta-analysis that reported such differences, and (c) the general popularity of cognitive therapy in research studies and among research-oriented professionals.

Meta-analysis confirms that there are wide variations among the ESs when research treatments are compared against no-treatment comparison groups (Smith et al., 1980). Smith et al. (1980)—the original large-scale meta-analysis of psychotherapy—found ESs that varied from about .40 to over 1.20. In their analysis, as well as in a follow-up analysis by Shapiro and Shapiro (1982), the strongest ESs were obtained by cognitive and cognitive behavioral therapies. This finding may have contributed to the perception that these latter therapies are more effective than other treatment models.

By the same token, in those analyses that have specifically addressed the question of the relative effects of cognitive therapies, there often has been a weak trend favoring these treatments (e.g., Beutler, Machado, & Neufeldt, 1994; Smith et al., 1980; Shapiro & Shapiro, 1982). However, it became apparent soon after meta-analyses became popular that different studies employed

varying levels of structure within the targeted treatments, provided different amounts of training to therapists in the different conditions, and used outcome measures that varied widely in reactivity. More importantly, these variations systematically were clustered such that studies of cognitive therapies were almost always characterized by having the most structured treatments, the most highly trained therapists, and the most reactive measures, relative to the studies of various comparison treatments. The observed differences in ESs disappeared when these artifacts were corrected—when the treatments were equally structured, when measures were adjusted for reactivity, and when the treatments were delivered with equal skill. These latter, null findings have been confirmed in all systematic meta-analytic comparisons of cognitive and noncognitive therapies that have corrected for these contaminating factors. Systematic and direct, head-to-head comparisons of cognitive and cognitive-behavioral models pitted against psychodynamic therapy, interpersonal therapy, experiential therapy, and other insight models, have failed to yield reliable and strong differences favoring the cognitive therapies. The first such comparison was conducted by Smith and Glass (1977) in a comprehensive analysis that included every study of counseling and psychotherapy available at the time. This first excursion into the use of meta-analysis yielded some differences in favor of cognitive and cognitive-behavioral therapy, findings that were repeated when the authors published their findings in a book (Smith et al., 1980). However, in the latter report, the authors corrected their findings by adjusting for the variations in the reactivity of the outcome measures and noted that under these circumstances, the differences disappeared. This led them and authors of later meta-analyses to accept what has come to be called, the Dodo bird verdict (Luborsky, Singer, & Luborsky, 1975). That is, all psychotherapies, structured or not and cognitive or not, produce essentially equivalent findings when tested against one another (e.g., Berman, Miller, & Masserman, 1985; Grissom, 1996; Shapiro & Shapiro, 1982; Wampold, 2001).

In perhaps the most carefully done meta-analytic assessment of the specific effects of cognitive therapies, Wampold, Mondin, Moody, Stich, Benson, and Ahn (1997) clustered all treatments into classes (e.g., cognitive, cognitive-behavioral, psychodynamic, etc.) and then compared each class of treatment with every other class and every treatment within a class with others from that same class. The best estimate of an overall, mean ES representing these comparisons was zero. Cognitive therapy accounted for less than 1% of the variations in outcomes among the therapies. An inspection of these effects broken down by specific disorders does not change the picture. Cognitive and cognitive-behavioral therapies rarely are shown to overshadow the effects of dynamic and relationship therapies. In fact, at least one recent meta-analysis compared long-term, psychodynamic therapies with cognitive therapy and concluded that, at least among patients with personality disorder and other chronic problems, dynamic insight-oriented therapies may be treatments of choice (Leichsenring & Rabung, 2008). This latter meta-analysis, at the very least, adds to the consistency of the evidence indicating that there is little advantage for cognitive therapy when compared with relationship and insight models of treatment, even among patients with serious anxiety disorders and depression.

ASSERTION 3: THE RELATIONSHIP BETWEEN THE PATIENT AND THERAPIST DETERMINES MOST OF THE MEANINGFUL OUTCOMES THAT CAN BE ATTRIBUTED TO PSYCHOTHERAPY

An alternative to the perspective that structured, research–based (EST) treatment method determines the preponderance of change in psychotherapy has been advanced by scholars who disagree with the emphasis that RCT research methods give to brand name models. This alternative asserts that the principle effects of psychotherapy are derived from the quality of the interpersonal alliance or working relationship that develops between the patient and the therapist, rather than from the brand of therapy used (Norcross, 2002). This is a persuasive argument, and like the belief that some treatments are better than others, has earned the devotion of a large group of scholars and even a larger number of clinical practitioners. And, like the evidence that psychotherapy is effective, there is a broad range of research that confirms that there is a positive correlation between the strength of the therapeutic alliance or relationship and the amount of therapeutic change observed (e.g., Duncan & Miller, 2000; Norcross & Lambert, 2006; Wampold, 2001; Westin, Novotny, &

Thompson-Brenner, 2004). However, these studies leave unanswered the question of how much the role the therapeutic relationship plays relative to the treatment model. This question has been addressed in meta-analytic reviews of psychotherapy, however.

A meta-analytic review of the systematic studies of the role played by the therapeutic alliance or relationship on outcomes reveals that the magnitude of this relationship between treatment alliance and outcome is relatively and surprisingly small. For methodological reasons, ESs in this literature are usually expressed as correlations, but we have translated them to *d* statistics in the following, using a table of the normal curve. This translation is designed to ease the comparison of these findings with the preceding ones. As an aside, the fact that the expressions of ESs through correlations always result in larger numerical values than when these same ESs are expressed as *d,* may partially account for why people often believe that relationship accounts for a very substantial portion of the outcome variance. In the following paragraphs, as we report both correlations and *d,* the reader will observe that the result is not nearly as impressive when expressed as the latter value as it is when expressed as the former.

Stevens, Hynan, and Allen (2000) reported a mean ES of $r = .11$ ($d = .03$) to describe the impact of the therapeutic relationship on outcome. Somewhat larger correlations of these variables have been reported by others, however. Horvath and Symonds (1991), for example, reported ESs ranging from $r = .21$ (d = .11) to $r = .26$ ($d = .17$), while Martin, Garske, and Davis (2000) reported a mean r of .23 ($d = .13$). These ESs, while consistent, are small and are of less predictive power than frequently attributed to them. They suggest that the therapeutic relationship is important but that it accounts for less than 7% of the variation among outcomes. This hardly qualifies as support for the truth of the relationship assertion.

What Can We Conclude?

The analyses summarized here offer a good argument that *all three of the previous assertions are myths*—Evidence for them is weak. To take a more extreme view of the results, one might argue that all three propositions are false or so nearly so that it makes little difference. This conclusion reflects a strong preponderance of the evidence and is reflected in many replicated meta-analytic reviews that have addressed these questions, only some of which have been reported here.

We have postulated about some of the reasons that may account for or partially account for why both scientists and practitioners often hold to these unsupported beliefs in spite of the disconnect with available evidence. Their beliefs may be the product of selective understanding or misunderstanding of the research findings or may (in the case of Assertion 3) be a misinterpretation of the differential meanings of correlations and *d*. Regardless of the reason, however, the fact that these beliefs are held by scientists who are studying psychotherapy, in particular, may impede the advance of more striking scientific findings by binding these scientists to using and advocating on behalf of research methods that do not lead to the optimal understanding of psychotherapy effects.

I do not offer these latter conclusions lightly. But, I have been forced to them as I have inspected carefully the consistency of the large body of research available (remember, I did not start writing this article with these beliefs, I developed them as I wrote). This struggle to understand what scientific evidence can do to improve clinical practice has led me to adopt a series of articles of faith, or if you prefer, working postulates, which may help clarify the implications of the foregoing findings.

Five Articles of Faith Related to Psychotherapy Research

From the 40 years of experience that I have devoted to the task of seeking truth in clinical practice, as a scientist, as a practitioner, as a teacher, as a consultant, and as a psychotherapy patient, I have drawn five basic conclusions. I have framed these as ''Articles of Faith'' as I continue to encounter disagreement about them.

Evidence Based on Strong Belief and Good Intention, Alone, Only Works for a Short While. Throughout the 100+ year history of psychotherapy, the principle test of the efficacy of any intervention has been the personal belief of the therapist, usually based on ''experience'' bolstered by the theory of a charismatic authority (Cummings & O'Donohue, 2008). Cummings and O'Donohue argue that when the opinion leaders of our field replace the search for discovery, we

become little different than a religion and the result is a tower of Babel around what is true and what was really meant by the original Guru. The nexus of this conflict is in the tendency to attribute to leaders a degree of infallibility and to mark truth by the number of followers one has rather than by the demonstrated validity of the assumptions that one pronounces. The faith given to these pronouncements, in the absence of more substantive and replicable evidence of their truth, has proved to be of limited lasting value. Faith in the absence of evidence is folly. One thinks, for example, of the many survivors of Hurricanes Katrina, Rita, and Ike who had faith that their houses would survive the oncoming storm; that the levees in New Orleans would hold; that they would be rescued; and that FEMA would come to their aid. Now, three years after Hurricane Katrina, hundreds of people with such faith are still dead, the levees are still points of danger, and thousands have yet to be able to rebuild their houses or relocate permanently.

In psychotherapy, interventions whose evidence relies on a top-down model of transmitting ''truth'' have enjoyed widespread appeal. Programs in *rebirthing therapy, reprogramming therapy, past lives therapy, recovered memory therapy,* and many others that are based on a rationale provided by a charismatic and believable leader, more than on scientific evidence, have come under fire when incidents occur to reveal their ability to harm others (Beutler, 2000; Lilienfeld, Lynn, & Lohr, 2003). In some cases, such as rebirthing therapy, the result has been death; in others, such as reprogramming therapy, it has been the psychological destruction of lives and families (Beutler, 2000; Public Broadcasting System, 1995). These public examples of the harmful effects of psychotherapy are often so striking that they become headlines that almost invariably harm the reputation of psychotherapists and psychotherapy.

For example, when rebirthing therapy was revealed in a newspaper and subsequent television stories, to have produced the death of a 10-year-old girl, the effects were widespread and all of psychotherapy was harmed (Mercer, Sarner, & Rosa, 2003). Psychotherapists of all kinds were forced to share the shame.

In addition to reliance on the authority of others or the appeal of a theory, a more pervasive and potentially insidious (e.g., Beutler, 2000; Lillifeld, 2006) standard is widely used. Namely, one's ''personal experience.'' Good intentions and personal beliefs based on one's idiosyncratic experience is by far the standard most used in determining the effectiveness among clinicians of all ilks. I heard these sentiments expressed recently by a member of the National Alliance for the Mentally Ill, whose own expertise was attributed to 10 years of experience with her own bipolar son. This individual, who is now the Head of a very influential residential treatment program for those with Bipolar Disorders, said, ''A degree doesn't mean anything about what a person knows. What counts is experience and I've been at this for 10 years.'' This anti-intellectualism and antiscientific comment is not too far removed from those of professional clinicians of any background who argue that their experience is more accurate and valuable in their treatment than anything that might arise from scientific or academic study.

Of course, there are some corrections within the public view that take place when the errors of experience are so strikingly obvious that it does not take a research study to bring their practice to a close and to turn the light on all of psychotherapy. Reliance on personal experience, and especially on the experience of others, no matter how strong one's good will and beneficent intentions, constitute weak support in the face of such public opinion.

But, of course, if experience really led to greater accuracy and validity of knowledge, then it would follow that there would be a gradual homogenization of knowledge with the passage of time. Gradually, with experience, we would reject ineffective theories and settle only on those that worked for specific populations. The plethora of ineffective and harmful psychotherapies (Lillifeld, 2006) that are built largely on experience, good intentions, and well meaning would rapidly be consolidated into a few that closely approximated the truth. We would not have ''Rebirthing Therapy'' or ''Reprogramming Therapy,'' nor their associated harmfulness (Beutler, 2000).

The Common Socially Derived Alternatives to Science, as the Basis of Clinical ''Truth,'' also Provide Weak Protection Against Ineffective and Mal- Practice. Historically, there have been three publically and legally used criteria to determine when a treatment is effective: (a) a community

standard of common practice; (b) a case law standard of a respectable minority; and (c) a healthcare standard of cost effectiveness (Beutler, Clarkin, & Bongar, 2000). The first two of these socially derived and commonly accepted standards have traditionally been applied to determine when a procedure can be considered to constitute malpractice. The third is one that has long been used in healthcare settings to determine the clinical value and reimbursement of a treatment or treatment program.

The standard of common practice requires that one's approach to treatment is acknowledged and acceptable to other practitioners in one's community. Indeed, it requires that the procedure be shown to be in common or daily use. In other words, there must be evidence that the treatment is popular (Black, 1990; Klerman, 1990). This standard relies heavily on evidence that other, similarly trained and experienced individuals, practice in similar ways and with similar treatments.

The standard of a respectable minority arose from case law, specifically out of concern that the standard of common practice was insensitive to emerging but not yet popular treatments. This standard recognized that the healthcare fields do not always have a consensual view of what is effective. This standard requires a demonstration that a "significant" minority of practitioners share a belief that has been articulated in legal procedures and that has defined a standard of practice (Furrow, 1980; Klerman, 1990).

Unfortunately, case law (Hood v Phillips, 1976) has defined this latter standard in a way that virtually ensures that everyone who is not protected by the common practices rule can be protected from malpractice claims under the "respectable minority" protection. This case law defines a "Respectable Minority" as being as few as six individuals who share a favorable opinion of the treatment, and a single written articulation of how the treatment is done, as being adequate evidence of its value (Beutler et al., 2000).

The third standard was the original one used to test the efficacy of managed healthcare programs and in modified form continues to be heavily weighted in contemporary discussions of a treatment's value (Aaron, 1996). It pits the number of people served by the procedure against the cost of distributing these services as a measure of effects. Thus, a "good enough" treatment is one that is delivered to many but costs nothing.

None of the three public criteria for assessing efficacy provide for a means for assessing objective evidence of change. And herein lies their failing. All three of the legalistic criteria require an analysis of how the treatment is delivered, rather than one that measures how much change it has produced. And, the criteria for experience is always more experience. There are no blinds or protections against the influence of unfounded opinion, self-fulfilling prophecy, or self-serving, and because of this, none of the criteria provides a true protection against a treatment that is harmful but widely used and cheap.

While there have been emerging improvements in both court standards and in healthcare policy deliberations, all as a function of becoming more reliant on the findings of objective and systematic scientific research (e.g., Daubert v Merrell Dow pharmaceuticals, 1993), the draw and attraction of these unsubstantiated criteria remain strong. Indeed, the value of scientific standards is verified by the very evidence that changes have taken place in these criteria. The increasing reliance on scientific findings has brought a concomitant increase in the stability and replicability to both legal and healthcare arenas.

Randomized Clinical Trials are a Viable Scientific Option for Addressing some Treatment Questions. When the NIMH decided to support the Treatment of Depression Collaborative Program in the 1980s, it was an innovative and interesting idea. I recall a meeting of the Society for Psychotherapy Research (SPR) in 1986—a meeting that I organized at Wellsley College—at which the keynote speaker, Dr. Gerald Klerman, who was then head of NIMH, introduced the use of RCT designs as a viable way to study psychotherapy. He emphasized that we must come to view psychotherapy like we do aspirin. That is, each form of psychotherapy must have known ingredients, we must know what these ingredients are, they must be trainable and replicable across therapists, and they must be administered in a uniform and consistent way within a given study.

In the service of those objectives, RCT research methods were initiated and rose to prominence as the

required methodology for determining one's status as an EST. The strength of the RCT methodology was that it could hone in on the "specific" ingredients of a treatment. From the beginning, it was implied or overtly asserted that these "specific" effects are those things that are embodied in the model of change used by the therapist (Elkin, 1994). All other influences were and are considered to be incidental to the treatment, and these included variations in therapist delivery, aspects of the working relationship, therapist skill, and nondiagnostic patient factors. Indeed, great effort has and is exerted to ensure that the influence of all of these factors is either eliminated or is held at a constant level. It is this logic of the RCT that demands that therapists be trained to a criterion of performance. In an ideal RCT for psychotherapy, all therapists within a given treatment would be identical (therapist variance is error variance) and all therapies would operate within an equivalent working relationship.

Some recent scholars (e.g., Addis & Cardemil, 12 2006), when faced with concerns that manualized training may stifle the creativity and individuality of therapists, have argued that RCTs are not designed to restrict the flexibility and personal creativity of the therapist (p. 151). But, in fact, this assertion is diametrically opposed to the basic RCT paradigm. The strength of RCT designs is that they control or eliminate variations of all known variables that are not specific parts of the treatment, and they attempt to program or control all aspects that are an active part of treatment. Most of the psychotherapies studied as ESTs consider relationship to be an important common factor or moderator but not a specific ingredient of the targeted intervention. To that degree, the scientific study of these treatments would either attempt to control the role of relationship by training all therapists to the same level of relationship ability or eliminate the influence of this variable by reducing its contribution to zero.

Clinicians, I believe, have a greater appreciation for 13 the specific effects of the therapeutic relationship than would be accorded by the rationale of RCT designs. Clinicians probably rarely attempt to ensure that the same working relationship occurs with each patient or to apply the treatment in the same way for all. Thus, among clinicians, the assertions of Addis and Cardimil (2006), that flexibility and creativity is allowed in controlled trial designs, would be consistent with their own beliefs and clinical practice. But, such flexibility would diminish the strength of the RCT design and violate its fundamental assumptions, which are based on allowing only a planned and systematic variation to occur in the independent variable (treatment) while controlling all others by holding them at a constant level. If such flexibility were encouraged, the power of the RCT would be lost. It is hard to imagine why scientists would strive, as they do, to train therapists within each treatment studied, to achieve a high level of reliability—the higher the better. What is a high reliability level intended to demonstrate if not that all therapists are behaving in closely similar ways? An ideal reliability of 1.0 would indicate equivalence. It is to achieve this ideal that scientists endeavor to train and retrain therapists and to drop outliers. Particularly "creative" therapists are either trained to recapture the original reliability standard or are dropped from the analysis as being nonrepresentative.

From the perspective of the most brand name psychotherapies, relationship quality is more rightly considered to establish a common platform which makes the application of the real treatment possible, rather than being a fundamental and specific part of the "treatment effect." Thus, it is usually studied after the fact, a procedure that is necessitated in an implicitly acknowledgement that the training has failed in its effort to hold this variable at the same level across therapists and patients. No serious consideration is given to the possibility that different patterns or relationship characterize effective applications of different procedures.

There was great resistance to Klerman's proposal that the research paradigm be translated into a narrow RCT model, at that 1986 SPR meeting, largely because doing so would limit psychotherapy to an assessment of what the therapist did with particular diagnostic groupings of patients (Klerman, 1986). Such an approach would ignore the personal characteristics and interpersonal compatibility of the therapist and patient involved. It would ignore the importance of the therapeutic relationship.

Notwithstanding these limitations of RCT designs, we have learned some interesting things since the

advent of RCTs in psychotherapy and are indebted to the EST movement that this methodology spawned. I am not referring here to the importance and value of the treatments that constitute the many lists of ESTs now available (Beutler, 2004; Beutler et al., 2003a; Chambless & Ollendick, 2001). The evidence cited in the foregoing paragraphs has not demonstrated the value of these lists for optimizing or even increasing the effectiveness of psychotherapy. The head-to-head comparisons of different therapies suggest that most manual-driven therapies are equivalently effective and not substantially different from most rationally derived therapies.

More interesting, I believe, is the rather paradoxical evidence that while most manual-driven treatments earn equivalent results, RCT studies have shown us that some psychotherapies are ineffective and even harmful (Beutler, 2005; Lilienfeld, 2007; Lilienfeld et al., 2003; Singer & Lalich, 1996). It appears to be easier to identify a bad treatment than a very good one, the latter falling prey to the Do-do bird. A surprisingly large number of well-known treatments has been found, in RCT studies, to be ineffective or even harmful. These treatment include such treatments as Drug Abuse and Resistance Education, Recovered Memory Therapy that is often used to treat female victims of rape, Grief Counseling for Bereavement, Expressive-Experiential therapies, and the most widely used treatment for acute effects of mass trauma, Critical Incident Stress Debriefing (Lilienfeld, 2007). When compared to no-treatment or placebo treatments, meta-analytic reviews of some widely accepted interventions earn ESs that are negative. That is, some of the treatments in common use make people get worse, even w*hen therapist effects are reduced to as low a level as possible*. That is, the average outcome for one receiving these treatments is deterioration.

If the foregoing tells us anything, it is that when some forms of psychotherapy are found to be effective, it may be *in spite of* the treatment, not because of it. The patient, the therapist, or the way they are paired, may offset the negative effects of the treatment techniques themselves, to facilitate change. That is, the beneficial effects of the therapeutic process may arise because of the resilient aspects of patients, the therapeutic qualities of people—things that cannot be randomly assigned to treatments—or from interventions that cannot be randomly trained. This realization led me to a fourth article of faith.

Some Research Questions are not Effectively Addressed with RCT Designs and are Best Answered by Naturalistic and Quasi-Experimental Studies. While the comparative results of ESTs have largely failed to be terribly impressive, scholars who apply RCT studies to psychotherapy continue to assure us that only random assignment studies are of sufficient scientific note as to provide believable evidence of psychotherapeutic efficacy (Chambless & Hollon, 1998). Many EST scholars, however, are coming to recognize that such weak results may indicate that the real influences in psychotherapy include effects that are associated with variables that are nonrandomly distributed aspects of the therapist, the relationship, and the patient (Castonguay & Beutler, 2006; Duncan & Miller, 2006). Whether or not a variable is capable of being randomly assigned should not dictate whether studies of these variables are considered to be sufficiently important as to warrant the attention of scientists. Indeed, there are many constructs that are central to nonpsychology sciences that are not appropriately or possibly studied through random assignment. Nor should randomization be a major criterion that determines the worth or merit assigned to a variable. Rather, the nature of the variable should be looked upon as a clue that can lead us to select among the available methodological procedures, those that are appropriate and sensitive to the kinds of characteristics that are being studied. Neither the Big Bang nor the theory of trans-species evolution have been subjected to randomized controlled trials, but few doubt their importance. Nor for that matter have natural disasters, terrorist events, and star movement been excluded from scientific study because they could not be randomly assigned.

In like manner, therapist and patient personalities, interpersonal values, therapist and patient gender, social skills and attachment levels, and the like, are not always capable of being randomly assigned and yet are of sufficient worth as to be given scientific consideration as being part of the specific effects of psychotherapy. The influence of some of these nonrandomized variables have been subjected to meta-analyses, and the ESs can

be compared to those obtained from variables that have been subject to RCT studies. For example, a meta-analysis by Beutler et al. (2003a) revealed that relationship factors (d = .17) and the personal and professional characteristics of therapists (d = .30) account for meaningfully more of the outcome variance than that associated with the intervention model used (mean d < .00). All of these findings suggest that many extra-intervention contributors to psychotherapy are worthy of being included within our definition of *psychotherapy*. They make stronger and equally consistent contributions to treatment outcome than the more formal aspects of the interventions themselves. Such observations underline the importance of the fifth article of faith.

Changing the Definitions of ''Psychotherapy'' and of ''Research-Informed Practice'' that is used in Research is Required to Advance our Understanding of their Importance. The articles of faith as articulated in the foregoing have led my research colleagues and I to shift the research definition both of ''psychotherapy'' and of RIP. The narrow view held by most EST research paradigms is not only unworkable when studying characteristics, qualities, and variables that are not appropriately or even capable of being assigned to people randomly, but are inconsistent with the way that such variables are conceptualized in clinical work. Such variables as are embodied in the person of the therapist, or that are captured within the patient's response dispositions, and those that index a degree of fit between the selected therapy and the patient must be considered to be potentially active ingredients of psychotherapy, itself. These variables are, or should be, central to developing effective treatments in clinical practice and should be given equal attention within the context of psychotherapy research. They deserve study as *part of* and *central* to and specific aspects of the psychotherapeutic process, not just as interesting but incidental correlates of what is considered a psychotherapy composed of disembodied procedures.

Accordingly, our research group has redefined psychotherapy for research purposes in order to be more consistent with the definitions operationalized within clinical practice. We have come to believe that separating the person of the therapist from the acts of psychotherapy—in the manner suggested by the medication metaphor proposed by Klerman—is unsupportable in psychotherapy research. If, as we have proposed, these aspects of character, preference, fit, and expectation, contribute more consistent and stronger predictive power in outcome assessments than the technical aspects of the interventions, then they are the treatment.

Specifically, we define psychotherapy, both in clinical and research applications, as: *The therapeutic management, control, and adaptation of patient factors, therapists' factors, relationship factors, and techniques factors that are associated with benefit and helpful change.*

Shifting the definitions from constructs derived from theories of psychopathology and psychotherapy, to the integration of patient, therapist, intervention, and relationship components, has led to a marriage among treatment methods (Nathan & Gorman, 2002); participant predictors, and empirically supported relationships (Norcross, 2002). Concomitantly, the change in the way that psychotherapy and research evidence is defined shifts us from relying on a narrow range of methods by which to extract ''truth'' and a similarly narrow range of models and patients (e.g., RCTs) to the investigation of one or more research-informed principles of effective therapy (Beutler et al., 2000; Castonguay & Beutler, 2006). It will also move the field to analyze the role of dimensional dynamics and interactions among therapist activities, patients and problem traits (severity, personality, etc.) and therapists, rather than maintaining the static and categorical view of the process that currently dominates the field.

To illustrate the differences that would be invoked by broadening these definitions of psychotherapy and RIP, the following section of this article will briefly summarize four studies, which I hope will illustrate how a broad range of controlled, quasi-experimental, and naturalistic scientific methods can be applied as a coordinated program of research whose results converge on important findings. While some of the results are interesting, I am presenting these studies for a broader purpose. Namely, they illustrate the interdigitation of using multiple scientific methods for the purpose of analyzing interactions among many complex variables. These are studies that my colleagues, students, and I have conducted. I offer them here, rather than examples from other investigators, for two

reasons: (a) I am intimately familiar with each of the studies, and (b) I have found no examples of others who have systematically utilized such an array of controlled and naturalistic methods in a converging sequence.

The first study to be reported was a prospective, quasi-RCT study that used archival data to identify and then cross-validate principles of strategic change and methods of measuring important variables (Beutler et al., 2000). The second study (Beutler et al., 2003b) illustrates the combination of an RCT and a regression analysis which illustrates some of the strengths of these methods when used together to focus on the interactions among therapist, intervention, relationship, and fit of the treatment and patient. The third study (B.E. Johannsen & L.E. Beutler, unpublished data) was a cross-cultural study that sought to validate two basic principles of change that had been identified in the first two studies. It utilized a quasi-experimental design to inspect the fit of therapy and patient factors. And, the fourth study (S. Kimpara, L. Henderson, & L.E. Beutler, unpublished data) was a naturalistic cross-validation of the clinical applicability of two treatment principles that had been identified in the earlier investigations.

Together, these studies provide a reasonable ES estimate of the gains in ESs that are associated with integrating multiple variable domains and a broadened definition of psychotherapy. These studies illustrate the application of multiple design elements to get a more comprehensive picture of optimal therapy than that which is possible using an RCT methodology, alone.

Study 1: Beutler et al. (2000) undertook a three-stage, quasi-RCT study of variables that predict and determine the effectiveness of psychotherapy. The study began with an exhaustive review of over 2,000 outcome research studies in order to define and then validate the role of patient and therapist characteristics, treatment dimensions, the fit of treatment to patient, and the therapeutic relationships that are associated with outcomes. This review also provided the data by which to extract from extant research findings 15 hypothetical principles that describe the relationship among these variables and outcomes.

The first phase of the study identified patient and treatment qualities that had been associated directly with outcomes as well as those that constituted well-matched dyads of patient and treatment. This phase also resulted in an articulation of clinically friendly principles that predicted how outcomes would emerge as these variables interacted with one another. In the second phase of the study, instruments were developed to measure the variables that had been the bases for these strategic hypotheses. These instruments were designed to tap patient factors, qualities of the therapeutic relationship, and the dimensions that constituted a good treatment fit. Patient qualities were tapped through independent clinical ratings; aspects of the interventions that were associated with good outcomes were assessed through ratings by experienced and trained clinicians; and aspects of good fit between patient and treatment were measured by combining the two sets of ratings representing patient and treatment characteristics.

A third phase of the study provided a direct test of the strategic hypotheses utilizing an archival data set of 289 subjects. These participants represented depressive spectrum and chemical abuse disorders and were drawn from four different RCT samples along with one naturalistic treatment sample. Seven different manualized models of psychotherapy, a manualized medication treatment, and a TAU condition represented the treatments utilized. All patients had been randomly assigned to one of a subset of the treatments, and within each data set, therapists were trained to criteria using one of the targeted therapy manuals and randomly assigned to an intervention. Patient entry data for this study were derived from the measures developed in the second phase of the study and were completed after trained (PhD) clinicians had listened to intake recordings and reviewed the preassignment personality and symptom measures taken at intake. The therapy procedures used, the quality of the working relationship, and the fit of the patient and treatment were extracted from ratings of early and late psychotherapy sessions using the therapy process measures developed in phase two. Outcomes were assessed by standard measures of psychological well-being (depression and anxiety) taken at pretherapy and posttherapy. Ratings of therapy activities and the fit of the therapy to patient characteristics were all applied to patients in the nine different treatments by trained and masked raters to ensure that all ratings were independent and uniform.

The data were analyzed by a series of structural equation models with posthoc analyses of specific relationships. The results provided support for 13 of the 15 original hypotheses. The 13 supported hypotheses were re-framed to assume the form of strategic principles that could be used to inform and guide the therapist and to provide assistance in developing a strategic plan for treatment. Five additional principles were derived from a consensual analysis of clinician ratings for dealing with dangerous patients and added to the total. Thus, 18 guiding principles were extracted from the findings, variously representing suggestions about developing a therapeutic relationship, assigning a context of therapy (treatment intensity, location, mode, and format), implementing common classes of interventions (directive and insight interventions and emotional regulation procedures), and adapting the intervention to accommodate moderating aspects of the patient's personality.

Specifically the findings determined that some aspects of treatment were directly related to patient outcomes and served as direct prognostic indicators. Likewise, some patient variables served as moderators of outcome and were found to indicate the ''fit'' of treatment and patient. For example, patient factors such as functional impairment, coping styles, levels of trait-like resistance to change, and level of distress were found to moderate corresponding treatment qualities (e.g., treatment intensity, insight-behavioral focus, therapist directiveness, and use of emotional confrontation) that had been identified in the literature review.

Of greatest importance, this study exemplified the application of findings from extant research to the extraction of new data from extant archival data sets in order to test predictors and fit of treatment and patient factors. It combined RCT and naturalistic designs and in that process gave some hints about the multiple and interactive qualities that affect outcome. By looking, at once, at therapy models, therapy procedures, patient characteristics, therapy context, and relationship factors, patterns among these variables emerged.

Study 2: Beutler et al. (2003b) studied 40 co-morbid depressed and chemically abusing patients using an RCT design. Patients were randomly assigned to one of three therapy models, including a prescriptive therapy (Beutler & Harwood, 2000) that was based on 10 of the 18 principles derived from Study 1. The treatments included a standardized cognitive-behavioral intervention (Wright, Beck, Newman, & Liese, 1993) and a narrative intervention (Moreira, Beutler, & Goncalves, 2003) in addition to the prescriptive procedure. Patients, treatments, and therapists (within treatments) were randomly assigned to one another, and patient-therapy matching dimensions derived from the earlier study were constructed from four dimensions of patient and treatment. Analyses were undertaken in two stages: (a) analysis of treatment model and (b) an analysis of patient, therapy, relationship, and treatment compatibility. The former analysis was based on analysis of variance procedures and the latter used linear regression models. While the three therapy models studied were relatively equivalent in efficacy, the patient, treatment, relationship, and treatment fit variables contributed independent variance to the benefits obtained. It was the fit of the treatment to the patient that accounted for the greatest degree of long-term change while treatment techniques predicted end of treatment status but tended to lose their effects in a relatively short period of time.

This study demonstrates the value of studying the complex interactions among factors from different domains. When considered only as therapist behaviors, all three therapy models produced similar effects. However, when therapy/therapist factors (d = .20), patient factors (d = .40), relationship factors (d = .40), and treatment fit factors (d = 1.40) were included within the definition of the treatments and analyzed as interaction and moderating variables, strong effects were observed, especially when the overall compatibility among patient qualities and the nature of the treatment were considered. The fit of the treatment to the particular patient accounted for the strongest effects on outcomes of all variable classes at one year after treatment.

Study 3: B.E. Johannsen & L.E. Beutler (unpublished data) applied a naturalistic design to a sample of 92 outpatients who were seen either in the United States or in Argentina. All patients were assigned to therapists and level of treatment fit with the therapist using random procedures. All outcomes were assessed using standardized premeasures and postmeasures. Patients were followed for three months or until they terminated treatment, whichever came first. Analyses of

two patient-therapist matching dimensions were conducted separately. The fit between patient coping style and therapist use of symptom-focused (among those with external coping styles) or insight (among those with internalizing coping styles) based interventions were strongly related to outcomes in both cultural groups. The better the match, the better the outcomes (d = .61).

The strategic fit between patient level of trait like resistance and therapist use of directive (for nonresistant patients) or nondirective (for resistant patients) procedures was related to outcomes among the Argentine patients, but not among the U.S. patients. The overall ES was d = .83. However, the U.S. sample failed to show a strong effect of treatment fit on this latter dimension. Following the suggestions uncovered in the earlier studies, the role of level of impairment was assessed as a further moderator of outcomes. Specifically, a relationship was only found among patients who were rated as being at least moderately distressed and impaired.

This study confirmed that patient variables serve as differential indicators for modifying aspects of psychotherapy to achieve a compatible ''fit'' among patients both in Argentine and U.S. samples. If a typical RCT analysis had been the sole procedure undertaken of these patients, with or without posthoc analyses of relationship factors, neither the effects of treatment procedures nor the moderating effects of patient factors would have been disclosed. By using a broad and inclusive definition of psychotherapy, a corresponding complex analysis was suggested and more detail was revealed about the optimization of psychotherapy effects.

Study 4: This is an example of a confirmation study in which treatment methods, treatment fit, and clinical utility were found in a recent naturalistic investigation of a homogeneous group of shy (avoidant and internalizing) individuals. In this study (S. Kimpara, L. Henderson, & L.E. Beutler, unpublished data), a structured treatment protocol had evolved and been tailored to work with shy and avoidant individuals based on clinical theory and experience. The treatment began with an eight week course of symptom focused, cognitive therapy and then followed by a 16–24 week course of psychodynamic psychotherapy.

The structure of the therapy used in this study happenstantially and independently corresponded with two predictive principles that had been extracted from the results of Study 1 and had been tested in Studies 2 and 3. Specifically, these studies suggested that an optimal treatment for internalizing (i.e., avoidant) individuals would consist of a symptom-oriented phase early in treatment followed by an insight-oriented phase of more uncertain duration. The availability of a treatment which followed this model provided for a natural occurring experiment and an opportunity to test these principles. Shy (internalizing and avoidant) patients were expected to benefit from treatment as a function of how closely the therapists followed the two-phase treatment.

A multiple regression and growth curve analysis of these data revealed that, as expected, treatment benefit was related to both compliance and the dominance of patient internalization tendencies (d = .76). The use of a natural experiment to confirm previously observed findings from more controlled designs confirmed the usefulness of employing multiple and flexible research methods for clarifying relationships among variables that constitute a broad and inclusive view of psychotherapy.

The convergence of findings is notable among these four studies, each of which utilized different research methods and designs, and all of which were based on different samples and employed different levels of treatment analysis. Collectively, the results confirm the validity of many of the principles originally defined in the review of extant research (Beutler et al., 2000). In each case, however, the direct analysis of the therapy(ies) was strikingly uninformative without considering the role of patient, intervention, and relationship factors. Only when the mix of these variables was incorporated within the definition of ''psychotherapy'' were we able to see a path to optimal treatment.

CONCLUSIONS

In this article, I have advanced the thesis that the way that psychotherapy is studied and defined in contemporary EST research is unnecessarily narrow and may, in fact, impede the search for optimal clinical effects. This weakness in contemporary research practices may have

led to over rate and over sell the extent of our knowledge about the effectiveness of psychotherapy.

By falling prey to an over reliance on a single methodology—randomized controlled trials—research results have failed to account for the degree of impact that is effected by psychotherapy. Scientists who consider any one method as a ''gold standard'' may inadvertently contribute to the oversimplification of psychotherapy findings by ignoring and devaluing the breadth of variables and factors that clinicians have, long ago, acknowledged to be as important to outcomes as the procedures that are used. By ignoring the role of nondiagnostic or extradiagnostic factors, relationship factors, and how treatments might fit with the patient's experience and problems, the RIP and EST movements may have ensured that the findings from research are weak and less than useful to clinical practice.

Drawing on a handful of studies that have focused on evaluating the level of fit between research-defined treatment and patient qualities, my students, colleagues, and I (Beutler et al., 2003a,b; B.E. Johannsen & L.E. Beutler, unpublished data; Satoko, Henderson, & Beutler, in preparation) have found uniformly moderate to large ESs (ranging from d = .30 to .91) to be associated with therapy variables. Such findings confirm that research investigations of psychotherapy would be well to maintain a flexible view of the therapeutic process; one that extends beyond what the therapist does to include when and how he or she does it.

In this study, I have identified five articles of faith that I believe support the conclusion that psychotherapy, in research as in practice, is a process that includes all variables within the network of systems that are and can be used to facilitate gains and benefits. In turn, research-informed practice, it is argued, must be more than RCT-informed practice and must draw findings from all relevant, reliable, and systematic scientific methods into a comprehensive understanding of treatment effects. Through examples of research that integrates multiple research and statistical methods into psychotherapy studies, I have attempted to lay the groundwork for a discourse among well-meaning scientists and practitioners about the nature of knowledge in this area. It is my hope that this discourse will expand and extend the role of science in clinical practice in ways that will advance both our knowledge and psychotherapeutic practice.

REFERENCES

Aaron, H. (1996). End of an era. *Brookings Review*, *14*(1), 35–37.

Addis, M. E., & Cardemil, E. V. (2007a). Psychotherapy manuals can improve outcomes. 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. 131–140). Washington, DC: American Psychological Association.

Addis, M. E., & Cardemil, E. V. (2007b). Dialogue: Convergence and contention. 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. 149–151). Washington, DC: American Psychological Association.

Berman, J. S., Miller, C., & Masserman, P. J. (1985). Cognitive therapy versus systematic desensitization: Is one treatment superior? *Psychological Bulletin*, *97*, 451–461.

Beutler, L. E. (2000). David and Goliath: When psychotherapy research meets health care delivery systems. *American Psychologist*, *55*, 997–1007.

Beutler, L. E. (2004). The empirically-validated treatments movement: A scientist-practitioner's perspective. *Clinical Psychology: Science and Practice*, *11*, 225–229.

Beutler, L. E., & Harwood, T. M. (2000). *Prescriptive therapy: A practical guide to systematic treatment selection*. New York: Oxford University Press.

Beutler, L. E., Williams, R. E., Wakefield, P. J., & Entwistle, S. R. (1995). Bridging scientist and practitioner perspectives in clinical psychology. *American Psychologist*, *50*, 984–994.

Beutler, L. E., Clarkin, J. F., & Bongar, B. (2000). *Guidelines for the systematic treatment of the depressed patient*. New York: Oxford University Press.

Beutler, L. E., Malik, M., Alimohamed, S., Harwood, T. M., Talebi, H., Noble, S., et al. (2003a). Therapist variables. In M. J. Lambert (Ed.), *Handbook of psychotherapy and behavior change* (5th ed., pp. 227–306). New York: John Wiley and Sons.

Beutler, L. E., Moleiro, C., Malik, M., Harwood, T. M., Romanelli, R., Gallagher-Thompson, D., et al. (2003b). A comparison of the Dodo, EST, and ATI indicators among co-morbid stimulant dependent, depressed patients. *Clinical Psychology & Psychotherapy*, *10*, 69–85.

Black, H. C. (1990). *Black's law dictionary* (6th ed.). St. Paul, MN: West.

Castonguay, L. G., & Beutler, L. E. (Eds.). (2006). *Principles of therapeutic change that work: Integrating relationship, treatment, client, and therapist factors*. New York: Oxford University Press.

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

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

Cummings, N. A., & O'Donohue, W. T. (2008). *Eleven blunders that cripple psychotherapy in America*. New York: Routledge Press.

Daubert v Merrell Dow pharmaceuticals. (1993). (92-102), 509, U. S. 579.

Duncan, B. L., & Miller, S. D. (2000). *The heroic client*. San Francisco, CA: Jossey-Bass.

Duncan, B. L., & Miller, S. D. (2006). Treatment manuals do not improve outcomes. 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. 140–149). Washington, DC: American Psychological Association.

Elkin, I. (1994). The NIMH treatment of depression collaborative research program: Where we began and where we are. In A. E. Bergin, & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 114–139). New York: John Wiley and Sons.

Emmelkamp, P. M., Bouman, T. K., & Blaauw, E. (1994). Individualized versus standardized therapy: A comparative evaluation with obsessive-compulsive patients. *Clinical Psychology and Psychotherapy*, *1*, 95–100.

Furrow, B. R. (1980). *Malpractice in psychotherapy*. Lexington, MA: Lexington Books.

Grissom, R. J. (1996). The magical number .7 ± .2: Meta-meta-analysis of the probability of superior outcome in comparisons involving therapy, placebo, and control. *Journal of Consulting and Clinical Psychology*, *64*, 973–982.

Hood v. Phillips. (1976). 537 s.w.2d 291 (Tex. Civ. App. 1976).

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.

Klerman, G. L. (1986). *Keynote address*. Delivered to the annual meeting of the Society for Psychotherapy Research, Wellesley, MS.

Klerman, G. L. (1990). The psychiatric patient's right to effective treatment: Implications of *Osheroff v.* Chestnut Lodge. *American Journal of Psychiatry*, *147*(4), 409–418.

Lilienfeld, S. O. (2007). Psychological treatments that cause harm. *Perspectives on Psychological Science*, *2*, 53–70.

Lilienfeld, S. O., Lynn, S. J., & Lohr, J. M. (2003). *Science and pseudoscience in clinical psychology*. New York: Guilford Press.

Luborsky, L., Singer, B., & Luborsky, L. (1975). Comparative studies of psychotherapies. *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.

Mercer, J., Sarner, L., & Rosa, L. (2003). *Attachment therapy on trial*. Westport, CT: Praeger Press.

Moreira, P., Beutler, L. E., & Goncalves, O. F. (2003). Narrative change in psychotherapy: Differences between good and bad outcome cases in cognitive, narrative, and prescriptive therapies. *Journal of Clinical Psychology*, *64*, 1181–1194.

Nathan, P. E., & Gorman, J. M. (Eds.). (2002). *A guide to treatments that work* (2nd ed.). New York: Oxford University Press.

Norcross, J. C. (Ed.). (2002). *Psychotherapy relationships that work: Therapist contributions and responsiveness to patient needs*. New York: Oxford University Press.

Norcross, J. C., & Lambert, M. J. (2006). The therapy relationship. 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. 208–218). Washington, DC: American Psychological Association.

Ollendick, T. H., & King, N. J. (2007). Empirically supported treatments typically produce outcomes superior to non-empirically supported treatments. 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. 308–317). Washington, DC: American Psychological Association.

Public Broadcasting System. (1995). *Frontline*. New York: Public Broadcasting Co.

Schulte, D., Kunzel, R., Pepping, G., & Schulte-Bahrenbert, T. (1992). Tailor-made versus standardized therapy of phobic patients. *Advanced Behavior Research and Therapy*, *14*, 67–92.

Shapiro, D. A., & Shapiro, D. (1982). Meta-analysis of comparative therapy outcome studies: A replication and refinement. *Psychological Bulletin*, *92*, 581–604.

Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, *32*, 752–776.

Smith, M. L., Glass, G. V., & Miller, T. I. (1980). *The benefits of psychotherapy*. Baltimore, MD: Johns Hopkins University Press.

Stevens, S. E., Hynan, M. T., & Allen, M. (2000). A meta-analysis of common factor and specific treatment effects across outcome domains of the phase model of psychotherapy. *Clinical Psychology: Science and Practice*, *7*, 273–290.

Thaler, M., & Lalich, J. (1996). *Crazy therapies*. San Francisco, CA: Jossey-Bass.

Wampold, B. E. (2001). *The great psychotherapy debate: Models, methods, and findings*. Hillsdale, NJ: Erlbaum.

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, Aall must have prizes@. *Psychological Bulletin*, *122*, 203–215.

Westin, D., Novotny, C. M., & Thompson-Brenner, H. (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin*, *130*, 631–663.

Wright, F. D., Beck, A. T., Newman, C. F., & Liese, B. S. (1993). Cognitive therapy of substance abuse: Theoretical rationale. In L. S. Onken, J. D. blaine, & J. J. Boren (Eds.), *Behavioral treatments for drug abuse and dependence* (NIDA Monograph No. 137). Washington, D.C.: National Institute of Drug Abuse.

Received October 8, 2008; revised November 24, 2008; accepted December 1, 2008.

Author Query Form

Journal: CPSP

Article: 1168

Dear Author,
During the copy-editing of your paper, the following queries arose. Please respond to these by marking up your proofs with the necessary changes/additions. Please write your answers on the query sheet if there is insufficient space on the page proofs. Please write clearly and follow the conventions shown on the attached corrections sheet. If returning the proof by fax do not write too close to the paper's edge. Please remember that illegible mark-ups may delay publication.

Many thanks for your assistance.

Query reference	Query	Remarks
1	**AUTHOR: A running head short title was not supplied; please check if this one is suitable and, if not, please supply a short title of up to 40 characters that can be used instead.**	
2	**AUTHOR: Shadish, Matt, Navaro, & Phillips, 2000 has not been included in the Reference List, please supply full publication details.**	
3	**AUTHOR: Lipsey and Wilson (1993) has not been included in the Reference List, please supply full publication details.**	
4	**AUTHOR: Schulte, Kunzel, Pepping, & Schulte-Bahrenberg, 1992 has been changed to Schulte, Kunzel, Pepping, and Schulte-Bahrenbert, 1992 so that this citation matches the Reference List. Please confirm that this is correct.**	
5	**AUTHOR: Addis & Cardermil, 2007 has been changed to Addis and Cardemil, 2007a,b so that this citation matches the Reference List. Please confirm that this is correct.**	
6	**AUTHOR: Ollendick & King, 2006 has been changed to Ollendick and King, 2007 so that this citation matches the Reference List. Please confirm that this is correct.**	
7	**AUTHOR: Beutler, Machado, & Neufeldt, 1994 has not been included in the Reference List, please supply full publication details.**	
8	**AUTHOR: Luborsky, Singer, & Luborsky, 1970 has been changed to Luborsky, Singer, and Luborsky, 1975 so that this citation matches the Reference List. Please confirm that this is correct.**	
9	**AUTHOR: Leichsenring & Rabung, 2008 has not been included in the Reference List, please supply full publication details.**	

10	AUTHOR: Westen, Novotny, & Thompson-Brenner, 2004 has been changed to Westin, Novotny, and Thompson-Brenner, 2004 so that this citation matches the Reference List. Please confirm that this is correct.
11	AUTHOR: Lillifeld, 2006 has not been included in the Reference List, please supply full publication details.
12	AUTHOR: Addis & Cardemil, 2006 has not been included in the Reference List, please supply full publication details.
13	AUTHOR: Please approve/amend inclusion of citation Klerman (1986).
14	AUTHOR: Beutler, 2005 has not been included in the Reference List, please supply full publication details.
15	AUTHOR: Singer & Lalich, 1996 has not been included in the Reference List, please supply full publication details.
16	AUTHOR: Lilienfield, 2007 has been changed to Lilienfeld, 2007 so that this citation matches the Reference List. Please confirm that this is correct.
17	AUTHOR: Wright, Newman, & Liese, 1993 has been changed to Wright, Beck, Newman, and Liese, 1993 so that this citation matches the Reference List. Please confirm that this is correct.
18	AUTHOR: Moreira, Beutler, Goncalves, 2008 has been changed to Moreira, Beutler, and Goncalves, 2003 so that this citation matches the Reference List. Please confirm that this is correct.
19	AUTHOR: happenstantially—Please check terminology.
20	AUTHOR: Satoko, Henderson, & Beutler, in preparation—Please update citation with publication details and include in the reference list or else change to unpublished data and provide author initials.
21	AUTHOR: Please check publisher name in Black (1990).
22	AUTHOR: Thaler & Lalich (1996) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List.
23	AUTHOR: Please check document title in Wampold et al. (1997).
24	AUTHOR: Please provide page range in Wright et al. (1993).
25	WILEY-BLACKWELL: Please check date of submission/revision/acceptance.

MARKED PROOF

Please correct and return this set

Please use the proof correction marks shown below for all alterations and corrections. If you wish to return your proof by fax you should ensure that all amendments are written clearly in dark ink and are made well within the page margins.

Instruction to printer	*Textual mark*	*Marginal mark*
Leave unchanged	··· under matter to remain	New matter followed by ⌐ or ⌐(x2)
Insert in text the matter indicated in the margin	⌐	
Delete	/ through single character, rule or underline or ⊢⊣ through all characters to be deleted	⌐/ or ⌐/(x2)
Substitute character or substitute part of one or more word(s)	/ through letter or ⊢⊣ through characters	new character / or new characters /
Change to italics	— under matter to be changed	⌐
Change to capitals	≡ under matter to be changed	≡
Change to small capitals	= under matter to be changed	=
Change to bold type	~ under matter to be changed	~
Change to bold italic	≈ under matter to be changed	≈
Change to lower case	Encircle matter to be changed	≠
Change italic to upright type	(As above)	⌐
Change bold to non-bold type	(As above)	⌐
Insert 'superior' character	/ through character or ⌐ where required	Y or Y under character e.g. Y^{3} or Y^{3}
Insert 'inferior' character	(As above)	⌐ over character e.g. Y_{2}
Insert full stop	(As above)	⊙
Insert comma	(As above)	,
Insert single quotation marks	(As above)	Y or Y and/or Y or Y
Insert double quotation marks	(As above)	Y or Y and/or Y or Y
Insert hyphen	(As above)	⊢
Start new paragraph	⌐	⌐
No new paragraph	⌐	⌐
Transpose	⌐	⌐
Close up	linking ⁀ characters	⁀
Insert or substitute space between characters or words	/ through character or ⌐ where required	Y
Reduce space between characters or words	| between characters or words affected	⊤