

PRACTICE-ORIENTED RESEARCH

Approaches and Applications

LOUIS CASTONGUAY, MICHAEL BARKHAM, WOLFGANG LUTZ,
AND ANDREW MCALEAVEY

There are many controversies in the field of psychotherapy. Numerous debates remain ongoing, for example, about what treatments are (or are not) effective for certain disorders and what variables are responsible for change. Although these debates are of great conceptual and clinical significance, they fade in comparison to the gravity of the schism that is at the core of clinical and counseling psychology. While these disciplines, as well as many training programs in other mental health professions, are based on the scientist-practitioner model, it is well documented that psychotherapists are not frequently and substantially influenced by empirical findings when they conduct their case formulations, treatment plan, and implementations (e.g., Cohen, Sargent, & Sechrest, 1986; Morrow-Bradley & Elliott, 1986).

There are a number of ways to explain the apparent indifference of clinicians toward psychotherapy research. To begin with, many scientific investigations are perceived as being limited in terms of their clinical relevance. The emphasis on internal validity, especially in traditional randomized controlled trials (RCTs), has sometimes come at a cost in terms of external validity. For instance, the focus on setting, a priori, the number of sessions and inclusion/exclusion criteria, among other constraints required for controlled research, may well reduce

error variance. However, the generalization of the findings to everyday practice is not always clear-cut (for further elaboration see Chapters 1, 3, and 14, this volume). It has also been argued that researchers pay limited attention to the concerns that therapists have when working with their clients (Beutler, Williams, Wakefield, & Entwistle, 1995). As described elsewhere (Castonguay, Boswell, et al., 2010), this could be viewed as a consequence or a reflection of “empirical imperialism” that has prevailed in many programs of research in which individuals who see very few clients per week decide what should be studied and how it should be investigated, to understand and facilitate the process of change.

The argument has also been made that clinicians would pay more attention to research findings if they were involved in research (e.g., Elliott & Morrow-Bradley, 1994). However, a number of obstacles can interfere with such involvement. Many therapists conducted research projects during graduate training that were unrelated to their clinical work. Similarly, not every clinician had the opportunity to work with an advisor who was conducting research while also treating psychotherapy clients of their own. Consequently, many clinicians lacked an early-career model based on conducting scientifically rigorous and clinically relevant studies that would then help them identify questions that could make a difference in their clinical work, or to help identify the most appropriate methods to investigate these questions. Full-time clinicians, even those

The authors are most grateful for the tremendous help provided by Soo Jeong Youn in preparing this chapter.

who were mentored by ideal scholars, are also confronted with pragmatic obstacles that can seriously interfere with an involvement in research, such as limited time, lack of resources, and difficulties in keeping up-to-date with methodological and statistical advances.

Needless to say, many have lamented over the gap between science and practice, and, over the six decades since the inception of the scientific-practitioner model (Raimy, 1950), several efforts have been made to foster and/or repair this concept (e.g., Soldz & McCulloch, 2000; Talley, Strupp, & Butler, 1994). The various avenues that are currently being promoted (and debated) to define evidence-based practice reflect a resurgence of the need to build stronger links between research and practice (e.g., Goodheart, Kazdin, & Sternberg, 2006; Norcross, Beutler, & Levant, 2006). Interestingly, it could also be argued that the current attention given to evidence-based practice has been triggered by the delineation and advocacy of empirically supported treatments (ESTs; Chambless & Ollendick, 2001). Although several scholars have warned that the promulgation of ESTs could deepen the schism between research and clinicians (e.g., Elliott, 1998), there seems to be no doubt that the EST movement has galvanized diverse efforts to foster the use of empirical information in the conduct of clinical tasks.

Directly related to the EST movement are the empirical investigations that have been conducted to test whether treatments shown to be effective under the stringent criteria of controlled trials also work when delivered in naturalistic settings. These effectiveness, as opposed to efficacy, studies are guided by the rationale that scientific advances will improve mental health care if it can be demonstrated that effective treatments (i.e., yielding large effect sizes) for specific and debilitating problems can be implemented and adopted in routine clinical care (Tai et al., 2010). A related effort has been the publication of important books and articles aimed at disseminating the research findings on ESTs, with the goal of offering a list of “treatments that work” (e.g., Nathan & Gorman, 2002). Complementing such publications are a large number of books describing how clinicians can apply specific ESTs. In fact, a number of these well known books are published versions of treatment manuals that have been used in clinical trials (e.g., Beck, Rush, Shaw, & Emery, 1979; Klerman, Weissman, Rounsaville, & Chevron, 1984). As argued elsewhere (Castonguay, Schut,

Constantino, & Halperin, 1999), such treatment manuals provide specific guidelines for interventions that can be extremely helpful to clinicians, as long as they are not imposed as the only form of therapy to be reimbursed. Nor that they are prescribed or used rigidly without being individualized to the needs of particular clients, and without consideration of other empirical data that can help foster process and outcome.

In response to the effort to bring science into practice via the validation and dissemination of specific treatments for particular disorders, came other initiatives emphasizing different variables and methodologies. These included the task forces on empirically supported therapeutic relationships (Norcross, 2011) and empirically based principles of change (Castonguay & Beutler, 2005a). The books that emerged from these task forces not only review the literature about variables related to the client and relationship, but also offer clinical guidelines derived from the empirical literature. In addition, noteworthy contributions (e.g., texts by Cooper [2008] and Lebow [2006]), have successfully taken on the challenge of presenting, without jargon, how research findings can be used in clinical practice.

Even though the efforts reported above focus on different variables and rely on different avenues of dissemination, they all share a top-down approach: that is, science is transmitted, and potentially adopted, via researchers informing therapists about the issues that have been studied and the lessons that can be derived from the findings. For example, in the United Kingdom some of these findings, derived from traditional RCTs and related meta-analytic studies, largely determine the national treatment guidelines to which practitioners and services are required to adhere. In this chapter, we refer to these efforts as manifestations of the paradigm of evidence-based practice. Although such efforts have and will continue to provide useful information to therapists, they nevertheless all reflect a more or less benign form of empirical imperialism.

One possible way to avoid or reduce empirical imperialism is for clinicians to be actively engaged in the design and/or implementation of research protocols. Such *practice-orientated research*, conducted not only for but also, at least in some way, by clinicians, reflects a bottom-up approach to building and using scientific knowledge. This approach is likely to create new pathways of connections between science and

practice, both in terms of process and outcome. By fostering a sense of shared ownership and mutual collaboration between researchers and clinicians (e.g., in deciding what data to collect and/or how to collect it), this actionable approach can build on complementary expertise, compensate for limitations of knowledge and experience, and thus foster new ways of conducting and investigating psychotherapy. By emerging directly from the context in which therapists are working, practice-oriented research is likely to be intrinsically relevant to their concerns and can optimally “confound” research and practice: that is, when the design of studies leads clinicians to perform activities that are simultaneously and intrinsically serving both clinical and scientific purposes.

The primary goal of this chapter is to describe three main approaches within the overarching paradigm of practice-oriented research: *patient-focused research*, *practice-based evidence*, and *practice research networks*. All three approaches share commonalities, the most notable being the collection of data within naturalistic settings. However, they also represent, in the order that they are presented in this chapter, a gradual variation on two crucial dimensions: first, in terms of the focus of research knowledge (from very specific to very broad), and second, in terms of active involvement of practitioners in the design, implementation, and dissemination of research. Although this chapter does not stand as a comprehensive review, it provides examples of psychotherapy studies that have been conducted within each of the three approaches highlighted and their application to practice. The chapter also briefly addresses some additional lines of inquiry that are aimed at fostering the link between research and practice.

We hasten to say that we do not view the strategies of accumulation and dissemination of empirical knowledge described in this chapter (i.e., practice-oriented research) as being superior to those typically associated with the evidence-based practice movement. Rather, we would argue for adopting a position of equipoise between these two complementary paradigms. Although traditional RCTs are often viewed as the gold standard within a hierarchy of evidence, this position has been challenged: “The notion that evidence can be reliably placed in hierarchies is illusory. Hierarchies place RCTs on an undeserved pedestal, for . . . although the technique has advantages it also has significant disadvantages”

(Rawlins, 2008).¹ And in relation to the potential of practice-based evidence, Kazdin (2008) has written that “[W]e are letting the knowledge from practice drip through the holes of a colander.” The *colander effect* is a salutary reminder of the richness of data that is potentially collectable but invariably lost every day from routine practice. A position of equipoise would advocate that neither paradigm alone—evidence-based practice or practice-oriented research—is able to yield a robust knowledge base for the psychological therapies. Furthermore, it is important to recognize that the methods typically associated with these approaches are not mutually exclusive. As we describe later, for example, RCTs have been designed and implemented within the context of practice research networks. Hence, rather than viewing these two approaches as dichotomous, a robust knowledge-base needs to be considered as a chiasmus that delivers *evidence-based practice and practice-oriented evidence* (Barkham & Margison, 2007).

PATIENT-FOCUSED RESEARCH

This section on patient-focused research has the goal of presenting one way of thinking about the scientist-practitioner gap from a scientist’s as well as a practitioner’s perspective. The main tool to achieve this goal is the careful study of patterns of patient change as well as tracking individual patients’ progress over the course of treatment and feeding back the actual treatment progress into clinical practice. Patient-focused research provides tools in order to support, but *not* replace, clinical decision-making with actual ongoing research data and specially developed decision support tools. The goal is, for example, to identify negative and positive developments early on in treatment and then to feed these back to therapists so they can combine science and practice immediately during the ongoing treatment. This is akin to physicians using lab test data and vital sign measures to manage physical ailments such as diabetes (see Lambert, 2010).

Importantly, the models discussed in this section are based on a generic approach to psychotherapy. Psychotherapies are viewed as a class of treatments defined by overlapping techniques, mechanisms, and proposed outcomes. Outcomes

¹Sir Michael Rawlins has been Chairman of the United Kingdom’s National Institute for Health and Clinical Excellence (NICE) since its inception in 1999.

are measured by summing items related to many disorders. Instead of identifying particular treatments for particular diagnoses as is the case in clinical trials, patient-focused research focuses more on the (real time) improvement of the actual treatment as implemented and the development of tools in order to achieve that task (Lutz, 2002). Overall, it supports a research perspective more focused on outcomes and the improvement of actual clinical practice based on empirical knowledge and less based on a debate about therapeutic schools (e.g., Goldfried, 1984; Grawe, 1997). Accordingly, the core of this approach requires research to be conducted on the course of patient change for individual clients/patients to learn about differences in patient change as well as subgroups of patients with specific patterns of change.

To date, the field of psychotherapy research has studied different types of psychopathology and accumulated a large amount of knowledge in terms of specific treatments for particular diagnostic subgroups (e.g., Barlow, 2007; Nathan & Gorman, 2002; Schulte, 1998). However, considerably less is known about different types of patient change. This situation is puzzling given that research has provided support for patient variability as a substantial source in explaining outcome variance, which Norcross and Lambert (2012) have estimated to be in the region of 30%. In contrast, treatment techniques have been reported as explaining only a small portion of the outcome variance (e.g., Lambert & Ogles, 2004; Wampold, 2001). Accordingly, careful examinations of how and when patients progress during treatment, or fail to do so, may both increase our understanding of psychotherapy and provide us with tools that could improve its effectiveness.

The following section is organized in three parts. First, a short introduction sets out the history of patient-focused research (dosage and phase models of therapeutic progress). Second, the main focus and themes of patient-focused research are described and discussed (rationally and empirically derived methods, nearest neighbors techniques, and new ways of detecting patterns of patient change and variability). And finally, the evidence-base for applying these methods to yield feedback to therapists is considered.

Dosage and Phase Models of Therapeutic Progress

The theoretical origins of patient-focused psychotherapy research, often described in the

literature as the “expected treatment response model,” are the dosage and phase models of psychotherapy. The dosage model of psychotherapeutic effectiveness established a positive, but negatively accelerating, relationship between the number of sessions (dose) and the probability of patient improvement (effect) such that increased number of sessions is associated with diminishing returns (Howard, Kopta, Krause, & Orlinsky, 1986). In subsequent work, Howard, Lueger, Maling, and Martinovich (1993) as well as Kadera, Lambert, and Andrews (1996) interpreted findings as representing rapid improvement early in treatment while in later phases increasing numbers of sessions were needed to reach a higher percentage of changed patients (see also Chapter 6, this volume). For instance, Howard et al. (1986), analyzing data on 2,431 patients from 15 studies, found that after 2 sessions 30% of patients had shown positive results. The percentages increased to 41% after 4 sessions, 53% after 8 sessions, and 75% after 26 sessions. In an extended analysis, Lambert, Hansen, and Finch (2001), using survival statistics and a more refined clinically significant change criteria, showed that these rates of improvement were overestimates of the speed of improvement and were dependent on patients’ pretreatment functioning. Their results showed that 50% of the patients who were in the dysfunctional range before treatment needed 21 sessions of treatment to reach the criteria for clinically significant change. However, for 70% of patients in the dysfunctional range to reach clinically significant change, more than 35 sessions were necessary. Further research has shown differential patient change rates by diagnosis and symptoms (Barkham et al., 1996; Kopta, Howard, Lowry, & Beutler, 1994; Maling, Gurtman, & Howard, 1995). In addition, Hansen, Lambert, and Forman (2002) reported that in clinical practice success rates are lower when treatment plans do not allow for enough sessions. Hence, a variety of factors will impact on the rate of change for each individual patient. An extension of this line of research can be seen in the good-enough level of change concept (e.g., Barkham et al., 2006; Stiles, Barkham, Connell, & Mellor-Clark, 2008; see later in this chapter).

The phase model further amplifies the dose-effect model by focusing on which specific dimensions of outcome are changing and in what temporal sequence (Howard et al., 1993). It proposes three sequential and progressive phases of the therapeutic recovery process and

assumes sequential improvement in the following areas of patient change: (1) *remoralization*, the enhancement of well-being; (2) *remediation*, the achievement of symptomatic relief; and (3) *rehabilitation*, the reduction of maladaptive behaviors, cognitions, and interpersonal problems that interfere with current life functioning (e.g., self-management, work, family, and partner relationships). In applying the dose-effect and phase models to therapeutic change, the decelerating curve of improvement can be related to the increasing difficulty of achieving treatment goals over the course of psychotherapy. Moreover, a causal relationship between changes in these dimensions was proposed with the phase model. That is, improvement in well-being is assumed to be necessary, but not sufficient, for a reduction of symptoms, which is assumed to be necessary for the subsequent enhancement in life functioning (cf. Stulz & Lutz, 2007).

In a replication study, Stulz and Lutz (2007) identified three patient subgroups on the basis of their development over the course of treatment in the dimensions of the phase model. In all of these subgroups, well-being increased most rapidly, followed by symptom reduction, while improvement in life-functioning was slowest. This finding supports the notion of differential change sensitivity for the three dimensions. Further, approximately two thirds of cases were consistent with the predicted temporal sequencing of phases (i.e., well-being to symptoms to functioning). However, a smaller but significant proportion of patients, approximately 30%, violated at least one of the two predicted sequences (e.g., moving directly from well-being to functioning). In addition, results suggested that the phase model seemed to be less powerful in describing treatment progress among more severely disturbed patients. A similar finding was also reported by Joyce, Ogrodniczuk, Piper, and McCallum (2002). In light of the earlier findings, further refinement focusing on differential change sequences between individuals is important.

Patient-Focused Research and Expected Treatment Response

The dosage and phase models define the process of recovery in psychotherapy for an average patient. However, patterns of improvement for individuals can vary significantly from the general trend (Krause, Howard, & Lutz, 1998). Thus, to accommodate this individuality, a model

could be helpful that estimates an expected course of recovery for individual patients based on their progress-relevant pretreatment characteristics. Indeed, this was the starting point for patient-focused psychotherapy research (Howard, Moras, Brill, Martinovich, & Lutz, 1996). Patient-focused research is concerned with the monitoring, prediction, and evaluation of individual treatment progress during the course of therapy by means of the repeated assessment of outcome variables, the evaluation of these outcome variables through decision rules, and the feedback of this information to therapists and patients (e.g., Lambert, Hansen, et al., 2001; Lutz, 2002). Such quality management efforts have been recognized not only as a promising method but as evidence-based practice that identifies patients at risk for treatment failure, supports adaptive treatment planning during the course of treatment, and, as a result, enhances the likelihood of positive treatment outcomes (Lambert, 2010; Shimokawa, Lambert, & Smart, 2010).

Patient-focused research asks how well a particular treatment works for the actual treated patient (i.e., whether the patient's condition is responding to the treatment he or she is currently engaged in). The evaluation of progress depends on the idiosyncratic presentation of the patient with respect to his or her expected treatment response. For example, minimal progress by Session 8 might be insufficient for many patients to consider their treatment as a success. However, for a highly symptomatic patient with comorbid levels of impairment (e.g., multiple symptoms as well as interpersonal problems) such moderate progress might be considered a success (Lutz, Stulz, & Köck, 2009). As a result, feedback systems to support clinical decision making in psychotherapy should include decision rules that are able to evaluate treatment progress based on the individual patient's status (Barkham, Hardy, & Mellor-Clark, 2010; Lambert, 2010; Lutz, 2002; see also Chapter 6, this volume).

Two distinct approaches to decision rules have been used to determine expected progress and to provide feedback (cf. Lambert, Whipple, et al., 2002; Lutz, Lambert, et al., 2006). One approach comprises rationally derived methods that are based on predefined judgments about progress using clinicians' ratings based on changes in mental health functioning over sessions of psychotherapy. The other approach comprises empirically derived methods that, in contrast, are based on statistically derived

expected treatment response (ETR) curves based on large available data sets that are respecified for each individual client.

Rationally Derived Methods

Rationally derived methods of patient-focused research use psychometric information based on standardized measures (e.g., the Brief Symptom Inventory [BSI]; Derogatis, 1993) to make an a priori definition about a patient's status and change. This then serves as a benchmark for his or her expected change and the evaluation of progress. A classic example of the rationally derived method can be seen in the concept of reliable and clinically significant change (Jacobson & Truax, 1991). The first component in this concept focuses on the actual amount of change achieved by the patient, which has to be greater than expected by measurement error of the instrument alone. The measurement error of an instrument depends on its reliability, hence the term *reliable change* (which comprises both *reliable improvement* and *reliable deterioration*). The second component, *clinically significant change* (or, more precisely, *clinically significant improvement*), occurs if a client who before treatment was more likely to belong to a patient sample is, at the final assessment, more likely to belong to a nonpatient sample (e.g., a community sample). Consequently, a patient has achieved *reliable and clinically significant improvement* if his or her score on the primary outcome measure meets both these criteria, indicating that the extent of improvement exceeds measurement error and the endpoint score is more likely to be drawn from a nonclinical population.

The following example of a rationally derived method used within a large feedback study is somewhat more complex. In a large-scale study funded by a German health insurance company comprising 1,708 patients within three regions of Germany, a rationally derived decision rule based on an extension of clinically significant change criteria was used (e.g., Lutz, Böhnke, & Köck, 2011). Feedback to the therapists was based on a patient's presentation at intake and on his or her amount of change by a certain session. This information was implemented into a graphical report, which was then fed back to clinicians who had the option to discuss these results via progress charts with patients.

To give feedback on initial patient status and patient progress to therapists at every assessment, all patients completed three instruments: the BSI,

the Inventory for Interpersonal Problems (IIP; Horowitz, Rosenberg, Baer, Ureño, & Villaseñor, 1988), and a disorder-specific instrument (e.g., a patient diagnosed with a depressive disorder would complete the Beck Depression Inventory; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). Patients were first classified into three categories by each instrument according to their initial impairment. For example, patients were categorized as initially "highly impaired" if their pretreatment score on that specific instrument was above the mean of an outpatient sample. Initially "moderately impaired" patients scored below the mean of that reference sample, but above the cutoff score of a nonpatient population for that instrument (e.g., Jacobson & Truax, 1991). Patients who scored below that cutoff score were categorized as "minimally impaired."

The feedback and evaluation of progress were based on the following decision rules. For the "minimally impaired" patients, each positive change resulted in a positive evaluation. For "moderately impaired" patients, change was considered positive only if improvement reached at least the predefined amount of the reliable change index (RCI) for that instrument. Finally, treatment change of "highly impaired" patients was viewed as positive only if patients fulfilled the criteria of reliable and clinically significant change. A negative reliable change was rated as deterioration independent of initial scores. The ratings for each of the three instruments were then integrated into a global score by summing them. Furthermore, the therapist could also be informed of the stability of treatment progress by reporting on progress over several administrations of the measures (for further details, see Lutz, Böhnke, & Köck, 2011; Lutz, Stulz, et al., 2009). The outcome findings of this study are briefly summarized in the subsequent section on feedback.

Empirically Derived Methods

Empirically derived methods define the expected treatment course based on previously treated patients with similar intake characteristics. These patient-specific databases are then used to determine the expected change for future patients. Furthermore, confidence or prediction intervals can be assigned around the predicted courses of improvement. Hence it is possible to provide an estimate of how much a patient's actual progress diverges from the expected course of change together with the probabilities of a successful outcome.

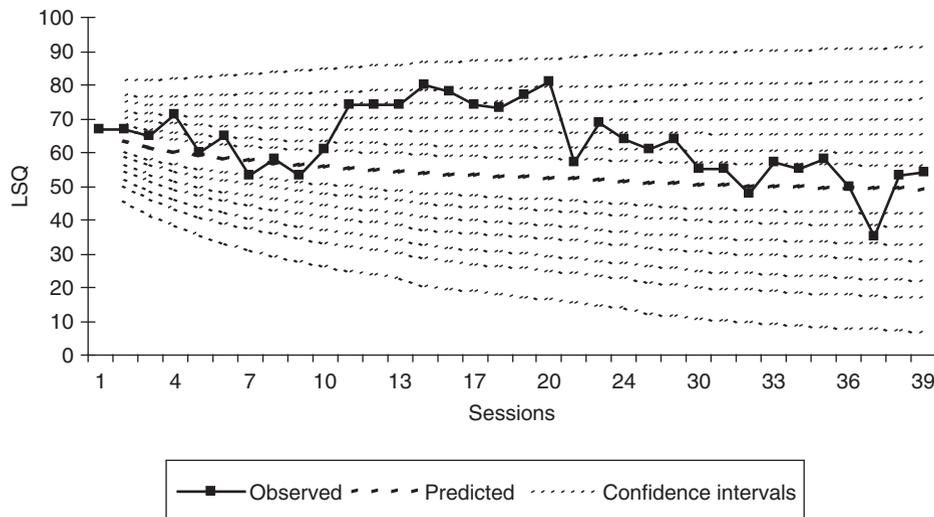


FIGURE 4.1 Predicted individual treatment response and confidence intervals (patient example) for the Outcome Questionnaire-30 (LSQ) and nearest neighbor predictive methods.

Source: Reprinted from “The Probability of Treatment Success, Failure, and Duration—What Can Be Learned From Empirical Data to Support Decision Making in Clinical Practice?” by W. Lutz, M. J. Lambert, S. C. Harmon, A. Tschitsaz, E. Schürch, and N. Stulz, 2006, in *Clinical Psychology and Psychotherapy*, 13, p. 227. Copyright 2006 by John Wiley & Sons, Inc.

In an application of empirically derived ETRs, Lutz, Martinovich, and Howard (1999) analyzed data from 890 psychotherapy outpatients and identified a set of seven intake variables that allowed prediction of individual change (e.g., initial impairment, chronicity, previous treatment, patient’s expectation of improvement). Figure 4.1 shows the ETR profile (predicted change based on intake variables) and the actual treatment progress of one selected patient with the Outcome Questionnaire-30 (OQ-30) as a dependent variable from an extended study with 4,365 patients (Lutz, Lambert, et al., 2006). To further explore the empirical decision system, different prediction intervals from 67% to 99.5% were considered around the predicted course of each patient. Using this schema, it was shown that the greater the number of actual scores a patient receives outside a confidence interval and the higher the interval, then the higher is the predictive validity of the actual score for the end of treatment.

In this way, actual treatment progress can be compared to the expected course of treatment and warning signals can be developed if a patient’s progress falls below a predefined failure boundary. As the number of observed values falling below this failure boundary increases, for example between Sessions 2 and 8, then the probability of

treatment failure increases. Also vice versa, as the number of observed values occurring above this failure boundary increases, then the probability of treatment success increases. Thus, the more and the further any extreme positive deviations are detected, then the higher is the probability for treatment success. Similarly, the more and the further any extreme negative deviations occur (e.g., early in treatment), then the higher the probability is for treatment failure (Lutz, Lambert, et al., 2006). These resulting percentages over the course of treatment can be employed as supporting tools by practitioners to adapt and potentially reevaluate their treatment strategy to enhance the patient’s actual outcome. For example, a deviation from the ETR profile in a specific session might result in a “warning” feedback signal to the therapists and supervisors or other clinicians involved in the case (e.g., Finch, Lambert, & Schaalje, 2001; Lambert, Whipple, et al., 2002; Lueger et al., 2001; Lutz, 2002). Different approaches to ETR models have been developed that provide information to understand individual patient progress and to assist in improving treatment strategies. For example, the application of ETR models has been extended to different diagnostic groups or symptom patterns as well as being applied to the study of therapist effects. The models have also been improved by

adding patient change information during the early course of treatment as predictors in order to have an adapted ETR model that is better able to predict patient change later in treatment (e.g., Lutz, Martinovich, Howard, & Leon, 2002; Lutz, Stulz, Smart, & Lambert, 2007). Two further extensions are presented here: One concerns how to identify subgroups of patients for developing ETRs, and the second concerns adjusting ETRs to different shapes or patterns of patient change.

Nearest Neighbors Techniques to Generate ETR Curves

To refine the prediction of ETR curves, Lutz et al. (2005) introduced an extended growth curve methodology that employs nearest neighbors (NN) techniques. This approach is based on research in areas other than psychotherapy in which large databases with many kinds of potentially relevant parameters (e.g., temperature and barometric pressure) recorded on a daily basis are used to make predictions of alpine avalanches (e.g., Brabec & Meister, 2001). This methodology was adapted by Lutz et al. (2005) in a sample of 203 psychotherapy outpatients seen in the United Kingdom to predict the individual course of psychotherapy based on the most similar previously treated patients (nearest neighbors). Similarity among patients was defined in terms of Euclidean distances between these variables. In a subsequent study, Lutz, Saunders, et al. (2006) tested the predictive validity and clinical utility of the approach in generating predictions for different treatment protocols (cognitive-behavioral therapy [CBT] versus an integrative CBT and interpersonal treatment [IPT] protocol). The NN method created clinically meaningful patient-specific predictions between the treatment protocols for 27% of the patients, even though no average significant difference between the two protocols was found. Using a sample of 4,365 outpatients in the United States, Lutz, Lambert, et al. (2006) further demonstrated the NN technique to be superior to a rationally derived decision rule with respect to the prediction of the probability of treatment success, failure, and treatment duration using the Outcome Questionnaire (OQ-45; e.g., Lambert, 2007).

In summary, these findings suggest that models of identifying similar patients could be an alternative approach to predicting individual treatment progress and to identifying patients at risk for treatment failure. It might be used in clinical settings either to evaluate the progress

of an individual patient in a given treatment protocol, or to determine what treatment protocol (e.g., CBT or IPT) or treatment setting (e.g., individual, family, or group) is most likely to result in a positive outcome based on similar already treated patients. Furthermore, if used in the context of a clinical team, the model could be used to identify therapists who are most effective in working with a particular group of already treated patients (nearest neighbors) who could then provide consultation on treatment plans or supervision for a trainee or novice therapist working with the new case.

New Ways of Detecting Patterns of Patient Change and Variability

The models discussed previously take into account differences in patient change but they are built on the assumption that there is one specific shape of change (e.g., log-linear) for all patients in the data set. Although this assumption makes sense in order to estimate a general trend over time, actual patient change may follow highly variable temporal courses and this variation might not just be due to measurement error, but rather be clinically meaningful (e.g., Barkham et al., 2006; Barkham, Stiles, & Shapiro, 1993; Krause et al., 1998). Growth mixture models (GMM) relax this single population assumption and allow for parameter differences across unobserved subgroups by implementing a categorical latent variable into a latent growth-modeling framework (e.g., Muthén, 2001, 2004). This technique assumes that individuals tend to cluster into distinct subgroups or patterns of patient change over time and allows the estimation of different growth curves for a set of subgroups. Such GMMs have been used to analyze psychotherapy data in naturalistic settings (Lutz et al., 2007; Stulz & Lutz, 2007; Stulz, Lutz, Leach, Lucock, & Barkham, 2007) and from randomized controlled trials (Lutz, Stulz, & Köck, 2009).

Figure 4.2 shows an application of a GMM in a sample of 192 patients drawn from the U.K. database mentioned earlier (Stulz et al., 2007). These patients completed the short-form versions of the Clinical Outcome in Routine Evaluation-Outcome Measure (CORE-SF; Cahill et al., 2006). In this example, shapes or patient clusters of early change (up to Session 6) have been identified to predict later outcome and treatment duration. Figure 4.2a displays the five different shapes of early change identified with the GMM. As can be seen, one cluster can be characterized

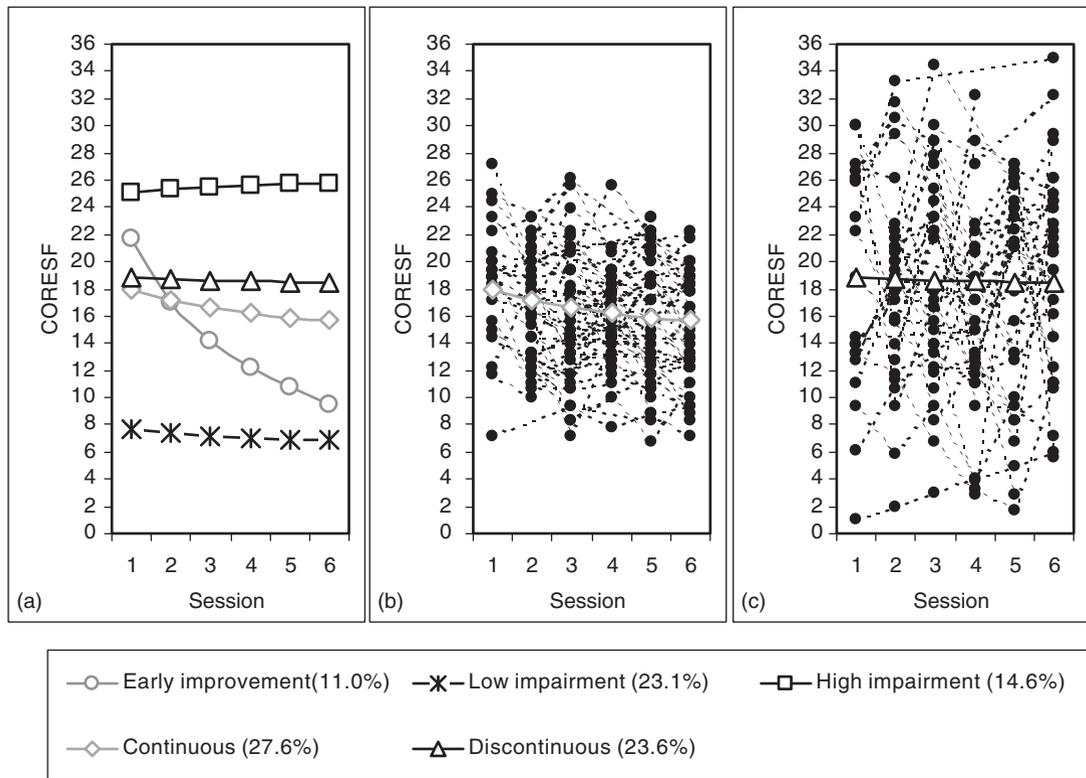


FIGURE 4.2 The five different shapes of early change (up to Session 6) identified in a sample of $N = 192$ psychotherapy outpatients using GMM (Figure 4.2a), and the observed individual growth curves for the continuous group (Figure 4.2b) and the discontinuous group (Figure 4.2c).

Source: Reprinted from “Methodological Background of Decision Rules and Feedback Tools for Outcomes Management in Psychotherapy,” by W. Lutz, N. Stulz, Z. Martinovich, S. Leon, and S. M. Saunders, 2009, *Psychotherapy Research*, 19(4), p. 507. Copyright 2009 Routledge.

by “early improvement” (11%). Patients in this cluster start with high scores on the CORE-SF but improve rapidly and substantially—more than 90% of those patients still show a substantial improvement at the end of treatment. A second cluster can be characterized by “high impairment” (23.1%) with little or no early patient change. The third cluster includes patients with “low impairment” (14.6%) who seem to have little or no early change until Session 6. The two remaining clusters in Figure 4.2a show two moderately impaired groups with similar average growth curves but, interestingly, very different individual treatment courses. Figures 4.2b and 4.2c display the plots of the actual individual treatment courses around the average growth curves in these two groups. Figure 4.2b presents the growth curves of the 27.6% of patients categorized into the “continuous” group who

showed modest session-to-session variation in the early phase of treatment. These can be contrasted to the individual growth curves displayed in Figure 4.2c. These patients (23.6%) were categorized into the “discontinuous” group as they demonstrated fairly substantial session-to-session variation. When using the reliable change criterion (Jacobson & Truax, 1991) to evaluate pretreatment to posttreatment change in these two groups, results revealed a higher rate of reliably improved patients in the “discontinuous” group than in the “continuous” group (44% versus 19%). Importantly, however, treatment duration was not different between the two groups ($M = 24.63$ versus $M = 25.55$ sessions, n.s.). Conversely, the rate of reliably deteriorated patients was also higher in the “discontinuous” patient group relative to the “continuous” group (13% versus 0%). The results from this study

suggest that instability during early treatment phases seems to result in higher chances for positive treatment outcomes but also higher risk for negative treatment outcomes.

To date, research on the advantages and disadvantages of rationally derived and empirical approaches has yielded mixed results. For example, Lambert, Whipple, et al. (2002) compared a rationally derived method to predict patient treatment failure with a statistical growth curve technique. The results showed broad equivalence between both methods but the empirical approach was somewhat more accurate. Other research also indicates that the empirically derived methods might be slightly superior (e.g., Lutz, Lambert, et al., 2006; Lutz, Saunders, et al., 2006). Irrespective of the selected approach (rationally derived or empirical), further research on differential patterns of change is necessary to clarify typical patterns for subgroups of patients as well as relating these empirical findings to clinical theories. Clinical theories that have a simple concept of treatment progress (i.e., a patient has a problem, a treatment approach is applied, the patient becomes healthy) appear oversimplistic and need to be adapted to take into account empirically defined change patterns. They could be further enhanced by considering related mediators and moderators causing different patterns of change that can then be used to guide or support clinical decisions (Kazdin, 2009).

Research on patterns of change is still in a preliminary phase. More studies are needed to further validate and replicate the findings obtained so far, and consideration needs to be given to the development of simpler methods. However, this research has the potential to provide therapists with decision guidelines that are individualized to each of their patients, especially early in treatment, as well as to identify and better understand the meaning of discontinuous treatment courses.

Provision of Feedback to Therapists and Patients

The above methods provide actuarial and predictive information on the course of treatment that has the potential to be used to enhance patient outcomes. At the practice level, the most apparent self-corrective function of routinely collected data derived from measurement systems is when it is used in the form of feedback to the practitioner, an area of research that has been espoused by the

APA Presidential Task Force on Evidence-Based Practice (2006). Indeed, despite the small differences in predictive accuracy, research on feedback appears to be a powerful tool for enhancing outcomes, especially for patients who are at risk of treatment failure (e.g., Carlier et al., 2012; Lambert, 2010; Lutz, Böhnke, & Köck, 2011; Newham & Page, 2010; Shimokawa et al., 2010). In this subsection, we consider the evidence base for using feedback routinely in clinical practice.

Recognizing Failing Patient Outcomes

The need for corrective feedback has been shown in comparisons between practitioners' and actuarial predictions of patient deterioration. Hannan et al. (2005) reported data from 48 therapists (26 trainees, 22 licensed) who were informed that the average rate of deterioration, defined as reliable deterioration on the OQ-45, was likely to be in the region of 8%. Given this base rate, the therapists were tasked with identifying, from a data set of 550 patients, how many would deteriorate by the end of treatment. Actual outcome data indicated that 40 clients (7.3%)—very close to the base rate—deteriorated by the end of therapy. Use of the actuarial predictive methods led to the identification of 36 of these 40 deteriorated cases. By contrast, the therapists predicted that a total of only 3 of the 550 clients would deteriorate, and only 1 of these had, in actuality, deteriorated at the end of therapy. Such data provides a powerful argument for investing in methods and procedures that enhance practitioners' treatment responses and planning in relation to patients who may be on course to fail in therapy.

Meta-Analyses and Reviews of Feedback

Carlier et al. (2012) carried out a review of 52 trials of feedback, 45 of which were based in mental health settings. The two largest subgroups of studies comprised those using global outcome measures ($N = 24$), of which 13 studies supported feedback, and depression measures ($N = 11$), of which 7 studies supported feedback. Overall, 29 of the 45 studies supported the superiority of providing feedback. Although providing a broad evidence base for feedback, this review lacked the precision afforded by a meta-analytic approach. A number of meta-analyses of outcomes feedback studies have been carried out (e.g., Knaup, Koesters, Schoefer, Becker, & Puschner, 2009; Lambert et al., 2003; Shimokawa et al., 2010). Lambert et al. (2003) completed a meta-analytic

review of three large-scale studies² in which the findings suggested that formally monitoring patient progress has a significant impact on clients who show a poor initial response to treatment. Implementation of a feedback system reduced client deterioration by between 4% and 8% and increased positive outcomes. Knaup et al. (2009) reviewed 12 studies³ and reported a small but significant positive short-term effect ($d = .10$; 95% CI .01 to .19). However, health gains were not sustained.

Lambert and Shimokawa (2011) carried out a meta-analysis of patient feedback systems relating to the Partners for Change Outcome Management System (PCOMS) and the OQ System. Three studies covering the PCOMS, drawn from two published reports (Anker, Duncan, & Sparks, 2009; Reese, Norsworthy, & Rowlands, 2009), indicated that the average client in the feedback group was better off than 68% of those in the treatment-as-usual group. Results indicated that patients in the feedback group had 3.5 times higher odds of experiencing reliable improvement while having half the odds of experiencing reliable deterioration. In terms of the OQ system, Lambert and Shimokawa (2011; see also Shimokawa et al., 2010) reanalyzed the combined data set ($N = 6,151$) from all six OQ feedback studies published to date.⁴ The three main comparisons were: no feedback, OQ-45 feedback, and OQ-45 plus clinical support tools (CST). Based on intent-to-treat analyses, the combined effects, using Hedges's g , of mean posttreatment OQ-45 scores for feedback only, patient/therapist feedback, and CST feedback were -0.28 , -0.36 , and -0.44 respectively. Shimokawa et al. (2010) concluded that all forms of feedback were effective in improving outcomes while reducing treatment failures (i.e., deterioration), with the exception of

patient/therapist feedback for reducing treatment failures.

Even though this area of research is still relatively recent, research on feedback in clinical practice is already an internationally studied area of investigation. A program of research in Australia on feedback, subsequent to the above review and meta-analyses, has considered the impact of providing feedback given at a specific time-point during the course of treatment and at follow-up (Byrne, Hooke, Newnham, & Page, 2012; Newnham, Hooke, & Page, 2010). Newnham et al. (2010) employed a historical cohort design to evaluate feedback for a total of 1,308 consecutive psychiatric and inpatients completing a 10-day CBT group. All patients (inpatients and day patients), whose diagnoses were primarily depressive and anxiety disorders, completed the World Health Organization's Wellbeing Index (WHO-5; Bech, Gudex, & Johansen, 1996) routinely during a 10-day cognitive-behavioral therapy group. The first cohort ($n = 461$) received treatment-as-usual. The second cohort ($n = 439$) completed monitoring measures without feedback, and for patients in the third cohort ($n = 408$), feedback on progress was provided to clinicians and patients midway through the treatment period. Feedback was effective in reducing depressive symptoms in patients at risk of poor outcome. In a 6-month follow-up study, Byrne et al. (2012) compared the no-feedback cohort with the feedback cohort. Feedback was associated with fewer readmissions over the 6-month period following completion of the therapy program for patients who, at the point of feedback, were on track to make clinically meaningful improvement by treatment termination. Importantly, the authors argued that the findings suggested feedback could result in cost saving in addition to being associated with improved outcomes following treatment completion for patients deemed on track during therapy.

Besides recently published meta-analyses or reviews (e.g., Carlier et al., 2012; Shimokawa et al., 2010), several advances have also been made to adapt feedback systems to different patient populations and settings. For example, Reese, Toland, Slone, and Norsworthy (2010) carried out a randomized trial comparing feedback with treatment-as-usual for couple psychotherapy ($N = 46$ couples) within a routine service setting. The setting was a training clinic and the therapists were practicum trainees. At the level of the individual client, rates of clinically significant change

²The three studies comprised Lambert et al. (2001); Lambert, Whipple, Bishop, et al. (2002); Whipple et al. (2003).

³The 12 studies comprised the following: Ashaye, Livingston, & Orrell (2003); Bauer (2004); Berking, Orth, & Lutz (2006); Brodey et al. (2005); Hawkins, Lambert, Vermeersch, Slade, & Tuttle (2004); Lambert et al. (2001); Lambert, Whipple, Vermeersch, et al. (2002); Marshall et al. (2004); Schmidt et al. (2006); Slade et al. (2006); Trudeau (2001); Whipple et al. (2003).

⁴The six studies comprised: Harmon et al. (2007); Hawkins et al. (2004); Lambert et al. (2001); Lambert, Whipple, Bishop, et al. (2002); Slade, Lambert, Harmon, Smart, & Bailey (2008); Whipple et al. (2003).

TABLE 4.1 Patients' Evaluations of the Quality Monitoring Project: Absolute Number and Percentage of Patients in the Respective Response Categories

Question	N*	Completely Right	Partially Right	Neither	Partially Wrong	Completely Wrong
		N (%)	N (%)	N (%)	N (%)	N (%)
I like the idea of having a project monitoring the quality of outpatient psychotherapy.	597	374 (62.6)	177 (29.6)	41 (6.9)	3 (0.5)	2 (0.3)
I find it important to monitor the results of psychotherapeutic treatments.	597	399 (66.8)	156 (26.1)	30 (5.0)	8 (1.3)	4 (0.7)
The time I needed to answer the questions was appropriate.	597	389 (66.8)	181 (30.3)	14 (2.3)	12 (2.0)	1 (0.2)
I had a hard time answering the questions because they affected me too much.	594	21 (3.5)	32 (5.4)	50 (8.4)	159 (26.8)	332 (55.9)
My therapist showed me the feedback and discussed it with me.	275	182 (66.2)	46 (16.7)	10 (3.6)	10 (3.6)	27 (9.8)
The information was very important to me.	258	85 (32.9)	105 (40.7)	35 (13.6)	13 (5.0)	20 (7.8)
The feedback was very helpful to me.	258	73 (28.3)	98 (38.0)	52 (20.2)	16 (6.2)	19 (7.4)
The feedback reflected my own assessment.	257	62 (24.1)	119 (46.3)	44 (17.1)	18 (7.0)	14 (5.4)
I felt like I could deal more responsibly with my psychological problems because of the feedback.	256	46 (18.0)	75 (29.3)	72 (28.1)	26 (10.2)	37 (14.5)

Note. *The questions in the lower part of the table were only given to patients in the feedback condition.

for the feedback versus nonfeedback groups were 48.1% and 26.3% respectively while for reliable change the rates were 16.7% and 5.3% respectively. This advantage to the feedback condition also held when the couple was used as the unit of analysis, with 29.6% of couples in the feedback condition meeting clinically significant change versus 10.5% (no feedback). Respective rates for reliable change only were 14.8% versus 5.3%.

Bickman, Kelley, Breda, de Andrade, and Riemer (2011) carried out a randomized trial of feedback for youths within naturalistic settings comprising 28 services across 10 states.⁵ Services were randomly assigned to a control condition comprising access to feedback every 90 days, or an experimental condition comprising weekly access to feedback. Because many of the youths in the 90-day condition ended treatment prior to their practitioners accessing the feedback, the authors considered this condition as a no-feedback control. Effect size (Cohen's *d*) advantages to the

feedback condition held regardless of the source of the outcome: .18 (youths), .24 (clinicians), and .27 (caregivers). The authors argued that although the effect sizes were small, they showed how outcomes could be improved without invoking new evidence-based treatment models. Feedback has also been evaluated in a nonrandomized study for substance-abuse patients (Crits-Cristoph et al., 2012). The design employed a two-phase implementation (Phase 1, weekly outcomes; Phase 2, feedback) with results showing advantages to the feedback phase. Crucially, however, these methods cannot be the sole basis for making clinical decisions—it can only be a support tool in aid of making clinical decisions, which always stays in the hands of the clinician.

In the feedback study carried out in Germany that was reported earlier, therapists received feedback for their patients several times during the course of treatment. Table 4.1 and Figure 4.3 show how they responded to the feedback provided. On approximately 70% of occasions, therapists made some use of the feedback either by taking some action or by drawing some consequence concerning their treatment formulation.

⁵The study originally comprised 49 services randomized, of which 10 control and 11 experimental services subsequently withdrew.

Due to the feedback, I...(multiple choices possible; 701 responses on $N = 394$ patients)

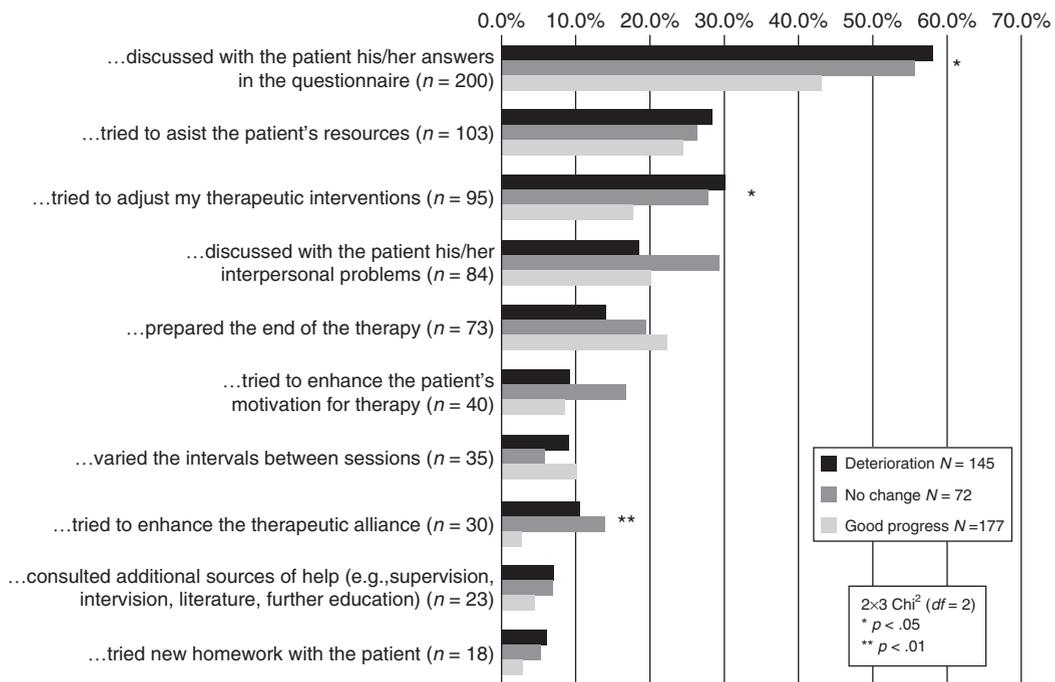


FIGURE 4.3 Consequences and modifications therapists did due to feedback related to patients' progress.

This is a high rate of action by therapists in response to the feedback information, especially given that most of the feedback was indeed positive feedback about the progress of patients. As can be seen in Figure 4.3, however, if patients showed negative progress early in treatment, then therapists, after receiving feedback, responded with a significant increase in the frequency of discussing the results with patients, adapting their treatment strategy, or trying to improve the therapeutic alliance (Lutz, Böhnke, Köck, & Bittermann, 2011). The positive evaluation from the patients participating in this study was also high, even when taking into account that not all of the patients responded. On almost all questions (see Table 4.1), the positive response rate exceeded 80%.

The above reviews, meta-analyses, and empirical reports provide an evidence-base that feedback to practitioners on patients shown not to be on-track enhances their outcomes across an increasing diversity of therapeutic modalities and patient populations. An early focus on university settings has broadened into wider samples of patient presentations. Research and development

foci have moved to considering the most effective clinical support tools to aid the practitioner's decision making in how best to respond to a patient who is not on track.

In addition, there have been calls for the development of a theory for feedback (see Bickman et al., 2011; Carlier et al., 2012), a call similar to those seeking a theoretical model for the impact of routine outcome measurement (see Greenhalgh, Long, & Flynn, 2005). From a viewpoint of differing paradigms, this area of work shows how both practice-oriented research and trials methodology can yield a robust evidence base for one area of clinical activity. Moreover, it shows how the former can provide a platform for more intensive trial work that might enable a more fine-grained investigation into the mechanisms and theory-building of how feedback achieves better patient outcomes for those patients deemed not to be on-track.

Summary

Patient-focused research has provided the field with new insights about how patients change,

with regard to the relationship between the amount of treatment received and outcome, as well as with respect to various patterns of progress, or lack thereof, experienced by different groups of clients. In addition, feedback on outcome progress has been found to be an effective tool to support treatment, especially for patients at risk of treatment failure (e.g., Carlier et al., 2012; Lambert, 2010; Lutz, Böhnke, & Köck, 2011; Newnham & Page, 2010; Shimokawa, et al., 2010). Furthermore, combining clinical support tools with such feedback has enhanced its effect (cf. Shimokawa et al., 2010). In this way research on outcomes feedback has two clinical implications: First, it allows therapists to track clinical progress on an individual level in order to determine, as early as possible, if a patient is moving in the right direction; and second, it has led to the delineation of decision support tools based on the variability and patterns in patient change. However it is important to emphasize, that clinically, outcome feedback can only serve as information to guide or support the decision-making process; the actual decisions of what goals or tasks to pursue, as well as when to continue, intensify, or terminate treatment remain to be made by the clinician and the patient.

PRACTICE-BASED EVIDENCE

This section focuses on a further form of practice oriented research, namely *practice-based evidence*, which is a reversal of the term *evidence-based practice*. Together, these two terms generate a chiasmus⁶—*evidence-based practice and practice-based evidence*—that has the potential for yielding a rigorous and robust knowledge base for the psychological therapies (Barkham & Margison, 2007). As the term suggests, practice-based evidence is rooted in routine practice and aims to reprivilege the role of the practitioner as a central focus and participant in research activity (for a detailed description, see Barkham, Stiles, Lambert, & Mellor-Clark, 2010). Although the approach shares much in common with patient-focused research, the hallmarks of repeated measurement and a primary focus on patients that underpin patient-focused research are not

sine qua non for practice-based evidence. Accordingly, practice-based evidence encompasses a broader, looser—less focused—collection of activities but takes its starting point as what practitioners do in everyday routine practice. At its heart, practice-based evidence is premised on the adoption and ownership of a *bona fide measurement system* and its implementation as standard procedure within routine practice. Implementation may be in the form of a pre- and posttherapy administration, repeated measurement intervals, or on a session-by-session basis. In terms of the yield of practice-based evidence, results can be considered at two broad levels: first, at the level of the individual practitioner whether working alone in private practice or within a community of practitioners in which the aim is to use data to *improve their practice*, and second, at a collective level in which the aim is to pool data such that it can contribute to and *enhance the evidence base* for the psychological therapies. With these two central aims, practice-based evidence delivers anew to the scientist-practitioner agenda.

This section provides illustrative examples of the yield of practice-based research by summarizing four key areas. First, a brief summary is provided of the development of selected measurement and monitoring systems, as representative of the field. Then illustrative findings are reported focusing on three successive levels of routine practice: the level of practitioners, then at the level of single services or providers, and finally, multiple services.

Measurement and Monitoring Systems

Although there are numerous features of practice-based evidence, the central component is the adoption and implementation of a measurement and monitoring system as part of routine practice. In contrast to stand-alone outcome measures, measurement and monitoring systems collect information on context and outcomes that are then used to improve practice and enhance the evidence base of the psychological therapies. The drive toward the adoption of measurement *systems* grew out of a developing trend for health insurance companies to seek evidence of outcomes and also from a growing frustration with the fragmented state regarding outcome measurement generally. The latter was illustrated in a review of 334 outcome studies from 21 major journals over a 5-year period (January 1983 to October

⁶Technically this is an *antimetabole* in which the same words are repeated in inverse order. All antimetaboles are, by definition, *chiasmic*, which is the generic term used to refer to this grammatical structure in which there is a criss-cross (Greek: Chi χ) of words.

1988) that showed 1,430 outcome measures were cited, of which 851 were used only once (Froyd, Lambert, & Froyd, 1996). In routine practice, decisions on the selection of outcome measures were determined by factors such as those used in trials, which were invariably proprietary measures carrying a financial cost, or determined by idiosyncratic, historical, or local influences.⁷ These factors combined to militate against building a cumulative body of evidence derived from routine practice settings that could complement the evidence derived from trials methodology.

Measurement Systems

Measurement systems began to be developed in the 1990s with the first outcomes management system being named COMPASS (Howard et al., 1996; Sperry, Brill, Howard, & Grissom, 1996). The COMPASS system—comprising evaluations of current state of well-being, symptoms, and life functioning—together with subsequent research reported by Lueger et al. (2001) provided the basis for other outcomes management systems that drew upon the phase model as a conceptual foundation. These included the Treatment Evaluation and Management (TEaM) instrument (Grissom, Lyons, & Lutz, 2002) and the Behavioral Health Questionnaire (Kopta & Lowry, 2002).

Subsequently other measurement systems have been developed. Examples of systems developed include:

- The Outcome Questionnaire-45 and associated measures (OQ-45; Lambert, Hansen, & Harmon, 2010; Lambert, Lunnen, Umphress, Hansen, & Burlingame, 1994): The OQ Psychotherapy Quality Management System has, at its heart, the OQ-45, which assesses three main components: symptoms, especially depression and anxiety; interpersonal problems; and social role functioning. For more information, see Lambert, Hansen, and Harmon (2010); also www.oqmeasures.com
- The Treatment Outcome Package (TOP; Kraus & Castonguay, 2010; Kraus, Seligman, & Jordan, 2005; Youn, Kraus, & Castonguay, 2012): The TOP comprises 58 items that assess 12 symptom and functioning domains: work functioning, sexual functioning, social conflict, depression, panic, psychosis, suicidal

ideation, violence, mania, sleep, substance abuse, and quality of life. In addition, the TOP measures demographics, health, stressful life events, treatment goals, and satisfaction with treatment. For further information, see www.OutcomeReferrals.com

- CelestHealth System for Mental Health and College Counseling Settings (CHS-MH; Kopta & Lowry, 2002): The Behavioral Health Measure (BHM) comprises four instruments that (a) assess complete behavioral health, (b) alert at the first session whether the client is at risk to do poorly in psychotherapy, and (c) evaluate the relationship between the therapist and the client. For more information, see www.celesthealth.com
- Partners for Change Outcome Management System (PCOMS; Miller, Duncan, Brown, Sparks, & Claud, 2003; Miller, Duncan, Sorrell, & Brown, 2005): The PCOMS comprises two 4-item scales: the Outcome Rating Scale (ORS; Miller et al., 2003) and the Session Rating Scale (SRS; Duncan & Miller, 2008). The ORS targets key components of mental health functioning while the SRS focuses on aspects of the therapeutic alliance. For more information, see www.heartandsoulofchange.com
- The Clinical Outcomes in Routine Evaluation system (CORE; Barkham, Mellor-Clark, et al., 2010; Mellor-Clark & Barkham, 2006; Evans et al., 2002): The CORE-OM (Barkham et al., 2001; Evans et al., 2002) is a pan-theoretical outcome measure comprising 34 items tapping the domains of subjective well-being, problems, functioning, and risk. It lies at the heart of the broader CORE System, which provides contextual information on the provision of the service received by the patient (Mellor-Clark & Barkham, 2006). A family of measures is available for differing uses and for specific populations and translations are available in 20 languages. For more information, see www.coreims.co.uk

Outcomes systems have also been developed for specific populations and treatment modalities. For example, the Contextualized Feedback Intervention Training (CFIT) has been developed for youths (Bickman, Riemer, Breda, & Kelley, 2006) and the Integrative Problem Centered Metaframeworks for family therapy (IPCM; Pinsoff, Breunlin, Russell, & Lebow, 2011).

Although each outcome system differs on any number of particular features, they reflect a

⁷For a summary of the status of outcome measures, see Whipple and Lambert (2011).

common aim, namely to measure and monitor patient outcomes *routinely* from which data is then used—fed back—to improve service delivery and patient outcomes. A resulting feature of practice-based evidence is, therefore, its ability to provide self-correcting information or evidence at the levels of practice and science within a short time frame. The following three subsections provide illustrative examples of the research yield at the levels listed earlier: practitioners, single services, and multiple services.

Practitioner Level: Effective Therapists and Therapist Effects

Reprivileging the therapist as a central focus of practice-based research redresses the balance in which the focus on treatments has long been dominant. Given that practitioners are the greatest resource—and cost—of any psychological delivery service, an equal investment in and prioritizing of practitioners is required to that already committed to the development and implementation of evidence-based treatments. The development of models of treatment based on the identification and observation of the practices of practitioners in the community who empirically obtain the most positive outcomes was a key recommendation of the APA Presidential Task Force on Evidence-Based Practice (2006). Research activity, especially trials, has predominantly used the patient rather than the practitioner as the primary unit upon which design features and analyses have been powered and premised. However, patients allocated to conditions within trials and observational studies are nested within therapists. This means that the outcomes of patients for any given therapist will be related to each other and likely different from those for patients seen by another (or other) therapist(s). Where a hierarchical structure is present but ignored in the analyses, assumptions about the independence of patient outcomes are violated, standard errors are inflated, *p*-values exaggerated, and the power of the trial reduced (e.g., Walwyn & Roberts, 2010).

A focus on what has come to be termed *therapist effects* developed following an article by Martindale (1978) and a subsequent meta-analysis in this area by Crits-Cristoph et al. (1991) as well as a critique of design issues (Crits-Christoph & Mintz, 1991). Wampold's (2001) text *The Great Psychotherapy Debate* followed, in which he concluded the impact of therapist effects as

being in the region of 8%. Subsequent reanalyses of therapist effects in the NIMH Treatment of Depression Collaborative Research Program (TDCRP; Elkin et al., 1989) by Elkin, Falconer, Martinovitch, and Mahoney (2006), and Kim, Wampold, and Bolt (2006) highlighted the problems of low power and of attempting to determine therapist effects from trials that were originally designed to assess treatment effects. Elkin et al.'s (2006) advice was clear, namely that therapist effects would be best investigated using (very) large samples drawn from managed care or practice networks—that is, routine settings. Subsequent reports on therapist effects and effective practitioners have been consistent with this advice (e.g., Brown, Lambert, Jones, & Minami, 2005; Kraus, Castonguay, Boswell, Nordberg, & Hayes, 2011; Lutz, Leon, Martinovich, Lyons, & Stiles, 2007; Okiishi, Lambert, Nielsen, & Ogles, 2003; Okiishi et al., 2006; Saxon & Barkham, 2012; Wampold & Brown, 2005).

A series of studies utilizing data from Pacific Care Behavioral Health, a managed behavioral health care organization, focused on various aspects of therapist effects and effectiveness (Brown & Jones, 2005; Brown et al., 2005; Wampold & Brown, 2005). Brown et al. (2005) evaluated the outcomes of 10,812 patients treated by 281 therapists between January 1999 and June 2004. Mean residual change scores, obtained by multiple regression, were used to adjust for differences in case mix among therapists. Raw change scores as well as mean residualized change scores were compared between the 71 psychotherapists (25%) identified as highly effective and the remaining 75% of the sample. During a cross-validation period—used as a more conservative estimate accounting for regression to the mean—the highly effective therapists achieved an average of 53.3% more change in raw change scores than the other therapists. Results could not be explained by case mix differences in diagnosis, age, sex, intake scores, prior outpatient treatment history, length of treatment, or therapist training/experience.

Wampold and Brown (2005) analyzed data comprising a sample of 581 therapists and 6,146 patients, the latter completing a 30-item version of the OQ-45. Multilevel modeling yielded a therapist effect of 5%, somewhat lower than the 8% the authors reported as an estimate from clinical trials. To explain this counterintuitive finding, they reasoned that the restricted severity range employed in trials, thereby leading to a

more homogeneous sample, yielded a smaller denominator when calculating the therapist effect.

The above studies focused on overall therapist effects. However, it might be that therapists are differentially effective depending on the specific focus of the clinical presentation. This question was addressed in an archive data set comprising services contracted with Behavioral Health Laboratories (BHL). Kraus et al. (2011) analyzed the outcomes of 6,960 patients seen by 696 therapists (i.e., 10 clients per therapist) in the context of naturalistic treatment in which the TOP was used. The specific aim was to investigate the effectiveness across the 12 domains in the TOP. With the exception of Mania, which had a low base rate, the reliability of the remaining 11 domains ranged from .87 to .94. Therapists were defined as effective, harmful, or neither based on categories of change using the criterion of the reliable change index (RCI) as follows: *effective* therapist if their average client reliably improved, *harmful* if their average client reliably deteriorated, and *ineffective/unclassifiable* if their average client neither improved nor worsened. Hence therapists were deemed effective or otherwise according to average change scores on each domain of the TOP, where a specific therapist could be classified as effective in treating depression (for example) and ineffective or harmful in treating substance abuse. In all, 96% of therapists were classified as effective in treating at least one TOP domain while classifications varied widely across 11 of the 12 domains (Mania was excluded due to low base rate.) Effective therapists displayed large positive treatment effects across domains (Cohen's $d = 1.00$ to 1.52). For example, in the domain of depression 67% of therapists were rated as effective (i.e., their average patient achieved reliable change in the domain of depression) with an average treatment effect size of 1.41. Harmful therapists demonstrated large, negative treatment effect sizes ($d = -0.91$ to -1.49). An important finding was that therapist domain-specific effectiveness correlated poorly across domains, suggesting that therapist competencies may be specific to domains or disorders rather than reflecting a core attribute or underlying therapeutic skill construct. This study highlights the distinction between seeking and analyzing competencies in specific domains of patient experience versus averaging the therapist effects by analyzing total scores across their patients. For a discussion of the advantages and

limitations of these two approaches in outcome monitoring, see McAleavey, Nordberg, Kraus, and Castonguay (2012).

The notion that some therapists are more effective than others caught attention with the use of the term *supershrink* (Ricks, 1974) in relation to a report on a very effective practitioner, with Bergin and Suinn (1975) labeling its opposite as *pseudoshrink*. Okiishi and colleagues addressed the concept of the exceptional therapist in consecutive studies (Okiishi et al., 2003, 2006). They utilized data from a large data pool in a university counseling center in which clients completed the OQ-45 on a regular basis. Both studies selected cases in which there were at least 3 data points. In addition, the initial study sampled practitioners who had seen a minimum of 15 clients each yielding a target sample of 56 therapists and 1,779 clients. This sample was extended in the second study and the criterion for the number of clients seen per therapist was increased to 30, yielding a target sample of 71 therapists and 6,499 clients (Okiishi et al., 2006). In this latter study, analyses focused on the average ranking of these 71 therapists based on their combined rankings according to their effectiveness (i.e., patient outcomes) and efficiency (i.e., number of sessions delivered). The authors examined and contrasted the top and bottom 10% of therapists (i.e., the ends of the distribution). The seven most effective therapists saw their clients for an average of 7.91 sessions, with clients making gains of 1.59 OQ points per session and resulting in a pre-posttherapy average change score on the OQ-45 of 13.46 ($SD = .76$). By contrast, the seven least effective therapists saw their clients for an average of 10.59 sessions, making gains of .48 OQ points per session and a pre-postaverage OQ-45 change score of 5.33 ($SD = 1.66$). Hence, these analyses suggested the most effective therapists achieved threefold gains for their patients compared with the least effective. Classifying clients seen by these most- and least-effective therapists according to the clinical significance of their change on pre-posttherapy scores (Jacobson & Truax, 1991; recovered, improved, no change, or deteriorated) showed therapists at the top end of the distribution had an average recovery rate of 22.4% with a further 21.5% improved while therapists at the bottom end of the distribution had a recovery rate of 10.61% with a further 17.37% improved. In addition, bottom-ranked therapists had a 10.56% deterioration rate while the equivalent percentage was 5.20% for top-ranked therapists.

The finding that some therapists achieve appreciably better outcomes than average highlights the naturally occurring variability in outcomes for therapists. For example, Saxon and Barkham (2012) investigated this phenomenon using a large U.K. data set in which clients completed the CORE-OM. Like Okiishi et al. (2006), the authors employed the recommendation of Soldz (2006) with each practitioner seeing a minimum of 30 patients. They investigated the size of therapist effects and considered how this therapist variability interacted with key case-mix variables, in particular, patient severity and risk. The study sample comprised 119 therapists and 10,786 patients. Multilevel modeling, including Markov chain Monte Carlo procedures, was used to derive estimates of the therapist effects and to analyze therapist variability.

The model yielded a therapist effect of 7.8% that reduced to 6.6% by the inclusion of therapist caseload variables. Effects for the latter rate varied between 1% and 10% as patients' scores reflecting their levels of subjective well-being, symptoms, and overall functioning became more severe. The authors concluded that a significant therapist effect existed, even when controlling for case mix, and that the effect increased as patient severity increased. Patient recovery rates, using Jacobson's criteria for reliable and clinically significant improvement, for individual therapists ranged from 23.5% to 95.6%. Overall, two-thirds of therapists ($n = 79$, 66.4%) could be termed as average in that the 95% confidence intervals surrounding their residual score crossed zero and could not, therefore, be considered different from the average therapist. The mean patient recovery rate for this group of therapists was 58.0%. For 21 (17.7%) therapists their outcomes were better than average with a mean patient recovery rate of 75.6%, while for 19 (16.0%) therapists their outcomes for patients were poorer than average with a mean recovery rate of 43.3%.

The studies by Okiishi et al. (2003, 2006) as well as by Saxon and Barkham (2012) highlight the considerable differences that exist in therapist effectiveness when comparing the two ends of the distribution of therapists. Although the majority of therapists cannot be differentiated from each other (i.e., are not significantly different from the average), differences between the extremes are real and meaningful for patients and, when considered in relation to the population of therapists as a whole, have significant implications for professional policy and practice. Variability

is a phenomenon that is inherent in all helping professions and it would seem important to understand the extent of this phenomenon in routine practice. Developing supportive ways of providing feedback at both the individual therapist and organizational level is an area that needs attention.

Single Service Level and Benchmarking in Routine Settings

A service or professional center providing psychological therapy will have, as a priority, a focus on its effectiveness, efficiency, quality, and cost, while patients, as consumers, will increasingly want to be assured they are seeking help from a professional agency (i.e., mental health center or service) that is effective. Current work has built on ideas dating back to the seminal work of, for example, Florence Nightingale (1820–1910)—who suggested a simple 3-point health-related outcome measure for her patients of relieved, unrelieved, and dead—and Ernest Codman (1869–1940), who implemented an “end results cards” system for collating the outcomes and errors on all patients in his hospital in Boston. However, while using a measurement system provides data on the actual service, practice data requires a comparator or standard against which to locate its own outcomes or other data. This requirement has led to the practice of benchmarking service data (for a summary, see Lueger & Barkham, 2010). Benchmarking can either involve comparisons with similar types of service (i.e., outcomes of other practice-based studies) or against the results of trials (i.e., assumed to be the gold standard). Persons, Burns, and Perloff (1988) provided an early example in which they compared cognitive therapy as delivered in a private practice setting with outcomes from two trials (Murphy, Simons, Wetzel, & Lustman, 1984; Rush, Beck, Kovacs, & Hollon, 1977). The authors concluded that their results were broadly consistent with those from the trials. Subsequently, although better described as effectiveness studies rather than practice-based, Wade, Treat, and Stuart (1998) as well as Merrill, Tolbert, and Wade (2003) provided examples of research that transported treatments into more routine settings and evaluated them using a benchmarking approach. Subsequent methods for determining benchmarks have been devised that enable comparisons with trials (see Minami, Serlin, Wampold, Kircher, & Brown, 2006; Minami, Wampold, Serlin, Kircher, & Brown, 2007).

Benchmarking as an approach has burgeoned across a range of service settings and patient populations whereby routine services and/clinics have been able to establish their relative effectiveness. Examples include the following: service comparisons year-on-year (e.g., Barkham et al., 2001; Gibbard & Hanley, 2008) and with national referential data (Evans, Connell, Barkham, Marshall, & Mellor-Clark, 2003), OCD in childhood (Farrell, Schlup, & Boschen, 2010) and in adults (Houghton, Saxon, Bradburn, Ricketts, & Hardy, 2010), group CBT (e.g., Oei & Boschen, 2009), psychodynamic-interpersonal therapy (e.g., Paley et al., 2008), CBT with adults (e.g., Gibbons et al., 2010; Westbrook & Kirk, 2005) and with adolescents (e.g., Weersing, Iyengar, Kolko, Birmaher, & Brent, 2006). These benchmarking studies, which are only a sample, all share a common aim of providing an evidence base regarding the effectiveness of interventions as delivered in routine services. But it is also possible to see specific themes by which studies can be grouped. These include, underrepresented approaches (e.g., non-CBT interventions), new or innovative interventions, and extensions to broader populations and/or settings.

In terms of *underrepresented* approaches, Gibbard and Hanley (2008), for example, reported a study employing data from a single service over a 5-year period using the CORE-OM, in which counselors delivered person centered therapy (PCT). In this study, a total of 1,152 clients were accepted into therapy and the data sample comprised 697 clients who completed CORE-OM forms at both pre- and posttherapy (i.e., 63% completion rate). Rates for reliable improvement⁸ calculated for each year separately ranged between 63.1% (second year) and 73.5% (third year) with an overall rate for the 5-year period of 67.7%. The authors concluded that PCT was an effective intervention in primary care. Moreover, based on a smaller subset of data ($n = 196$), they concluded that PCT was also effective for moderate to severe problems of longer duration.

Similarly, Paley et al. (2008) reported the outcomes of a single service delivering psychodynamic-interpersonal (PI) therapy. Full data was available for 62 of the 67 patients who were referred by either their general practitioner or psychiatrist to receive psychotherapy. Outcomes were obtained for the CORE-OM and

the BDI and were then benchmarked against data reported from other practice-based studies. The pre-posttherapy BDI effect size for the PI service was .76 compared with a benchmark of .73 derived from CBT delivered at the Oxford-based CBT clinic (Westbrook & Kirk, 2005). When only those clients who initially scored above clinical threshold were considered, the pre-posttherapy effect size for PI therapy was .87 compared with an effect size of 1.08 for the CBT routine clinic. Rates of reliably and significant improvement (Jacobson & Truax, 1991) were identical for both services at 34%, indicating broad equivalence in outcomes of the contrasting therapeutic approaches in routine settings. Both these studies illustrate the effectiveness of interventions in routine practice settings that are underrepresented when national bodies determine treatment interventions of first choice. Addressing this issue requires either the necessary funding to secure an evidence-base sufficient to satisfy national bodies (e.g., NICE) or, more radically, a reevaluation of how we define the nature of evidence.

Studies of *new interventions or applications* of evidence-based interventions within a novel package are exemplified by, for example, Richards and Suckling (2009) who reported on data from a single service following a U.K. government initiative to improve access to evidence-based psychological therapies (see Layard, 2006). This government initiative provided initial investment in new posts and training focused on low-intensity (i.e., self-help) and high-intensity (i.e., traditional) cognitive-behavioral therapy. One feature of its implementation was the requirement for patients at each session to complete the PHQ-9 (Kroenke, Spitzer, & Williams, 2001) and GAD-7 (Spitzer, Kroenke, Williams, & Löwe, 2006). This enabled outcome data to be reported on all patients (i.e., including those leaving treatment unilaterally). Of 2,795 patients assessed, 2,017 received more than one session, and change indices were reported for the 689 patients who completed treatment (i.e., 43% of those receiving more than 1 session; 24% of those assessed). Pre-posttreatment effect sizes, using the posttreatment SD as the denominator, were 1.38 for depression and 1.38 for anxiety (here recalculated as 1.46 and 1.52 respectively using the pre-treatment SD as the denominator).

The authors benchmarked these outcomes against a number of published data drawn from trials, practice-based studies, and reviews. Although they concluded that the outcomes fell within the range of comparator studies, the ranges

⁸The authors used reliable improvement rather than reliable and clinically significant improvement.

for the comparator benchmark effect sizes were sufficiently large as to lack precision: 0.80 to 1.46 for depression, and 0.73 to 2.1 for anxiety. However, the specific pre- to posttherapy effect sizes for depression and anxiety based on the PHQ-9 and GAD-7 respectively were very similar to data reported by Stiles, Barkham, Mellor-Clark, and Connell (2008a) in which a subset of counselors self-reporting to use CBT ($N = 1,045$ patients) for mixed diagnosis yielded an effect size of 1.34 (here adjusted to 1.30 when based on pretherapy SD). It would be important to note, however, that although the Richards and Suckling study contains many of the features of practice-based evidence as outlined earlier, it was reporting a government-funded initiative and was therefore well resourced, targeted CBT, and used mandated condition-specific outcome measures. Hence it better reflects a hybrid of both evidence-based practice and practice-based evidence.

In relation to extending interventions originally tested in trials into *new clinical populations or settings* in routine settings, there are numerous studies using a benchmarking approach. For example, Oei and Boschen (2009) adopted this approach in evaluating group-based CBT as delivered in a community setting. The authors compared the pre-posttreatment effect size of 0.64 obtained from the community setting with the pre-post effect sizes reported from other studies, including that of Westbrook and Kirk (2005), and considered the outcomes to be broadly similar. Other researchers have employed a similar strategy across a wide range of situations: for example, an outpatient setting for adolescents presenting with OCD (Farrell et al., 2010), CBT in a pain clinic (Morley, Williams, & Hussain, 2008), and cognitive-analytic therapy in routine practice (Marriott & Kellett, 2009). Common to them all has been the intention of extending the findings of efficacy trials into routine practice settings with the aim of building a broader evidence base.

Overall, results have shown interventions in routine practice to be effective but, in most cases, to fall short of the gains achieved in efficacy trials. One comparison between practice-based data and a selection of trials in depression showed the outcomes of trials to be approximately 12% superior to those of practice-based studies based on mean pre-posttreatment change (Barkham et al., 2008). A similar difference was found when methods of reliable and clinically significant improvement were applied. However, the differences were larger when comparing pre-posttherapy effect

sizes. This result was likely due to the tendency for trials to have restrictions on patient inclusions leading to reduced variance (and consequently higher effect sizes). Regardless of which metric is used, superior outcomes, on average, in trials compared with practice-based studies is what would be expected given the highly selective and more protective environment in which trials are implemented. However, while average outcomes from routine settings may, in general, fall short of efficacy studies, the likely greater variation in individual therapist outcomes in practice-based studies might mean that there are therapists in routine setting who achieve consistently better outcomes than the top ranked therapists in trials. The above results suggest the need for policy makers to consider a broader evidence base than provided by trials alone.

Although benchmarking appears an attractive and relatively low-cost strategy for securing a comparator for practice-based studies, it is not without problems. First, comparisons often do not take into account the differential dose received by patients in routine services compared with trials. For example, a review of the clinical trials literature showed between 57.6% and 67.2% of patients improved within an average of 12.7 sessions. By contrast, naturalistic data showed that the average number of sessions received in a national database of more than 6,000 patients was less than five sessions with the rate of improvement in this sample being approximately 20% (Hansen et al., 2002).

Second, an intrinsic difficulty with adopting a benchmarking strategy concerns the selection of the benchmark used, which invariably differs across studies, thereby making comparisons study-specific. For example, two studies cited earlier using benchmarking approaches for CBT for depression with adults used different benchmarks: one with a specific focus on depression (Gibbons et al., 2010) and the other employing a more generic sample (Westbrook & Kirk, 2005). Although both studies employed the BDI as one of their primary outcome measures, in their selection of a benchmarking study Westbrook and Kirk used data from Persons, Bostrom, and Bertagnolli (1999) and the NIMH TDCRP (Elkin et al., 1989) while Gibbons et al. (2010) used unpublished data derived from DeRubeis et al. (2005). In both studies there were specific reasons for the selection of the benchmarks used that made the direct comparisons credible, but at the expense of adopting a common comparator.

Other studies have used benchmarks from, for example, three or more comparator studies. A logical strategy is to use the most recent study or studies as benchmarks, thereby generating successively new (and different) comparators over time. To reduce the arbitrary decision of which benchmark to select, one strategy would be to benchmark any new study against the cumulative body of previous studies. Houghton and colleagues (2010) adopted this strategy and determined a benchmark from nine published studies on OCD. Adopting such an approach would mean that as new studies are published and added to the body of evidence, the benchmark should become increasingly robust (i.e., less vulnerable to the influence of any single study) and have ever-increasing tighter confidence intervals. Another approach would be extending the procedures applied to depression by Minami et al. (2007) to other clinical presentations. Either way, the aim would be to use a benchmark that better represents the collective body of trials evidence rather than any single trial alone.

Multiple Services: Effectiveness of Psychological Therapies in Routine Settings

In this subsection, we review evidence derived from the investigation of data from multiple services in which the aim is to move beyond statements that are specific to any single service toward considering data as contributing to the knowledge base about psychological therapies in general. Specifically, we focus on analyses of routine data sets that address issues of the overall effectiveness of therapies as well as how much therapy is considered enough.

Meta-Analyses of Practice-Based and Effectiveness Studies

In a meta-analysis of practice-based studies addressing common mental health problems, Cahill, Barkham, and Stiles (2010) reported the average pre-posttreatment effect size from 10 studies for the treatment of depressive symptomatology, using a fixed effect model, to be 1.29 (95% CI 1.26 to 1.33). However, using a fixed effect model restricts the extent to which the findings can be generalized in that the results pertain only to the sample of therapists used in the studies. Using a random effects model allows the results to be generalized to the population of patients as a whole and this analysis

yielded an overall pre-post effect size of 1.14 (95% CI 0.96 to 1.32). Applying Jacobson and Truax's (1991) criteria of clinical change to seven studies using the CORE-OM yielded a mean rate of 56% of patients meeting the threshold for reliable and clinically significant improvement, with an additional 18% achieving reliable improvement only, and 25% showing no reliable change.

Stewart and Chambless (2009) carried out a meta-analysis of 56 effectiveness studies—opposed to practice-based studies—of CBT across five adult anxiety disorders: panic disorder, social anxiety disorder, PTSD, generalized anxiety disorder, and obsessive-compulsive disorder. The authors rated studies on nine dimensions reflecting clinical representativeness in relation to: settings, referrals, therapists, structure, monitoring, no pretherapy training of therapists, no randomization, clinically representative patients, and allowance of medication. Studies were included if they scored 3 or more on a 9-point scale. Further, inspection of the supplementary data shows only 2 of the 56 studies received a maximum rating (i.e., 9) as indicating being clinically representative on all dimensions (<http://dx.doi.org/10.1037/a0016032.supp>). Indeed, the authors stated: “The real-world mental health practitioner may not agree that studies included in this meta-analysis are clinically representative” (p. 601). Pre-post effect sizes for each condition were compared to three benchmark studies (i.e., trials) targeting each specific presenting condition. Results indicated the pre-posttreatment effect sizes for the effectiveness studies in three of the conditions (social anxiety $ES = 1.04$; generalized anxiety disorder $ES = 0.92$; and obsessive-compulsive disorder $ES = 1.45$) to be within the range of the efficacy benchmarks, while findings for PTSD ($ES = 2.59$) exceeded those of trials. Only the results for panic disorder ($ES = 1.02$) were below the lower range of the benchmark studies.

Overall, results from both these meta-analyses indicate that interventions delivered in routine practice are effective, yielding mainly large pre-posttherapy effect sizes. Hence, with the exception of the finding that PTSD as delivered in effectiveness studies exceeded those of trials, the pre-posttreatment effect sizes all fell within the range 0.90 to 1.45; that is, the average patient at posttreatment was better than between 83% and 93% of people prior to treatment.

Practice-Based Studies of Comparative Treatment Outcomes

While the above meta-analytic studies have focused on the effects for clinically defined presentations, empirical studies have addressed the comparative effectiveness of differing schools of interventions (e.g., Stiles, Barkham, Twigg, Mellor-Clark, & Cooper, 2006; Stiles, Barkham, Mellor-Clark, et al., 2008a). Stiles et al. (2006) utilized data drawn from 58 services in the United Kingdom from which a subset of 1,309 patients were selected who, according to the therapists self-reported accounts, received either cognitive-behavioral therapy (CBT), person-centered therapy (PCT), or psychodynamic therapy (PDT). Patients completed the CORE-OM at the beginning and end of their treatment and therapists indicated which treatment approaches were used. Hence, results focused only on patients who completed treatment. Comparisons were made between six subgroups: three comprising CBT, PCT, or PDT only (i.e., pure), and three treated with one of these plus one additional approach used by the therapist (e.g., integrative, supportive, art). These latter three groups were designated as CBT + 1, PCT + 1 or PDT + 1 respectively. All six groups averaged marked improvement with an overall pre-posttherapy effect size of 1.36. Treatment approach and degree of purity (i.e., “pure” versus “+1”) each accounted for statistically significant but comparatively very small proportions of the variance in CORE-OM scores.

Stiles et al. (2008a) replicated this study drawing on an original data pool comprising 33,587 patients, which, after excluding incomplete data from either patients and/or therapists, yielded complete data on 12,162 patients. The study focused on the same six subgroups as in the previous study yielding a sample of 5,613 patients. All six groups began treatment with equivalent CORE-OM scores, and all averaged marked improvement with an overall pre-posttherapy effect size of 1.39. Figure 4.4 presents notched boxplots of the extent of change across the six groups. Distributions of change scores were all similar. Although the authors signaled caution because of limited treatment specification, non-random assignment, incomplete data, and other issues, the study was criticized on these same grounds (see Clark, Fairburn, & Wessley, 2008), criticisms that in turn were rebutted (see Stiles, Barkham, Mellor-Clark, & Connell, 2008b). The debate exemplifies the tensions surrounding

the use and interpretation of data from routine practices. However, Stiles and colleagues concluded that these routine treatments were effective for patients who complete them but that those who fail to complete, or indeed even begin, treatment deserve attention by researchers and policymakers.

The impact of using a sample of completer patients versus an intention-to-treat sample has a direct impact on the reporting of outcomes in routine practice. Using the CORE database comprising 33,587 patients, Barkham, Stiles, Connell, and Mellor-Clark (2012) calculated rates of reliable and clinically significant improvement, as defined by Jacobson and Truax (1991), based on completer and intention-to-treat samples drawn from this data set. The particular focus was on the different ways in which the total sample can be defined as follows: all patients referred to the service (referred), only those patients assessed (assessed), only those patients accepted into treatment and attending at least one session (attenders), or only patients completing treatment (completers). The rates, as would be expected, were highly dependent on which definition of the sample was used and, to a lesser extent, on the statistic used for determining improvement or recovery (i.e., reliable and clinically significant improvement, or case versus not case).

The recovery rate for the completer sample using the criteria of reliable and clinically significant improvement was 58.3%. However, this rate fell to 36.7% when calculated on an intent-to-treat sample based on those patients who were assessed and attended at least one session. Moreover, the attender sample comprised only 50% of the original full sample. Although criticisms might be leveled at studies using only completer samples, they do provide a test of the treatment as received by the patient and having an agreed ending between patient and therapist. By contrast, intent-to-treat analyses in routine settings are likely to reflect a combination of treatment and service effects. However, more than anything else, these findings attest to the need to invest in efforts to retain patients in therapy. The delivery of efficacious treatments is a necessary but not sufficient condition for maximizing the personal and social benefits of the psychological therapies.

Dose-Effect Relations

A continuing question within psychotherapy research has been the issue of how much therapy

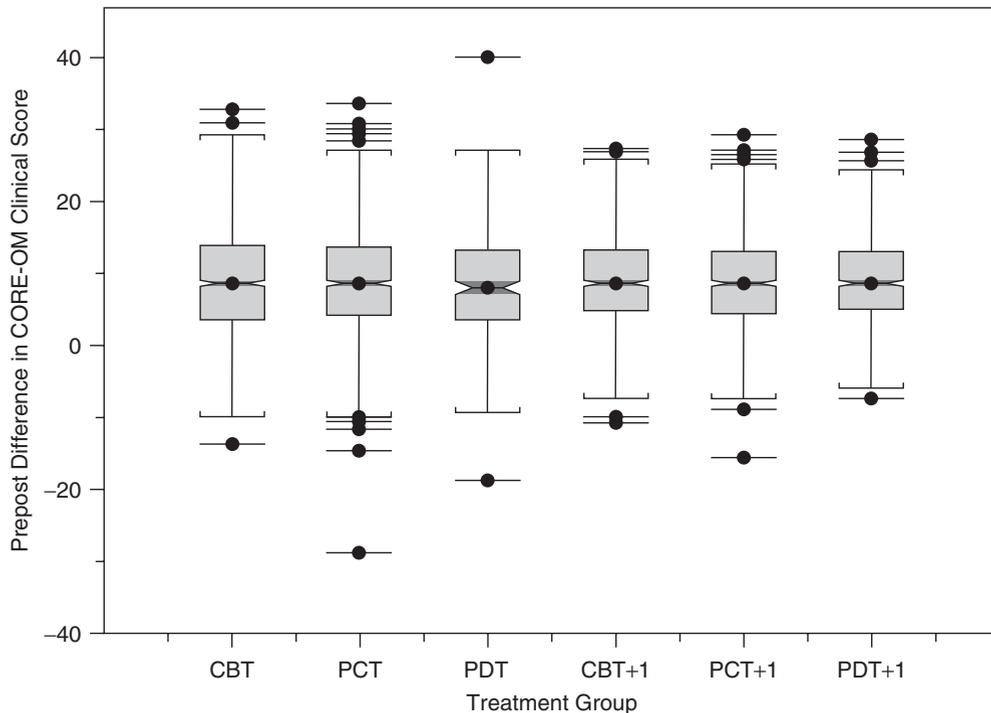


FIGURE 4.4 Notched boxplots showing pre-/postdifferences in CORE-OM clinical scores.

Note. The notch shows the 95% confidence interval around the median. The boxes show the middle 50% of the distribution. The whiskers show the range, except that observations falling 1.5 times the interquartile range or more away from the top or bottom of the box are considered outliers and are shown separately. CBT, Cognitive, behavioral, or cognitive/behavioral therapy ($n = 1,045$); PCT, person-centered therapy ($n = 1,709$); PDT, psychodynamic/psychoanalytic therapy ($n = 261$); CBT+1, CBT combined with one other therapy ($n = 1,035$); PCT+1, PCT combined with one other therapy ($n = 1,033$); PDT+1, PDT combined with one other therapy ($n = 530$).

Reprinted from “Effectiveness of Cognitive-Behavioural, Person-Centred, and Psychodynamic Therapies in UK Primary Care Routine Practice: Replication with a Larger Sample,” by W.B. Stiles, M. Barkham, J. Mellor-Clark, & J. Connell, 2008, *Psychological Medicine*, 38, pp. 677–688. Cambridge, United Kingdom: Cambridge University Press.

is enough. As previously mentioned, Howard et al. (1986) characterized the path of client improvement as a negatively accelerating function of treatment length, which has usually been interpreted as reflecting the diminishing strength of each successive session. Successive studies using data from naturalistic studies have been carried out using measurement of client dysfunction on a session-by-session basis and a binomial classification of dysfunction—recovered or not recovered—at each session (e.g., Anderson & Lambert, 2001; Hansen & Lambert, 2003; Hansen, Lambert, & Forman, 2002; Harnett, O’Donovan, & Lambert, 2010; Kadera, Lambert, & Andrews, 1996). The results from these studies suggest a range of 11 to 18 sessions as the number of sessions needed to achieve clinically significant change for half of the sample starting therapy.

The usual interpretation in psychotherapy research has considered the aggregate curves as representing an average of individual dose–effect curves that are also negatively accelerated, modeled as log-linear functions of session number. There are, however, other possibilities. Barkham et al. (1996) observed that session-by-session plots of improvement in particular symptoms for up to 16 sessions tended to look more or less linear and noted that, in the dose–effect studies by Howard et al. (1986), clients had varying lengths of treatment, so that different aggregations of clients were represented at each successive point. To encompass these results, Barkham and colleagues suggested that problems might be assumed to improve at a steady (i.e., linear) rate across sessions until it reaches a *good enough level* (GEL), at which point the

client and therapist either redeploy therapeutic efforts to other problems or discontinue therapy.

Subsequent work has provided support for the GEL model. Barkham et al. (2006) studied clients ($N = 1,868$) seen for a variety of problems in routine primary care mental health practices, who attended 1 to 12 sessions, had planned endings, and completed the CORE-OM at the beginning and end of their treatment. The percentage of clients achieving reliable and clinically significant improvement (RCSI) on the CORE-OM did not increase with the number of sessions attended. For clients who began treatment above the CORE-OM clinical cutoff ($n = 1,472$), the RCSI rate ranged from 88% for clients who attended 1 session down to 62% for clients who attended 12 sessions ($r = -.91$).

A replication study was carried out by Stiles et al. (2008) who examine rates of improvement in psychotherapy in United Kingdom primary care settings as a function of the number of sessions attended. Included in the study were adult clients who returned valid assessments at the beginning and the end of their treatment, had planned endings, began treatment above the clinical cutoff score, and were seen for 20 or fewer sessions ($N = 9,703$). Clients' average assessment scores improved substantially across treatment, with a pre-post treatment effect size of 1.96 with 62.0% achieving reliable and clinically significant improvement (RCSI). Clients' mean pre-post treatment change were broadly constant regardless of treatment duration (in the range of 0 to 20 sessions) and the RCSI rate decreased slightly with treatment duration, as fewer clients fell below the cutoff at longer durations. The results were consistent with therapists and clients tending to make appropriately responsive decisions about treatment duration (see also Chapter 8, this volume).

In a comparison between the competing predictions of the dose-effect and good enough level models, Baldwin, Berkeljon, Atkins, Olsen, and Nielsen (2009) examined the relationship between rate of change and total dose in 4,676 psychotherapy patients who received individual psychotherapy. Patients attended 6.46 sessions on average ($SD = 4.14$, range = 3–29, $Mdn = 5$). Results were most consistent with the GEL model with rate of change being related to total dose of treatment. That is, small doses were related to relatively fast rates of change, whereas large doses were related to slow rates of change.

In a further study, Reese, Toland, and Hopkins (2011) sought to replicate the GEL model and explore if session frequency moderates the influence that the number of sessions has on the rate of change in psychotherapy. They used an archival naturalistic data set with a U.S. university counseling center sample ($n = 1,207$), with treatment progress measured using the OQ-45. Results were consistent with the GEL model (i.e., clients who attended fewer sessions evidenced faster rates of change). Findings also indicated that the rate of change was influenced by session frequency (i.e., clients who attended more sessions on average per week demonstrated more rapid improvement).

In light of the above findings, session frequency needs to be considered at the level both of practice but also in terms of definitions of dose within research. Practice-based evidence can be used to establish clinical decisions and, consequently, policy about the needed length of therapy—something that clinical trials have failed to address because of the restriction of defining a set dosage when, in fact, some individuals will need more and some less than the fixed dose.

Summary

The studies reported in this section have set out both the yield and potential of an approach that goes part way to redressing the balance with trials methodology as well as promoting a strategy for ensuring better capture and use of data from routine practice. Crucially, however, it reprivileges the practitioner as both a focus and an active agent in the research endeavor by investigating their effects and effectiveness, both as individual therapists and also as the key resource defining any psychological service. Collectively, findings from this approach indicate therapist effects to account for between approximately 5% and 8% of outcome variance, with some therapists yielding considerably better patient outcomes than other therapists. Where evaluations of individual services have been carried out and benchmarked against other studies, findings suggest they are effective but invariably not to the same extent as in trials. And comparisons between differing treatment approaches have yielded broadly equivalent outcomes while investigations into how much therapy is enough suggests that fixed durations of treatment may not be the best way of allocating valuable resources.

PRACTICE RESEARCH NETWORKS

There is no doubt that practitioners are at the center of, and substantially contribute to, the investigations that are conducted within both the patient-focused and practice-based research approaches. Clinicians are directly involved in data collection, and this data collection is aimed at understanding or improving the practice of psychotherapy, including their own, in naturalistic environments. The level of engagement of clinicians in research tends to be even greater in the third approach of practice-oriented research: practice research networks (PRNs). As stated by Parry, Castonguay, Borkovec, and Wolfe (2010), PRNs:

“[H]ave been alternately defined as a group of practicing clinicians that cooperates to collect data and conduct research study” (Zarin, Pincus, West, & McIntyre, 1997), and “large numbers of practicing clinicians and clinical scientists brought together in collaborative research on clinically meaningful questions in the naturalistic setting for the sake of external validity and employing rigorous scientific methodology for the sake of internal validity” (Borkovec, 2002, p. 313).

When based on a partnership of practitioners and researchers, PRNs involve, optimally, collaboration on all aspects of investigation: from the generation of ideas to the design, implementation, and publication of studies. This collaboration aims to foster a sense of equality, shared ownership, and mutual respect between researchers and clinicians, and promoting diversity of scholarship (i.e., different ways of understanding and investigating complex phenomena). It also capitalizes on the complementary expertise, knowledge, and experiences of each stakeholder to provide unique opportunities for two-way learning in order to conduct studies that are both clinically relevant and scientifically rigorous. Also characteristic of PRNs is the consideration of respective needs and resources of both clinicians and researchers in designing protocols that balance issues of internal validity and feasibility, which can sometimes compete and at other times be synergistic. By having practitioners involved in deciding what studies should be conducted

and how they should be implemented, PRNs also provide practitioners with an active voice in “setting the research agenda” (Zarin et al., 1997) and a vehicle for shaping the empirical evidence upon which practice could be based. For more discussion on the defining features of PRNs, as well as strengths, weaknesses, and challenges associated with research conducted within them, see Garland, Hurlburt, and Hawley (2006), McMillen, Lenze, Hawley, and Osborne (2009), Parry et al. (2010), and Zarin et al. (1997).

The first PRNs were created in the 1950s to meet the needs of specific healthcare fields, such as primary care in rural areas, and nursing (see Bradley, Sexton, & Smith, 2005; McMillen et al., 2009; Zarin et al., 1997). However, it was not until four decades later that behavioral or psychological health PRNs began to emerge. Currently, there are several PRN infrastructures that are devoting at least part of their focus to research related to psychotherapy. Although these infrastructures vary considerably, they can be clustered into three categories identifying the main groups of clinicians involved: PRNs based primarily on *professional organizations*, *specific disorders*, and *common clinical settings*. This section presents brief descriptions of some of the PRNs within each of these clusters as well as some examples (it is not a comprehensive survey) of psychotherapy studies conducted in each.

As the three approaches of practice-oriented research presented in this chapter focus on the utilization and delivery of care in naturalistic settings, it will be no surprise that some of the studies conducted in PRNs address topics also investigated in patient-focused and/or practice-based research. However, in part because the ideas for studies in PRNs tend to rely more on a diversity of voices, reflecting both clinicians’ and researchers’ perspectives, investigations carried out within PRNs cover a wider range of topics than those conducted in the two approaches described earlier.

Professional Organization PRNs

A number of PRNs have been created or sponsored by professional organizations representing several fields of mental health, including psychiatry, psychology, social work, marriage and family therapy, and art therapy. Although the studies conducted within these PRNs cover a wide range of interests, they do reflect an apparent desire for research on actual practice that is shared by therapists of different training backgrounds.

The American Psychiatric Institute for Research and Education

The American Psychiatric Institute for Research and Education's PRN (APIRE-PRN) was originally conceived as a nationwide network of psychiatrists, collaborating on data collection and conducting research on a variety of issues related to clinical services delivery, health services, and health policy (West & Zarin, 1995). It was established in 1993 in response to the understanding that more research relevant to practice was needed in psychiatry. The psychotherapy studies conducted within this PRN have been mostly observational, short in duration, as well as simple to implement and adaptable to different office practices in order to avoid interfering with clinical routines. While we cover some of the most psychotherapy-relevant studies here, it should be noted that recent studies associated with this PRN have involved large-scale investigations that are based on randomly selected psychiatrists from the American Psychiatric Association Physician Masterfile, rather than on the more limited (and thus less representative) samples of psychiatrists that participated in earlier investigations.

Drawing from a broad range of patients, treatments, and treatment settings, the infrastructure of the PRN allows for multiple areas of research. One such area is how patients' characteristics affect clinical care. For example, Herbeck et al. (2004) explored variations in the use of second-generation antipsychotic medications by race and found that, even after controlling for potential confounds, African Americans were less likely to be prescribed second-generation antipsychotic medications, considered to be the treatment of choice by psychiatrists, compared to Whites.

Another study, by Duffy et al. (2008), assessed the feasibility and helpfulness in adding a depression severity monitoring measure to 19 psychiatric practices. The results suggested that the inclusion of the measure was both feasible and helpful for treatment: 93% of psychiatrists reported utilizing the instrument, and in 40% of patient visits it led to treatment changes such as dosage, medication change, or recommendation for psychotherapy and medication.

Data collected through the APIRE-PRN has also allowed researchers to conduct studies on issues related to therapeutic processes. For example, Herbeck et al. (2005) examined various factors related to treatment noncompliance in patients with substance use disorders. The results suggested that several variables, such as presence

of a personality disorder, low Global Assessment of Functioning (GAF; American Psychiatric Association, 2000) scores, side effects from medications, and seeing psychiatrists at a discounted service fee, were associated with an increased probability of treatment noncompliance.

Trends in psychiatric practice are another major area investigated in this PRN. West, Wilk, Rae, Narrow, and Regier (2003) compared the fees that psychiatrists receive for providing a psychotherapy session (with a medical evaluation) with three medication management sessions and found that psychotherapy sessions were reimbursed \$74.52 less than medication management sessions. Wilk, West, Rae, and Reiger (2006) studied the patterns and predictors of psychotherapy use in psychiatric patients and found that two-thirds of patients received psychotherapy as part of their treatment and more than half of these were with their treating psychiatrist.

A study conducted by Zarin, Johnson, and West (2005) addressed an issue related to clinical relevance of traditional research: the comparability of clients in RCTs and general practice. They compared patients diagnosed with bipolar and schizophrenia across several demographic, clinical (such as principal diagnoses, comorbidity, medical history), and treatment (medications and dosages) variables. Overall, the results supported the conclusion that the patients and treatments in RCTs are not representative of clinical practice: PRN patients were predominantly white, female, and older, with more comorbid diagnoses compared to RCT patients. Approximately 40% of the patients with schizophrenia and 50% of the bipolar patients in the PRN would not have met inclusion criteria for an RCT.

The Pennsylvania Psychological Association

Originally created by a full-time academician (Thomas Borkovec) and a full-time clinician (Stephen Ragusea), the Pennsylvania Psychological Association PRN (PPA-PRN) has been devoted to the conduct of psychotherapy research in outpatient clinics in Pennsylvania. In all investigations of this PRN, clinicians and researchers have fully collaborated in the specification of the goals, design, implementation, and dissemination of the findings. Based on the coordinated work of three committees (core assessment, study protocol, and ethics), the first study was conducted for the sole purpose of evaluating the feasibility of a research infrastructure in naturalistic outpatient psychotherapy (Borkovec, Echmendia, Ragusea,

& Ruiz, 2001). This led to the adoption of a standardized outcome measure by more than 50 private practitioners, who used it as an assessment tool in routine care. Although clearly stating that their intention “was not to draw specific conclusions from the results of the PRN study for either theoretical or applied purposes” (p. 159), members of the PRN reported a number of findings that could (with additional studies aimed at determining mechanisms of change, as well as cause and effect relationships) lead to clarification of how therapy works and how it can be improved. Specifically, in addition to indicating that significant improvement took place during treatment across a number of outcome dimensions (e.g., symptoms, interpersonal problems, and life functioning), the results suggest that some client and therapist characteristics were associated with improvement. For example, while clients’ initial expectancy for a positive outcome was predictive of greater change, interpersonal distress was negatively related to improvement. Interestingly, female clients showed more therapeutic change than male clients regarding family relationships, and clients of male therapists showed greater changes in intimacy than clients of female therapists. Moreover, while a higher number of clients in a therapist’s caseload was predictive of worse outcome, better outcome was associated with the number of sessions received (Borkovec et al., 2001; Ruiz et al., 2004). A number of lessons for future PRN efforts were also derived from this study, such as the need to provide incentives to participants, as well as securing grants from funding agencies and psychological associations.

A second investigation was developed and carried out by 13 clinicians of varying theoretical orientations who, for a period of 18 months, invited all of their new clients (adults, adolescents, and children) to participate, except when psychotherapists judged such participation to be clinically contra-indicated. The study, involving a total of 146 clients, was a randomized clinical trial aimed at examining two general questions. First, clinicians wanted to know what clients and therapists find the most helpful and hindering during therapy. The second goal was to determine whether the provision of such feedback at the end of every session improves the effectiveness of therapy. The study led to the collection and coding of close to 1,500 events that were reported as being helpful or detrimental during sessions. Among the results obtained, the therapists’ efforts to foster clients’ awareness of their emotions, thoughts, and behaviors were perceived as

particularly helpful by both clients and therapists. Events leading to the strengthening of the alliance were also reported by therapists as particularly helpful. Furthermore, therapists identified as particularly hindering some omissions that they themselves committed (e.g., failure to provide needed structure) during sessions (Castonguay, Boswell, et al., 2010). The difficulty of collecting outcome measures at the end of treatment, however, prevented the researchers/clinicians from answering the second question of the study. After the completion of this study, participating therapists were interviewed regarding their experience of designing and implementing the research protocol within their private practice (Castonguay, Nelson, et al., 2010). A qualitative analysis of these interviews revealed a number of benefits. The therapists reported, for example, that receiving feedback from clients about helpful and hindering events allowed them to adjust their interventions toward clients’ needs, and that writing down their thoughts after each session provided clients with the opportunity to consolidate therapeutic material. They also reported that their involvement in the study fostered new learning, as well as a sense of community with other professionals with shared goals. Therapists also reported that clients perceived their research participation as intrinsically meaningful in that it provided them with an opportunity to contribute to scientific knowledge. However, there were a number of difficulties reported, such as having to depart from their clinical routine, remembering detailed procedures of the research protocol, and finding time to complete and review questionnaires. At times, research tasks were perceived as interfering with clients’ needs (e.g., when an explanation of the study took away time to establish rapport early in therapy).

A number of recommendations for future PRN studies were also generated by the analyses of the interviews, including the design of research procedures that are as simple and clear as possible, direct and easily available consultation to address problems of data collection, and strategies aimed at increasing clients’ and therapists’ motivation (such as financial incentives for clients to help reduce the problem of post-treatment data collection mentioned above). Perhaps the most important recommendation for future PRN endeavors was the need to develop and conduct studies in which the empirical and clinical goals are confounded. That is, “studies for which it is impossible to fully distinguish whether the nature of the questions investigated, tasks implemented,

or the data collected are empirical or clinical” (Castonguay, Nelson, et al., 2010, p. 352).

Additional Professional Organization PRNs

At least six other professional organizations have sponsored PRNs: American Association for Marriage and Family Therapy (AAMFT), American Counseling Association (ACA), Association for Cognitive Analytic Therapy, American Psychological Association (APA), Art Therapy Practice, and the National Association of Social Workers (NASW). At this point in time, the studies that have been conducted in these PRNs have been mostly surveys (often nationally representative), which primarily examine practitioner and client characteristics, practice settings, as well as treatment and services patterns (e.g., Bradley et al., 2005; Huet, Springham, & Evans, 2008; Northey, 2002; Simmons & Doherty, 1995; Smith, Whitaker, & Weismiller, 2006). Some of the findings of these surveys have revealed important information. For instance, Smith, Whitaker, and Weismiller (2006), from the NASW PRN, reported that more than 25% of the clients treated by social workers were diagnosed with a substance use disorder, but fewer than half of the sample’s social workers had received specific training in treatment of substance abuse during the previous 12-month period. The authors interpreted these findings as suggesting that social workers may not be receiving adequate training in the types of problems they need to treat. It should also be mentioned that the ACA-PRN has begun to examine outcomes of treatments provided by counselors. Smith, Sexton, and Bradley (2005) reported preliminary outcomes from a study of 143 clients seen by 26 counselors in routine practice. In this study, clients completed the OQ-45 before and after treatment, and scores were examined for differences. Though the authors described this study as preliminary, and a complete description of findings was not presented, they did find significant overall improvement in symptom distress after counseling for those clients who were significantly distressed at the start of treatment. This study is noteworthy since a substantial amount of work that organization-based PRNs have conducted to date has been focused on documenting practices in routine care, rather than addressing outcomes. Studies like this one suggest that these PRNs, given their wide reach, may be able to provide meaningful information as

to rates of treatment success as well as potential obstacles to therapeutic improvement.

Disorder-Specific PRNs

Several PRN initiatives have focused on specific clinical problems, such as substance abuse, disruptive behavioral problems in children, child maltreatment, eating disorders, and autism. Like the PRN infrastructures described in the previous sections, they have led to a variety of studies in terms of topics and complexity. This variety no doubt reflects variation in the longevity and resources across these different PRNs. It also represents the wide range of interests that drives clinicians and researchers in their inquiries about psychotherapy-related issues.

The National Drug Abuse Treatment Clinical Trials Network

The National Drug Abuse Treatment Clinical Trials Network (CTN) is a particularly well-developed PRN infrastructure. Created by the National Institute on Drug Abuse as a way “to bring drug abuse researchers into the real world of the treatment clinic while creating opportunities that allow practitioners to participate in treatment research” (Hanson, Leshner, & Tai, 2002, p. 69), the CTN has not only conducted numerous studies on psychotherapy and behavioral health interventions for drug abuse, but has also developed a model infrastructure for promoting high-quality experimental research in a large, clinic-based network (Tai et al., 2010). This PRN has produced 271 published journal articles at the time of this writing (National Drug Abuse Treatment Clinical Trials Network, 2011).

Explicit collaboration between researchers and providers, including clinicians and directors of treatment programs, is built into the infrastructure of this PRN. This is accomplished at both regional and national levels. In the CTN, regional “nodes” include university-based researchers and clinic-based practitioners who meet, discuss, plan, and conduct research projects within a particular part of the country. Nodes often collaborate in CTN-funded studies that require at least three nodes to participate, increasing external validity. As described in Tai et al. (2010), research projects that are conducted through this infrastructure are approved and supported nationally by NIDA, which provides input on research design and oversees the study progress. A centralized management corporation provides

an overall structure to this PRN by coordinating clinical research training, given by providers and researchers, as well as managing data and providing statistical expertise. The governance of the CTN is assumed by a steering committee that comprises an equal representation of providers and researchers with representation from each node. Exchanges between researchers and providers have also been promoted through more than 20 special interest groups, in which many research ideas are developed and issues in need of research are advocated. It is noteworthy that while much of the research conducted in the CTN is largely indistinguishable from traditional clinical research, what makes it a PRN is the active participation of practice sites, their directors, and their clinicians in the design and implementation of each project. This demonstrates the potential for PRNs to conduct highly rigorous studies that emphasize both internal and external validity, as well as the potential for clinicians to value, utilize, and conduct such research.

As of 2011, 52 primary research studies have been conducted or funded, including over 20 RCTs along with several surveys and process studies (National Drug Abuse Treatment Clinical Trials Network, 2011). Although many of the studies conducted by this PRN have tested the efficacy of pharmacologic interventions, several have provided large-scale, field trial clinical investigations of psychotherapeutic and behavioral treatments. For instance, Robbins et al. (2011) conducted a multisite RCT to evaluate treatment-as-usual (TAU) and brief strategic family therapy (BSFT) for adolescent substance users. In the main outcome of this study (i.e., urinalysis, which only assesses the previous week at posttreatment), the researchers found no significant differences between treatment types. However, they did find evidence of better retention, treatment engagement, and parent-reported family functioning for the BSFT condition, as well as a significant difference between groups in self-reported drug use.

In another study, Carroll et al. (2006) trained drug counselors in motivational interviewing (MI) techniques, and integrated these techniques into a single session of intake and initial assessment for drug abuse patients. This study showed that the MI procedures increased retention rates for the first month of treatment, an important factor in drug counseling. Ball et al. (2007) also tested a brief motivation enhancement therapy (MET) against usual counseling and found significant

effects on retention. Both studies also suggested that there was meaningful variability between sites in the CTN as well as differences in the types of substances being treated in terms of the effectiveness of the motivational interventions.

Additionally, the CTN has produced a large amount of research on the effectiveness of empirically supported treatments when delivered to particularly underserved populations, which are not always well represented in large-scale RCTs (Burlaw et al., 2011). Studies have included a specific focus on, among others, African Americans (e.g., Montgomery, Burlaw, Kosinski, & Forcehimes, 2011), Spanish speakers (e.g., Carroll et al., 2009), and American Indians (e.g., Forcehimes et al., 2011).

The variety of methods and research topics of studies conducted within the CTN is expansive and clearly illustrates the potential for high-quality research in naturalistic settings. Studies in the CTN have focused on therapist characteristics (e.g., Suarez-Moreles et al., 2010), client pretreatment characteristics (e.g., Hartzler, Donovan, & Huang, 2010), dissemination of efficacious treatments into the community (e.g., Walker et al., 2010), brain imaging (e.g., Upadhyay et al., 2010), surveys (e.g., Pinto, Yu, Spector, Gorroochurn, & McCarty, 2010), and statistical and research methodology (e.g., Morgan-Lopez, Saavedra, Hien, & Fals-Stewart, 2010). Such breadth of research topics suggests that there are many areas of research in which clinicians are potentially interested.

It is important to note that in addition to this research, work from the CTN has also touched on practitioners' views of participating in research. Knudsen, Ducharme, and Roman (2007) surveyed practitioners in the CTN and found that counselors' perceptions of increased stress due to research (e.g., larger workload, emotional stress) were associated with intentions to discontinue working at their clinic. In contrast, the perception that research was helpful to patients and the clinic (e.g., increasing retention rates, increasing sense of teamwork) was positively associated with counselors' intentions to stay. This has clear implications for any research studies developed in PRNs in the future.

Practice and Research: Advancing Collaboration

A researcher-practitioner partnership (the Practice and Research: Advancing Collaboration [PRAC]), focusing on the treatment of disruptive

behavioral problems (DBP) in children, has been developed by Garland and her colleagues at the University of California at San Diego (Garland, Hurlburt, & Hawley, 2006). The primary goal of this partnership is to understand the therapeutic processes and the associated clinical outcomes in the treatment of DBP, in a naturalistic setting. As described in Garland et al., clinicians representing six clinical sites collaborated with researchers in the planning (e.g., selection and adaptation of measures, refinement of the methodological design, increase of feasibility and clinical relevance), implementation (e.g., recruitment of participants, problem solving), and interpretation of a large study. This study, involving more than 80 therapists of various professional backgrounds and more than 200 children, examined how principles of change underlying evidence-based practice (EBP) for DBP were consistent with common practices for this problem in usual care. Based on observational coding of psychotherapy sessions with children and their family conducted in community outpatient clinics, the specific aims were to describe the therapeutic process in terms of treatment strategy and relationship variables, and examine the relationship between practice elements (evidence-based and practitioner-based) and outcome. The study also aimed to investigate pre-treatment characteristics of clients and therapists. Recently published results (Brookman-Frazee, Haine, Baker-Ericzen, Zoffness, & Garland, 2010) indicate that, even though EBP elements were used in almost all the 1,215 sessions that were coded, the overall intensity (i.e., how extensively the treatment element was delivered) of these elements was low for both caregivers and children. The implication was that the use of EBP elements was brief. Brookman-Frazee et al. (2010) also reported that the child's age (older) and therapists identifying themselves as cognitive behavioral or behavioral in orientation were significantly associated with higher EBP use.

Similar to the procedures carried out in the PPA-PRN, a qualitative analysis was conducted to learn how the clinicians and researchers viewed their collaboration at the beginning and during the PRAC project, as well as their recommendations for future collaborative projects (Garland, Plemmons, & Koontz, 2006). Both challenges (such as tensions related to group dynamics, conceptual issues, and communication) and positive experiences (perceived trust in others, increase in knowledge, and changes in their views of research

and practice) were reported. In particular, an emphasis on clinical concerns and observations in the formulation and development of research questions was highlighted as a characteristic of an ideal collaboration.

The Healthy Families America Research Practice Network

The Healthy Families America Research Practice Network (HFA RPN) is aimed at the prevention of child maltreatment. To this end, it has grouped academic researchers with community-based evaluators. As of 2007, HFA RPN comprised, as part of its researchers and practitioners council, 25 researchers and 15 clinicians (Galano & Schellenbach, 2007). Although this infrastructure is focused on prevention, it shares many of the same goals as the PRNs focused on treatment and psychotherapy based interventions (e.g., facilitate communication between researchers and clinicians, integration of research based practices in naturalistic settings, identification of most clinically relevant questions, and strategies to investigate them). It has emerged, in part, because of similar perceived limitations of traditional research in social sciences. As in a few of the PRN infrastructures described above, it has led to a common standardized database, as well as the conduct of a large study involving 100 sites across nine states and aimed, in part, at implementing evidence-based intervention strategies. Interestingly, the building of infrastructure and conduct of research within community interventions involves issues that are not typically addressed in psychotherapy, but that might be worth considering in our attempt to improve mental health care, such as the involvement of policy makers and human services agencies.

Additional Disorder-Specific PRNs

Some researchers have also employed a practice research network approach as a way to recruit practicing clinicians to conduct studies on specific disorders. For example, Westen, Shedler, Durrett, Glass, and Martens (2003) used this approach in order to randomly select and contact clinicians who were members of either the American Psychological Association or American Psychiatric Association, and asked them to provide data on their patients for a national study looking at psychopathology in adolescents. Thompson-Brenner, Boisseau, and Satir (2010) used a similar approach and sample of clinicians when conducting a study looking at eating disorders in

adolescents. Other researchers have established collaborations with community providers to promote cooperative and symbiotic efforts in the dissemination of evidence-based practices and decrease the research-practice gap. Stahmer and colleagues (Brookman-Frazee, Stahmer, Searcy, Feder, & Reed, 2012; Stahmer, Brookman-Frazee, Lee, Searcy, & Reed, 2011), for example, developed the BRIDGE (Bond-Regulate-Interact-Develop-Guide-Engage) Collaborative, a group comprised of researchers, providers, parents, and funding agencies, aimed at implementing interventions in community early intervention settings for infants and toddlers who are at risk of autism spectrum disorders (ASD). Similarly, researchers and community stakeholders collaborated to develop a mental health intervention protocol and corresponding therapist training model for school-age children with ASD treated in community mental health settings. Results of a pilot study indicated that community mental health therapists can be trained in a short period of time to implement (with fidelity and promising effectiveness) interventions drawn from evidence-based practice for this clinical population (Brookman-Frazee, Drahotka, & Stadnick, 2011).

PRNs Based on Common Settings

Another type of PRN is aimed at conducting studies that may address various clinical problems and may involve therapists of different training backgrounds. In contrast with those described earlier, these PRNs have been developed to better understand and improve practice in specific clinical settings. Two examples of such PRNs are presented next. One examines services provided in college counseling centers, and the other operates in a psychotherapy training clinic.

The Center for Collegiate Mental Health

The Center for Collegiate Mental Health (CCMH) is an infrastructure that has been created to foster mutually beneficial bridges between a large number of collaborators (mental health treatment providers, psychological and information science researchers, industry leaders, and university administrators), all of whom are invested in the collection of data to enhance the mental health services provided to college students (Castonguay, Locke, & Hayes, 2011; Hayes, Locke, & Castonguay, 2011; Locke, Bieschke, Castonguay, & Hayes, 2012). This infrastructure now includes

more than 150 college counseling centers across the United States. These centers are using the same instruments in routine practice, allowing for a collection and “real-time” processing of a massive flow of standardized (and Institutional Review Board-approved) data. The first instrument, the Standardized Data Set (SDS) covers a range of basic issues such as client/counselor demographics, mental health history, and living situation. The second measure (the Counseling Center Assessment of Psychological Symptoms [CCAPS]; Locke et al., 2011; McAleavey et al., 2012) is a multidimensional assessment of clients’ difficulties and treatment outcome. A short version of the CCAPS has now been validated to allow for the purposes of repeated assessments (Locke, McAleavey, et al., 2012). While routine data collection is primarily administrated at the center level (each counseling center has autonomy to set their own schedule of administrations), research studies and access to data of the national sample is centralized at CCMH.

A pilot study of more than 19,000 clients seen at counseling centers in the fall semester of 2008 led to the examination of the clinical difficulties and use of counseling services by college students of ethnic and/or sexual minorities. One of the questions investigated, for example, was whether students who represent a double minority status, racial/ethnic and sexual, experience more intense and/or specific types of clinical problems as compared to students who are members of only one minority group (Hayes, Chun-Kennedy, Edens, & Locke, 2011). Another question examined was whether eating disorder and body image problems, both predominant foci of psychological services provided in counseling centers, should be viewed as a problem predominantly experienced by white women or if college students from a variety of backgrounds express similar concerns (Nelson, Castonguay, & Locke, 2011). Two other investigations found that different groups of sexual orientation minorities (e.g., gay, lesbian, and bisexual clients) and gender minorities (e.g., transgender clients) experience different types of psychological difficulties, some of them being extremely severe (Effrig, Bieschke, & Locke, 2011; McAleavey, Castonguay, & Locke, 2011).

A second wave of studies has begun to address issues related to therapeutic change. For example, Boswell, McAleavey, Castonguay, Hayes, and Locke (2012) have investigated how the particular client pre-treatment characteristic of prior treatment history affects outcomes in

routine treatment. Despite their expectations, the authors found that clients who had previously been in counseling, but not previously taken psychiatric medications or been hospitalized for mental illness, were slower to respond to a course of counseling and evidenced less overall improvement than other clients. Lockard, Hayes, McAleavey, and Locke (2012) have also reported on a counseling outcome study stemming from the CCMH PRN infrastructure. In this study the authors used data from clients at two clinics on one university campus and compared them to a group of nonclinical peers from the same university. These authors report that the mean level of academic distress among those students in counseling decreased significantly over 6 weeks of counseling, while it remained stable in the nonclinical group. This finding provides support for the effect of counseling and also demonstrates the value of collaboration between counselors at different clinics.

Also within the context of the CCMH infrastructure, a qualitative study was recently conducted on the experience of both doctoral students and their clients in using the CCAPS on a repeated basis (Martin, Hess, Ain, Nelson, & Locke, 2012). As in the PPA-PRN and the PRAC studies mentioned above, obstacles, such as limited time available to devote to collecting data, were reported, but so were benefits. For example, 62% of clients reported that completing the CCAPS helped them think differently about their problems. Similarly, 64% of the therapists reported using the outcome scores to modify the case formulation of their clients' difficulties. In addition, the therapists, all of whom were trainees, reported discussing their clients' CCAPS scores during supervision. This study provides an example of how the collection and use of data in naturalistic settings can simultaneously serve research, clinical, and training purposes (Castonguay, 2011).

The Pennsylvania State University Training Clinic PRN

With the goal of fostering a seamless integration of research, practice, and psychotherapy training, the Department of Psychology at Penn State University has transformed its training clinic into a PRN. This has been achieved by incorporating four major components into its training program (see Castonguay et al., 2004; Parry et al., 2010): a core outcome battery, the Treatment Outcome Package (TOP; Kraus et al., 2005); standardized

diagnostic assessment procedures; a selection committee for the evaluation of research proposals (including representatives from the faculty, clinical staff, students, and practitioners from the community); and an innovative agreement with the office of research protection to efficiently streamline the Institutional Review Board (IRB) review process. This infrastructure has enabled several students to find themselves in a situation in which they are seeing clients, meeting their clinical hour requirements, and collecting their masters and/or dissertation data. Studies by students, postdoctoral fellows, and faculty members are conducted in this infrastructure only if they are judged to be clinically relevant, to reflect the clinic's mission of promulgating and integrating rigorous (and pluralistic) research within clinical services and training, and judged to be minimally invasive to the functioning of the clinic. At this point in time, more than 10 studies have been launched; a few of them are briefly described next.

Boswell, Castonguay, and Wasserman (2010) reported on a study of training, psychotherapeutic intervention use, and perceived session quality in this PRN. The authors had PRN therapists complete measures of training variables and then had clients and therapists complete measures of psychotherapy techniques use and session outcome after each session of psychotherapy. This process study tracked 19 therapists with 42 clients across an average of 10 sessions per client using quantitative self-report methods.

Interestingly, the authors found that several training-relevant variables (e.g., self-identified psychotherapeutic orientation, years in training, degree status, practicum orientation) were not related to the kinds of psychotherapy techniques reported in sessions. This was surprising, and suggests that the types of techniques used in therapy may not be a simple function of training. In addition, these authors found a complex relationship between technique use and session quality as rated by clients, suggesting that, overall, clients reported the most benefit from sessions involving greater frequency of techniques associated with interpersonal therapy and behaviors considered to be "common factors" across therapy orientations. In addition, an interesting result was found for clients who received especially frequent common factors (compared to other clients of the same therapist) from therapists who reported a particularly high frequency of common factors behaviors (compared to other therapists). These clients indeed reported diminished session quality

when their therapists used more CBT techniques than usual. This suggests an important lesson for trainees and experienced therapists: The implementation of CBT interventions should proceed with careful consideration of the current context of treatment, especially when therapy with a client has relied heavily on common factors (the basic helping skills and therapeutic environment), because the client is likely to experience these as less helpful when compared with other clients.

In another study, Nordberg, Castonguay, Fisher, Boswell, & Kraus (2008) examined outcomes of treatment in this PRN. Using the TOP, these authors sought to replicate and extend earlier work carried out in naturalistic settings of psychotherapy (see Stulz et al., 2007). Specifically, using growth mixture modeling of outcome data in treatment, this study found support for the identification of three latent classes of clients with depressive symptoms: low-distress slow responders, high-distress slow responders, and high-distress fast responders. This study used archival information derived through the PRN's unique data generation process to identify depression chronicity as a potential discriminator between the two high-distress groups.

Summary

Reflecting clinicians' wide arrays of interests, a very broad range of topics has been investigated in psychotherapy PRN studies, such as the clinical problems and utilization of services by ethnic and/or sexual minorities, and the impact of clients' characteristics on prescribed care, treatment compliance and outcome. Different procedures related to clinical care have also been examined, including the feasibility and impact of assessment measures on routine practice, types of treatments and interventions used by therapists, the link between therapists' interventions and the impact of the session on clients, as well as helpful and hindering events in sessions. Several studies have also investigated the outcome of therapy from a variety of angles, such as the examination of patterns of change in a training clinic, effectiveness of treatment-as-usual in clinical practice, as well as the effectiveness (via randomized trials conducted in naturalistic settings) of specific interventions and empirically supported treatments for underserved populations. These issues represent a large array of questions that are pertinent to the practice and effectiveness of psychotherapy. While some of these issues

are consistent with the interests of academic researchers and/or the priorities of funding agencies, others more directly reflect the concerns and questions of those practitioners working within the demands and pressures of everyday routine practice. Such convergence and diversity are no doubt indicative of the complexity of psychotherapy, as well as showing the likely benefit of recognizing the merits of different research strategies, and the relevance of their findings, to further understand and improve therapeutic practice.

OTHER INITIATIVES TO CLOSE THE SCIENCE-PRACTICE GAP

There have been numerous and diverse attempts to bridge the science-practice gap in addition to the three approaches discussed earlier (practice-based evidence, patient-focused research, and practice research networks). Here we briefly present a few major trends in this work.

A number of authors have identified ways to facilitate practicing psychotherapists conducting research in the context of their own practice, (e.g., Goldfried, 1984; Kazdin, Kratochwill, & VandenBos, 1986). The clinical case study is perhaps the original model for clinicians interested in contributing to empirical knowledge. In recent years, there has been an increasing call for such in-depth analyses of single cases as an essential form of knowledge (e.g., Barlow & Nock, 2009; Borckardt et al., 2008; Dattilio, Edwards, & Fishman, 2010; Iwakabe & Gazzola, 2009; Kazdin, 2008, 2010). This call has also come as the available methods of case study analysis have expanded to include numerous diverse approaches to maximize empirical and theoretical knowledge derived from single cases. Stiles (2007, 2010) has proposed that a unique benefit of case study research is that it can identify ways in which existing theories are inadequate and need to be developed. Stiles proposed methods by which psychotherapists can implement a case study with the explicit goal of building new and modified theories that could later be tested with other methods (2007), as well as ways in which these case studies can constitute practice-based evidence (2010). Elliott and colleagues (Elliott, 2002; Elliott et al., 2009) described a hermeneutic single-case efficacy design, which was developed explicitly to test whether psychotherapy has been an active and meaningful contributor to client

improvement. This method relies on quantitative as well as qualitative assessment strategies to make this interpretive decision. Numerous models like these have been developed specifically to enable case-based research in applied settings of psychotherapy (e.g., Beeson & Robey, 2006; Borkardt et al., 2008; Galassi & Gersh, 1993; Hayes, 1981; Mayott-Blum et al., 2011). However, it is likely the case that countless clinical theories have not been adequately documented using such case study research methods, and many of the clinical lessons that could be learned from such case studies are not rapidly or systematically integrated into researchers' studies. This is perhaps a particular instance of the colander effect mentioned earlier: numerous pieces of knowledge slipping through the grates.

Aside from case studies, there has also been considerable encouragement offered to practicing clinicians in conducting research within their everyday clinical work (e.g., Goldfried, 1984). Clinicians/researchers such as Jacqueline Persons (e.g., Persons et al., 1988; Persons, Roberts, Zalecki, & Brechwald, 2006) and David Burns (e.g., Burns & Nolen-Hoeksema, 1991, 1992) have provided the field with exemplary models of how process and outcome data can be collected in private practice in order to shed light on therapeutic change. Research in applied settings has also included large-scale quantitative evaluations of treatment programs (e.g., Hardy, Weatherford, Locke, DePalma, & D'luso, 2011; Steinfeld, Coffman, & Keyes, 2009). In addition, there have been calls for the application of controlled qualitative research methods, primarily through detailed analysis of cases and aggregating this information within a practice, to improve conceptualization of clients' presenting problems and strengthen effectiveness of future services (e.g., Silverstein, Auerbach, & Levant, 2006). Moreover, researchers have described tools to aid clinicians conducting research, including quantitative methods for analyzing and better understanding naturalistic psychotherapy data (e.g., Crosbie, 1993; Jones, 2003; Speer & Greenbaum, 1995).

There has been a developing trend toward increasing the clinical relevance of research, based on the belief that much research, especially internally valid RCTs, may not suffice to inform the clinical situation (e.g., Persons & Silberschatz, 1998). For instance, journals have been launched that are dedicated to promoting clinicians' participation in and/or consumption

of research. These have included the *Journal of Clinical Psychology: In Session* and *Pragmatic Case Studies in Psychotherapy*, to name just two. Other journals, such as *Psychotherapy*, have recently put out specific calls for papers meeting the criteria for a "practice review," a literature review with clinical implications and discussion (e.g., Davis & Hayes, 2011) and "empirically based case studies," which incorporate clinical material with predefined quantitative process and outcome data.

Additional methods for increasing the clinical relevance of research have included other initiatives to seek out clinicians' feedback on existing research and to direct future research. For instance, Goldfried (2011) described a feedback mechanism for clinicians to provide input on the problems and difficulties encountered when attempting to implement ESTs in their practice. Experienced clinicians have also been invited to provide their perspective about the current state of research in psychotherapy. That is, how findings confirm what they already knew, are inconsistent with what they believed or have observed, and provide them with new information or perspectives about therapy. In addition, their views have been sought about dimensions, issues, or questions that have not been covered in the empirical literature but would be important to investigate (e.g., Castonguay, Adam-Term, et al., 2010). Similarly, as described earlier, Martin et al. (in press) and Garland et al. (2006), as well as Castonguay et al. (2011), have provided qualitative analyses of psychotherapists' experiences participating in research projects.

There has been an additional trend in the literature towards exploring ways to optimally train graduate students for careers that balance research and practice. For instance, Stricker and Trierweiler (1995) suggested that if clinicians adopted the role of a "local clinical scientist," they could incorporate empirical findings from research while developing local expertise relating to their own clients through repeated and systematic clinical work. Goldfried (1984) suggested that training programs should in fact explicitly aim to encourage their students to engage in research once out of training, in order to close the divide between research and practice.

Recently, the Association for Psychological Science has sponsored an effort to further encourage graduate programs to emphasize research in training. As described by Baker, McFall, and Shoham (2009), the Psychological

Clinical Science Accreditation System is designed to accredit only those programs that train graduates who “can generate and apply psychological clinical science effectively” (p. 88). This system may have significant effects on the way that many clinical psychologists approach research and practice. There have been other endeavors at closing the science-practice gap through training, including a current initiative of the American Psychological Association’s Division 12 (Society for Clinical Psychology) to make recommendations to training programs about how to implement evidence-based practice (Beck et al., 2012). Among the sources of knowledge that have been considered for such recommendations are the efforts in health fields, such as medicine, in order to foster the translation and implementation of research evidence into clinical practice (see Gray, 2004; Guyatt & Rennie, 2002; Straus, Glasziou, Richardson, & Haynes, 2011). Regardless of how the gap between research and practice is addressed, it is highly likely that training will be a key component.

CONCLUSION: CONVERGENCE AND FUTURE DIRECTIONS

There are a number of convergences between the three main research approaches, as well as with the other avenues noted above, that have attempted to integrate science and practice. Crucially, these studies are conducted in naturalistic settings rather than highly controlled research environments. As such, a priority is placed on external validity. In addition, with the exception of some descriptive and survey studies, the research conducted within each approach of the paradigm of practice-oriented research is based on the adoption and implementation of a standardized measurement system as part of routine practice (e.g., CCAPS, CORE, OQ, TOP). Furthermore, while the degree to which clinicians are involved in the design of the studies varies within and across the various approaches, four unifying goals of the paradigm of practice-oriented research can be identified. First, to provide practitioners with the opportunity to be *active participants* in scientific endeavors and have ownership of the data collected. This goal represents a fundamental shift away from traditional efficacy research. Second, to use data as it is being collected to inform their interventions during therapy, thereby seamlessly integrating or *confounding science and practice*. Third, to rely on this data, collected

individually or aggregated from groups of clinicians, to examine questions that *they perceive as relevant to their practice*. Hence, a key driver is the desire for local knowledge such that the results have a direct bearing and relevance on everyday practice. And fourth, to allow practitioners to *contribute to the accumulation of rigorous knowledge* aimed at better describing, understanding, and ultimately improving psychotherapy. This goal is consistent with the aims of more traditional psychotherapy research guided and conducted by full-time researchers and demonstrates the complementarity of evidence-based practice and practice-oriented research.

Interestingly, although a wide range of research topics has been investigated, a number of themes have emerged as the foci of studies in two, or all three, approaches of practice-oriented research described in this chapter. Among these themes are the following: therapist effects, dose-effect relationship, differential trajectories of change, description of routine practice and its effectiveness, as well as the investigation and comparison of therapeutic approaches. Moreover, a variety of research methodologies have been used, more or less extensively, in different types of practice-oriented research, including descriptive, correlational, and experimental designs. In addition, the application of multilevel modeling to large data sets has become an increasing hallmark of practice-oriented research.

It should be mentioned that with the conduct of practice-oriented research there frequently comes a number of methodological and pragmatic challenges. For example, there are considerable hurdles to implementing the routine monitoring of individual treatment progress that is at the core of many patient-focused and practice-based studies. These include administrative costs, the need to bring clinicians and administrative staff on board, the need for sufficient expertise to produce timely decision algorithms and reports, and training in how to interpret and evaluate the information provided. Ongoing consultation with researchers and among clinicians, help from research assistants in collecting and managing data, as well as a considerable amount of time devoted to the design and learning of research protocols are among the strategies that have been emphasized to prevent or resolve obstacles encountered in PRN studies. Like most meaningful professional tasks, the conduct of research for and by clinicians requires additional work and is facilitated by resources and funding that

are not always available in routine treatment settings.

As with any type of research, including RCTs, a number of limitations have been identified with studies conducted within the paradigm of practice-oriented research (see McMillen et al., 2009; Parry et al., 2010; Stiles et al., 2008a; Zarin et al., 1997). Some of these, especially with regard to internal validity, are not likely to be addressed in future investigations. For example, it is neither possible nor advisable for clinicians to have every prospective new client assigned to repeated and blind assessments before and after treatment, so that reliable judgment of the diagnoses of research participants can be ascertained. However, improvements could be made to address current limitations. For example, practice-based research can be criticized for the lack of assessment of therapists' treatment adherence and competence, thereby precluding confident statements about what interventions therapists used and how well they implemented them in studies investigating or comparing different forms of therapy. Although no doubt costly in terms of time and expertise, checklists or observer assessment of therapists' delivery of therapy should be considered in future studies, perhaps building on the therapist and client reports of techniques conducted session-by-session that have been collected in some PRN studies (e.g., Boswell et al., 2010). Crucially, however, such procedures need to be integrated into routine practice, thereby delivering on the confounding of research and practice activities.

Practice-oriented research has offered contributions to the field that are beneficial to therapists and their clients and should be recognized both by psychotherapy scholars, irrespective of their methodological preferences, and by policy makers. At a clinical level, for instance, repeated assessment and immediate feedback of a patient's mental health functioning during the course of therapy can alert therapists to patient non-response or negative response, support decisions on treatment planning and strategies (e.g., when and how to repair alliance ruptures), and help determine when treatment has been sufficient. These tools can be used by practitioners of all theoretical orientations and levels of experience when making complex and individualized decisions in their day-to-day practice.

At a scientific and policy-making level, the studies reported in this chapter set out both the yield and potential of an overarching paradigm

that goes part way to redressing the balance with trials methodology as well as promoting a strategy for ensuring better capture and use of data from routine practice; that is, addressing the collider effect (Kazdin, 2008). As demonstrated in studies described in the patient-focused section, for example, the quality of patient care has been improved for those patients who have been deemed to be off-track as a result of a program of trials carried out in routine settings and extended to routine settings as one component of good practice (see Lambert, 2010). As also illustrated in the practice-based section, the combination of multilevel modeling to reflect the hierarchical structure of the data (i.e., patients nested within practitioners, who are nested within services, see Adelson & Owen, 2012), together with very large data sets and intelligent data capture methods has reprivileged practitioners and services in generating robust evidence of their effectiveness. In addition, these data can be used to address service delivery issues not amenable to trials methodology. For example, therapist effects may be related to practitioners' abilities to retain a patient in treatment rather than dropping out, which may in turn be more crucial than differences between treatment orientations for those patients remaining in treatment. Similarly, dosage in trials is fixed whereas the weight of evidence from studies based in routine practice settings suggests that some patients require fewer sessions while others require more and that patients themselves may be the best arbiter of the question "how much therapy is enough." Also based on multilevel modeling methods, studies reported in the practice-research network section suggest that while some types of events can be perceived as particularly helpful (even when accounting for variability across sessions, patients, and therapists), some specific types of interventions (e.g., CBT techniques) can be associated with negative impact when used by particular therapists with particular clients.

Taken as a whole, these examples indicate that studies conducted using the three research approaches described in this chapter are increasingly able to capture and analyze data that reflect the complex structure and processes involved in delivering psychological therapies in the real world. This evidence base from routine practice needs to be considered in conjunction with RCT evidence by bodies informing national policies to realize the full potential from the chiasmus of evidence-based practice and practice-oriented evidence.

More issues remain to be investigated and further scientific, as well as clinical, advances are likely to be achieved with increased utilization of a diversity of research methods. Among the numerous recommendations that can be made for future research, a few appear particularly worthy of attention. First, practice-oriented collaborators might be encouraged to pursue studies exploring the interaction of participant (i.e., client and therapist) characteristics, relationship variables (e.g., alliance), and technical factors (common to several forms of psychotherapy or unique to particular approaches). Examining the moderating and mediating roles that some of these elements may have on treatment outcome could well capture intricate details of the change process in applied settings. Second, and complementing these complex quantitative analyses, we would also suggest that particular emphasis be given to extensive qualitative analyses of significant episodes during therapy. Third, large practice-oriented infrastructures have been described as optimal contexts to implement sophisticated experimental methods (e.g., additive, dismantling, and parametric designs) to simultaneously test the effectiveness of new approaches of psychotherapy and examine cause-effect relationships between interventions and outcome (see Borkovec & Castonguay, 1998), thereby offering another avenue for the seamless integration of clinical and scientific pursuits.

The same recommendations for future investigations can, and have been, made toward research conducted outside of the approaches described in this chapter (see Castonguay & Beutler, 2005b). Such convergence should be viewed as a warning sign of a possible false dichotomy between traditional (i.e., evidence-based) and practice-oriented research paradigms. Not only do these lines of research share important goals including, as mentioned above, the quest to better understand and improve psychotherapy, but they can also focus on similar issues and use identical research methods, including randomized clinical trials. This is exemplified in the trials designs used in the research on feedback (e.g., Harmon et al., 2007). There are, obviously, important distinctions between these two strategies in how studies are designed and conducted. For example, the starting point of traditional effectiveness studies—a relatively practice-oriented type of traditional trials-based research—is a manualized and well-controlled treatment that is then imported into the naturalistic environment.

By contrast, in most studies conducted within practice-oriented research, the starting point is ongoing psychotherapy that is then studied and sometimes manipulated, via randomized trials, to gradually improve its potency. However, perhaps the most important distinction is the guiding force of the investigations conducted. Whereas most traditional investigations are guided, and often funded, by the research programs of academicians, studies in practice-oriented infrastructure are based on the active participation of clinicians in collaborative research endeavors. This means that the recommendations for future research mentioned above should be viewed as tentative suggestions. Presenting them as priorities or imperatives would amount to empirical imperialism: researchers telling clinicians what to study and how to study it.

In terms of future directions, the most important issue in relation to practice-oriented research is not what studies need to be conducted, but what can be done to facilitate the collaboration of researchers and clinicians in designing and conducting studies in which they wish to invest their time and energy. On this, we can only provide a glimpse of ideals that could be pursued. On a pragmatic level, it would be beneficial to foster the following initiatives. First, embed practice-oriented research during training. This could be facilitated by implementing repeated measurements and the use of feedback within clinical training and supervision, as well as by providing opportunities for students to do research that is not only clinically relevant but that interfaces with their clinical experiences. As noted elsewhere, it could be argued “that simultaneous, seamless, and repeated integration of science and practice activities as early as possible in a psychotherapist’s career might create an intellectual and emotional (hopefully secure) attachment to principles and merits of the Boulder model” (Castonguay, 2011, p. 135). Second, “ask and tell” by surveying clinicians about what they want to know and what kind of study they would like to build and implement with others, and then publish the results of these studies to inform and stimulate the field into action. Third, work locally but collaborate globally. This could be done, for example, by creating large networks that connect smaller groups of clinicians and researchers collecting data on the same variables at different sites (see Borkovec, 2002; Castonguay, 2011). Fourth, use data that is already available—that is, archived or secondary data sets. Many studies can be carried

out by taking advantage of archival data open to researchers (e.g., NIDA, CCMH).

Finally, we would argue that the engagement of practice-oriented collaborators could be most fruitfully carried forward by the adoption of three guiding principles. First, cover the colander that leaks clinically based knowledge. To maximize the ability of practice-oriented research to generate new and actionable findings, let alone to foster its collaborative spirit, it would be most beneficial to conduct studies that address clinicians' questions and that are designed, in part, on their knowledge and observations. Second, as far as possible, avoid constructing studies requiring tasks that are not immediately relevant to the conduct of therapy. In other words, we need to think beyond the "bridge" between science and practice. Indeed, rather than trying to connect science and practice, as if they stand on different river banks, we should strive to confound the two activities to create a new, unified landscape of knowledge and action. And third, make the research for and by clinicians count. There is a robust argument to be made to funders and policy makers to ensure that the evidence derived from practice-oriented research contributes in equal measure to the development of national, local, and professional guidelines. Methodologists within the wider discipline of public health have argued that "if the health professions and their sponsors want more widespread and consistent evidence-based practice, they will need to find ways to generate more practice-based evidence that explicitly addresses external validity and local realities" (Green & Glasgow, 2006, p. 128). We would also argue that this is a two-way street. Although it is clear that our understanding and conduct of psychotherapy can be improved by the scientific contributions of practice-oriented research, clinicians are more likely to engage in designing, implementing, and disseminating studies if there is clear evidence that the merit and impact of these studies will be fairly considered and duly recognized by scholars, researchers, and policy makers.

REFERENCES

- Adelson, J. L., & Owen, J. (2012). Bringing the psychotherapist back: Basic concepts for reading articles examining therapist effects using multilevel modeling. *Psychotherapy, 49*, 152–162.
- American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev.). Washington, DC: American Psychiatric Association.
- American Psychological Association. (2006). Evidence-based practice in psychology: APA Presidential task force on evidence-based practice. *American Psychologist, 61*, 271–285.
- Anderson, E. M., & Lambert, M. J. (2001). A survival analysis of clinically significant change in outpatient psychotherapy. *Journal of Clinical Psychology, 57*, 875–888.
- 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 treatment setting. *Journal of Consulting and Clinical Psychology, 77*, 693–704.
- Ashaye, O. A., Livingston, G., & Orrell, M. W. (2003). Does standardized needs assessment improve the outcome of psychiatric day hospital care for older people? A randomized controlled trial. *Aging and Mental Health, 7*, 195–199.
- Baker, T. B., McFall, R. M., & Shoham, V. (2009). Current status and future prospects of clinical psychology: Toward a scientifically principled approach to mental and behavioral health care. *Psychological Science in the Public Interest, 9*, 67–103.
- Baldwin, S. A., Berkeljon, A., Atkins, D. C., Olsen, J. A., & Nielsen, S. L. (2009). Rates of change in naturalistic psychotherapy: Contrasting dose-effect and good-enough level models of change. *Journal of Consulting and Clinical Psychology, 77*, 203–211.
- Ball, S. A., Martino, S., Nich, C., Frankforter, T. L., Van Horn, D., Crits-Christoph, P., . . . Carroll, K. M. (2007). Site matters: Multisite randomized trial of motivational enhancement therapy in community drug abuse clinics. *Journal of Consulting and Clinical Psychology, 75*, 556–567.
- Barkham, M., Connell, J., Stiles, W. B., Miles, J. N. V., Margison, J., Evans, C., & Mellor-Clark, J. (2006). Dose-effect relations and responsive regulation of treatment duration: The good enough level. *Journal of Consulting and Clinical Psychology, 74*, 160–167.
- Barkham, M., Hardy, G. E., & Mellor-Clark, J. (Eds.). (2010). *Developing and delivering practice-based evidence: A guide for the psychological therapies*. Chichester, United Kingdom: Wiley.
- Barkham, M., & Margison, F. (2007). Practice-based evidence as a complement to evidence-based practice: From dichotomy to chiasmus. In C. Freeman & M. Power (Eds.), *Handbook of evidence-based psychotherapies: A guide for research and practice* (pp. 443–476). Chichester, United Kingdom: Wiley.
- Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., . . . McGrath, G. (2001). Service profiling and outcomes benchmarking using the CORE-OM: Towards practice-based evidence in the psychological therapies.

- Journal of Consulting and Clinical Psychology*, 69, 184–196.
- Barkham, M., Mellor-Clark, J., Connell, J., Evans, R., Evans, C., & Margison, F. (2010). The CORE measures & CORE system: Measuring, monitoring, and managing quality evaluation in the psychological therapies. In M. Barkham, G. E. Hardy, & J. Mellor-Clark (Eds.), *Developing and delivering practice-based evidence: A guide for the psychological therapies* (pp. 175–219). Chichester, United Kingdom: Wiley.
- Barkham, M., Rees, A., Stiles, W. B., Shapiro, D. A., Hardy, G. E., & Reynolds, S. (1996). Dose-effect relations in time-limited psychotherapy for depression. *Journal of Consulting and Clinical Psychology*, 64, 927–935.
- Barkham, M., Stiles, W. B., Connell, J., & Mellor-Clark, J. (2012). The meaning of treatment effectiveness in routine NHS primary care psychological therapy services. *Psychology and Psychotherapy: Theory, Research and Practice*, 85, 1–16.
- Barkham, M., Stiles, W. B., Connell, J., Twigg, E., Leach, C., Lucock, M., . . . Angus, L. (2008). Effects of psychological therapies in randomized trials and practice-based studies. *British Journal of Clinical Psychology*, 47, 397–415.
- Barkham, M., Stiles, W. B., Lambert, M. J., & Mellor-Clark, J. (2010). Building a rigorous and relevant knowledge-base for the psychological therapies. In M. Barkham, G. E. Hardy, & J. Mellor-Clark (Eds.), *Developing and delivering practice-based evidence: A guide for the psychological therapies* (pp. 21–61). Chichester, United Kingdom: Wiley.
- Barkham, M., Stiles, W. B., & Shapiro, D. A. (1993). The shape of change in psychotherapy: Longitudinal assessment of personal problems. *Journal of Consulting and Clinical Psychology*, 61, 667–677.
- Barlow, D. (Ed.). (2007). *Clinical handbook of psychological disorders*. New York, NY: Guilford Press.
- Barlow, D. H., & Nock, M. K. (2009). Why can't we be more idiographic in our research? *Perspectives on Psychological Science*, 4, 19–21.
- Bauer, S. (2004). Outcome monitoring and feedback: Strategies for optimisation of inpatient psychotherapy. Doctoral Thesis (PhD), Faculty of Social and Behavioral Sciences, University of Tübingen, Germany, 2004.
- Bech, P., Gudex, C., & Johansen, K. S. (1996). The WHO (Ten) well-being index: Validation in diabetes. *Psychotherapy and Psychosomatics*, 65, 183–190.
- Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York, NY: Guilford Press.
- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 561–571.
- Beck, J. G., Castonguay, L. G., Chronis-Tuscano, A., Klonsky, E. D., McGinn, L., & Youngstrom, E. (2012, November). *Giving students a fish versus teaching students to fish: Evidence-based training principles and practical guidelines*. Symposium presented at the 46th Annual Convention of the Association for Behavioral and Cognitive Therapies. National Harbor, MD.
- Beeson, P. M., & Robey, R. R. (2006). Evaluating single-subject treatment research: Lessons learned from the aphasia literature. *Neuropsychological Review*, 16, 161–169.
- Bergin, A. E., & Suinn, R. M. (1975). Individual psychotherapy and behavior therapy. *Annual Review of Psychology*, 26, 509–556.
- Berking, M., Orth, U., & Lutz, W. (2006). Effects of systematic feedback to the therapist on patient progress. An empirical study in a cognitive-behavioral inpatient setting (in German). *Zeitschrift für Klinische Psychologie und Psychotherapie*, 35, 21–29.
- 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.
- Bickman, L., Kelley, S. D., Breda, C., de Andrade, A. R., & Riemer, M. (2011). Effects of routine feedback to clinicians on mental health outcomes of youths: Results of a randomized trial. *Psychiatric Services*, 62, 1423–1429.
- Bickman, L., Riemer, M., Breda, C., & Kelley, S. D. (2006). CFIT: A system to provide a continuous quality improvement infrastructure through organizational responsiveness, measurement, training, and feedback. *Report on Emotional & Behavioral Disorders in Youth*, 6, 86–87, 93–94.
- Borckardt, J. J., Nash, M. R., Murphy, M. D., Moore, M., Shaw, D., & O'Neil, P. (2008). Clinical practice as natural laboratory for psychotherapy research: A guide to case-based time-series analysis. *American Psychologist*, 63, 77–95.
- Borkovec, T. D. (2002). Training clinic research and the possibility of a national training clinics practice research network. *Behavior Therapist*, 25, 98–103.
- Borkovec, T. D., & Castonguay, L. G. (1998). What is the scientific meaning of “Empirically Supported Therapy”? *Journal of Consulting and Clinical Psychology*, 66, 136–142.
- Borkovec, T. D., Echemendia, R. J., Ragusea, S. A., & Ruiz, M. (2001). The Pennsylvania practice research network and future possibilities for clinically meaningful and scientifically rigorous psychotherapy research. *Clinical Psychology: Science and Practice*, 8, 155–168.
- Boswell, J. F., Castonguay, L. G., & Wasserman, R. H. (2010). Effects of psychotherapy training and intervention use on session outcome. *Journal of Consulting and Clinical Psychology*, 78, 717–723.
- Boswell, J. F., McAleavey, A. A., Castonguay, L. G., Hayes, J. A., & Locke, B. D. (2012). The effect

- of previous mental health service utilization on change in counseling clients' depressive symptoms. *Journal of Counseling Psychology*, 59, 368–378.
- Brabec, B., & Meister, R. (2001). A nearest-neighbor model for regional avalanche forecasting. *Annals of Glaciology*, 32, 130–134.
- Bradley, L. J., Sexton, T. L., & Smith, H. B. (2005). The American counseling association practice research network (ACA-PRN): A new research tool. *Journal of Counseling & Development*, 83, 488–491.
- Brodey, B. B., Cuffel, B., McCulloch, J., Tani, S., Maruish, M., Brodey, I., & Unutzer, J. (2005). The acceptability and effectiveness of patient reported assessments and feedback in a managed behavioral healthcare setting. *American Journal of Managed Care*, 11, 774–780.
- Brookman-Frazee, L. I., Drahota, A., & Stadnick, N. (2011). Training community mental health therapists to deliver a package of evidence-based practice strategies for school-age children with autism spectrum disorders: A pilot study. *Journal of Autism and Developmental Disorders*. doi: 10.1007/s10803-011-1406-7
- Brookman-Frazee, L., Haine, R. A., Baker-Ericzen, M., Zoffness, R., & Garland, A. F. (2010). Factors associated with the use of EBP strategies in usual care youth psychotherapy. *Administration and Policy in Mental Health*, 37, 254–269.
- Brookman-Frazee, L., Stahmer, A., Searcy, L. K., Feder, J., Reed, S. (2012). Building a research-community collaborative to improve preventative care for children at-risk for autism spectrum disorders. *Journal of Community Psychology*, 40, 715–734.
- Brown, G. S., & Jones, E. R. (2005). Implementation of a feedback system in a managed care environment: What are patients teaching us? *Journal of Clinical Psychology/In Session*, 61, 187–198.
- Brown, G. S., Lambert, M. J., Jones, E. R., & Minami, T. (2005). Identifying highly effective therapists in a managed care environment. *American Journal of Managed Care*, 8, 513–520.
- Burlew, A. K., Larios, S., Suarez-Morales, L., Holmes, B. W., Venner, K. L., & Chavez, R. (2011). Increasing ethnic minority participation in substance abuse clinical trials: Lessons learned in the national institute on drug abuse's clinical trials network. *Cultural Diversity and Ethnic Minority Psychology*, 17, 345–356.
- Burns, D. D., & Nolen-Hoeksema, S. (1991). Copying styles, homework compliance, and the effectiveness of cognitive-behavioral therapy. *Journal of Consulting and Clinical Psychology*, 59, 305–311.
- Burns, D. D., & Nolen-Hoeksema, S. (1992). Therapeutic empathy and recovery from depression in cognitive-behavioral therapy: A structure equation model. *Journal of Consulting and Clinical Psychology*, 60, 441–449.
- Byrne, S. L., Hooke, G. R., Newnham, E. A., & Page, A. C. (2012). The effects of progress monitoring on subsequent readmission to psychiatric care: A six-month follow-up. *Journal of Affective Disorders*, 137, 113–116.
- Cahill, J., Barkham, M., & Stiles, W. B. (2010). Systematic review of practice-based research on psychological therapies in routine clinic settings. *British Journal of Clinical Psychology*, 49, 421–454.
- Cahill, J., Barkham, M., Stiles, W. B., Twigg, E., Rees, A., Hardy, G. E., & Evans, C. (2006). Convergent validity of the CORE measures with measures of depression for clients in brief cognitive therapy for depression. *Journal of Counseling Psychology*, 53, 253–259.
- Carlier, I. V. E., Meuldijk, D., Van Vliet, I. M., Van Fenema, E., Van der Wee, N. J. A., & Zitman, F. G. (2012). Routine outcome monitoring and feedback on physical or mental health status: evidence and theory. *Journal of Evaluation in Clinical Practice*, 18, 104–110.
- Carroll, K. M., Ball, S. A., Nich, C., Martino, S., Frankforter, T. L., Farentinos, C., . . . Woody, G. E. for the National Institute on Drug Abuse Clinical Trials Network. (2006). Motivational interviewing to improve treatment engagement and outcome in individuals seeking treatment for substance abuse: A multisite effectiveness study. *Drug and Alcohol Dependence*, 81, 301–312.
- Carroll, K. M., Martino, S., Ball, S. A., Nich, C., Frankforter, T. L., Anez, L., . . . Farentinos, C. (2009). A multisite randomized effectiveness trial of motivational enhancement therapy for Spanish-speaking substance users. *Journal of Consulting and Clinical Psychology*, 77, 993–999.
- Castonguay, L. G. (2011). Psychotherapy, psychopathology, research and practice: Pathways of connections and integration. *Psychotherapy Research*, 21, 125–140.
- Castonguay, L. G., Adam-Term, R., Cavanagh, T., Hill, B., Johnson, B., Magnavita, J., . . . Spayd, C. (2010). *What do clinicians think about research?* Structured discussion presented at the meeting of the Society for Psychotherapy Research, Asilomar, California.
- Castonguay, L. G., & Beutler, L. E. (Eds.). (2005a). *Principles of therapeutic change that work*. New York, NY: Oxford University Press.
- Castonguay, L. G., & Beutler, L. E. (2005b). Common and unique principles of therapeutic change: What do we know and what do we need to know? In L. G. Castonguay & L. E. Beutler (Eds.), *Principles of therapeutic change that work*. New York, NY: Oxford University Press.
- Castonguay, L. G., Boswell, J. F., Zack, S., Baker, S., Boutselis, M., Chiswick, N., . . . Grosse Holtforth, M. (2010). Helpful and hindering events in psychotherapy: A practice research network study. *Psychotherapy: Theory, Research, Practice, and Training*, 47, 327–344.
- Castonguay, L. G., Locke, B. D., & Hayes, J. A. (2011). The center for collegiate mental health: An example of a practice-research network in

- university counseling centers. *Journal of College Student Psychotherapy*, 25, 105–119.
- Castonguay, L., Nelson, D., Boutselis, M., Chiswick, N., Damer, D., Hemmelstein, N., . . . Borkovec, T. (2010). Clinicians and/or researchers? A qualitative analysis of therapists' experiences in a practice research network. *Psychotherapy: Theory, Research, Practice, and Training*, 47, 345–354.
- Castonguay, L. G., Pincus, A. L., Arnett, P. A., Roper, G., Rabian, R., & Borkovec, T. B. (2004, November). *Psychology training clinic as a research practice network: Integrating research and clinical practice in graduate school*. Paper presented at the annual meeting of the North American Society for Psychotherapy Research, Springdale, Arizona.
- Castonguay, L. G., Schut, A. J., Constantino, M. J., & Halperin, G. S. (1999). Assessing the role of treatment manuals: Have they become necessary but non-sufficient ingredients of change? *Clinical Psychology: Science and Practice*, 6, 449–455.
- Chambless, D. L., & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology*, 52, 685–716.
- Clark, D. M., Fairburn, C. G., & Wessely, S. (2008). Psychological treatment outcomes in routine NHS services: a commentary on Stiles et al. (2007). *Psychological Medicine*, 38, 629–634.
- Cohen, L. H., Sargent, M. M., & Sechrest, L. B. (1986). Use of psychotherapy research by professional psychologists. *American Psychologist, Special Issue: Psychotherapy Research*, 41, 198–206.
- Cooper, M. (2008). *Essential research findings in counseling and psychotherapy: The facts are friendly*. London, United Kingdom: Sage.
- Crits-Christoph, P., Baranckie, K., Kurcias, J., Beck, A., Carroll, K., Perry, K., . . . Zitrin, C. (1991). Meta-analysis of therapist effects in psychotherapy outcome studies. *Psychotherapy Research*, 1, 81–91.
- Crits-Christoph, P., & Mintz, J. (1991). Implications of therapist effects for the design and analysis of comparative studies of psychotherapy. *Journal of Consulting and Clinical Psychology*, 54, 20–26.
- Crits-Christoph, P., Ring-Kurtz, S., Hamilton, J. L., Lambert, M. J., Gallop, R., McClure, B., . . . Rotrosen, J. (2012). A preliminary study of the effects of individual patient-level feedback in outpatient substance abuse treatment programs. *Journal of Substance Abuse Treatment*, 42, 301–309.
- Crosbie, J. (1993). Interrupted time-series analysis with brief single-subject data. *Journal of Consulting and Clinical Psychology*, 61, 966–974.
- Dattilio F. M., Edwards, D. J. A., & Fishman, D. B. (2010). Case studies within a mixed methods paradigm: Toward a resolution of the alienation between researcher and practitioner in psychotherapy research. *Psychotherapy: Theory, Research Practice, Training*, 47, 427–441.
- Davis, D. M., & Hayes, J. A. (2011). What are the benefits of mindfulness? A practice review of psychotherapy-related research. *Psychotherapy*, 48, 198–208.
- Derogatis, L. R. (1993). *BSI. Brief symptom inventory. Administration, scoring, and procedures manual* (4th ed.). Minneapolis, MN: National Computer Systems.
- 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.
- Duffy, F. F., Chung, H., Trivedi, M., Rae, D. S., Regier, D. A., & Katzelnick, D. J. (2008). Systematic use of patient-rated depression severity monitoring: Is it helpful and feasible in clinical psychiatry? *Psychiatric Services*, 59, 1148–1154.
- Duncan, B. L., & Miller, S. D. (2008). *The Outcome and Session Rating Scales: The revised administration and scoring manual, including the Child Outcome Rating Scale*. Chicago: Institute for the Study of Therapeutic Change.
- Effrig, J., Bieschke, K. J., & Locke, B. D. (2011). Examining victimization and psychological distress in transgender college students. *Journal of College Counseling*, 14, 143–157.
- Elkin, I., Falconner, L., Martinovich, Z., & Mahoney, C. (2006). Therapist effects in the national institute of mental health treatment of depression collaborative research program. *Psychotherapy Research*, 16, 144–160.
- 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. *Archives of General Psychiatry*, 46, 971–992.
- Elliott, R. (1998). Editor's introduction: A guide to the empirically supported treatments controversy. *Psychotherapy Research*, 8, 115–125.
- Elliott, R. (2002). Hermeneutic single-case efficacy design. *Psychotherapy Research*, 12, 1–21.
- Elliott, R., & Morrow-Bradley, C. (1994). Developing a working marriage between psychotherapists and psychotherapy researchers: Identifying shared purposes. In P. F. Talley, H. H. Strupp, & S. F. Butler (Eds.), *Psychotherapy research and practice: Bridging the gap* (pp. 124–142). New York, NY: Basic Books.
- Elliott, R., Partyka, R., Wagner, J., Alperin, R., Dobrenski, R., Messer, S. B., . . . Castonguay, L. G. (2009). An adjudicated hermeneutic single case efficacy design study of experiential therapy for panic/phobia. *Psychotherapy Research*, 19, 543–557.
- Evans, C., Connell, J., Barkham, M., Margison, F., Mellor-Clark, J., McGrath, G., & Audin, K. (2002). Towards a standardised brief outcome measure: Psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry*, 180, 51–60.

- Evans, C., Connell, J., Barkham, M., Marshall, C., & Mellor-Clark, J. (2003). Practice-based evidence: Benchmarking NHS primary care counselling services at national and local levels. *Clinical Psychology & Psychotherapy, 10*, 374–388.
- Farrell, L. J., Schlup, B., & Boschen, M. J. (2010). Cognitive-behavioral treatment of childhood obsessive-compulsive disorder in community-based clinical practice: Clinical significance and benchmarking against efficacy. *Behaviour Research and Therapy, 48*, 409–417.
- Finch, A. E., Lambert, M. J., & Schaalje, B. G. (2001). Psychotherapy quality control: The statistical generation of expected recovery curves for integration into an early warning system. *Clinical Psychology and Psychotherapy, 8*, 231–242.
- Forcehimes, A. A., Venner, K. L., Bogenschutz, M. P., Foley, K., Davis, M. P., Houck, J. M., . . . Begaye, P. (2011). American Indian methamphetamine and other drug use in the southwestern United States. *Cultural Diversity and Ethnic Minority Psychology, 17*, 366–76.
- Froyd, J. E., Lambert, M. J., & Froyd, J. D. (1996). A review of practices of psychotherapy outcome measurement. *Journal of Mental Health, 5*, 11–16.
- Galano, J., & Schellenbach, C. J., (2007). Healthy families America research practice network: A unique partnership to integrate prevention sciences and practice. *Journal of Prevention and Integration in the Community, 34*, 39–66.
- Galassi, J. P., & Gersh, T. L. (1993). Myths, misconceptions, and missed opportunity: Single-case designs and counseling psychology. *Journal of Counseling Psychology, 40*, 525–531.
- Garland, A. F., Hurlburt, M. S., & Hawley, K. M. (2006). Examining psychotherapy processes in a services research context. *Clinical Psychology: Science and Practice, 13*, 30–46.
- Garland, A. F., Plemmons, D., & Koontz, L. (2006). Research-practice partnership in mental health: lessons from participants. *Administration and Policy in Mental Health and Mental Health Services, 33*, 517–528.
- Gibbard, I., & Hanley, T. (2008). A five-year evaluation of the effectiveness of person-centred counselling in routine clinical practice in primary care. *Counselling and Psychotherapy Research, 8*, 215–222.
- Gibbons, C. J., Fournier, J. C., Stirman, S. W., DeRubeis, R. J., Crits-Christoph, P., & Beck, A. T. (2010). The clinical effectiveness of cognitive therapy for depression in an outpatient clinic. *Journal of Affective Disorders, 125*, 169–176.
- Goldfried, M. R. (1984). Training the clinician as scientist-professional. *Professional Psychology: Research and Practice, 15*, 477–481.
- Goldfried, M. R. (2011). Generating research questions from clinical experience: Therapists' experiences in using CBT for panic disorder. *Behavior Therapist, 34*, 57–62.
- Goodheart, C. D., Kazdin, A. E., & Sternberg, R. J. (2006). *Evidence-based practice: Where practice and research meet*. Washington, DC: American Psychological Association.
- Grawe, K. (1997). Research-informed psychotherapy. *Psychotherapy Research, 7*, 1–19.
- Gray, G. E. (2004). *Evidence-based psychiatry*. Washington, DC: American Psychiatric Association.
- Green, L. W., & Glasgow, R. E. (2006). Evaluating the relevance, generalizability, and applicability of research: Issues in external validation and translational methodology. *Evaluation & the Health Professions, 29*, 126–153.
- Greenhalgh, J., Long, A. F., & Flynn, R. (2005). The use of patient reported outcome measures in routine clinical practice: Lack of impact or lack of theory? *Social Science & Medicine, 60*, 833–843.
- Grissom, G. R., Lyons, J. S., & Lutz, W. (2002). Standing on the shoulders of a giant: Development of an outcome management system based on the dose model and phase model of psychotherapy. *Psychotherapy Research, 12*, 397–412.
- Guyatt, G. H., & Rennie, D. (Eds.). (2002). *Users' guides to the medical literature*. Chicago, IL: AMA Press.
- Hannan, C., Lambert, M. J., Harmon, C., Nielsen, S. L., Smart, D. W., Shimokawa, K., & Sutton, S. W. (2005). A lab test and algorithms for identifying clients at risk for treatment failure. *Journal of Clinical Psychology: In Session, 61*, 155–163.
- Hansen, N., & Lambert, M. (2003). An evaluation of the dose-response relationship in naturalistic treatment settings using survival analysis. *Mental Health Services Research, 5*, 1–12.
- Hansen, N. B., Lambert, M. J., & Forman, E. M. (2002). The psychotherapy dose-response effect and implications for treatment delivery services. *Clinical Psychology: Science and Practice, 9*, 329–343.
- Hanson, G. R., Leshner, A. I., & Tai, B. (2002). Putting drug abuse research to use in real-life settings. *Journal of Substance Abuse Treatment, 23*, 69–70.
- Hardy, J. A., Weatherford, R. D., Locke, B. D., DePalma, N. H., & D'lusio, N. T. (2011). Meeting the demand for college student concerns in college counseling centers: Evaluating a clinical triage system. *Journal of College Student Psychotherapy, 25*, 220–240.
- Harmon, S. C., Lambert, M. J., Smart, D. W., Hawkins, E. J., Nielsen, S. L., Slade, K., & Lutz, W. (2007). Enhancing outcome for potential treatment failures: Therapist/client feedback and clinical support tools. *Psychotherapy Research, 17*, 379–392.
- Harnett, P., O'Donovan, A., & Lambert, M. J. (2010). The dose response relationship in psychotherapy: Implications for social policy. *Clinical Psychologist, 14*, 39–44.
- Hartzler, B., Donovan, D. M., & Huang, Z. (2010). Comparison of opiate-primary treatment seekers with and without alcohol use disorder. *Journal of Substance Abuse Treatment, 39*, 114–123.

- Hawkins, E. J., Lambert, M. J., Vermeersch, D. A., Slade, K., & Tuttle, K. (2004). The effects of providing patient progress information to therapists and patients. *Psychotherapy Research, 14*, 308–327.
- Hayes, J. A., Chun-Kennedy, C., Edens, A., & Locke, B. D. (2011). Do double minority students face double jeopardy? Testing minority stress theory. *Journal of College Counseling, 14*, 117–126.
- Hayes, J. A., Locke, B. E., & Castonguay, L. G. (2011). The center for collegiate mental health: Practice and research working together. *Journal of College Counseling, 14*, 101–104.
- Hayes, S. C., (1981). Single case experimental design and empirical clinical practice. *Journal of Consulting and Clinical Psychology, 49*, 193–211.
- Herbeck, D. M., Gitek, D. J., Svikis, D. S., Montoya, I. D., Marcus, S. C., & West, J. C. (2005). Treatment compliance in patients with comorbid psychiatric and substance use disorders. *American Journal on Addictions, 14*, 195–207.
- Herbeck, D. M., West, J. C., Ruditis, I., Duffy, F. F., Fitek, D. J., Bell, C. C., & Snowden, L. R. (2004). Variations in use of second-generation antipsychotic medication by race among adult psychiatric patients. *Psychiatric Services, 55*, 677–684.
- Horowitz, L. M., Rosenberg, S. E., Baer, B. A., Ureño, G., & Villaseñor, V. S. (1988). Inventory of interpersonal problems: Psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology, 56*, 885–892.
- Houghton, S., Saxon, D., Bradburn, M., Ricketts, T., & Hardy, G. E. (2010). The effectiveness of routinely delivered cognitive behavioural therapy for obsessive-compulsive disorder: A benchmarking study. *British Journal of Clinical Psychology, 49*, 473–489.
- Howard, K. I., Kopta, M., Krause, M. S., & Orlinsky, D. E. (1986). The dose-effect relationship in psychotherapy. *American Psychologist, 41*, 159–164.
- Howard, K. I., Lueger, R. J., Maling, M. S., & Martinovich, Z. (1993). A phase model of psychotherapy: Causal mediation of outcome. *Journal of Consulting and Clinical Psychology, 61*, 678–685.
- Howard, K. I., Moras, K., Brill, P., Martinovich, Z., & Lutz, W. (1996). The evaluation of psychotherapy. *American Psychologist, 52*, 1059–1064.
- Huet, V., Springham, N., & Evans, C. (2008). Art therapy practice research network. <http://www.baat.org/atprn.html>
- Iwakabe, S., & Gazzola, N. (2009). From single-case studies to practice-based knowledge: Aggregating and synthesizing case studies. *Psychotherapy Research, 19*, 601–611.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12–19.
- Jones, W. P. (2003). Single-case time series with Bayesian analysis: A practitioner's guide. *Measurement and Evaluation in Counseling and Development, 36*, 28–39.
- Joyce, A. S., Ogrodniczuk, J., Piper, W. E., & McCallum, M. (2002). A test of the phase model of psychotherapy change. *Canadian Journal of Psychiatry, 47*, 759–766.
- Kadera, S. W., Lambert, M. J., & Andrews, A. A. (1996). How much therapy is really enough: A session-by-session analysis of the psychotherapy dose-effect relationship. *Journal of Psychotherapy Practice and Research, 5*, 132–151.
- Kazdin, A. E. (2008). Evidence-based treatment and practice: New opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care. *American Psychologist, 63*, 146–159.
- Kazdin, A. E. (2009). Understanding how and why psychotherapy leads to change. *Psychotherapy Research, 19*, 418–428.
- Kazdin, A. E. (2010). *Single-case research designs: Methods for clinical and applied settings* (2nd ed.). New York, NY: Oxford University Press.
- Kazdin, A. E., Kratochwill, T. R., & VandenBos, G. R. (1986). Beyond clinical trials: Generalizing from research to practice. *Professional Psychology: Research and Practice, 17*, 391–398.
- Kim, D-M, Wampold, B. E., & Bolt, D. M. (2006). Therapist effects in psychotherapy: A random-effects modeling of the National Institute of Mental Health Treatment of Depression Collaborative Research Program data. *Psychotherapy Research, 16*, 161–172.
- Klerman, G. L., Weissman, M. M., Rounsaville, B. J., & Chevron, E. S. (1984). *Interpersonal therapy of depression*. New York, NY: Basic Books.
- Knaup, C., Koesters, M., Schoefer, D., Becker, T., & Puschner, B. (2009). Effect of feedback of treatment outcome in specialist mental healthcare: Meta-analysis. *British Journal of Psychiatry, 195*, 15–22.
- Knudsen, H. K., Ducharme, L., J., & Roman, P. M. (2007). Research participation and turnover intention: An exploratory analysis of substance abuse counselors. *Journal of Substance Abuse Treatment, 33*, 211–217.
- Kopta, S. M., Howard, K. I., Lowry, J. L., & Beutler, L. E. (1994). Patterns of symptomatic recovery in time-unlimited psychotherapy. *Journal of Consulting and Clinical Psychology, 62*, 1009–1016.
- Kopta, S. M., & Lowry, J. L. (2002). Psychometric evaluation of the behavioral health questionnaire-20: A brief instrument for assessing global mental health and the three phases of psychotherapy outcome. *Psychotherapy Research, 12*, 413–426.
- Kraus, D., & Castonguay, L. G. (2010). Treatment outcome package (TOP)—Development and use in naturalistic settings. In M. Barkham, G. E. Hardy, & J. Mellor-Clark (Eds.), *Developing and delivering practice-based evidence: A guide for the psychological*

- therapies (pp. 155–174). Chichester, United Kingdom: Wiley.
- Kraus, D. R., Castonguay, L., Boswell, J. F., Nordberg, S. S., & Hayes, J. A. (2011). Therapist effectiveness: Implications for accountability and patient care. *Psychotherapy Research, 21*, 267–276.
- Kraus, D. R., Seligman, D. A., & Jordan, J. R. (2005). Validation of a behavioral health treatment outcome and assessment tool designed for naturalistic settings: The treatment outcome package. *Journal of Clinical Psychology, 61*, 285–314.
- Krause, M. S., Howard, K. I., & Lutz, W. (1998). Exploring individual change. *Journal of Consulting and Clinical Psychology, 66*, 838–845.
- Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*, 606–613.
- Lambert, M. J. (2007). Presidential address: What we have learned from a decade of research aimed at improving psychotherapy outcome in routine care. *Psychotherapy Research, 17*, 1–14.
- Lambert, M. J. (2010). *Prevention of treatment failure: The use of measuring, monitoring, and feedback in clinical practice*. Washington, DC: American Psychological Association.
- Lambert, M. J., Hansen, N. B., & Finch, A. E. (2001). Patient-focused research: Using patient outcome data to enhance treatment effects. *Journal of Consulting and Clinical Psychology, 69*, 159–172.
- Lambert, M. J., Hansen, N. B., & Harmon, S. C. (2010). Outcome questionnaire system (the OQ system): Development and practical applications in health-care settings. In M. Barkham, G. E. Hardy, & J. Mellor-Clark (Eds.), *Developing and delivering practice-based evidence: A guide for the psychological therapies* (pp. 141–154). Chichester, United Kingdom: Wiley.
- Lambert, M. J., Lunnen, K., Umphress, V., Hansen, N., & Burlingame, G. M. (1994). Administration and scoring manual for the outcome questionnaire (OQ-45.1). Salt Lake City, UT: IHC Center for Behavioral Healthcare Efficacy.
- Lambert, M. J., & Ogles, B. M. (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. *Psychotherapy, 48*, 72–79.
- Lambert, M. J., Whipple, J. L., Bishop, M. J., Vermeersch, D. A., Gray, G. V., & Finch, A. E. (2002). Comparison of empirically derived and rationally-derived methods for identifying patients at risk for treatment failure. *Clinical Psychology and Psychotherapy, 9*, 149–164.
- Lambert, M. J., Whipple, J. L., Hawkins, E. J., Vermeersch, D. A., Nielsen, S. L., & Smart, D. W. (2003). Is it time for clinicians to routinely track patient outcome?: A meta-analysis. *Clinical Psychology: Science & Practice, 10*, 288–301.
- Lambert, M. J., Whipple, J. L., Smart, D. W., Vermeersch, D. A., Nielsen, S. L., & Hawkins, E. J. (2001). The effects of providing therapists with feedback on client progress during psychotherapy: Are outcomes enhanced? *Psychotherapy Research, 11*, 49–68.
- Lambert, M. J., Whipple, J. L., Vermeersch, D. A., Smart, D. W., Hawkins, E. J., Nielsen, S. L., & Goates, M. K. (2002). Enhancing psychotherapy outcomes via providing feedback on client progress: A replication. *Clinical Psychology and Psychotherapy, 9*, 91–103.
- Layard, R. (2006). The case for psychological treatment centres. *British Medical Journal, 332*, 1030–1032.
- Lebow, J. (2006). *Research for the psychotherapist: From science to practice*. New York, NY: Routledge.
- Lockard, A. J., Hayes, J. A., McAleavey, A. A., & Locke, B. D. (2012). Change in academic distress: Examining differences between a clinical and nonclinical sample of college students. *Journal of College Counseling, 15*, 233–246. doi: 10.1002/j.2161-1882.2012.00018.x
- Locke, B. D., Bieschke, K. J., Castonguay, L. G., & Hayes, J. A. (2012). The center for collegiate mental health (CCMH): Studying college student mental health through an innovative research infrastructure that brings science and practice together. *Harvard Review of Psychiatry, 20*, 233–245.
- Locke, B. D., Buzolitz, J. S., Dowis, J. D., Lei, P., Boswell, J. F., McAleavey, A. A., . . . Hayes, J. A. (2011). Development of the counseling center assessment of psychological symptoms. *Journal of Counseling Psychology, 58*, 97–109.
- Locke, B. D., McAleavey, A. A., Zhao, Y., Lei, P. W., Hayes, J. A., Castonguay, L. G., . . . Lin, Y. C. (2012). Development and initial validation of the counseling center assessment of psychological symptoms-34 (CCAPS-34). *Measurement and Evaluation in Counseling and Development, 45*, 151–169.
- Lueger, R. J., & Barkham, M. (2010). Using benchmarks and benchmarking to improve quality of practice and services. In M. Barkham, G. E. Hardy, & J. Mellor-Clark (Eds.), *Developing and delivering practice-based evidence: A guide for the psychological therapies* (pp. 223–256). Chichester, United Kingdom: Wiley.
- Lueger, R. J., Howard, K. I., Martinovich, Z., Lutz, W., Anderson, E. E., & Grissom, G. (2001). Assessing treatment progress of individual patients using expected treatment response models. *Journal of Consulting and Clinical Psychology, 69*, 150–158.
- Lutz, W. (2002). Patient-focused psychotherapy research and individual treatment progress as scientific groundwork for an empirical based clinical practice. *Psychotherapy Research, 12*, 251–273.

- Lutz, W., Böhnke, J., & Köck, K. (2011). Lending an ear to feedback systems: Evaluation of recovery and non-response in psychotherapy in a German outpatient setting. *Community Mental Health Journal*, *47*, 311–317.
- Lutz, W., Böhnke, J. R., Köck, K., & Bittermann, A. (2011). Diagnostik und psychometrischer Verlaufsrückmeldungen im Rahmen eines Modellprojektes zur Qualitätssicherung in der ambulanten Psychotherapie [Diagnostic and psychometric feedback within a pilot project on quality assurance in outpatient psychotherapy]. *Zeitschrift für Klinische Psychologie und Psychotherapie*, *40*, 283–297.
- Lutz, W., Lambert, M. J., Harmon, S. C., Tschitsaz, A., Schürch, E., & Stulz, N. (2006). The probability of treatment success, failure and duration—What can be learned from empirical data to support decision making in clinical practice? *Clinical Psychology and Psychotherapy*, *13*, 223–232.
- Lutz, W., Leach, C., Barkham, M., Lucock, M., Stiles, W. B., Evans, C., . . . Iverson, S. (2005). Predicting change for individual psychotherapy clients based on their nearest neighbors. *Journal of Consulting and Clinical Psychology*, *73*, 904–913.
- Lutz, W., Leon, S. C., Martinovich, Z., Lyons, J. S., & Stiles, W. B. (2007). Therapist effects in outpatient psychotherapy: A three-level growth curve approach. *Journal of Counseling Psychology*, *54*, 32–39.
- Lutz, W., Martinovich, Z., & Howard, K. I. (1999). Patient profiling: An application of random coefficient regression models to depicting the response of a patient to outpatient psychotherapy. *Journal of Consulting and Clinical Psychology*, *67*, 571–577.
- Lutz, W., Martinovich, Z., Howard, K. I., & Leon, S. C. (2002). Outcomes management, expected treatment response and severity adjusted provider profiling in outpatient psychotherapy. *Journal of Clinical Psychology*, *58*, 1291–1304.
- Lutz, W., Saunders, S. M., Leon, S. C., Martinovich, Z., Kosfelder, J., Schulte, D., . . . Tholen, S. (2006). Empirically and clinically useful decision making in psychotherapy: Differential predictions with treatment response models. *Psychological Assessment*, *18*, 133–141.
- Lutz, W., Stulz, N., & Köck, K. (2009). Patterns of early change and their relationship to outcome and follow-up among patients with major depressive disorders. *Journal of Affective Disorders*, *118*, 60–68.
- Lutz, W., Stulz, N., Martinovich, Z., Leon, S., & Saunders, S. M. (2009). Methodological background of decision rules and feedback tools for outcomes management in psychotherapy. *Psychotherapy Research*, *19*, 502–510.
- Lutz, W., Stulz, N., Smart, D. W., & Lambert, M. J. (2007). Die Identifikation früher Veränderungsmuster in der ambulanten Psychotherapie [Patterns of early change in outpatient therapy]. *Zeitschrift für Klinische Psychologie und Psychotherapie*, *36*, 93–104.
- Maling, M. S., Gurtman, M. B., & Howard, K. I. (1995). The response of interpersonal problems to varying doses of psychotherapy. *Psychotherapy Research*, *5*, 63–75.
- Marriott, M., & Kellett, S. (2009). Evaluating a cognitive analytic therapy service; practice-based outcomes and comparisons with person-centred and cognitive-behavioural therapies. *Psychology and Psychotherapy: Theory, Research and Practice*, *82*, 57–82.
- Marshall, M., Lockwood, A., Green, G., Zajac-Roles, G., Roberts, C. & Harrison, G. (2004). Systematic assessments of need and care planning in severe mental illness. *Cluster randomised controlled trial. British Journal of Psychiatry*, *185*, 163–168.
- Martin, J. L., Hess, T. R., Ain, S. C., Nelson, D. L., & Locke, B. D. (2012). Collecting multi-dimensional client data using repeated measures: Experiences of clients and counselors using the CCAPS-34. *Journal of College Counseling*, *15*, 247–261. doi: 10.1002/j.2161-1882.2012.00019.x
- Martindale, C. (1978). The therapist-as-fixed-effect fallacy in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *46*, 1526–1530.
- Mayotte-Blum, J., Slavin-Mulford, J., Lehmann, M., Pesale, F., Becker-Matero, N., & Hilsenroth, M. (2011). Therapeutic immediacy across long-term psychodynamic psychotherapy: An evidence-based case study. *Journal of Counseling Psychology*, *59*, 27–40.
- McAleavey, A. A., Castonguay, L. G., & Locke, B. D. (2011). Sexual orientation minorities in college counseling: Prevalence, distress, and symptom profiles. *Journal of College Counseling*, *14*, 127–142.
- McAleavey, A. A., Nordberg, S. S., Hayes, J. A., Castonguay, L. G., Locke, B. D., & Lockard, A. J. (2012, September 3). Clinical validity of the Counseling Center Assessment of Psychological Symptoms-62 (CCAPS-62): Further evaluation and clinical applications. *Journal of Counseling Psychology*. Advance online publication. doi: 10.1037/a0029855
- McAleavey, A. A., Nordberg, S. S., Kraus, D. R., & Castonguay, L. G. (2012). Errors in treatment outcome monitoring: Implications for real-world psychotherapy. *Canadian Psychology*, *53*, 105–114.
- McMillen, J. C., Lenze, S. L., Hawley, K. M., & Osbourne, V. A. (2009). Revisiting practice based research networks as a platform for mental health services research. *Administration and Policy in Mental Health*, *36*, 308–321.
- Mellor-Clark, J., & Barkham, M. (2006). The CORE system: Developing and delivering practice-based evidence through quality evaluation. In C. Feltham & I. Horton (Eds.), *Handbook of counselling and psychotherapy* (2nd ed., pp. 207–224). London, United Kingdom: Sage.

- Merrill, K. A., Tolbert, V. E., & Wade, W. A. (2003). Effectiveness of cognitive therapy for depression in a community mental health center: A benchmarking study. *Journal of Consulting and Clinical Psychology, 71*, 404–409.
- Miller, S. D., Duncan, B. L., Brown, J., Sparks, J. A., & Claud, D. A. (2003). The outcome rating scale: A preliminary study of the reliability, validity, and feasibility of a brief visual analog measure. *Journal of Brief Therapy, 2*, 91–100.
- Miller, S. D., Duncan, B. L., Sorrell, R., & Brown, G. S. (2005). The partners for change outcome system. *Journal of Clinical Psychology: In Session, 61*, 199–208.
- Minami, T., Serlin, R. C., Wampold, B. E., Kircher, J. C., & Brown, G. S. (2006). Using clinical trials to benchmark effects produced in clinical practice. *Quality & Quantity, 42*, 513–525.
- Minami, T., Wampold, B. E., Serlin, R. C., Kircher, J. C., & Brown, G. S. (2007). Benchmarks for psychotherapy efficacy in adult major depression. *Journal of Consulting and Clinical Psychology, 75*, 232–243.
- Montgomery, L., Burlew, A. K., Kosinski, A. S., & Forcehimes, A. A. (2011). Motivational enhancement therapy for African American substance users: A randomized clinical trial. *Cultural Diversity and Ethnic Minority Psychology, 17*, 357–365.
- Morgan-Lopez, A. A., Saavedra, L. M., Hien, D. A., & Fals-Stewart, W. (2010). Estimating statistical power for open-enrollment group treatment trials. *Journal of Substance Abuse Treatment, 40*, 3–17.
- Morley, S., Williams, A., & Hussain, S. (2008). Estimating the clinical effectiveness of cognitive behavioural therapy in the clinic: Evaluation of a CBT informed pain management programme. *Pain, 137*, 670–680.
- Morrow-Bradley, C., & Elliott, R. (1986). Utilization of psychotherapy research by practicing psychotherapists. *American Psychologist, Special Issue: Psychotherapy Research, 41*, 188–197.
- Murphy, G. E., Simons, A. D., Wetzel, R. D., & Lustman, P. J. (1984). Cognitive therapy and pharmacotherapy, singly and together, in the treatment of depression. *Archives of General Psychiatry, 41*, 33–41.
- Muthén, B. O. (2001). Second-generation structural equation modeling with a combination of categorical and continuous latent variables. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 291–332). Washington, DC: American Psychological Association.
- Muthén, B. O. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *Handbook of quantitative methodology for social sciences* (pp. 345–368). Newbury Park, CA: Sage.
- Nathan, P. E., & Gorman, J. M. (2002). (Eds.). *A guide to treatments that work* (2nd ed.). New York, NY: Oxford University Press.
- National Drug Abuse Treatment Clinical Trials Network. (2011, December 21). *Dissemination library*. Retrieved from <http://ctndisseminationlibrary.org/>
- Nelson, D. L., Castonguay, L. G., & Locke, B. D. (2011). Challenging stereotypes of eating and body image concerns among college students: Implications for diagnosis and treatment of diverse populations. *Journal of College Counseling, 14*, 158–172.
- Newnham, E. A., Hooke, G. R., & Page, A. C. (2010). Progress monitoring and feedback in psychiatric care reduces depressive symptoms. *Journal of Affective Disorders, 127*, 139–146.
- Newnham, E. A., & Page, A. C. (2010). Bridging the gap between best evidence and best practice in mental health. *Clinical Psychology Review, 30*, 127–142.
- Norcross, J. C. (2011). *Psychotherapeutic relationships that work* (2nd ed.). New York, NY: Oxford University Press.
- Norcross, J. C., Beutler, L. E., & Levant, R. F. (2006). *Evidence-based practice in mental health: Debate and dialogue on the fundamental questions*. Washington, DC: American Psychological Association.
- Nordberg, S. S., Castonguay, L. G., Fisher, A. J., Boswell, J. F., & Kraus, D. (2008, June). *Measuring the shape of change: Results from three years of repeated assessment in a university clinic*. Paper presented as part of a symposium at the annual meeting of the Society for Psychotherapy Research. Barcelona, Spain.
- Northey, W. F., Jr. (2002). Characteristics and clinical practices of marriage and family therapists: A national survey. *Journal of Marital and Family Therapy, 28*, 487–494.
- Oei, T. P. S., & Boschen, M. J. (2009). Clinical effectiveness of a cognitive behavioral group treatment program for anxiety disorders: A benchmarking study. *Journal of Anxiety Disorders, 23*, 950–957.
- Okiishi, J. C., Lambert, M. J., Eggett, D., Nielson, S. L., Dayton, D. D., & Vermeersch, D. A. (2006). An analysis of therapist treatment effects: Toward providing feedback to individual therapists on their patients' psychotherapy outcome. *Journal of Clinical Psychology, 62*, 1157–1172.
- Okiishi, J., Lambert, M. J., Nielsen, S. L., & Ogles, B. M. (2003). Waiting for supershrink: An empirical analysis of therapist effects. *Clinical Psychology & Psychotherapy, 10*, 361–373.
- Paley, G., Cahill, J., Barkham, M., Shapiro, D. A., Jones, J., Patrick, S., & Reid, E., (2008). The effectiveness of psychodynamic-interpersonal psychotherapy in routine clinical practice: A benchmarking comparison. *Psychology and Psychotherapy: Theory, Research and Practice, 85*, 157–175.
- Parry, G., Castonguay, L. G., Borkovec, T. D., & Wolfe, A. W. (2010). Practice research networks and psychological services research in the UK and USA. In M. Barkham, G. E. Hardy, & J. Mellor-Clark (Eds.), *Developing and delivering practice-based*

- evidence: *A guide for the psychological therapies*. Hoboken, NJ: Wiley.
- Persons, J. B., Bostrom, A., & Bertagnolli, A. (1999). Results of randomized controlled trials of cognitive therapy for depression generalize to private practice. *Cognitive Therapy and Research*, *23*, 535–548.
- Persons, J. B., Burns, D. D., & Perloff, J. M. (1988). Predictors of dropout and outcome in cognitive therapy for depression in a private practice setting. *Cognitive Therapy and Research*, *12*, 557–575.
- Persons, J. B., Roberts, N. A., Zalecki, C. A., & Brechwald, W. A. G. (2006). Naturalistic outcome of case-formulation-driven cognitive-behavioral therapy for anxiety depressed outpatients. *Behaviour Research and Therapy*, *44*, 1041–1051.
- Persons, J. B., & Silberschatz, G. (1998). Are results of randomized controlled trials useful to psychotherapists? *Journal of Consulting and Clinical Psychology*, *66*, 126–135.
- Pinsoff, W., Breunlin, D. C., Russell, W. P., & Lebow, J. (2011). Integrative problem-centered metaframeworks therapy II: Planning, conversing, and reading feedback. *Family Process*, *50*, 314–336.
- Pinto, R. M., Yu, G., Spector, A. Y., Gorroochurn, P., & McCarty, D. (2010). Substance abuse treatment providers' involvement in research is associated with willingness to use findings in practice. *Journal of Substance Abuse Treatment*, *39*, 188–194.
- Raimy, V. (Ed.). (1950). *Training in clinical psychology*. New York, NY: Prentice-Hall.
- Rawlins, M. D. (2008). De testimonio: On the evidence for decisions about the use of therapeutic interventions. The Harveian Oration of 2008. Royal College of Physicians. Reprinted in *The Lancet* (2008), *372*, 2152–2161.
- Reese, R. J., Norsworthy, L., & Rowlands, S. (2009). Does a continuous feedback system improve psychotherapy outcome? *Psychotherapy: Theory, Research, Practice, & Training*, *46*, 418–431.
- Reese, R. J., Toland, M. D., & Hopkins, N. B. (2011). Replicating and extending the good enough level model of change: Considering session frequency. *Psychotherapy Research*, *21*, 608–619.
- Reese, R. J., Toland, M. D., Slone, N. C., & Norsworthy, L. (2010). Effect of client feedback on couple psychotherapy outcomes. *Psychotherapy Theory, Research, Practice, Training*, *47*, 616–630.
- Richards, D. A., & Suckling, R. (2009). Improving access to psychological therapies: Phase IV prospective cohort study. *British Journal of Clinical Psychology*, *48*, 377–396.
- Ricks, D. F. (1974). Supershrink: Methods of a therapist judged successful on the basis of adult outcomes of adolescent patients. In D. F. Ricks, M. Roff, & A. Thomas (Eds.), *Life history in research in psychopathology*. Minneapolis: University of Minnesota Press.
- Robbins, M. S., Feaster, D. J., Horigian, V. E., Rohrbaugh, M., Shoham, V., Bachrach, K., . . . Szapocznik, J. (2011). Brief strategic family therapy versus treatment as usual: Results of a multi-site randomized trial for substance using adolescents. *Journal of Consulting and Clinical Psychology*, *79*, 713–727.
- Ruiz, M. A., Pincus, A. L., Borkovec, T. B., Echemendia, R., Castonguay, L. G., & Ragusea, S., (2004). Validity of the inventory of interpersonal problems (IIP-C) for predicting treatment outcome: An investigation with the Pennsylvania practice research network. *Journal of Personality Assessment*, *83*, 213–222.
- Rush, A. J., Beck, A. T., Kovacs, M., & Hollon, S. (1977). Comparative efficacy of cognitive therapy and pharmacotherapy in the treatment of depressed outpatients. *Cognitive Therapy and Research*, *1*, 17–38.
- Saxon, D., & Barkham, M. (2012). Patterns of therapist variability: Therapist effects and the contribution of patient severity and risk. *Journal of Consulting and Clinical Psychology*, *80*, 535–546.
- Schmidt, U., Landau, S., Pombo-Carril, M. G., Bara-Carril, N., Reid, Y., Murray, K., . . . Katzman, K. (2006). Does personalized feedback improve the outcome of cognitive-behavioural guided self-care in bulimia nervosa? A preliminary randomized controlled trial. *British Journal of Clinical Psychology*, *45*, 111–121.
- Schulte, D. (1998). *Therapieplanung [Treatment planning]* (2nd ed.). Göttingen: Hogrefe.
- Simmons, D. S., & Doherty, W. J. (1995). Defining who we are and what we do: Clinical practice patterns of marriage and family therapists in Minnesota. *Journal of Marital and Family Therapy*, *21*, 403–407.
- Shimokawa, K., Lambert, M. J., & Smart, D. W. (2010). Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *Journal of Consulting and Clinical Psychology*, *78*, 298–311.
- Silverstein, L. B., Auerbach, C. F., & Levant, R. F. (2006). Using qualitative research to strengthen clinical practice. *Professional Psychology: Research and Practice*, *37*, 351–358.
- Slade, K., Lambert, M. J., Harmon, S. C., Smart, D. W., & Bailey, R. (2008). Improving psychotherapy outcome: The use of immediate electronic feedback and revised clinical support tools. *Clinical Psychology & Psychotherapy*, *15*, 287–303.
- Slade, M., McCrone, P., Kuipers, E., Leese, M., Cahill, S., Parabiaghi, A., . . . Thornicroft, G. (2006). Use of standardised outcome measures in adult mental health services. *British Journal of Psychiatry*, *189*, 330–334.
- Smith, H. B., Sexton, T. L., & Bradley, L. J. (2005). The practice research network: Research into practice, practice into research. *Counselling and Psychotherapy Research*, *5*, 285–290.

- Smith, M. J. W., Whitaker, T., & Weismiller, T. (2006). Social workers in the substance abuse treatment field: A snapshot of service activities. *Health & Social Work, 31*, 109–115.
- Soldz, S. (2006). Models and meanings: Therapist effects and the stories we tell. *Psychotherapy Research, 16*, 173–177.
- Soldz, S., & McCullough, L. (Eds.). (2000). *Reconciling empirical knowledge and clinical experience: The art and science of psychotherapy*. Washington, DC: American Psychological Association.
- Speer, D. C., & Greenbaum, P. E. (1995). Five methods for computing significant individual client change and improvement rates: Support for an individual growth curve approach. *Journal of Consulting and Clinical Psychology, 63*, 1044–1048.
- Sperry, L., Brill, P. L., Howard, K. I., & Grissom, G. R. (1996). *Treatment outcomes in psychotherapy and psychiatric interventions*. New York, NY: Brunner/Mazel.
- Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine, 166*, 1092–1097.
- Stahmer, A. C., Brookman-Frazee, L., Lee, E., Searcy, K., & Reed, S. (2011). Parent and multidisciplinary provider perspectives on earliest intervention for children at risk for autism spectrum disorders. *Infants & Young Children, 24*, 344–363.
- Steinfeld, B. I., Coffman, S. J., & Keyes, J. A. (2009). Implementation of an evidence-based practice in a clinical setting: What happens when you get there? *Professional Psychology: Research and Practice, 40*, 410–416.
- Stewart, R. E., & Chambless, D. L. (2009). Cognitive-behavioral therapy for adult anxiety disorders in clinical practice: A meta-analysis of effectiveness studies. *Journal of Consulting and Clinical Psychology, 77*, 595–606.
- Stiles, W. B. (2007). Theory-building cases studies of counselling and psychotherapy. *Counselling and Psychotherapy Research, 7*, 122–127.
- Stiles, W. B. (2010). Theory-building case studies as practice-based evidence. In M. Barkham, G. E. Hardy, & J. Mellor-Clark (Eds.), *Developing and delivering practice-based evidence: A guide for the psychological therapies* (pp. 91–108). Chichester, United Kingdom: Wiley-Blackwell.
- Stiles, W. B., Barkham, M., Connell, J., & Mellor-Clark, J. (2008). Responsive regulation of treatment duration in routine practice in United Kingdom primary care settings. *Journal of Consulting and Clinical Psychology, 76*, 298–305.
- Stiles, W. B., Barkham, M., Mellor-Clark, J., & Connell, J. (2008a). Effectiveness of cognitive-behavioural, person-centred, and psychodynamic therapies in UK primary care routine practice: Replication with a larger sample. *Psychological Medicine, 38*, 677–688.
- Stiles, W. B., Barkham, M., Mellor-Clark, J., & Connell, J. (2008b). Routine psychological treatment and the Dodo Verdict: A rejoinder to Clark et al. (2007). *Psychological Medicine, 38*, 905–910.
- Stiles, W. B., Barkham, M., Twigg, E., Mellor-Clark, J., & Cooper, M. (2006). Effectiveness of cognitive-behavioural, person-centred, and psychodynamic therapies as practiced in UK national health service settings. *Psychological Medicine, 36*, 555–566.
- Straus, S. E., Glasziou, P., Richardson, W. S., & Haynes, R. B. (2011). *Evidence-based medicine: How to practice and teach EBM* (4th ed.). New York, NY: Churchill Livingstone.
- Stricker, G., & Trierweiler, S. J. (1995). The local clinical scientist: A bridge between science and practice. *American Psychologist, 50*, 995–1002.
- Stulz, N., & Lutz, W. (2007). Multidimensional patterns of change in outpatient psychotherapy: The phase model revisited. *Journal of Clinical Psychology, 63*, 817–833.
- Stulz, N., Lutz, W., Leach, C., Lucock, M., & Barkham, M. (2007). Shapes of early change in psychotherapy under routine outpatient conditions. *Journal of Consulting and Clinical Psychology, 75*, 864–874.
- Suarez-Morales, L., Martino, S., Bedregal, L., McCabe, B. E., Cuzmar, I. Y., Paris, M., Jr., . . . Szapocznik, J. (2010). Do therapist cultural characteristics influence the outcome of substance abuse treatment for Spanish-speaking adults? *Cultural Diversity and Ethnic Minority Psychology, 16*, 199–205.
- Tai, B., Straus, M. M., Liu, D., Sparenborg, S., Jackson, R., & McCarty, D. (2010). The first decade of the national drug abuse treatment clinical trials network: Bridging the gap between research and practice to improve drug abuse treatment. *Journal of Substance Abuse Treatment, 38*(Suppl. 1), S4–S13.
- Talley, P. F., Strupp, H. H., & Butler, S. F. (1994). (Eds). *Psychotherapy research and practice: Bridging the gap*. New York, NY: Basic Books.
- Thompson-Brenner, H., Boisseau, C. L., & Satir, D. A. (2010). Adolescent eating disorders: Treatment and response in a naturalistic study. *Journal of Clinical Psychology, 66*, 277–301.
- Trudeau, L. S. (2001). *Effects of a clinical feedback system on client and therapist outcomes in a rural community mental health center*. Doctoral Thesis (PhD), Institute for Social and Behavioral Research, Iowa State University, Iowa.
- Upadhyay, J., Maleki, N., Potter, J. S., Elman, I., Rudrauf, D., Knudsen, J., . . . Borsook, D. (2010). Alterations in brain structure and functional connectivity in prescription opioid-dependent patients. *Brain, 133*, 2098–2114.
- Wade, W. A., Treat, T. A., & Stuart, G. L. (1998). Transporting an empirically supported treatment for panic disorder to a service clinic setting: A benchmarking strategy. *Journal of Consulting and Clinical Psychology, 66*, 231–239.

- Walker, N. R., Rosvall, T., Field, C. A., Allen, S., McDonald, D., Salim, Z., . . . Adinoff, B. H. (2010). Disseminating contingency management to increase attendance in two community substance abuse treatment centers: Lessons learned. *Journal of Substance Abuse Treatment, 39*, 202–209.
- Walwyn, R., & Roberts, C. (2010). Therapist variation within randomised trials of psychotherapy: Implications for precision, internal and external validity. *Statistical Methods in Medical Research, 19*, 291–315.
- Wampold, B. E. (2001). *The great psychotherapy debate: Models, methods, and findings*. Mahwah, NJ: Erlbaum.
- 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.
- Weersing, V. R., Ivengar, S., Kolko, D. J., Birmaher, B., & Brent, D. A. (2006). Effectiveness of cognitive-behavioral therapy for adolescent depression: A benchmarking investigation. *Behavior Therapy, 37*, 36–48.
- West, J. C., Wilk, J. E., Rae, D. S., Narrow, W. E., & Regier, D. A. (2003). Economic grand rounds: Financial disincentives for the provision of psychotherapy. *Psychiatric Services, 54*, 1582–1583, 1588.
- West, J. C., & Zarin, D. A. (1995). Practice-relevant research findings: APA's psychiatric research network. *Behavioral Healthcare Tomorrow, 4*(3), 38–39.
- Westbrook, D., & Kirk, J. (2005). The clinical effectiveness of cognitive behaviour therapy: Outcome for a large sample of adults treated in routine practice. *Behaviour Research and Therapy, 45*, 1243–1261.
- Westen, D., Shedler, J., Durrett, C., Glass, S., & Martens, A. (2003). Personality diagnoses in adolescence: DSM-IV axis II diagnoses and an empirically derived alternative. *The American Journal of Psychiatry, 160*, 952–966.
- Whipple, J. L., & Lambert, M. J. (2011). Outcome measures for practice. *Annual Review of Clinical Psychology, 7*, 87–111.
- Whipple, J. L., Lambert, M. J., Vermeersch, D. A., Smart, D. W., Nielsen, S. L., & Hawkins, E. J. (2003). Improving the effects of psychotherapy: The use of early identification of treatment failure and problem solving strategies in routine practice. *Journal of Counseling Psychology, 58*, 59–68.
- Wilk, J. E., West, J. C., Rae, D. S., & Regier, D. A. (2006). Patterns of adult psychotherapy in psychiatric practice. *Psychiatric Services, 57*, 472–476.
- Youn, S. J., Kraus, D. R., & Castonguay, L. G. (2012). The treatment outcome package: Facilitating practice and clinically relevant research. *Psychotherapy, 49*, 115–122.
- Zarin, D. A., Johnson, J. L., & West, J. C. (2005). Challenges to evidence-based medicine: A comparison of patients and treatment in randomized controlled trials with patients and treatments in the APA practice research network. *Social Psychiatry and Psychiatric Epidemiology, 40*, 27–35.
- Zarin, D. A., Pincus, H. A., West, J. C., & McIntyre, J. S. (1997). Practice-based research in psychiatry. *American Journal of Psychiatry, 154*, 1199–1208.