

7

ASSESSMENT AND EVALUATION IN CLINICAL PRACTICE

ALAN E. KAZDIN

There is a long-standing hiatus between research and clinical practice. Among the many issues in the context of psychotherapy is that treatment research is conducted in well-controlled laboratory settings and conditions that depart from many conditions in clinical practice (e.g., Borkovec & Rachman, 1979; Heller, 1971; Kazdin, 1978). Efficacy and effectiveness research have been distinguished to reflect these differences (Hoagwood, Hibbs, Brent, & Jensen, 1995). Efficacy studies are conducted in controlled settings and under conditions that depart from clinical practice. Effectiveness studies are conducted in clinical settings with a diverse set of patient, therapist, and treatment administration characteristics. Evidence-based treatments (EBTs) are based almost exclusively on studies in highly controlled settings, and this fact has been repeatedly discussed as a concern about the generalizability of the findings to clinical practice (e.g., Persons & Silberschatz, 1998; Westen, Novotny, & Thompson-Brenner, 2004).

Whether the substantive findings from research ought to serve as a basis for clinical practice has been argued repeatedly. There is, in my view, a

This work was supported, in part, by a grant from the National Institute of Mental Health (MH59029).

much greater problem in relation to the gap between treatment research and clinical practice: The problem pertains to the sharply different ways of evaluating information in research and clinical settings and the use of the information to draw conclusions. Mental health researchers and clinicians are well familiar with the methods of treatment research, with emphasis on evaluating treatment effects on the basis of mean differences among various treatment and control conditions. The way in which treatment is evaluated in the tradition of quantitative research and hypothesis testing has no useful parallel in evaluating whether a particular patient seen in treatment gets better or has improved. Also, multiple patients in a group study are evaluated with standardized measures, none of which may capture their particular problems very well. Standardization in this way is very useful for research but not so clearly relevant for therapy in practice.

Clinical practice focuses on the individual rather than on groups. The case study has dominated clinical work; the therapist describes in a narrative way possible causes of the problem, how he or she formulated or conceived the case, the patient's clinical course, and the treatment provided. Typically, therapists in clinical practice do not evaluate cases in a systematic way. They may complete evaluations in relation to reimbursement or required clinical and hospital documentation. However, the information clinicians usually use to evaluate patient progress over the course of treatment is based on their views, experiences, and impressions. When these impressions are not systematically codified, they constitute the familiar anecdotal case study, in which the therapist constructs narrative information to draw inferences and to make connections in relation to possible etiologies, treatment course, and intervention effects.

In this chapter, I discuss the importance of systematic evaluation in clinical practice, illustrate steps that therapists can use, and highlights critical clinical and research issues that need to be addressed to provide the necessary underpinnings of evaluation. Although I focus on concrete steps to conduct evaluation, I also raise broader issues that pertain to features of clinical training that may unwittingly undermine evaluation. The focus of the chapter is on *systematic evaluation*, by which I mean evaluation by methods or instruments that begin with a construct and seek to operationalize that construct in ways that can be scrutinized, validated, and replicated by others in similar circumstances.

THE NEED FOR SYSTEMATIC EVALUATION IN CLINICAL WORK

Systematic evaluation is critically important in clinical practice. First, the therapist cannot assume that any given treatment will be effective in any given case. It is important to monitor treatment effects in an ongoing way to make decisions about continuing or terminating treatment and altering treat-

ment on the basis of how the patient is responding. Some patients make rapid changes quite early in treatment, so-called sudden therapeutic gains (Tang & DeRubeis, 1999); others do not make expected changes and do not respond to treatment, so-called signal-alarm cases (Lambert et al., 2003). Of course, change can occur in some areas but not in others, and change may occur at different rates in the different areas. Systematic evaluation permits finer delineation of therapeutic changes than would be possible with more global clinical judgments and unsystematic assessment.

Second, systematic evaluation is intended to add to clinical evaluation or judgment. There is no need to abandon clinical judgment. However, the need for systematic evaluation stems in part from the limitations of judgment. A discussion of clinical judgment begins with the selectivity of perception, cognitive heuristics, and the utility of clinical predictions. These are weighty concepts, but they point to therapist limitations in perception and cognition in gathering information and drawing conclusions. Systematic measures have their own artifacts and biases, but these can be evaluated and even corrected or taken into account in systematic ways.

Third, clinicians are wont to note the complexity of clinical cases. Clients bring multiple problems to treatment, the problems change as treatment begins, and new problems emerge. Complexity is an argument for systematic evaluation as well. Are the goals of treatment being achieved? Which goals, and to what degree? Are there new goals? Is the patient actually functioning better in everyday life? Systematic evaluation can improve decision making in light of the complexity of the case.

Fourth, systematic information obtained with individuals in clinical practice can greatly contribute to the knowledge base. Systematic assessment and the accumulation of cases can yield new insights even when experimental designs cannot be invoked. Over time, as cases accumulate, analyses can identify client characteristics that may influence outcomes and the course of change in treatment among individuals with different types of problems. Examples of contributions to the knowledge base from accumulated clinical information are evident in private practice (e.g., Clement, 1999), clinics (e.g., Fonagy & Target, 1994), and research settings (e.g., Lambert, Hansen, & Finch, 2001; Lambert et al., 2003). Patient information gathered systematically over the course of treatment can be useful for both generating and testing hypotheses.

STEPS FOR ASSESSMENT AND EVALUATION IN CLINICAL WORK

Suggestions that therapists engage in the systematic evaluation of the individual case are not new and include recommended ways of assessing and reporting cases to improve clinical care and contribute to the knowledge

base (e.g., Cone, 2000; Fishman, 2001; Hayes, Barlow, & Nelson, 1999; Kazdin, 1993, 1996; Meier, 2003). These recommendations focus on bridging the methodological gap between clinical work and research. Evaluation can include multiple components including assessment, research design, and data analyses and interpretation. All are relevant to clinical work. However, **assessment**, as a component of evaluation, includes the most pivotal step for improving clinical work. For present purposes **systematic assessment** refers to the use of measures that provide replicable information and that have evidence in their behalf in relation to various types of reliability and validity as pertinent to their use (e.g., test-retest reliability if used repeatedly, concurrent validity in relation to symptoms or functioning beyond the measurement device).

Systematic assessment and evaluation of the effects of treatment in clinical practice have as their **main goal** to foster high-quality patient care. Introducing systematic assessment to clinical practice is not merely the addition of a few measures to supplement clinical judgment. Several steps are essential, as summarized in Table 7.1 and elaborated in the sections that follow.

Specifying and Assessing Treatment Goals

Identifying treatment goals is a prerequisite for the selection of measures for assessment and evaluation. Therapy can have many different goals (e.g., reduction of symptoms, improved functioning at home and at work), and these are tailored to individual clients. Prioritizing the goals is important to permit initial assessment and to make treatment decisions. Treatment goals may vary over time on the basis of changing priorities and progress in treatment. For example, excessive dieting and maladaptive food consumption may serve as the initial treatment focus for a young adult referred for an eating disorder. The focus may later shift toward less immediate but no less important domains such as body image, management of stress, and relations with peers.

The notion of goals may unwittingly suggest that treatment always is aiming toward something concrete or a specific end. **Therapy may have as a goal helping individuals cope, vent, or tell their stories.** Making the goals of treatment explicit is important whether or not the goals are concrete. **Explicitness is a condition for assessment of progress over the course of treatment.**

Specifying and Assessing Procedures and Processes

Ideally, clinical evaluation specifies the means of achieving the goals. **The means may refer to the procedures used in treatment—that is, what the therapist does and what he or she asks the client to do in or outside of the sessions.** Alternatively, the means may reflect emergent processes or rela-

TABLE 7.1
Five Key Steps for Systematic Evaluation in Clinical Practice

Step	Description
1. Specify and assess treatment goals	Explicitly identify the initial focus of treatment and the goals or changes that are desired. Select or develop a measure that reflects the current status of the individual on these characteristics (e.g., symptoms, functioning).
2. Specify and assess procedures and processes	Explicitly identify the means or processes (procedures, tasks, activities, and experiences) that are expected to lead to therapeutic change. Measure the extent to which these means or their performance, execution, or implementation are achieved during treatment.
3. Select measures	Identify or develop the instruments, scales, or measures that will be used to assess progress over the course of treatment. Identifying the measure of process or procedures depends heavily on whether the procedures are straightforward (e.g., execution of tasks in the session) or emergent processes (e.g., alliance, bonding) that require separate measures.
4. Assess on multiple occasions	Measure performance on the measure of functioning toward which treatment is directed before treatment begins and then on a regular, ongoing basis over the course of treatment. Ongoing assessment may be every session, every other session, or some other regimen that allows the therapist to see any patterns or trends over time.
5. Evaluate the data	Display the information obtained from the assessment to examine changes, patterns, or other features of progress that can directly inform treatment decisions (e.g., changing or ending treatment, shifting the focus of treatment). Graphic displays are especially useful.

tionship issues (e.g., experiencing emotions, developing a therapeutic alliance). Specifying procedures or processes is not an end in itself. The primary goal is to use the information to benefit the client on the basis of how well treatment was implemented and the ends that were achieved.

Ongoing assessment of client progress may reveal that there is no therapeutic change. Assessment of procedures or processes that the therapist believes are important may provide useful information regarding how to proceed. The information may reveal that treatment procedures (e.g., addressing certain topics, engaging in role-play during the sessions) were not implemented very well or that processes within sessions (e.g., developing an alliance, dealing with a particular conflict) did not succeed. Hence, it is reasonable to try different strategies to alter these processes.

Assessment may reveal that the processes have been evoked fairly well but that no therapeutic changes are evident. Of course, patients do not change at the same rate. When enough time has elapsed to question whether change is still likely to occur is not known. (Indeed, data to guide clinical work on this question could readily emerge from systematic data accumulated from clinical practice.) In any case, when the patient has not changed or change is

not progressing well, it might be reasonable to try a different treatment. In advance, the therapist needs assurances that he or she tried the procedures or that the processes he or she identified were successful.

Selecting Measures

The next step is to operationalize the constructs by noting the specific measure or measures the therapist will use. Selecting measures requires decisions about the source of information (e.g., self- or clinician report) and assessment method (e.g., objective measures of personality or psychopathology, client diaries, card sorts, interviews, direct observation, biological markers or indexes). In principle, available measures include the full range of psychological instruments.

A few measures are now available that have been well tested in clinical work. For example, the **Outcome Questionnaire—45** (OQ-45; Lambert et al., 1996) is a self-report measure designed to measure client progress (e.g., weekly) over the course of treatment and at termination. The measure requires approximately 5 minutes to complete and provides information on four domains of functioning: symptoms of psychological disturbance (primarily depression and anxiety), interpersonal problems, social role functioning (e.g., problems at work), and quality of life (e.g., facets of life satisfaction). Total scores across the 45 items present a global assessment of functioning; the subscales target more specific areas. Research has evaluated different types of reliability and validity, with more than 10,000 patients included in the various reports (see Lambert et al., 2001, 2003).

Another example is the **COMPASS Outpatient Treatment Assessment System** (Howard, Brill, Lueger, & O'Mahoney, 1992; Lueger et al., 2001), a measure that includes 68 items in three broad scales: Current Well-Being (e.g., health, adjustment, stress, life satisfaction), Current Symptoms (e.g., various symptoms for psychiatric diagnoses), and Current Life Functioning (e.g., work, leisure, family, self-management). Careful psychometric evaluation in the context of clinical application supports the use of the measure in outpatient treatment with adults.

The OQ-45 and the COMPASS System provide a fixed set of items that are quite broad and cover domains likely to be relevant for most adults who come to treatment. **Goal Attainment Scaling, alternatively, is an assessment strategy that individualizes treatment goals.** The measure is based on collaboration between the patient and therapist at the outset of treatment about the goals and expectations of treatment (Kiresuk & Garwick, 1979; Kiresuk, Smith, & Cardillo, 1994). This measure has been widely used, applied, and validated and is illustrated in a case example later in this chapter.

These three measures are major options that have been carefully studied and have wide applicability to patients seen in outpatient or inpatient treatment. Other measures useful in clinical work that draw on a variety of

different assessment methods have been identified elsewhere (e.g., Alter & Evens, 1990; Faulkner & Gray Health Care Information Center, 1997; Meier, 2003). Rating scales are a particularly useful format that allow an endless array of options to be developed and evaluated (see Aiken, 1996). Diverse measures have been developed and formatted to facilitate their use in clinical settings (Clement, 1999; Wiger, 1999).

In addition to measurement of client functioning, measurement of treatment means or processes is important as well. The specific types of treatment and putative processes or features leading to change dictate the assessment focus. The therapist proposes (hypothesizes) that specific means are central to therapeutic change. If these means (e.g., quality of the relationship with the client, completion of specific homework assignments) can vary with treatment administration, their assessment is likely to be useful. The assessment priority is evaluating clinical outcomes and systematically collecting information on whether the client is changing over the course of treatment.

Assessing on Multiple Occasions

The major change that is needed in clinical practice is ongoing, continuous assessment during the course of treatment. Ongoing assessment can be used to chart where the client is at the beginning of treatment and to see whether changes are made over time. Several data points are needed not only to assess the mean level of functioning over time but also to give an idea of variability and trends on the measure. There are many opportunities for flexible application of continuous assessment. Ideally, but perhaps unrealistically, pretreatment assessment would include two or three assessment occasions to provide a baseline to help evaluate subsequent progress. Also, at initial assessment the client's level of performance on the measure may be at an extreme because of stress or crisis, and marked changes from the first to second assessment occasion can be expected because of statistical regression, passing of the crisis, and repeat testing (see Kazdin, 2003). Assessment before beginning the intervention may even show improvement in the client's status and hence has implications for reevaluating the goals of treatment, the means to obtain them, and the measures to evaluate progress.

The initial assessment provides descriptive information (baseline) about the client's level of performance and its variation. Perhaps only one assessment occasion is feasible, or indeed no assessment may be feasible because of the urgent nature of the treatment. For most psychotherapy clients, it is not clear that intervention is absolutely essential at the first contact. Usually, assessment can begin while efforts to manage the situation are under way. When treatment begins immediately, it may be possible to obtain a retrospective baseline in which the client and others in his or her life provide an estimate of the client's recent functioning. Apart from baseline assessment, evaluation during the treatment phase is pivotal to gauge whether any changes

are being made and whether the magnitude and rate of these changes are important clinically.

Evaluating the Data

Data evaluation refers to the use and interpretation of the assessment information. Two issues emerge in clinical care. First, one must decide whether change has occurred, is reliable, and departs from the fluctuations one would expect without the intervention. Second, are the changes important, and do they make a difference in the patient's life?

Ongoing assessment provides data before and during the course of treatment that serve as the basis for evaluating whether the changes are reliable and beyond routine fluctuations. Several methods are available to evaluate the reliability of the information (Kazdin, 2003). Of all methods, graphic display (e.g., a simple line graph) is particularly useful for seeing the pattern in the data obtained over time. Nonstatistical data evaluation methods (changes in means, levels, slope, latency of change over time) are used extensively in single-case research (applied behavior analysis) focused on interventions for diverse client and community populations and can be used for clinical evaluation (see Kazdin, 1982; Parsonson & Baer, 1978). Nonstatistical data evaluation does not require complex computations but follows directly from graphic presentation of the data. Other methods of graphing than a simple line graph (e.g., stem-and-leaf plots, box plots), usually used for multiple subjects from group research, might also be used to plot multiple data points from individual clients (see Meier, 2003; Rosenthal & Rosnow, 1991). If a patient's data are entered regularly on a database or office management system, then graphical presentation and simple slope or trend lines (e.g., regression lines) can be plotted automatically, as illustrated later.¹

In addition, there is interest in evaluating whether the changes made in treatment are clinically significant—that is, whether they make a difference to the client. Several measures of clinical significance have been used in treatment research, including whether level of functioning at the end of treatment falls within the normative range of individuals functioning adequately to well in everyday life, whether the individual makes a change that is large (e.g., in standard deviation units) on the measure, and whether the individual no longer meets criteria for a diagnosis that were met at intake (see Kazdin, 1999; Kendall, 1999). These measures all have interpretative

¹Statistical tests are available as well to consider changes over time and whether these changes are reliable according to the usual standards of research (Kazdin, 1982). I mention these only to note their availability. The value of identifying whether a particular change is or is not statistically significant is questionable in research (see Kazdin, 2003). Few would lobby for the use of statistical significance in clinical applications with individual patients. Yet some means of identifying whether the change is reliable is needed.

problems insofar as there are little or no data showing that someone who has made a change labeled as clinically significant is in fact functioning palpably better in the world (Kazdin, 2001). That said, these measures, especially the one in which a large change in the measure is required, have been applied in clinical work (Lambert et al., 2003; Lueger et al., 2001).

Some effort is needed to evaluate whether the treatment goals are approached or attained and whether the changes make a difference. With some clinical problems (e.g., panic attacks, tics), elimination of the problem can be taken as a clinically important change. With other problems (e.g., obesity), clinically important change may involve reduction of the problem to levels that improve health consequences (risk). These types of clinical problems are exceptions in the course of psychotherapy, however. Whether impairment declines, symptoms improve, marriages are better, and the experience of loss is alleviated are matters of degree, and whether the amount of change is helpful or enough is difficult to discern. Indeed, the same amount of change on a measure (e.g., of marital satisfaction) or set of measures may be experienced quite differently among patients and may have a varied impact on everyday life (e.g., whether they remain or do not remain married).

CASE ILLUSTRATION

The following case study highlights the steps outlined in previous sections and illustrates how they can be applied. The description emphasizes assessment and evaluation, rather than the details of the intervention itself.

Brief Background

Gloria was a European American, 39-year-old woman who referred herself for outpatient treatment. She was married and had two children (ages 16 and 17). She and her husband were both college educated; on the basis of her education and income and her husband's occupation, they were middle class. Gloria was not employed outside of the home; her husband was a manager of a computer software firm. She and her husband had been married for 18 years.

Gloria scheduled an appointment because she said she was depressed and needed to talk to someone. During the initial interview at the clinic, Gloria saw a male therapist, who asked, with open-ended questioning, about the reasons she sought treatment, sources of satisfaction and dissatisfaction in her life, relationships with significant others, symptoms, and related matters. Toward the end of this discussion, the therapist queried Gloria about what she expected and wished to obtain from treatment.

During the interview, Gloria indicated that she had been treated for depression on two separate occasions in her life, once during college and

once after the birth of her first child. On each occasion she was placed on medication. She reported some relief but also complained about side effects and did not feel really helped overall with her problems. Currently, she said, she was depressed again. She reported feeling "empty and lost" about her life and marriage. She said that she lacked meaning and direction in her life. She felt alienated from her husband and her children. In the case of her husband, she felt great emotional distance because of years of reduced intimacy, joint activity, and time together. Her children were very important to her, but she felt they did not need her very much now that they were teenagers. Gloria identified as her own goals for treatment simply feeling better about her life, not being depressed, having some direction, and improving relations with her spouse and others.

Assessment

In the initial interview, the therapist introduced systematic assessment after the open-ended discussion provided an initial formulation of the focus. He explained that the assessment procedures would help make the goals and directions of treatment more explicit and quantify the domains that were to be the foci. The therapist used three measures. The first measure was adapted from Goal Attainment Scaling (Kiresuk et al., 1994), which was developed decades ago as a general method to evaluate outcomes of mental health treatment, has been widely applied and tested, and has extensive information on psychometric properties, training, and use. The scale identifies individualized goals of treatment to reflect the domains pertinent to the patient. The therapist adapted the method to focus on the domains Gloria identified as important. Toward the end of the interview, the therapist identified four major concerns, themes, or areas as a beginning for them to work on (a) depressive thoughts and feelings, (b) little involvement in meaningful and fulfilling activities, (c) disengagement from her family, and (d) lack of supportive contacts outside of the home. The therapist and Gloria discussed these to see if they captured Gloria's experience, because they did not follow exactly from her original formulation of the problems.

For each theme, the therapist asked Gloria to help construct statements that they then graded to indicate different levels of functioning. The goal was to compose a 4-point scale for each theme in which 1 = worsening of the problem, 2 = no change in current functioning or feelings, 3 = some improvement, and 4 = attainment of goal of functioning and feeling on this domain. The therapist referred to this as the Gloria Scale (G Scale for short) and explained that it would help guide them during treatment. The therapist conveyed the concept of the 4-point scale, but Gloria provided the content of each of the statements. The therapist asked her to describe a way to characterize her current functioning or where she was now, what it would be like if she became worse, what some improvement might look like, and how things would

TABLE 7.2
Four Themes and Items for the Gloria Scale

Theme	Items
Depressive thoughts and feelings	<ol style="list-style-type: none"> I feel more depressed and dejected than I did before I started treatment. I feel about the same level of depression and dejection as I did when I started. I feel a little better about my mood, and things are not as bad as before. I feel a lot better, I do not think about my feelings as negative, and I have more energy to get out and do things.
Involvement in meaningful and fulfilling activities	<ol style="list-style-type: none"> I really feel paralyzed about doing anything any more. I am not doing anything differently now or anything special I like compared with when I started treatment. I feel better that I have some direction and focus in what to do. I am totally involved in some things, such as a career, that give me good feelings about life.
Disengagement from family	<ol style="list-style-type: none"> I do less with my husband and children than before and don't seem to care about doing things with them. Things really have not changed about my feelings. Everyone at home does his or her own thing, and my husband and I mostly just eat meals together. My husband and I are a little better. We go out once in a while and are a couple again. My husband and I are really "together." We are intimate in many ways, and I can feel that he cares for me.
Supportive contacts outside the home	<ol style="list-style-type: none"> I am isolated from people in general, including my relatives who live in town. Once in a while I see someone when I shop or at a school event with my children. We chat a bit, but nothing beyond the superficial. I meet someone to have coffee with or to go to an event or shopping with during the day. I meet a few people by myself or some couples that my husband and I can get together with, and we do this on a regular basis.

Note. The theme areas were derived from an open-ended interview with Gloria. The specific statements were generated by her to reflect what it would be like to become worse, to remain the same, to improve a little, and to achieve her goal for that theme. These alternative outcomes reflect Items 1 through 4 respectively under each theme. Each time she completed this measure, Gloria selected the statement under each theme that was closest to how she had felt during the previous week. From *Research Design in Clinical Psychology* (4th ed., p. 320), by A. E. Kazdin, 2003, Boston: Allyn & Bacon. Copyright 2003 by Allyn & Bacon. Reprinted with permission.

be if she really made the kind of change she wanted. Table 7.2 presents the four themes and the graded statements Gloria and the therapist constructed. For the assessment on the G Scale, she was instructed to select the statement under each theme area that characterized how she had felt during the previous week. After they had developed the scale in the initial session, the thera-

pist asked Gloria specifically if the second statement of each theme area really captured her current feelings; she reported that it did.

The therapist described two other measures and gave them to Gloria to complete. The Beck Depression Inventory (BDI; Beck, Steer, & Garbin, 1988) addressed the severity of her depressive symptoms. This measure includes 21 items; for each item, the client selects 1 of 5 statements reflecting differing severity of depressive symptoms (each item is scored 1–3). Gloria also completed the Quality of Life Inventory (QOLI; Frisch, 1998), a self-report scale that assesses overall quality of and satisfaction with life in 17 domains (e.g., love relationship, home, learning, recreation, friendships, philosophy of life, work, health, neighborhood). The weighted score is used for each domain on the basis of the client's rating of the importance of that domain in his or her life (0 = *not at all important*, 2 = *very important*) then multiplied by the satisfaction derived from that domain (−3 = *very dissatisfied* to +3 = *very satisfied*). The BDI and QOLI took approximately 20 minutes total to complete.

The initial contact with Gloria lasted about 2 hours. The interview and development of the G Scale took about 1.5 hours, and completion of the BDI and QOLI took the rest of the time. Gloria was scheduled to return the following week and was asked to come 20 minutes before the session. Before the second session began, she completed the BDI and QOLI and brought them to the therapist.

The therapist began the session by asking Gloria to select one statement from each of the four theme areas that they had discussed. The material had now been typed in a format similar to that of Table 7.2. They briefly discussed whether the areas were still important to her and whether she felt that their last interview had missed critical material. The therapist conveyed that the initial goals were a place to begin and that the information within the sessions and from the assessments would be important to make any midcourse corrections as needed.

At this point, the therapist described the treatment and said that it would take place on a weekly basis. The therapist selected treatment that combined cognitive therapy with interpersonal psychotherapy. Cognitive therapy focused on Gloria's maladaptive cognitions about herself related to her depressive affect, poor self-esteem, feelings that life was not worthwhile, and internal attributions regarding her views of herself (see Beck, Rush, Shaw, & Emery, 1979). Interpersonal psychotherapy focused on her interpersonal relations, her roles and the sources of satisfaction and emotion associated with each, and her feelings about herself as a spouse and parent (see Klerman, Weissman, Rounsaville, & Chevron, 1984). The therapy integrated these treatments and included assignments (e.g., shared activities with her husband) carried out between the sessions.

Each week, Gloria came about 20 minutes before the session to complete the G Scale, BDI, and QOLI. At the first and second sessions, the full scales were administered. However, there were symptoms and domains within

the two standardized scales that were not problematic or not seemingly relevant to Gloria. The therapist constructed abbreviated versions of the BDI (15 items) and the QOLI (13 domains) by eliminating selected items, and they used these versions throughout treatment. From each scale, the therapist quantified Gloria's performance with a summary total score for each measure to examine whether any systematic pattern or change was evident as treatment progressed. The assessment information is graphed in Figure 7.1. The two assessment occasions before treatment were delineated as baseline (Weeks B1 and B2 in the figure). The therapist added a linear regression line to each graph to characterize the slope that best fit the data. Overall, the individual data points and regression line suggest that Gloria showed improvement over time.

Although the overall scores are useful in summary form, the mean for all of the items of a given measure (e.g., BDI) with an individual case can suffer the same liability as means in group research—namely, they can obscure critical information. In Gloria's case, the G scale and the QOLI indicated that she had made little progress in her relationship with her husband. The relationship issues emerged more fully in the treatment session of Week 14. At the beginning of the session, the therapist indicated that he thought this would be a good time to discuss at length the original goals of treatment and how she had been doing on the basis of the assessment information and Gloria's appraisal of treatment.

Gloria indicated that she had felt much better about herself and her life. Her thoughts about her life, what she saw as important, and her direction were much better. Over the course of treatment, she had initiated a number of activities. She had begun a class at a local university and now planned to obtain a degree in nursing. She had developed greater interaction with a neighbor, a woman similar to her in age, whom she met almost daily to engage in routine activities (e.g., exercise, shopping). Also, from her class she met a few people whom she enjoyed. Finally, time with her children was more enjoyable. In general, she felt much better about her overall well-being. At the same time, she felt that her relationship with her husband had not been helped at all by treatment. Although she and her husband had gone out on a couple of dates, she felt that this was merely time together with no connection or closeness. She said she loved her husband and could not imagine being without him but that there seemed to be none of the closeness or contact they had experienced in the past. The therapist suggested that they focus more on this issue for a few sessions but that the immediate goal would be for her and her husband to consider joint steps toward improving their marriage. They then used the same method as that for developing the G Scale to identify theme areas within her marriage that were significant and to set anchor points and added this measure to Gloria's routine assessments.

Treatment continued for 5 more weeks. Gloria no longer completed the BDI and QOLI weekly, instead completing them every other week. Weekly

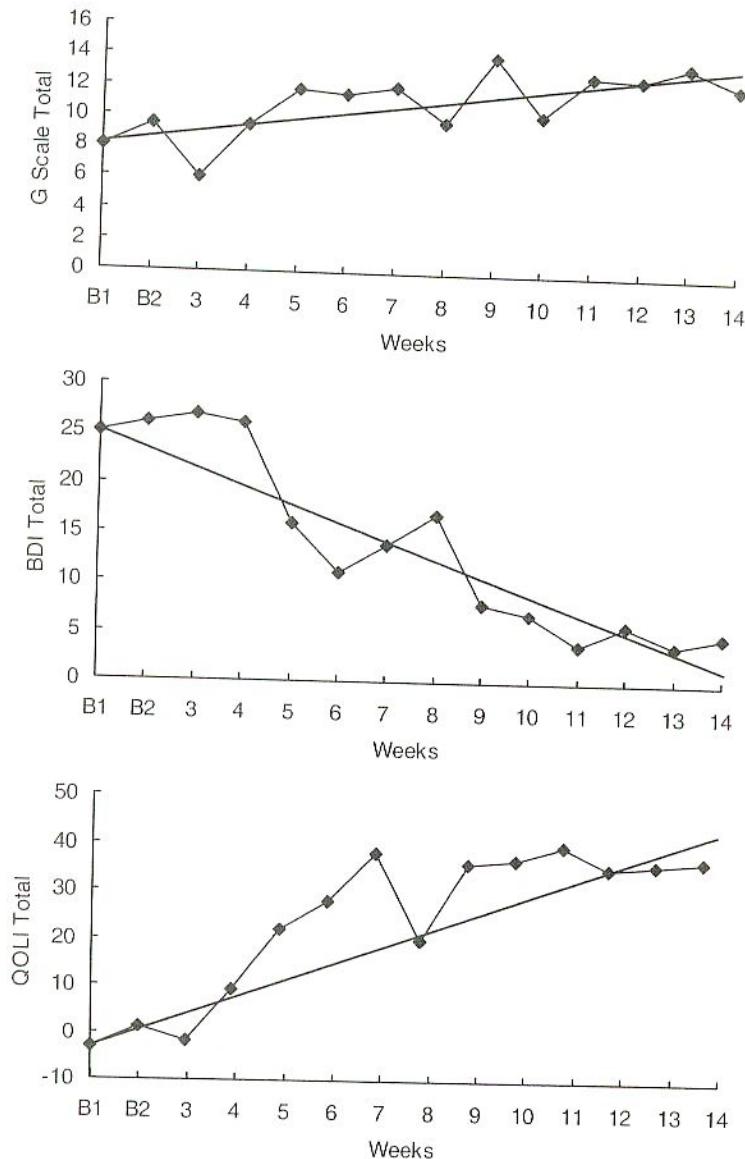


Figure 7.1. Session-by-session scores for Gloria on three measures including the Gloria Scale (G Scale, upper panel), the modified Beck Depression Inventory (BDI, middle panel), and the modified Quality of Life Inventory (QOLI, lower panel). The scores for each measure (BDI and QOLI) include items that were selected as relevant to the client and do not reflect the complete scales. Data are presented for 14 sessions (weeks). The first week was devoted entirely to interview and assessment. The second week began with completion of assessments followed by the initiation of treatment. The first two weeks (B1, B2) refer to baseline or pretreatment assessment. Given the direction of scoring of the measures, improvement would be reflected in increases for the G Scale and QOLI and a decrease in the BDI. Fitted to each graph is a linear regression line over the course of all data points. From *Research Design in Clinical Psychology* (4th ed., p. 322), by A. E. Kazdin, 2003, Boston: Allyn & Bacon. Copyright 2003 by Allyn & Bacon. Reprinted with permission.

assessment continued for the original G Scale and for the scale related specifically to her marital relationship. After 5 weeks, Gloria and her husband agreed to begin marital counseling, and Gloria ceased individual treatment.

Discussion

There are useful features of this case that illustrate evaluation in clinical practice. First, the therapist made efforts to make the initial goals of treatment explicit and to quantify them. Second, the therapist used systematic, clinically relevant, and user-friendly assessments. The assessment procedures included a highly individualized scale, the Gloria Scale. The therapist altered the two standardized measures (BDI, QOLI) to address the specific domains that seemed relevant to the client. Finally, he charted the information to evaluate Gloria's progress and used this information both to alter treatment and to suggest the need for further assessment and treatment.

Many limitations are apparent as well. First, the three measures overlapped in method (all self-report), and all were conducted in the treatment setting. Other methods (e.g., spouse ratings, daily log of activities during the week) may have been useful as well. Second, the therapist might have identified the issue of marital dissatisfaction earlier and given it a more central role early in treatment. Perhaps the marital issues became more salient as treatment progressed precisely because the client felt progress in other domains. Third, the regression lines must be interpreted cautiously. The lines suggest an overall improvement and hence are very useful in conveying a pattern, but the changes over time cannot be interpreted as being the result of treatment. Finally, the case description does not give an idea of how well Gloria adhered to the tasks in the session or performed the homework assignments. Whether she adhered to assignments and whether the therapist felt that cognitive issues and interpersonal functioning were suitably addressed within each session are important considerations.

Overall, the case illustration highlights the use of systematic assessment and evaluation. The assessment was individualized but also included standardized measures pertinent to the treatment focus. The measures required 20 minutes to complete before each session. Other measures are available that encompass multiple areas of functioning but require much less time for administration (e.g., 5 minutes for the OQ-45). Also, the data obtained during treatment were useful not only in evaluating progress but also in making decisions about the focus of treatment over time.

ISSUES AND LIMITATIONS IN SYSTEMATIC ASSESSMENT AND EVALUATION

Systematic assessment and evaluation can be readily implemented in clinical work to improve the therapist's inferences about whether change

occurs, the importance or extent of the change, and even the likelihood that treatment is responsible for the change. Yet there remain some obstacles to implementation.

Methodological Issues

Few measures are available that are clinically feasible and validated for clinical work. *Feasible* means user friendly and brief; *validated* refers to all of the usual concepts of validity but also to validity over time with multiple and repeated assessments. Progress has been made in identifying measures that can be used in individual therapy, that can be applied widely across clinical settings and clients, and that can provide information that contributes to the knowledge base more generally (e.g., Barkham et al., 2001; Kordy, Hannöver, & Richard, 2001).

Another issue pertains to evaluation of what factors might contribute to change. Valuable additions to evaluation efforts include assessment of the therapeutic process (e.g., relationship and alliance), of the various activities or exercises that treatment may depend on, and of other aspects of treatment. Clearly, the initial priority is to identify whether a given patient does improve and improves adequately. As this aspect of evaluation becomes more routine, the therapist might make more effort to evaluate possible mediators of change.

Finally, how to evaluate the data obtained from the measures raises several issues. Descriptive statistics (e.g., changes in means, slope) can be used for inferential purposes. Various data management programs can be used to enter data, to provide graphic displays, and to document progress in user-friendly ways (e.g., OPTAIO, 1997). Yet what decisions ought to be made from the data, and on the basis of what criteria? The patient may have made a clinically significant and important change, but with only rare exceptions researchers do not yet know how changes on measures used to evaluate treatment translate to actual changes in everyday life (Kazdin, 2001). Research is needed to understand the amount and type of change in treatment that constitute effects that genuinely benefit individual patients.

Clinical Issues and Concerns

Several concerns and objections may emerge in clinical practice about the utility of systematic assessment and evaluation. First, therapists are often concerned that assessment may interfere with the therapeutic relationship. The therapist is responsible for treatment; adding to that the role of the assessor or evaluator may, on conceptual grounds, mix roles and be viewed as antitherapeutic. Yet the presumption that evaluation harms is arguable; indeed, alternative assumptions are plausible as well (e.g., that not evaluating the patient can permit harm to occur, that evaluation may have no impact,

that evaluation may help). How clients perceive systematic evaluation probably depends on the therapist's views of evaluation and presentation of the evaluation objectives and methods. If evaluation is presented as a matter of course, as central to treatment, and as purposeful, then the client's views are likely to be positive.

Second, measures of a clinical problem may oversimplify the problem. Yet for a measure to be useful, it need not capture all there is about the construct. A measure provides a key sign, correlate, or sample of the problem—that is, an operational definition. Therapists usually are not interested in measures, but rather in constructs, or the characteristics the measures were designed to assess. They use measures even though they do not cover the entire scope of the problem.

A third and related concern is that assessment seemingly ignores the individuality or uniqueness of the client. Yet systematic assessment can be quite individualized, as exemplified in the case example. The clinician can decide with the patient which domains of functioning are most relevant and can build the assessment devices to reflect those domains. Clinical practice, unlike the usual research context, permits individualization of both assessment and treatment. Standardized assessment may still play a critically important role in therapy and can complement individualized assessment in critical ways. A client's profile on a standardized measure and his or her standing relative to a normative group of peers of the same age, sex, and ethnicity, for example, can provide meaningful data that may also guide treatment. The standardized nature of a measure is not a threat to patient individuality, but rather an opportunity to examine that individuality against a broader backdrop.

Fourth, an objection to evaluation in clinical work is based on the dynamic nature of treatment. In much of psychotherapy, there is not a single, simple patient problem that remains constant. Indeed, over half of clients seen in therapy add new target complaints over the course of treatment (Sorenson, Gorsuch, & Mintz, 1985). The changing focus of treatment and the multifaceted nature of the foci are not arguments against assessment. Rather, they make systematic assessment all the more important. It is critical to identify changes in problem domains and priorities from the standpoint of the patient and therapist. Therapists and clients can set new goals and present or withdraw assessments to reflect these changes.

If the goal of clinical work is to help patients and to address the concerns of a specific individual, here and now, then the case for systematic evaluation is easily made. In fact, the case does not need to be made for systematic assessment and evaluation. Just the opposite—in clinical work, where the individual patient is so important and direct benefits are the goal, unsystematic evaluation is difficult to justify. There are clearly urgent circumstances in which intervention must proceed immediately (e.g., disasters, suicide attempts in progress). The important exceptions certainly preclude

collecting baseline data, but they do not preclude evaluating the impact after the crises have abated.

IMPEDIMENTS TO CLINICAL EVALUATION

Introducing systematic evaluation into clinical practice can be readily accomplished. The steps are not too complex or onerous, assessment tools are available, and the yield from these tools has been shown to be useful in large-scale evaluations with many individual cases. Systematic evaluation is not in the mindset of most clinicians, clinic directors, residency and internship training directors, and clinical supervisors who oversee treatment. Clinicians are often faulted for their disinterest in evaluation in clinical work. Yet clinical training, whether in psychology, psychiatry, social work, or counseling, does not equip individuals to evaluate their cases in user-friendly and methodologically sound ways.

Psychotherapy Outcome Research

Psychotherapy outcome research is dominated by randomized controlled clinical trials (RCTs). Such trials are recognized to have special status with regard to testing interventions effects. My comments are not intended to impugn such trials. However, pivotal features of these trials make them not very relevant for clinical practice. Endless discussions have been provided about the conditions of testing treatment (e.g., efficacy studies) and how the results may not be generalizable to clinical practice. I wish to convey a point different from this now well-worn path: The methods, as well as the results, of RCTs are not generalizable to clinical practice. Methodological features of RCTs make them largely of little relevance to clinical work and unwittingly may impede evaluation in clinical practice.

RCTs of psychotherapy are characterized by pre- and posttreatment assessment and comparison of means between conditions (e.g., treatment and control groups) using statistical analyses of the data. These methodological features have no useful counterpart for treatment evaluation in clinical practice. In clinical work, therapists do not wish to give a fixed regimen of treatment and see how the patient has done after treatment has ended. To be sure, they do want to know how the patient is doing at the end of treatment, but they care just as much about assessment during treatment so they can make changes if and as needed or indeed stop the treatment because of early gains. RCTs do not provide methods that can be extended to address the priorities of clinical care. Therapists in any of the clinical disciplines who have been trained in research methods can extend few or none of the methods they have learned to clinical work.

A relatively recent development in treatment outcome studies is referred to as *patient-oriented research* (Howard, Moras, Brill, Martinovich, &

Lutz, 1996; Lambert et al., 2003). The key to patient-oriented research is ongoing assessment and monitoring of individual patients from the beginning to the end of treatment and use of the information to chart progress and make decisions about treatment. Unlike RCTs, patient-oriented research does not involve an extensive battery of pre- and posttreatment assessments. Rather, therapists conduct assessment at each session with a brief measure that captures functioning in diverse domains. The OQ-45 (Lambert et al., 1996) is one such brief measure. The measure requires only minutes to complete and provides information about multiple domains. The therapist evaluates the treatment by discerning the extent to which the client makes a change or fails to make a change. During treatment, different criteria can be used to guide treatment decisions (see Lambert et al., 2001, 2003). Patient-oriented research greatly reduces the gap between research and practice. The methods used in clinical research and practice become one and the same. Patient-oriented research provides a methodology that could be added to training to help would-be therapists become more interested in systematic evaluation of the treatment they provide.

Training in Alternative Methodologies

My comments on RCTs focused narrowly on the common model for psychotherapy outcome research. The comments apply more broadly to quantitative research methods involving group designs, null hypothesis testing, and statistical evaluation. But there are other research methods routinely omitted from training that offer great promise for and applicability to clinical work. The following paragraphs highlight three alternatives.

First, qualitative research methods are extremely relevant to clinical work. Qualitative research has its own methodology, including strategies for assessment, design, and data evaluation.² Qualitative research seeks knowledge in ways that are systematic, replicable, and cumulative. The methods look at phenomena in ways that are intended to reveal many of those facets of human experience that the quantitative tradition has been designed to circumvent—the human experience, subjective views, and how people represent (perceive, feel) and react to their situations in context. For example, quantitative research has elaborated many of the factors that contribute to or are associated with homelessness. Predictors of homelessness, the relative weight of these predictors, and the short- and long-term effects of homelessness (e.g., medical, psychiatric) on adults, children, and families have been elaborated in quantitative research. A qualitative study is likely to focus on the

²In clinical work, *qualitative* is sometimes used to refer to descriptive, anecdotal, and case study material. That is, the term has been inappropriately adopted to refer to any nonquantitative evaluation. This is a misuse—*qualitative* is not a synonym for loose or unsystematic data or “my opinions.” Indeed, it is an antonym for these characteristics. Qualitative research is rigorous, scientific, disciplined, and replicable and can both test and generate theory.

experience of being homeless, the details of the frustrations, and the conflicts and demands the experience raises in ways that are not captured by quantitative studies (e.g., Lindsey, 1998). Clinical psychology and related disciplines that have clinical practice as a career path rarely train students in qualitative research methods. This is unfortunate, because the methodology and its many rich variations provide options that would be well suited to promote understanding of the individual experience of patients, to systematically codify treatment changes, and to do so in replicable ways.

Second, single-case experimental designs have features that are readily adaptable to clinical work. Single-case experiment designs emerged from laboratory research with humans and other animals to study such basic processes as learning and performance. They have been used quite extensively in an area referred to as *applied behavior analysis* in which interventions are used to address goals of clinical dysfunction, education, rehabilitation, health, and scores of domains of functioning in everyday life (e.g., the home, business, and industry; see Kazdin, 2001). The designs are rigorous and can yield causal inferences, as that term is used in science. Among their key features, single-case experiments consist of multiple observations with one or a few cases, in sharp contrast to typical group experiments in which few observations (e.g., pretreatment, posttreatment) are made with multiple subjects. User-friendly variations of the designs provide tools that can be used to evaluate and to improve patient care in clinical work (e.g., Hayes et al., 1999; Kazdin, 1981, 2003). As with qualitative research, these designs are rarely included in clinical training in the mental health professions.

Finally, the case study, with all of its problems, can provide useful information and even permit strong inferences (see Sechrest, Stewart, Stickle, & Sidani, 1996). Among the issues with this methodology is understanding what sorts of influences compete with drawing inferences about events that happen with the case, how these influences might be combated or made implausible, and when and how can inferences be drawn as a result. The anecdotal case study as traditionally conceived and implemented is not the only alternative. There is much a clinician can do to bolster the quality of the inferences he or she draws about patient change and the reasons for change, but these techniques are rarely taught (Kazdin, 1981). Learning about the strengths of the case study and the underlying thought processes that can increase its yield would be enormously helpful in clinical training and ultimately in patient care.

I mention three methodologies here in addition to the quantitative tradition. These other three methodologies are much more readily adaptable to the questions and conditions of clinical practice than the methods in which mental health professionals are routinely trained. It is quite easy to point to a seeming antidata and antievaluation mindset among those in clinical practice. But probing a little deeper conveys that the training necessary for therapists to do evaluation in clinical work has not been and is not being provided.

General Comments

Training in psychology and science more generally is designed to teach substance (content areas of interest), methods (assessment, design), and a broader approach to science. This broader approach is a way of thinking about phenomena and systematizing the information one obtains to draw inferences. The thought processes reflect concerns about ways of operationalizing critical constructs, posing hypotheses about interventions and processes leading to change, and testing assumptions about interventions and their impact. Assessment, research design, and evaluation are not alien to clinical practice. Invariably, practitioners are drawing inferences, actively or passively making decisions regarding what they perceive, and so on. Introducing systematic assessment into clinical work brings these practices in harmony with tenets of science (e.g., testing hypotheses, operationalizing critical concepts, fostering replication). The special feature is to use evaluation concepts and practices to advance the therapeutic progress of individual clients.

The priority of the client and concerns for client well-being have been used as arguments for not evaluating treatment progress systematically. I would argue a fortiori that when the goal is to help someone and to address the needs, concerns, or desperation often evident in clinical work, assessment, evaluation, and drawing informed inferences are more important than ever. Much of the thinking underlying training appears to foster dichotomies in which research tenets and priorities (e.g., careful and systematic observation, collection of replicable data, focus on group designs) are contrasted with the priorities of clinical practice (e.g., concern for individuals and their unique circumstances, narrative and in-depth evaluation that is more qualitative). There is no need for these dichotomies, and patient care is the victim when they are fostered.

CONCLUSION

This chapter has advocated and illustrated systematic evaluation in clinical practice as a means to improve the quality of clinical care. Several steps for systematic evaluation were discussed, including specifying and assessing treatment goals, specifying and assessing procedures and processes, selecting measures, assessing on multiple occasions, and evaluating the data. A clinical case illustrated the use of these steps to convey the use of systematic evaluation in clinical work. Methodology and evaluation are not just for empirical research; they are for instances in which the clinician wants to know whether there is a change, difference, or effect and to isolate possible reasons.

An important point of departure for evaluation in clinical practice is the goal of the benefit of the individual patient. Assessment methods must

be able to accommodate a wide range of clinical problems and situations. There are two major challenges. First, psychologists must develop practices and procedures that can easily be integrated into clinical practice. Much progress has been made on this front; a few well-validated measures are available for clinical use. Technological innovations using various everyday gadgets (e.g., scheduling devices, cell phones with cameras) are likely to increase assessment options. The challenge of implementing evaluation in clinical work is not related to the paucity of tools, even though more and better tools will always be welcome.

Currently, the training of clinicians is a huge impediment to integrating evaluation into practice. Clinical evaluation needs to be conveyed as pivotal to patient care. Patient care and high-quality clinical work demand that clinicians use the best treatments and evaluation tools to evaluate their impact. The personal judgment and experience of a therapist, although clearly valuable, are not a substitute for the collection of systematic information and the use of that information in making critical decisions.

There is a seeming resistance among clinicians to conduct systematic evaluation or to use their scientific thinking for clinical work. In this regard, I very much favor the so-called medical models. In the context of medical problems, it would be poor practice, if not unethical, to conduct an evaluation without using medical tests and a thorough workup to provide solid information about the problem, where possible. A physician who has experience with many patients with similar symptoms is of great benefit. It would be of even more benefit if that physician drew on the amazing array of systematic assessments (e.g., blood work, scans of various sorts) that can rule in or rule out problems and that can be used to monitor whether the intervention, once initiated, is having any effect. One would not think of administering chemotherapy or surgery without evaluating over time the impact and durability of their effects. Unsystematic and loose assessment has its place too; "How are you feeling? Can you get around very much? How was your week?" are all part of bedside manner and could be recorded systematically, but usually are not. In the clinical practice of psychotherapy, there is little evaluation that is systematic. The stereotype of clinicians is that they enter clinical work in part because they care for people and are less interested in data and research. Let us hope this stereotype is a straw person. Clinicians want evaluation in clinical practice precisely because they care about the individual patient.

REFERENCES

- * Aiken, L. R. (1996). *Rating scales and checklists: Evaluating behavior, personality, and attitude*. New York: Wiley.

- Alter, C., & Evens, W. (1990). *Evaluating your practice: A guide to self-assessment*. New York: Springer.
- Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., et al. (2001). Service profiling and outcomes benchmarking using the CORE-OM: Toward practice-based evidence in the psychological therapies. *Journal of Consulting and Clinical Psychology*, 69, 184–196.
- Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford Press.
- Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, 8, 77–100.
- Borkovec, T., & Rachman, S. (1979). The utility of analogue research. *Behaviour Research and Therapy*, 17, 253–261.
- Clement, P. W. (1999). *Outcomes and incomes: How to evaluate, improve, and market your practice by measuring outcomes in psychotherapy*. New York: Guilford Press.
- Cone, J. D. (2000). *Evaluating outcomes: Empirical tools for effective practice*. Washington, DC: American Psychological Association.
- Faulkner & Gray Health Care Information Center. (1997). *The 1997 behavioral outcomes and guidelines sourcebook*. New York: Author.
- (*) Fishman, D. B. (2001). From single case to database: A new method for enhancing psychotherapy, forensic, and other psychological practice. *Applied & Preventive Psychology*, 10, 275–304.
- Fonagy, P., & Target, M. (1994). The efficacy of psychoanalysis for children with disruptive disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, 33, 45–55.
- (*) Frisch, M. B. (1998). Quality of life therapy and assessment in health care. *Clinical Psychology: Science and Practice*, 5, 19–40.
- Hayes, S. C., Barlow, D. H., & Nelson, R. O. (1999). *The scientist-practitioner: Research and accountability in clinical and educational settings* (2nd ed.). New York: Pergamon Press.
- Heller, K. (1971). Laboratory interview research as an analogue to treatment. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change: An empirical analysis* (pp. 126–153). New York: Wiley.
- Hoagwood, K., Hibbs, E., Brent, D., & Jensen, P. J. (1995). Efficacy and effectiveness in studies of child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology*, 63, 683–687.
- Howard, K. I., Brill, P. L., Lueger, R. J., & O'Mahoney, M. T. (1992). *Integra Outpatient Tracking Assessment*. Radnor, PA: Integra.
- (*) Howard, K. I., Moras, K., Brill, P. L., Martinovich, Z., & Lutz, W. (1996). Efficacy, effectiveness, and patient progress. *American Psychologist*, 51, 1059–1064.
- Kazdin, A. E. (1978). Evaluating the generality of findings in analogue therapy research. *Journal of Consulting and Clinical Psychology*, 46, 673–686.

- Kazdin, A. E. (1981). Drawing valid inferences from case studies. *Journal of Consulting and Clinical Psychology*, 49, 183–192.
- Kazdin, A. E. (1982). *Single-case research designs: Methods for clinical and applied settings*. New York: Oxford University Press.
- Kazdin, A. E. (1993). Evaluation in clinical practice: Clinically sensitive and systematic methods of treatment delivery. *Behavior Therapy*, 24, 11–45.
- Kazdin, A. E. (Ed.). (1996). Special section: Evaluation in clinical practice. *Clinical Psychology: Science and Practice*, 3, 144–181.
- (*) Kazdin, A. E. (1999). The meanings and measurement of clinical significance. *Journal of Consulting and Clinical Psychology*, 67, 332–339.
- (*) Kazdin, A. E. (2001). Almost clinically significant ($p < .10$): Current measures may only approach clinical significance. *Clinical Psychology: Science and Practice*, 8, 455–462.
- Kazdin, A. E. (2003). *Research design in clinical psychology* (4th ed.). Needham Heights, MA: Allyn & Bacon.
- Kendall, P. C. (Ed.). (1999). Special section: Clinical significance. *Journal of Consulting and Clinical Psychology*, 67, 283–339.
- (*) Kiresuk, T. J., & Garwick, G. (1979). Basic Goal Attainment Scaling procedures. In B. R. Compton & B. Gallaway (Eds.), *Social work processes* (Rev. ed., pp. 412–420). Homewood, IL: Dorsey.
- (*) Kiresuk, T. J., Smith, A., & Cardillo, J. E. (1994). *Goal Attainment Scaling: Applications, theory, and measurement*. Hillsdale, NJ: Erlbaum.
- Klerman, G. L., Weissman, M. M., Rounsaville, B. J., & Chevron, E. (1984). *Interpersonal psychotherapy of depression*. New York: Basic Books.
- Kordy, H., Hannöver, W., & Richard, M. (2001). Computer-assisted feedback-driven quality management for psychotherapy: The Stuttgart-Heidelberg Model. *Journal of Consulting and Clinical Psychology*, 69, 173–183.
- (*) Lambert, M. J., Hansen, N. B., & Finch, A. E. (2001). Client-focused research: Using client outcome data to enhance treatment effects. *Journal of Consulting and Clinical Psychology*, 69, 159–172.
- Lambert, M. J., Hansen, N. B., Umphress, V., Lunnen, K., Okiishi, J., Burlingame, G., et al. (1996). *Administration and scoring manual for the Outcome Questionnaire (OQ 45.2)*. Wilmington, DE: American Professional Credentialing Services.
- (*) 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 and Practice*, 10, 288–301.
- Lindsey, E. W. (1998). The impact of homelessness and shelter life on family relationships. *Family Relations*, 47, 243–252.
- (*) Lueger, R. J., Howard, K. L., 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.
- Meier, S. T. (2003). *Bridging case conceptualization, assessment, and intervention*. Thousand Oaks, CA: Sage.

- OPTAIO. (1997). *Provider's desktop*. San Antonio, TX: Psychological Corporation.
- Parsonson, B. S., & Baer, D. M. (1978). The analysis and presentation of graphic data. In T. R. Kratochwill (Ed.), *Single-subject research: Strategies for evaluating change* (pp. 101–165). New York: Academic Press.
- Persons, J. B., & Silberschatz, G. (1998). Are results of randomized controlled clinical trials useful to psychotherapists? *Journal of Consulting and Clinical Psychology*, 66, 126–135.
- Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis* (2nd ed.). New York: McGraw-Hill.
- Sechrest, L., Stewart, M., Stickle, T. R., & Sidani, S. (1996). *Effective and persuasive case studies*. Cambridge, MA: Human Services Research Institute.
- Sorenson, R. L., Gorsuch, R. L., & Mintz, J. (1985). Moving targets: Patients' changing complaints during psychotherapy. *Journal of Consulting and Clinical Psychology*, 53, 49–54.
- Tang, T. Z., & DeRubeis, R. J. (1999). Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology*, 67, 894–904.
- (*) Westen, W., Novotny, C. M., & Thompson-Brenner, H. (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin*, 130, 631–663.
- Wiger, D. E. (1999). *The psychotherapy documentation primer*. New York: Wiley.

Evidence-Based Psychotherapy

*Where Practice and
Research Meet*

Edited by

Carol D. Goodheart
Alan E. Kazdin
Robert J. Sternberg

(2006)

American Psychological Association • Washington, DC