Pragmatic Case Studies in Psychotherapy, http://pcsp.libraries.rutgers.edu
Volume 3, Module 4, Article 2, pp. 37-49, 10-01-07 [copyright by author]

Commentary on Story of "Hope":
Successful Treatment of Obsessive Compulsive Disorder

Systematic Evaluation to Improve the Quality of Patient Care: From Hope to Hopeful

ALAN E. KAZDIN [a,b,c]

[a] President-Elect, American Psychological Association; Department of Psychology, Yale University
[b] Correspondence concerning this article should be addressed to Alan E. Kazdin, Department of Psychology, Yale University, PO Box 208205, New Haven, CT 06520-8205.
Email: alan.kazdin@yale.edu
[c] The author's work is supported in part by a grant from the National Institute of Mental Health (MH59029).

ABSTRACT

Clement provides a case study that illustrates an excellent approach toward treatment. The approach conveys concretely the interplay of research and clinical application, assessment and systematic evaluation, and benchmarking of case progress based on similar cases in clinical applications and in research studies. The present comment underscores the important role of systematic evaluation as exemplified in this case report and as a key to high quality patient care. In addition, evaluation accumulated across multiple cases can contribute to the knowledge base both by testing and generating hypotheses about treatment. Evaluation is placed in the context of current discussions about evidence-based treatments and the ongoing debates about clinic practice and research.

Key words: evaluation in clinical practice; quality of patient care; accountability

---

Clement (2007) provides a detailed description of psychotherapy for a woman (named Hope) who is seen for an extended period. The case is used to illustrate a very special approach toward treatment. In the psychotherapy biz, "approach" typically refers to a conceptual view (e.g., psychodynamic, cognitive-behavioral, humanistic, multimodal, or more often than not eclectic). To be sure, this report encompasses various approaches in this sense, but much more significant is the approach toward treatment that includes the weaving of clinical work, evaluation, and empirically based patient care.

There are several characteristics of the approach to treatment. This commentary highlights critical features of the approach and places them in the context of contemporary issues about treatment. Systematic evaluation is a key feature of the approach and is advocated in these comments as a way to improve the quality of patient care and to draw on clinical work to add significantly to the knowledge base about treatment. Although the original article details a case

and her treatment, the present comments are restricted to the broader contribution that this case illustrates.

KEY CHARACTERISTICS OF THE APPROACH

The approach to treatment illustrated by this case and evident in Clement's other publications (e.g., Clement, 1999) includes several remarkable features. Salient among these are:

- Systematic evaluation of the individual over time to gauge treatment effects;
- Benchmarking of change in the patient by quantitative comparison with other cases;
- Reference back and forth from research to identify interventions to include in treatment and to evaluate the impact of treatment;
- Use of standardized metrics (effect sizes) to evaluate change; and
- Accumulation of a data over time from multiple cases.

There are of course other facets of the approach such as an effort to tailor the treatment to features of the client, to draw on multiple interventions or components of multiple interventions in developing the individualized treatment package, to use clinical judgment to make decisions about what to do in treatment, when, and so on. These latter features are not unique to this case or approach, although the approach can bring them into an empirical arena. For example, almost all of us say we:

- Individualize treatment to the "needs" of the client, although it is difficult to find a defensible, replicable, or empirically based way of doing that;
- Draw on "what works" from many different treatments to develop an individualized treatment package, although our particular selection of what is included or omitted and the relative proportion of each have no clear defensible or empirical basis and are not clearly reliable among clinicians treating the same individual; and
- Use our judgment for decision making about when to begin and end a particular component of treatment and how well the client is doing, although research (e.g., on clinical judgment, cognitive heuristics) at the very least raises questions regarding how this is accomplished (Garb, 2005).

It is the absence of a defensible method of evaluating therapy in clinical work that has made these traditional claims and others of the same ilk subject to criticism and immune from verification. Once one introduces systematic evaluation, the common claims we provide as

therapists now enter into an empirical arena. For example, when we say we individualize therapy, data on patient change provide information that the treatment in fact may have made a difference. When the data are accumulated across multiple cases, additional comments might be made about the impact of strategies common to different cases and whether some techniques or decisions are associated with particular outcomes. Data such as these can be used to draw inferences and to suggest important lines of work to be followed in research.

CONTEXT AND BROADER THEMES

It is useful to situate the approach Clement illustrates in the context of clinical psychology and current debates and controversies. There is a long-standing divide between clinical research and practice in relation to psychotherapy. For present purposes it is useful to distinguish two related facets of this debate. First, there is the substantive issue of what treatments ought to be used. By substantive I refer to the actual treatment that is used in the context of therapy, i.e., what techniques, procedures, and interventions to effect change. Second, there is an evaluative issue that lamentably is not a matter of open (or even closed) debate. By evaluative, I refer to how the gains or progress in treatment are assessed, i.e., what techniques, procedures, or measures are used.

The hiatus of research and practice often mixes these for good reason or focuses only on the treatment technique. The substantive facet of the debate is illustrated in the use of evidence-based treatments (EBTs). The debate includes many points but key among them is the concern that EBTS are not ready to use in clinic settings in large part because of the way these treatments are investigated. Given the highly controlled circumstances of research, will the results generalize to clinical practice (e.g., Hunsley, 2007; Persons & Silberschatz, 1998; Westen, Novotny, & Thompson-Brenne, 2004)? The discussion of the relation of treatment in well controlled laboratory settings and clinical work is not new and flares up regularly (e.g., Borkovec & Rachman, 1979; Heller, 1971; Kazdin, 1978).

The substantive issue (what therapy to use) is obviously related to the evaluative issue of how the treatments have been studied. Randomized controlled clinical trials (RCTs) are critically important to evaluating treatment — whether for cancer, heart disease, or depression. In clinical practice, it is often difficult for us to move from group results, mean differences between treatment and control groups, and statistically significant differences, to what to do with this individual patient in treatment right now who does not quite seem like that patient in the group study. It is always the case that an effective treatment is not necessarily effective with everyone on all occasions, so it is easy to say that this case before me now is unlike the case in the RCT and perhaps unlikely to respond like those in that study. Also, every case in front of me will invariably look different from the participants included in a study in part because method sections of research papers are relatively superficial in how they describe people, and the “real” human in front of me has all sorts of unique features, foibles, and characteristics, any one of which might moderate (influence) therapeutic change.

We learn early in our training that absent other data and information, the mean from a group is a good estimate of what will happen to individuals in that group — in other words, it is rational and reasonable to say that a person with an anxiety disorder is likely to respond in a fashion that is similar to the group in the RCT that had the same or closely related problem. It is easier to convey this point outside of the context of psychotherapy.

Consider life span predicted for a given individual. Assume for a moment we wish to predict the life span of a male who was born in 1970, who smokes cigarettes, has high blood pressure, does not exercise, and has a few other not so great habits. Using an available, even if not rigorously scientific calculator designed for this purpose, we learn that such an individual is expected to live into his mid 60s (https://www.livingto100.com/lifecalc.html?accept.php). The mean is likely to be wrong, not only because everyone with this profile will not hit the same number (e.g., 66 years of age), but also because critical other variables that were omitted (family history; good, bad, and ugly cholesterol levels), and "chance" (e.g., an unfortunate accident on the way to buy cigarettes) can change everything and routinely do. Although the mean is invariably wrong, it is much less likely to be as wrong as an estimate made on the basis of judgment that does not systematically integrate the data or that is based on impression alone.

There are of course separate issues here including obtaining systematic information and then integrating the information based on what predicts a particular outcome and the extent of its contribution to that outcome. The evaluative issue about the clinical-research divide encompasses the use of systematic measures and their integration in patient care. Leaving aside the substantive issue of what intervention is actually used in treatment, there is the matter of evaluating patient care or therapeutic change. Here the contrast of clinical work and research is particularly sharp. Clinical work infrequently uses objective measures to evaluate outcome; research uses them routinely.

In the context of therapy, we want to evaluate patient progress. Of course, every time we see a patient we are evaluating progress, but I mean *systematic* evaluation as that term is used in psychology. By systematic evaluation I refer to the use of psychological or other measures (e.g., in applications to physical health) that have or in principle could have established reliability and validity and provide replicable information about the status of patient functioning. By replicable, I mean that two clinicians working separately would obtain very similar results on the measure and would see the information in roughly the same way. There is no need to threaten the use of clinical judgment or to say these measures replace clinical judgment. That is important to state to make even more salient the other side; clinical judgment is no substitute for systematic evaluation.

THE CASE FOR SYSTEMATIC EVALUATION IN CLINICAL PRACTICE

Suggestions for systematic evaluation of the individual case in the context of treatment are not new. The suggestions include ways of assessing and reporting cases to improve clinical care and contribute to the knowledge base (e.g., Bloom & Fischer, 1982; Cone, 2000; Fishman,

2001; Hayes, Barlow, & Nelson, 1999; Kazdin, 1993, 1996; Meier, 2003). What is rare is actually seeing how such measures are used in practice and how, as in Clement's work, they can be seamless.

Quality of Patient Care.

The key argument favoring systematic evaluation pertains to the primary purpose of practice — high quality patient care. First, the effectiveness of treatment in any given case cannot be assumed. Whether one uses an EBT or one's own brand of individualized eclectic treatment, one cannot be sure in principle or practice that the treatment will be effective. In relation to EBTs, for example, aspirin, by-pass surgery, plastic surgery, chemotherapy, anti-depressant medication — all firmly established in specific contexts — cannot be depended on invariably to produce the desired outcome. Evaluation is all the more important in patient care because the individual is so important and because we do not have a guarantee of the result.

Second, it is important to monitor treatment effects in an ongoing way to make decisions about continuing or terminating treatment and altering treatment based on how well the patient is responding. Now well documented is the fact that some patients make rapid changes quite early in treatment, so called sudden therapeutic gains, as for example, shown in patients with depression or anxiety (e.g., Hofmann, Schulz, Meuret, Moscovitch, & Suvak, 2006; Tang & DeRubeis, 1999); others may not make expected changes and not respond to treatment (so-called signal-alarm cases [Lambert et al., 2003]). And of course there are the gradations in between and cases where change occurs in some areas of functioning but not in others or at different rates among the different areas. Systematic assessment permits finer delineation of therapeutic changes than would be the case with more global clinical judgments and unsystematic assessment.

Third, systematic evaluation is intended to add to clinical 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. There are stark, undeniable limitations in perception and cognition in gathering information and drawing conclusions (e.g., Gilovich, Griffin, & Kahneman, 2002; Pohl, 2004; Roediger & McDermott, 2000). Systematic measures have their own artifacts and biases, but these can be evaluated and even corrected or taken into account in systematic ways.

Fourth, as clinicians we often underscore the complexity of clinical cases. Clients bring multiple problems to treatment, the problems change as treatment begins, and new problems emerge (e.g., Sorenson, Gorsuch, & Mintz, 1985). The complexity of treatment cases (e.g., comorbidity, disruption of the family) has been used as one argument about why one might not use EBTs, although therapeutic change is not undermined or impeded by complex and comorbid cases (Doss & Weisz, 2006; Kazdin & Whitley, 2006). Leaving aside EBTs, clinical cases are complex and I, as other clinicians, am surprised daily on this score. That said, complexity is an argument for systematic evaluation. Are the goals of treatment being achieved? Will my version of treatment or some EBT work in this situation with its many nuances? As the case becomes

more complex and multifaceted, and the therapeutic changes more subtle, nuanced, or impeded, a fortiori does systematic evaluation become important.

Do we really need systematic evaluation? Why not just ask the clients how they are doing? Has our science moved so far away from people, that we just can't ask, "Are you better?" Research can be brought to bear on these questions. Briefly, yes we must ask the clients, but we also must tie ourselves to the mast. Our own judgments have all sorts of heuristic filters, as I mentioned already. The clients have their own issues. Adults who have received psychotherapy are generally satisfied with their treatment, fairly independently of the specific treatment they receive (Consumers Union, 1995). More pertinent, satisfaction with treatment is not highly related to therapeutic change (Ankuta & Ables, 1993; Lambert, Salzer, & Bickman, 1998; Pekarik & Wolff, 1996). Satisfaction with treatment is important, but it is readily distinguishable conceptually and apparently empirically from improvements on those domains that prompted treatment. Yes, we do need systematic evaluation in part to separate satisfaction with achieving any of the goals of treatment.

From my perspective the case for systematic evaluation is made on the priority of and our commitment to patient care. Leaving aside the substantive issue of what treatments (e.g., EBTs) are used in clinical practice, what objection could there be to evaluation in clinical practice? Yesteryear arguments (e.g., "assessment will harm the therapeutic relationship," or, "I [as therapist] know when progress and change have occurred") are still invoked but will pass as new cohorts of therapists emerge and findings and methods of therapy research diffuse into therapist training. There are situations in which patient evaluation is not likely to be appropriate. On the battlefield (e.g., real battlefield) and in the crises of trauma (e.g., tsunamis, hurricanes, and earthquakes) one cannot easily begin with systematic evaluation. This is not the common application of treatment — patient care requires jumping in (but see Basoglu, Salcioglu, Livanou, Kalender, & Acar, 2005).

Contribution of Clinical Practice to our Knowledge Base

Apart from improving patient care, evaluation in clinical practice can make novel, important, and scientifically sound contributions to the knowledge base. The accumulation of cases over time, each of which is systematically evaluated, can yield new insights about treatment process and outcome. The fact that the conditions are not controlled, in the sense of experimentation, does not preclude their value in adding significantly to the knowledge base. Two illustrations can make this concrete. First, a key research focus is identifying moderators of treatment, that is, those characteristics (e.g., of treatment, patients, therapists, setting, and contexts) that may influence therapeutic change. Therapists have scores of fascinating and clinically informed hypotheses about what makes a difference and often act on them without knowing that they actually do (e.g., Kazdin, Siegel, & Bass, 1990). Moderators could readily be studied from accumulated cases in clinical work. One could readily ask from such work, are there special variables in the case sheets, demographics, or notes that influence responsiveness to treatment? Systematic data on progress and outcome permit evaluation of these questions. Second, the accumulation of data can generate hypotheses. A database from a given clinical practice might suggest treatments that work or do not work or suggest critical moments or phases

of treatment associated with change. Specifically, data obtained in clinical practice exactly like Clement (2007) illustrates in the case of Hope and in his other work (e.g., Clement, 1999) could contribute to knowledge directly, could generate hypotheses to be studied in research, and in the process make treatment research more aligned with and relevant to clinical practice.

A critical deficit in our clinical training is in the evaluation of clinical cases. It is possible, for example, to draw causal inferences from cases studies, to bring to bear information from the case, to help control rival hypotheses, to test and generate hypotheses, and to provide outstanding and astounding findings that will add to the knowledge base (see Sechrest, Stewart, Stickle, & Sidani, 1996). There are ways of arranging the clinical situation that even constitute quasi-experiments, a term legitimized by group research (e.g., Kazdin, 1981). Our training feeds a clinical-research split by not conveying how evaluation can help patients and the accumulation of knowledge.

ARE METHODS TO EVALUATE PSYCHOTHERAPY AVAILABLE FOR USE IN CLINICAL WORK?

On the assumption that evaluation is desirable, are there any methods or measures that can be used beyond those developed by individual practitioners? Before answering one ought not to rule out measures that are developed by individual practitioners. First, psychologists know how to develop measures and are in an excellent position to do that. Clement's report conveys one such measure, and he has developed a broader set with wide applicability (Clement, 1999). Second, the "home-made" measures that an individual clinician develops will be readily available for evaluation if and as needed. That is, the accumulated information can even be used to improve and evaluate the measure.

Not everyone wants to develop a measure. Also, measures for use in clinical work have special requirements that begin with feasibility and user (clinician, client) friendliness, before we even address reliability and validity. Actually, methods have emerged that are reliable, valid, and user friendly. Within the past decade *patient-oriented research* has emerged as a complement to the usual RCT (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 using the information to chart progress and make decisions about treatment. Unlike RCTs, an extensive battery of pre- and posttreatment assessments is not provided. Rather, assessment is conducted each session with a brief measure that captures functioning in diverse domains.

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 five minutes to complete and provides information on four domains of functioning, including 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 of course target the more specific areas. Treatment is evaluated by discerning the extent to which each individual makes a change or fails to make a change. During treatment different criteria can be used to guide treatment decisions. Research has evaluated different types of reliability and validity. Over 10,000 patients have been included in the various reports (see Lambert et al., 2001, 2003).

Another example is the COMPASS Outpatient Treatment Assessment System (Howard, Brill, Leuger, & O'Mahoney, 1992; Lueger et al., 2001), a measure that includes 68 items comprising 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. [Editor's note: The latest version of the "COMPASS" assessment system is now called the "Polaris-MH Outcomes Management System." For a recent description and discussion of it, see Grissom & Lyons, 2006, with accompanying commentaries by Stickle, 2006, and Lueger, 2006.].

The two prior examples provide measures with a fixed set of items, even though they are quite broad and cover domains likely to be relevant for most adults who come to treatment. Another example is Goal Attainment Scaling and consists of an assessment strategy that individualizes treatment goals. The method is based on collaborating with the patient at the outset of treatment about the goals and expectations of treatment (Kiresuk & Garwick, 1979; Kiresuk, Smith, Cardillo, 1994). Among the advantages of this method is the ability to both individualize assessment and treatment and to also combine data across multiple cases.

General Comments

Only a decade ago, few measures were available with research in clinical settings. Now that measures are available, and I have highlighted only a few, their use ought to be much more strongly encouraged in clinical training and practice. We are in an unnecessary battle with reimbursement agencies about why this or that treatment is being used. More pressure will be evident to ask why an evidence-based treatment is not being used when one is available. After that is likely to follow, what data do we have about our individual patients to justify the treatment? In advance of having our care more managed, there is an opportunity for us to improve clinical care with individuals and to make the case for what we do based on quality of our care.[1]

[1] This is not the place to take up the weighty issues of reimbursement. If treatment currently is not reimbursed at rates sufficient to cover the costs of treatment delivery, adding an evaluation component will be seen as a move in the wrong direction. That is, will this not just add to time of treatment delivery and provide yet another cost on the backs of practitioners? Among the arguments against this: Evaluation need not take added time, and psychologists can uniquely make the case that evaluation is central to treatment delivery.

Over the long term, our field would benefit enormously from the knowledge that is lost everyday from clinical practice. Examples of contributions to the knowledge base from accumulating 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). Progress has been made in identifying measures that can be used in individual therapy, can be applied widely across clinical and service settings and clients, and can accumulate information that contributes to the knowledge base more generally (e.g., Barkham et al., 2001; Kordy, Hannöver, & Richard, 2001; Sperry, Brill, Howard, & Grissom, 1996). How unfortunate for patients and for our science to lose the opportunities afforded by systematic evaluation.

CONCLUSIONS

Clement presents a case study to illustrate his approach toward treatment. The features of the case, including the scope of the clinical issues, the range of techniques, and the thoughtful decision making on his part are all engaging to say the least. I have refrained from commenting on the case to consider what I believe to be a remarkable approach toward patient care that this case and Clement's prior publications reflect. The approach includes assessment and evaluation, reference back and forth to the research to identify viable treatments and to benchmark therapeutic change of individual patients, and the accumulation of data over time for individual patients and across multiple patients. I have emphasized systematic evaluation because I take this as the key ingredient to improve patient care.

Within psychotherapy, there is a history of avoiding systematic evaluation in patient care almost in principle. Of course there are wonderfully stark exceptions where clinical work has relied on evaluation, but salient instances are from researchers who were also clinicians and clinicians who were also researchers (e.g., Howard et al., 1996). In the context of daily clinical work, systematic evaluation of patient problems and progress over the course of treatment is rare. Moreover, there is no infrastructure to support systematic evaluation. For example, clinical psychology, psychiatry, and social work training programs, accreditation requirements, licensing criteria, and reimbursement requirements do not include or require systematic evaluation of patient progress. There is not likely to be a groundswell of patient interest in moving in the direction of improved evaluation. Patients are not likely to appreciate the issue and its importance and apparently are happy with what we do, whether it is evidence-based, evaluated, or effective.

From my perspective, the primary, guiding issue in clinical work is quality of patient care. For that, we want the best available information about patient functioning and progress. In the context of medical care, if there are systematic measures available that have reliability or validity, I want my physician to use them. I concede that I am delighted with the subjective indices of health when she says to me, "You are looking well, you look pretty healthy, and it seems as if you are still ambulatory, out of pain, and not grossly obese." Despite comments to the contrary from key relatives, I am human and responsive to such statements. Moreover, although

her judgments are subjective, they probably are valuable and predictive. Indeed if any one of the statements varied from those I noted, that might signal a significant problem (e.g., "You look horrible" or "The extra 50 lbs all seem to be in your neck."). How wonderfully important to add to these subjective impressions analyses of blood work, blood pressure, heart and pulse rate, urinalysis, and then further tests that one of these might prompt. These latter tests and measures are not perfect just because they are objective and have high (but hardly flawless) reliability and validity of various types (see Gigerenzer, 2002). Indeed, most readers will have a story where at least one objective medical test was erroneous.

Medical analogies in the psychotherapy business are viewed as suspicious; they are either in bad taste or categorically ruled out as irrelevant to what we do in the context of psychotherapy. In some Hegelian process where our roots (e.g., philosophy, biology, medicine) served as a thesis, it made brief sense to reject all of this as part of the antithesis. Now that our profession has formal identity, recognition, and widespread impact, it would be advantageous to leap into the synthesis stage — we share features of any clinical endeavor, all the while retaining our uniqueness. Evaluating impact with the best possible measures to help make decisions for immediate care is relevant in virtually any context in which a service is delivered and in which some outcome is desired. The fact that psychotherapy (and medicine) deals with people's lives makes this of unparalleled importance.

In relation to psychological services, clinical work, and patient care, one overriding question ought to guide us: How can we provide the best care available to this patient in front of us right now? For any answer to ignore the substantive findings from research or the methodological tools to that will aid decision making is difficult to imagine. Our experience and judgment *are* critical. However, as psychologists we have an extra reason or two to be humble about those. Critical areas of our field (e.g., cognitive heuristics, reality monitoring) convey that judgment, decision making, views, immediate perceptions, and recollections are not the mere accretion of unprocessed and unmodified experience. Supplementing our judgment with measures of patient progress provides an opportunity to improve patient care and improve our decision making. That the accumulated information across multiple cases could add remarkably to the knowledge base is a separate bonus.

In illustrating systematic evaluation, Clement presented the case of a woman called Hope, a pseudonym no doubt but a well chosen name for the report. The approach that the case illustrates includes systematic assessment and evaluation to benefit the patient. We as a field have not been moving toward this kind of work. In graduate school and internship training, in setting standards of psychological practice, and in suggesting to the public what to look for in their care, systematic evaluation of patient progress in treatment is not included or advocated, and this has not changed appreciably. Clement is fully aware of this situation and perhaps in his case wanted to offer us Hope.

REFERENCES

Ankuta, G.Y., & Abeles, N. (1993). Client satisfaction, clinical significance, and meaningful change in psychotherapy. *Professional Psychology: Research and Practice, 24*, 70-74.

Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., Benson, L., Connell, J., Audin, K., & McGrath, G. (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.

Basoglu, M., Salcioglu, E., Livanou, M., Kalender, D., & Acar, G. (2005). Single-session behavioral treatment of earthquake-related posttraumatic stress disorder: A randomized waiting list controlled trial. *Journal of Traumatic Stress, 18*, 1-11.

Bloom, M., & Fischer, J. (1982). *Evaluating practice: Guidelines for the accountable professional.* Englewood Cliffs, NJ: Prentice Hall.

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.

Clement, P.W. (2007). Story of "Hope:" Successful treatment of obsessive-compulsive disorder. *" Pragmatic Case Studies in Psychotherapy [Online]*, Vol. 3(4), Article 1, 1-36. Available: http://hdl.rutgers.edu/1782.1/pcsp_journal

Cone, J.D. (2000). *Evaluating outcomes: Empirical tools for effective practice*. Washington, DC: American Psychological Association.

Consumers Union. (1995). Mental health: Does therapy help? *Consumer Reports, 60 (November)*, 734-739.

Doss, A.J., & Weisz, J.R. (2006). Syndrome co-occurrence and treatment outcomes in youth mental health clinics. *Journal of Consulting and Clinical Psychology, 74*, 416-425.

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 Psychiatry, 33*, 45-55.

Garb, H.N. (2005). Clinical judgment and decision making. *Annual Review of Clinical Psychology, 1*, 67-89.

Gigerenzer, G. (2002). *Calculated risks: How to know when the numbers deceive you*. New York: Simon and Schuster.

Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment.* Cambridge: Cambridge University Press.

Grissom, G. & Lyons, J. (2006). Pragmatic case studies and evidence-based treatment: Research and clinical applications of a computerized outcomes management system. *Pragmatic Case Studies in Psychotherapy [Online]*, Vol. 2(3), Article 1, 1-28. Available: http://hdl.rutgers.edu/1782.1/pcsp_journal

Hayes, S.C., Barlow, D.H., & Nelson, R.O. (1999). *The scientist-practitioner: research and accountability in clinical and educational setting* (2nd ed). New York: Pergamon.

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 & Sons.

Hofmann, S.G., Schulz, S.M., Meuret, A.E., Moscovitch, D.A., & Suvak, M. (2006). Sudden gains during therapy of social phobia. *Journal of Consulting and Clinical Psychology, 74,* 687-697.

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.

Hunsley, J. (2007). Addressing key challenges in evidence-based practice in psychology. *Professional Psychology: Research and Practice, 38,* 113-121.

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. (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., Siegel, T., & Bass, D. (1990). Drawing upon clinical practice to inform research on child and adolescent psychotherapy. *Professional Psychology: Research and Practice, 21,* 189-198.

Kazdin, A.E., & Whitley, M.K. (2006). Comorbidity, case complexity, and effects of evidence-based treatment for children referred for disruptive behavior. *Journal of Consulting and Clinical Psychology, 74,* 455-467.

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.

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., Huefner, J.C., & Reisinger, C.W. (1996). *Administration and scoring manual for the Outcome Questionnaire (0q 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.

Lambert, W., Salzer, M.S., & Bickman, L. (1998). Clinical outcome, consumer satisfaction, and ad hoc ratings of improvement in children's mental health. *Journal of Consulting and Clinical Psychology, 66*, 270-279.

Lueger, R.J. (2006). Technology to support the clinical management of psychotherapy cases: Commentary on the Polaris-MH. *Pragmatic Case Studies in Psychotherapy [Online]*, Vol. 2(3), Article 3, 1-7. Available: http://hdl.rutgers.edu/1782.1/pcsp_journal

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.

Meier, S.T. (2003). *Bridging case conceptualization, assessment, and intervention.* Thousand Oaks, CA: Sage.

Pekarik, G., & Wolff, C.B. (1996). Relationship of satisfaction to symptom change, follow-up adjustment, and clinical significance. *Professional Psychology: Research and Practice, 27*, 202-208.

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.

Pohl, R.F. (Ed.) (2004). *Cognitive illusions: A handbook on fallacies and biases in thinking, judgment, and memory*. New York: Psychology Press.

Roediger, H.L., III & McDermott, K.B. (2000). Distortions of memory. In E. Tulving & F.I.M. Craik (Eds.). *The Oxford handbook of memory* (pp. 149-162). New York: Oxford University Press.

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.

Sperry, L., Brill, P.L., Howard, K.I., & Grissom, G.R. (1996). *Treatment Outcomes in Psychotherapy and Psychiatric Interventions*. New York: Brunner/Mazel.

Stickle, T.R. (2006). Empirical evaluation of the utility and effectiveness of clinically flexible application in evidence-based treatments. *Pragmatic Case Studies in Psychotherapy [Online]*, Vol. 2(3), Article 2, 1-7. Available: http://hdl.rutgers.edu/1782.1/pcsp_journal

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-Brenne, H. (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin, 130*, 631-663.