



<b>Title</b>	How do we find out what works for whom? Evaluating the efficacy and effectiveness of psychotherapy
<b>Authors(s)</b>	McMorrow, Eva, Carr, Alan
<b>Publication date</b>	2003-01
<b>Publication information</b>	McMorrow, Eva, and Alan Carr. "How Do We Find out What Works for Whom? Evaluating the Efficacy and Effectiveness of Psychotherapy" 2, no. 26 (January, 2003).
<b>Publisher</b>	Irish Association for Counselling and Psychotherapy
<b>Item record/more information</b>	<a href="http://hdl.handle.net/10197/5521">http://hdl.handle.net/10197/5521</a>

Downloaded 2024-05-19 20:08:48

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd\_oa)



© Some rights reserved. For more information

**McMorrow, E. & Carr, A. (2003). How do we find out what works for whom?  
Evaluating the efficacy and effectiveness of psychotherapy. *Éisteacht:  
Journal of the Irish Association of Counselling and Therapy*, 2 (26), 24-31.**

**How do we find out what works for whom?**

**Evaluating the Efficacy and Effectiveness of Psychotherapy**

**Eva McMorrow**

**Midland Health Board**

**Alan Carr**

**Department of Psychology, University College Dublin**

**Submitted to,** Editor of Eisteacht,

**Word count:** 5700 words with references and table

**Biographical note:** Eva Morrow is a psychologist with the Midland Health Board. Alan Carr is Director of the clinical psychology training programme at UCD.

**Correspondence address:** Alan Carr, Director of the Clinical Psychology Training Programme, Dept of Psychology, Science Building, University College Dublin, Belfield, Dublin 4, Ireland. email: [alan.carr@ucd.ie](mailto:alan.carr@ucd.ie), Phone: +353-1-7162390, FAX: +353-1-7162846

**ABSTRACT**

Controlled randomized clinical trials of psychotherapy have traditionally been used to test the efficacy of specific forms of psychotherapy for specific problems. The value of findings from such efficacy studies for practicing psychotherapists has been questioned because these studies involve clients and therapy procedures that are radically different from those typically used in routine clinical practice. Opponents of efficacy research have proposed health service-based effectiveness research as a more valuable alternative to efficacy research. Arguments for and against rigorously controlled efficacy research on the one hand, and 'real-world' effectiveness research on the other are explored in this paper.

## **INTRODUCTION**

In this era of evidence based practice, there is an increasing focus on evaluation of psychotherapy outcome and the use of this information as a guide for routine clinical practice (Carr, 2000; Nathan, Stuart and Dolan, 2000). The impetus for this is clinical, scientific, and economic. Clinicians wish to offer patients effective treatment. Scientists wish to understand how effective psychological change may be facilitated. Health care agencies want to know how to get value for money in the delivery of psychotherapeutic services. However, there are distinct differences between psychotherapeutic processes commonly evaluated in controlled research studies and the treatment processes typical of routine clinical practice. In controlled research, brief manualized treatments are typically offered to help homogenous groups of clients address circumscribed, acute psychological problems. In routine clinical practice, flexible open-ended treatment programmes are commonly offered to help heterogeneous groups of clients cope with complex, chronic psychological difficulties. The value of findings from controlled treatment outcome studies which involve problems, populations, therapists and interventions that differ markedly from those typically used in health service settings has been questioned. The debate has led to arguments for and against tightly controlled efficacy research on the one hand and 'real-world', effectiveness research on the other. In this paper some of the key issues in this debate will be explored.

## **EFFICACY AND EFFECTIVENESS**

Efficacy research studies aim to investigate whether particular treatments work under controlled conditions while effectiveness studies aim to investigate the generalisability, feasibility and cost-effectiveness of treatments of proven efficacy (Jacobson and Christensen, 1996). Efficacy research usually takes place under tightly controlled

experimental conditions in a university clinic and the methodology of choice is generally the randomised clinical trial (RCT). Effectiveness research, on the other hand, is usually carried out in health services where the use of stringent exclusion criteria, random assignment of cases to treatment and control conditions, and manualized treatments are not always feasible. Efficacy research aims to maximize internal validity while effectiveness research aims to maximize external validity (Nathan, Stuart and Dolan, 2000). Internal validity refers to the degree to which one can draw conclusions about the causal nature of relationships found between treatments and outcomes. External validity refers to the degree to which the causal relationship found between treatments and outcomes is generalisable to cases and settings other than those in the study. A summary of some of the more important differences between efficacy and effectiveness research is given in Table 1.

### **Efficacy Research: Randomized Controlled Trials**

Advocates of efficacy research maintain that psychological treatments must first be evaluated under highly controlled conditions before moving on to effectiveness research where the aim is to check whether treatments of known efficacy in the university clinic can be transported into the health services and delivered in a cost-effective way (Chambless and Hollon, 1998; Hollon, 1996; Jacobson and Christensen, 1996; Kendall, 1998; Persons and Silberschatz, 1998). Efficacy research conforms to rigorous guidelines on research design, sample selection, assessment procedures, treatment implementation and data analysis to ensure that internal validity is maximised. Examples of such criteria are given in reports by the Task Force on Promotion and Dissemination of Psychological Procedures of Division 12 (Clinical Psychology) of the American Psychological Association (Task Force, 1995; Chambless et al., 1996, 1998). The argument entailed by such guidelines is that to

determine if a specific treatment is efficacious for a specific disorder, then a randomized controlled trial should be conducted and replicated. In such a trial, a group of homogeneous cases that meet the diagnostic criteria for a specific disorder and who have no other co-morbid diagnoses should be randomly assigned to treatment and control conditions. The treatment should be delivered by trained therapists following an explicit treatment manual for a specified number of sessions. Before and after treatment and at a later follow-up time point, the cases in the treatment and control groups should be assessed by an assessment team who are unaware of whether cases were in the treatment or control group. For these assessments, reliable and valid psychological assessment procedures should be used which allow the symptoms of the disorder targeted by the treatment to be evaluated. After treatment and at follow-up cases should be classified, on the basis of their scores on the assessment instruments, as those showing clinically significant improvement and non-improvers. The statistical significance of intergroup differences in improvement rates or mean post-treatment scores should be determined using appropriate inferential statistics. When a randomized controlled trial of this sort shows that a specific treatment leads to a greater improvement rate than that shown by a control group and this result is replicated at least once, then a treatment may be said to be empirically supported.

### **Effectiveness Research**

Those who argue for effectiveness research make a number of important proposals (Fonagy and Target, 1996; Garfield, 1996, 1998; Goldfried and Wolfe, 1998; Persons, 1991; Seligman, 1995). They insist that treatment outcome research should be conducted, not in university clinics, with carefully selected homogeneous groups of clients recruited through media advertisements, but in routine health service settings with heterogeneous

groups of patients referred for treatment. Many of these patients will have multiple problems including the one of interest to the research team. Thus, in a study of major depression, patients may have co-morbid personality disorders or anxiety disorders. Proponents of effectiveness research accept that often it may be impractical or unethical to randomly assign cases to treatment and control groups. However, the comparability of treatment and control group cases on baseline demographic and diagnostic variables may be evaluated, post-hoc, using appropriate inferential statistics. Where baseline intergroup differences occur, procedures such as analysis of covariance may be used to statistically control for the effects of these baseline differences. In effectiveness research, experienced clinicians deliver treatments informed by a variety of theoretical perspectives in a way that is tailored to the unique needs of each patient. This, it is argued, is more appropriate than delivering a manualized treatment in a rigid way to all clients regardless of individual differences. Effectiveness researchers, like efficacy researchers argue that statistically significant intergroup differences in clinically significant improvement rates should be the basis for deciding if a treatment has been shown to work. However, they propose that it is not enough to simply measure symptomatic improvement. Quality of life and psychosocial adjustment should also be evaluated to determine if the treatment had an impact on clients' overall lifestyles.

### **CRITICISMS OF EFFICACY RESEARCH**

The overriding criticism of efficacy outcome research is that its findings lack external validity. Concluding that a treatment can produce change under controlled conditions in a university clinic does not guarantee that the same treatment will produce change in a health service setting (Chambless and Hollon, 1998; Persons and Silberschatz, 1998).



## Client Characteristics

In efficacy research, cases are usually screened using systems such as the DSM IV (American Psychiatric Association, 1994) or the ICD 10 (World Health Organization, 1992) in order to produce diagnostically homogeneous groups for treatment and evaluation. To be included in a RCT, typically cases must meet a specific set of diagnostic criteria, such as those for a major depressive disorder. Cases that do not meet these criteria or who meet these criteria but have additional co-morbid DSM IV axis I diagnoses or axis II personality disorders are excluded. This process of only including 'pure' cases in treatment efficacy studies allows the impact of specific treatments on specific diagnoses to be reliably evaluated (Chambless and Hollon, 1998).

Those who object to this process, claim that rigorous screening with stringent inclusion and exclusion criteria results in studies of cases characterized by single, easily treated presentations, and these are not representative of more complex cases with multiple co-morbid axis I and II diagnoses typically seen in clinical practice (Beutler, 1998; Kazdin and Weisz, 1998; Kendall, 1998; Silver and Silver, 1983; Tuma and Pratt, 1982; Weisz, Donnerberg, et al., 1995).

In response to this criticism Chambless and Hollon (1998) propose that it is possible to define inclusion criteria so as to include co-morbid complex presentations representative of cases seen in routine clinical practice. For example Frank, Kupfer, et al. (1990) evaluated the treatment of patients with recurrent depression and Linehan, Armstrong, et al. (1991) evaluated the efficacy of dialectical behaviour therapy for parasuicidal patients with borderline personality disorder. Jacobson and Christensen (1996) suggest stratifying clients according to problem complexity and then randomly assigning them to treatments within this stratification to investigate the efficacy of particular treatments with different degrees of problem complexity.

Seligman (1995) has argued that in routine clinical practice clients “actively shop” for a therapist or treatment that appeals to them, whereas in an RCT, they are randomly assigned to a treatment condition and a therapist. Thus, participants in RCTs are not representative of patients in routine clinical practice, insofar as they may have a different level of motivation or degree of investment in the therapy process. Jacobson and Christensen (1996) argue that the effects of selecting versus being assigned to a therapist or treatment programme may be evaluated within the context of RTC’s. Such a design would involve having two conditions, one where the clients are given information about various therapists and allowed to choose one, and another condition where the clients are allocated to a therapist and not allowed to actively seek one out.

### **Treatment Manuals**

Treatment manuals are a central part of therapy efficacy studies (Nathan, Stuart, et al., 2000). A treatment manual is a statement of the theory, principles and procedures which define a psychological treatment programme and includes both prescribed and proscribed interventions along with guidelines on treatment duration (Chambless and Ollendick, 2001). Manuals guarantee a high degree of standardisation of treatment procedures across different therapists and across the same therapists with different cases (Crits-Christoph and Mintz, 1991; Eifert, Schulte, et al., 1997). Their focus on specific treatment techniques and goals increases internal validity, facilitates replication of studies by different groups of investigators, and makes it easier to train therapists (Wilson, 1996).

Critics of treatment efficacy research have highlighted a number of problems with treatment manuals. Manualized treatments evaluated in RTC’s are not representative of treatments used in routine clinical practice which are more intuitive, self-corrective and flexible (Beutler, 1998; Edelson, 1994; Eifert, Schulte, et al., 1997; Goldfried and Wolfe,

1998; Persons and Silberschatz, 1998). In a major US survey of clinical psychologists Addis and Krasnow (2000) found that 47% of respondents had never used manuals. Seligman (1995) maintains that manuals prevent therapists from engaging in self-corrective processes thereby restricting the process of adapting interventions to clients' individual needs. Clinical procedures specified in treatment manuals are usually based on a single theoretical model, whereas in routine clinical practice most clinicians offer eclectic or integrative treatment programmes based on a number of theories. Treatment manuals usually define the duration and number of therapy sessions whereas in routine practice clinicians show considerable flexibility in matching the intensity and duration of therapy to clients' needs.

In response to these criticisms, advocates of efficacy research offer a number of cogent counter arguments in support of the use of treatment manuals in efficacy research. Jacobson and Christensen (1996) argue that RCTs can be used to investigate the relative efficacy of treatments based on single and multiple theoretical models; therapies of fixed duration and varied duration; and manualized treatments and those used in routine clinical practice characterized by an intuitive, self-corrective and flexible style. There is a growing body of evidence which shows that the use of treatment manuals usually leads to better outcomes for clients (Luborsky, McLellan, et al., 1985, 1997; Schulte, Kunzel, et al., 1992; Weisz, Donenberg, et al., 1995). The flexibility with which manualized procedures are used seems to have little effect on treatment outcome. In a comparative study of the impact of behavioural marital therapy, Jacobson et al., (1989) found that the flexible use of manualized procedures did not lead to a better treatment outcome than rigid adherence to manualized behavioural marital therapy. However, not all therapists have equally good outcomes with all forms of manualized therapy. In a study of manualized psychodynamic and cognitive behavioural therapy, Shapiro et al., (1989) found that of four therapists, each

of whom treated six clients, one therapist had much better results than the other three with cognitive behavioural therapy.

Wilson (1996) points out that practice based on treatment manuals is more resistant to errors caused by inaccurate individual case formulation. Routine clinical practice can lead to too much flexibility, and a consequent reduction in the quality of the active ingredients of treatment because experienced clinicians are no less protected from cognitive biases in drawing inferences about behaviour and in making judgements about people than non-professionals (Chapman and Chapman, 1969; O'Donohue and Szymanski, 1994, Starr and Katkin, 1969; Wilson, 1996). Silverman (1996) also supports the use of treatment manuals because he argues that the underlying commonalties among individuals with particular sorts of problems or diagnoses are more important than individual differences.

### **Randomization**

In efficacy research cases are randomly assigned to treatment and control groups. This is done to avoid bias, such as inadvertently putting the 'easy' cases in the treatment group and the 'difficult' cases in the control group and thereby increasing the chances of proving that the treatment being evaluated is effective (Heppner, Kivlighan and Wampold, 1992). Shadish and Ragsdale (1996) found that studies in which cases were randomly assigned to treatment and control groups consistently yielded greater treatment effects and less variable treatment effects than studies in which cases were assigned to groups in a non-random way. This empirical finding underlines the importance of randomization. However, the process of randomization, a cornerstone of efficacy research, has been the subject of much criticism.

In RCT's it is never possible to control for all possible extraneous variables and to be 100% confident that the only important difference between the two groups is that one received the treatment being evaluated and the other did not (Heppner et al., 1992; Howard, Krause, and Lyons, 1993; Krause and Howard, 1999). Treatment and control groups may differ on important therapist and client variables, such as therapist skill or the social support available to clients. Improvement rates after therapy may be due to these factors rather than the fact that the one group received a particular treatment.

Goldfried and Wolfe (1998) argue that cases with similar diagnosis may differ on the factors that caused their problems, and failure to take this into account may invalidate the results of RCT's. For example, negative beliefs about the self may be the main causal factor for depression in some cases, whereas in others, the primary difficulty may be relationship problems. If these different etiological factors were not incorporated into the design of an RCT comparing the efficacy of cognitive and interpersonal therapy, the random assignment of depressed clients to these two conditions might result in some participants being matched to a more appropriate treatment than others. Probably cognitive therapy would be more appropriate for clients with negative beliefs about the self and interpersonal therapy those with relationship difficulties.

In RCT's that compare the efficacy of two types of treatment, groups may differ on non-specific therapeutic processes such as the quality of the therapeutic relationship and it may be this rather than technical aspects of the two treatment that accounts for different improvement rates.

On the basis of an extensive review of psychotherapy process and outcome research, Lambert (1992) estimated that 30% of the variance in therapy outcome is due to the therapeutic relationship; 40% is due to characteristics of clients and their social contexts outside of therapy; 15% is due to specific therapeutic techniques; and 15% is due

to hope and the expectation of improvement. Lambert's findings, and the criticisms of randomization raised in this section, suggest that ideally efficacy research should incorporate important therapist factors, client factors and relationship factors into the design of RCT's carefully matching treatment and control groups on factors within each of these domains (Garfield, 1998, Lambert and Bergin, 1994). Where pre-treatment matching is not possible, for example in the case of therapeutic relationship factors, then the comparability of groups on such variables may be determined after treatment. If significant intergroup differences on such variables are found, their impact on treatment outcome may be statistically controlled through procedures such as analysis of covariance.

### **Extent and Durability of Treatment Effects**

Treatment efficacy research has traditionally determined the extent and durability of the impact of treatment by evaluating the statistical significance of intergroup differences after treatment and a number of months later on measures of symptomatology. For example, to determine the impact of a treatment for depression it would be acceptable in an efficacy study to check the statistical significance of differences between treatment and control groups on mean scores on the Beck Depression Inventory or improvement rates indexed by the number of cases that no longer meet the DSM IV diagnostic criteria for major depressive disorder after therapy and at six months follow-up. This approach to evaluating the impact of specific treatments on specific symptoms has two advantages. First, the specific symptoms targeted by the therapy being evaluated (rather than other variables) are always measured in RCT's, so statements can be made about the effects of therapy on the target problem after therapy and at follow-up. Second, inferential statistics are always used in RCT's to out rule the possibility that intergroup differences after treatment

and at follow-up on measures of symptoms targeted by the therapy are a chance occurrence.

This approach to evaluating the extent and durability of treatment effects is not without its critics. RCT's in focusing almost exclusively on symptomatic change often neglect to evaluate the impact of therapy on clients' daily functioning and quality of life (Jacobson, 1988; Kazdin and Weisz, 1998; Seligman, 1995). Concerning the durability of treatment effects, Kazdin and Weisz (1998) have called for outcome evaluation not only at the end of treatment and a number of months later, but also at long-term follow-up, one or two years after therapy to provide information on the maintenance and robustness of therapeutic change and to document the course of change once treatment has ended.

In response to these criticisms it is becoming more common for efficacy studies to assess general functioning in addition to changes in specific symptoms and to conduct long-term follow-up evaluation (Blatt, Zuroff, et al., 2000; Borkovec and Castonguay, 1998; Kazdin and Weisz, 1998). For example, Blatt et al., (2000) in a study of cognitive therapy, interpersonal therapy, and antidepressant medication evaluated the level of life adjustment as well as symptom severity, and included an 18-month follow-up evaluation in the study design. Clients in both psychotherapy groups reported more positive life adjustment than clients in the placebo or medication group after therapy and a year and a half later.

## **CONCLUSIONS**

From this review of the arguments for and against efficacy research it is clear that there are ways in which RCT's could be modified to make results from them more generalizable to routine clinical practice. Clients more representative of those seen in routine clinical practice could be included in RCT's: that is, referred rather than media solicited clients with chronic, complex, co-morbid presentations. To control for therapist,

client, and relationship factors, treatment and control groups could be matched, *a priori*, or *post-hoc*, on relevant variables within each of these domains. To address reservations about brief, single-theory manualized interventions, RCT's could be conducted to evaluate the relative efficacy of rigidly applied manualized interventions based on single theories, on the one hand, and manuals for pan-theoretical, non-time-limited, flexible treatment programmes on the other. Finally, to address reservations about the way the impact and durability of treatment effects are measured, more comprehensive assessment protocols which evaluate overall adjustment as well as symptomatic remission could be used in RCT's and long-term follow-up assessments conducted.

Weston and Morrison (2001) suggest that two types of outcome studies should be conducted: traditional RCT's and effectiveness studies in which the impact of therapy is tested on larger, more generalisable, naturalistic samples with the types of design features suggested in the previous paragraph. Conducting these effectiveness studies would involve the setting up of practice-research-networks consisting of clinicians in the community who would deliver manualized treatments to routinely referred clients. Comprehensive assessments would be conducted by clinicians, clients and independent assessors before and after therapy and at repeated follow-up time-points over several years. Similar recommendations for conducting multi-site health services based treatment effectiveness research have been made by others (Klein and Smith, 1999; Sobell, 1996).

These large effectiveness research projects would not only throw light on *whether* specific therapies are effective for specific problems when delivered in routine health care settings, but they could also offer a context within which to explore *how* particular forms of therapy work and how they might be improved. That is, these type of effectiveness projects offer a context within which process and profiling studies may be conducted.



A particularly useful approach to process research is the events paradigm, which focuses on critical incidents of interaction between therapists and clients (Rice and Greenberg, 1984; Stiles, Shapiro, and Elliott, 1986). To conduct this type of research within the context of large health service based treatment effectiveness projects, therapy sessions are regularly audio or video taped and archived. Using this archive, critical incidents within therapy sessions are identified by therapists, clients or observers and qualitatively analysed to determine precisely how therapists and clients interact during these events. The impact of these critical events on treatment outcome is then determined by statistically evaluating the relationship between the frequency, duration and timing of these events and treatment outcome. Where critical events are associated with positive outcomes, treatment manuals may be modified so as to incorporate strategies that facilitate these change promoting critical events.

With profiling studies, subgroups of cases in large effectiveness studies are compared on pre-treatment, post-treatment and follow-up, client, therapist and treatment process variables and this information is used to modify treatment manuals (Borkovec and Castonguay, 1998; Donenberg, 1999; Krause and Howard, 1999; Persons and Silberschatz, 1998). For example, pre-treatment characteristics on which cases that have and have not responded to a manualized treatment differ significantly may be identified. By pinpointing these baseline characteristics on which the responders differ from non-responders, hypotheses about how treatment for non-responders might be improved may be developed and tested. For example, if non-responders to individual cognitive behavioural therapy have significantly greater social isolation and family difficulties at baseline, then the therapy manual might be revised to include additional modules on social skills training and family intervention for these types of cases. The effectiveness of this

revised manual for these types of cases may then be evaluated in an RCT or effectiveness study.

Modified efficacy studies which incorporate the design features outlined in the opening paragraph of this section, conducted in health service settings, and which incorporate process and profiling sub-studies, have a very important role to play in helping to identify and refine effective psychotherapies.

## REFERENCES

- Addis, M.E., and Krasnow, A.D. (2000). A national survey of practising psychologists' attitudes toward psychotherapy treatment manuals. *Journal of Consulting and Clinical Psychology*, **65**, 331-339.
- American Psychiatric Association (1994). *Diagnostic and Statistical Manual of the Mental Disorders (Fourth Edition) DSM –IV*. Washington, DC: APA.
- Beutler, L.E. (1998). Identifying empirically supported treatments: What if we didn't? *Journal of Consulting and Clinical Psychology*, **66**, 113-120.
- Blatt, S.J., Zuroff, D.C., Bondi, C.M., Sanislow, C.A.III. (2000). Short and long -term effects of medication and psychotherapy in the brief treatment of depression: Further analysis of data from the NIMH TDCRP. *Psychotherapy Research*, **10**, 215-234.
- Borkovec, T.D., and Castonguay, L.G. (1998). What is the scientific meaning of empirically supported therapy? *Journal of Consulting and Clinical Psychology*, **66**, 136-142.
- Carr, A. (Ed.) (2000). *What Works With Children and Adolescents? A Critical Review of Research on Psychological Interventions with Children Adolescents and their Families*. London: Routledge.
- Chambless, D.L., and Hollon, S.D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, **66**, 7-18.
- Chambless, D.L., Sanderson, W.C., Shoham, V., Bennett Johnson, S., Pope, K.S., Crits-Christoph, P., Baker, M., Johnson, B., Woody, S.R., Sue, S., Beutler, L., Williams, D.A., and McCurry, S. (1996). An update on empirically validated therapies. *Clinical Psychologist*, **49**, 5-18.

- Chambless, D.L., and Ollendick, T.H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology*, **52**, 685-716.
- Chapman, L.J., and Chapman, J.P. (1969). Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology*, **74**, 193-204.
- Crits-Christoph, P., and Mintz, J. (1991) Implications of therapist effects for the design and analysis of comparative studies of psychotherapies. *Journal of Consulting and Clinical Psychology*, **59**, 20-26.
- Donenberg, G.R. (1999). Reconsidering “between-group psychotherapy outcome research and basic science: Applications to child and adolescent psychotherapy outcome research. *Journal of Clinical Psychology*, **55**, 181-190.
- Edelson, M. (1994). Can psychotherapy research answer this psychotherapist’s questions? In P.F. Talley, H.H. Strupp, and S.F. Butler (Eds.), *Psychotherapy research and practice: Bridging the gap* (pp.60-87). New York: Basic Books.
- Eifert, G.H., Schulte, D., Zvolensky, M.J., Lejuez, C.W., and Lau, A.W. (1997). Manualized behaviour therapy: Merits and challenges. *Behavior Therapy*, **28**, 499-509.
- Fonagy, P., and Target, M. (1996). Should we allow psychotherapy research to determine clinical practice? *Clinical Psychology: Science and Practice*, **3**, 245-250.
- Frank, E., Kupfer, D.J., Perel, J.M., Cornes, C., Jarrett, D.B., Mallinger, A.G., Thase, M.E., McEachran, A.B., and Grochocinski, V.J. (1990). Three- year outcomes for maintenance therapies in recurrent depression. *Archives of General Psychiatry*, **47**, 1093-1099.
- Garfield, S.L. (1996). Some problems associated with “validated” forms of psychotherapy. *Clinical Psychology: Science and Practice*, **3**, 218-229.

- Garfield, S.L. (1998). Some comments on empirically supported treatments. *Journal of Consulting and Clinical Psychology*, **66**, 121-125.
- Goldfried, M.R. and Wolfe, B.E. (1998). Toward a more clinically valid approach to therapy research. *Journal of Consulting and Clinical Psychology*, **66**, 143-150.
- Heppner, P.P., Kivlighan, D.M., and Wampold, B.E. (1992). *Research design in counselling*. California: Brooks/Cole Publishing.
- Howard, K.I., Krause, M.S., and Lyons, J.S. (1993). When clinical trials fail: A guide to disaggregation. In L.S. Onken, J.D. Blaine, and J.J. Boren (Eds.), *Behavioural treatments for drug abuse and dependence* (NIDA Research Monograph No. 137, pp. 291-302). Washington, DC: National Institute for Drug Abuse.
- Howard, K.I., Moras, K., Brill, P.L., Martinovich, Z., and Lutz, W. (1996). Evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. *American Psychologist*, **51**, (10) 1059-1064.
- Jacobson, N.S. (Ed.). (1988). Defining clinically significant change. *Behavioral Assessment*, **10** (2).
- Jacobson, N.S., and Christensen, A. (1996). Studying the effectiveness of Psychotherapy. How well can clinical trials do the job? *American Psychologist*, **51**, (10) 1031-1039.
- Jacobson, N.S., Schmalings, K.B., Holtzworth-Munroe, A., Karr, J.L., Wood, L.R., and Follette, V.M. (1989). Research-structured vs. clinically flexible versions of social learning based marital therapy. *Behaviour Research and Therapy*, **27**, 173-180.
- Kazdin, A.E., and Weisz, J.R. (1998). Identifying and developing empirically supported child and adolescent treatments. *Journal of Consulting and Clinical Psychology*, **66**, 19-36.
- Kendall, P.C. (1998). Empirically supported psychological therapies. *Journal of Consulting and Clinical Psychology*, **66**, (1) 3-6.

- Klein, D.F., and Smith, L.B. (1999). Organisational requirements for effective clinical effectiveness studies. *Prevention and Treatment*, **2**, Article 0002a. posted March 21 1999 Retrieved August 21, 2000, From the World Wide Web:  
<http://journals.apa.org/prevention/volume2/pre0020002a.html>
- Krause, M.S., and Howard, K.I. (1999). Between group psychotherapy outcome research and basic science: Revisited. *Journal of Clinical Psychology*, **55**, 159-169.
- Lambert, M.J. (1992). Psychotherapy outcome research: Implications for integrative and eclectic therapies. In J.C. Norcross and M.R. Goldfried (Eds.), *Handbook of psychotherapy integration* (pp.94-129). New York: Basic Books.
- Lambert, M.J., and Bergin, A.E. (1994). The effectiveness of psychotherapy. In S.L. Garfield and A.E. Bergin (Eds.), *Handbook of psychotherapy and behaviour change* (4 th ed., pp.143-189). New York: Wiley.
- Linehan, M.M., Armstrong, H.E., Suarez, A., Allmon, D., and Heard, H.L. (1991). Cognitive-behavioural treatment of chronically parasuicidal borderline patients. *Archives of General Psychiatry*, **48**, 1060-1064.
- Luborsky, L., McLellan, A.T., Diguier, L., Woody, G., and Seligman, D.A. (1997). The psychotherapist matters: Comparisons of outcomes across twenty-two therapists and seven patient samples. *Clinical Psychology: Science and Practice*, **4**, 53-63.
- Luborsky, L., McLellan, A.T., Woody, G.E., O'Brien, C.P., and Auerbach, A. (1985). Therapist success and its determinants. *Archives of General Psychiatry*, **42**, 602-611.
- Nathan, P.E., Stuart, S.P., and Dolan, S.L. (2000). Research on psychotherapy efficacy and effectiveness: Between Scylla and Charybdis? *Psychological Bulletin*, **126**, (6), 964-981.

- O'Donohue, W., and Szymanski, J. (1994). How to win friends and influence clients: Popular but problematic ideas that impair treatment decisions. *The Behavior Therapist*, **17**, 29-33.
- Persons, J.B. (1991). Psychotherapy outcome studies do not accurately represent current models of psychotherapy: A proposed remedy. *American Psychologist*, **46**, 99-106.
- Persons, J.B., and Silberschatz, G. (1998). Are results of randomised controlled trials useful to psychotherapists? *Journal of Consulting and Clinical Psychology*, **66**, (1) 126-135.
- Rice, L.N., and Greenberg, L.S. (1984). *Patterns of change*. New York: Guildford Press.
- Schulte, D., Kunzel, R., Pepping, G., and Schulte-Bahrenberg, T. (1992). Tailor-made versus standardized therapy of phobic patients. *Advances in Behaviour Research and Therapy*, **14**, 67-92.
- Seligman, M.E.P. (1995). The effectiveness of psychotherapy: The *Consumer Reports* study. *American Psychologist*, **50**, 965-974.
- Shadish, W.R., and Ragsdale, K. (1996). Random versus non-random assignment in controlled experiments: Do you get the same answer? *Journal of Consulting and Clinical Psychology*, **64**, 1290-1305.
- Shapiro, D.A., Firth-Cozens, J., and Stiles, W.B. (1989). Therapists' differential effectiveness: A Sheffield Psychotherapy Project addendum. *British Journal of Psychiatry*, **154**, 383-385.
- Silver, L.B., and Silver, B.J. (1983). Clinical practice of child psychiatry: A survey. *Journal of the American Academy of Child Psychiatry*, **22**, 573-579.
- Silverman, W.H. (1996). Cookbooks, manuals and paint-by-numbers: Psychotherapy in the 90's. *Psychotherapy*, **33**, 207-215.

- Sobell, L.C. (1996). Bridging the gap between scientists and practitioners: The challenge before us. *Behavior Therapy*, **27**, 297-320.
- Starr, B.J. and Katkin, E.S. (1969). The clinician as an aberrant actuary: illusory correlation and the incomplete sentences blank. *Journal of Abnormal Psychology*, **74**, 670-675.
- Stiles, W.B., Shapiro, D.A., and Elliott, R. (1986). Are all psychotherapies equivalent? *American Psychologist*, **50**, 995-1002.
- Task Force on Promotion and Dissemination of Psychological Procedures. (1995). Training in and dissemination of empirically- validated psychological treatments: Report and recommendations. *Clinical Psychologist*, **48**, 3- 23.
- Tuma, J.M., and Pratt, J.M. (1982). Clinical child psychology practice and training: A survey. *Journal of Clinical Child Psychology*, **11**, 27-34.
- Weisz, J.R., Donnerberg, G.R., Han, S.S., and Weiss, B. (1995). Bridging the gap between laboratory and clinic in child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology*, **63**, 688-701.
- Weston, D., and Morrison, K. (2001). A multidimensional meta-analysis of treatments for depression, panic, and generalized anxiety disorder: An empirical examination of the status of empirically supported therapies. *Journal of Consulting and Clinical Psychology*, **69**, 875-899.
- Wilson, G.T. (1996). Manual-based treatments: The clinical application of research findings. *Behavior Research Therapy*, **34**, 295-315.
- World Health Organization (1992.) *The ICD-10 Classification of Mental and Behavioural Disorders*. Geneva: WHO.



**Table 1. Comparison of efficacy and effectiveness research**

Efficacy Research	Effectiveness Research
<ul style="list-style-type: none"> <li>▪ Aims to determine whether an intervention has a significant therapeutic effect in an experimentally controlled environment</li> </ul>	<ul style="list-style-type: none"> <li>▪ Aims to determine the generalisability, feasibility and cost effectiveness of an intervention in the environment where the treatment is usually delivered</li> </ul>
<ul style="list-style-type: none"> <li>▪ Studies are carried out in a university setting</li> </ul>	<ul style="list-style-type: none"> <li>▪ Studies are carried out in a health services setting</li> </ul>
<ul style="list-style-type: none"> <li>▪ Cases are randomly assigned to treatment and control conditions within the context of a randomized controlled trial (RCT)</li> </ul>	<ul style="list-style-type: none"> <li>▪ Random assignment of cases to treatment and control conditions is rarely used, but post-therapy comparisons on baseline variables is used to check the similarity of groups in treatment and control conditions</li> </ul>
<ul style="list-style-type: none"> <li>▪ Diagnostically and demographically homogeneous groups with strict inclusion and exclusion criteria are used</li> </ul>	<ul style="list-style-type: none"> <li>▪ Heterogeneous groups are used</li> </ul>
<ul style="list-style-type: none"> <li>▪ Outcome is assessed under conditions of high internal validity</li> </ul>	<ul style="list-style-type: none"> <li>▪ Outcome is assessed under conditions of high external validity</li> </ul>
<ul style="list-style-type: none"> <li>▪ Clients often volunteer or are recruited through the media into RCTs, but thereafter cannot choose their therapist or treatment</li> </ul>	<ul style="list-style-type: none"> <li>▪ Clients are generally referred by family doctors and routine referral agents into effectiveness studies and can in some settings 'shop-around' for a preferred therapist or treatment programme</li> </ul>
<ul style="list-style-type: none"> <li>▪ Clients typically have acute problems</li> </ul>	<ul style="list-style-type: none"> <li>▪ Clients typically have chronic problem</li> </ul>
<ul style="list-style-type: none"> <li>▪ Clients typically have focal problems</li> </ul>	<ul style="list-style-type: none"> <li>▪ Clients typically have complex sets of multiple problems</li> </ul>
<ul style="list-style-type: none"> <li>▪ Clients typically have a single DSM IV axis I diagnosis</li> </ul>	<ul style="list-style-type: none"> <li>▪ Clients generally have multiple co-morbid DSM IV axis I diagnoses</li> </ul>
<ul style="list-style-type: none"> <li>▪ Clients with DSM IV axis II personality disorders are usually excluded from RCTs of treatments for axis I disorders</li> </ul>	<ul style="list-style-type: none"> <li>▪ Clients generally have DSM IV axis II personality disorders</li> </ul>
<ul style="list-style-type: none"> <li>▪ Treatment is often brief and of fixed duration</li> </ul>	<ul style="list-style-type: none"> <li>▪ Treatment duration is flexible</li> </ul>
<ul style="list-style-type: none"> <li>▪ Treatment is often pre-planned and</li> </ul>	<ul style="list-style-type: none"> <li>▪ Treatment is rarely planned in detail and</li> </ul>

- involves strict adherence to treatment manuals
  - Treatment is usually based on a single theory
  - Participants are often treated by either inexperienced trainee therapists or expert clinicians who specialize in the treatment being studied
  - Clients usually receive extensive pre-therapy preparation
  - treatment manuals are rarely used
  - Treatment is often informed by a number of theoretical orientations
  - Clients are usually treated by experienced practising clinicians who work with a wide range of clients and use a rang of interventions
  - Clients rarely receive pre-therapy preparation
-