Athens 1997

Horst Kächele (Ulm)

Tel.: (0731) 502--5661 Fax: (0731) 502-5
e-mail: kaechele@sip.medizin.uni-ulm.de

Efficacy and Effectiveness: Two Worlds of Research On Outcome in Psychotherapy

Horst Kächele

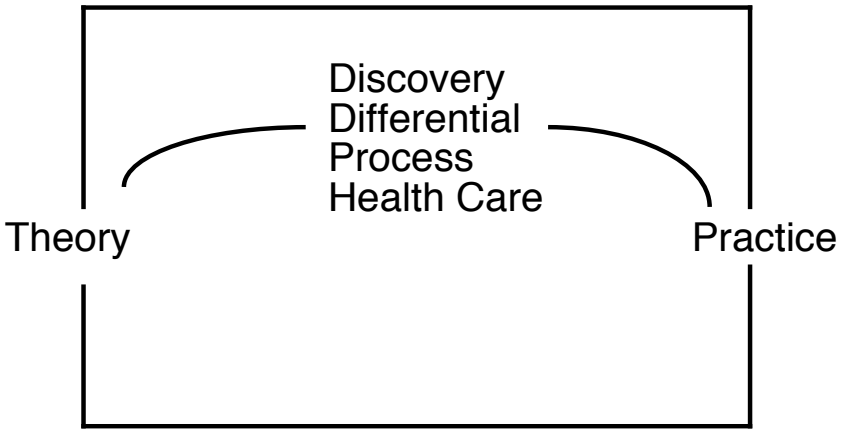
Athens 1997

abstract

Psychotherapy research aims at the scientific evaluation of existing practice and at the discovery of new fields of application. The early phases of psychotherapy research were marked by scientific justification and societal legitimation. These questions changed with extension of possible indications, with growing differentiation of treatment procedures and with the progressive implementation of psychotherapy within the health system. The early approach "does psychotherapy work at all" has been replaced by the questions "what works for whom" and "how does psychotherapy work".
The outcome question is raised in two modalities that represent two phases of treatment research. Describing this phase model of psychotherapy research helps to better evaluate the validity issue present on going discussions.

Korrespondenz-Adresse:
Prof. Dr. med. Horst Kächele
Abteilung für Psychotherapie und Psychosomatische Medizin
Universität Ulm
Am Hochsraess 8, 89081 Ulm

Psychotherapy research at any given time has multiple functions and tasks to perform; it aims at the scientific evaluation of existing practice and at the discovery of new fields of application
figure 1



The early period of psychotherapy research was marked by scientific justification and societal legitimation. Its questions changed with extension of possible indications, with growing differentiation of treatment procedures and with the progressive implementation of psychotherapy within the health system. The early approach "does psychotherapy work at all" has been replaced by the question "what works for whom" (Roth & Fonagy 1997) and "how does what kind of psychotherapy work".
Futhermore it has become obvious that the findings from systematic outcome research are directed at different audiences - e.g. at psychotherapists who conduct the treatment in question as well as to health professionals from related, often competitive disciplines. Research findings are addressed at those who benefit directly (e.g. patients or their relatives) as well as at those who fund the costs (e.g. insurance companies) or are responsible for adequate health policies (e.g. politicians, unions). The diverse groups may have totally different expectations (Strupp & Hadley 1977). Therefore outcome research has to provide a variety of information to satisfy the needs of the different interest groups.

The findings of many investigations consistently support the conclusion that overall psychotherapeutic treatments rightly have become an integral part of the medical system. This success opens a view for raising new issues. We have to attend more care to the validity aspects of experimental process-outcome research demonstrating efficacy and of field studies on the treatment as applied in usual routine demonstrating effectiveness. It is on this distinction that this contribution hopes to make some contribution..
To prepare for discussion I shall use a phase model of types of psychotherapy research that no necessary temporal order:

Phase I: descriptive study *(process research)*
Phase II: experimental analogue study *(process-outcome research)*
Phase III: clinical trial *(efficacy research)*
Phase IV: field study *(effectiveness research)*
Phase V: patient oriented treatment study *(efficiency research)*

Phase I descriptive study on the therapeutic process

We are now faced with the seeming paradox that, in spite of the overwhelming and certainly impressive evidence for the most frequently practiced forms of therapy, we are faced with many critical voices complaining that the many outcome studies have not contributed to a better understanding of therapeutic mechanisms (Grawe 1988). It is within this context that the very material of the therapeutic process is rediscovered and the detailed analysis of single cases once more achieves a prominent status (Dahl et al. 1988; Greenberg 1991; Greenberg & Pinsof 1986). This move entails increasingly focusing on microprocesses of the treatments requiring new assessment procedures and a better articulation of moment to moment events that may significantly influence treatment outcome (Kächele 1992).

Phase II experimental analogue study

The beginning of the experimental analogue derived from the classic studies of Greenspoon (1955) in the area of operant conditioning. Hans Strupp presumably was the first psychodynamic treatment researcher that used an

experienmtal analogue approach in his satudies on the therapist´s technique (1960). The experimental mind of behaviorists have cared for a immense output on these kind of studies (Kiesler 1971); there are really very few psychoanalytic minded researchers following the model set by Strupp. The studies on the use of free association are a laudable exemption (see Heckmann et al. 1987).

Phase III: clinical trial *(efficacy research)*

This type of research has received utmost attention as practically thousands of so called Randomized Controlled Trial (RCT)´s have been performed. Grawe et al. (1994) identified in 1985 more than 3.5 thousand of such studies. Using a number of sensible criteria they still ended up with 897 studies on adult real patients. The clinical trial consists in a comparison of two or more groups of patients in psychotherapy or in control conditions that were allotted by randomization to one of the conditions. Measurement of outcome is done using multiple perspectives and multiple measurements.

Ken Howard has recently given a succinct appreciation of the core features of RANDOMIZED CLINICAL TRIALS that emphasize the aspect of internal validity:

Methodology

1. Random assignment in a particular study never makes groups exactly comparable -- does give each group an equal chance at being best.
2. RCT are treatment focused they seeks main effects that are meant to be applicable to all or most patients; they can have mean differences favoring one treatment when the other treatment is actually better in terms of percent improved.
3. RCT show outcome overlap when effect size < 3.0; e.g. some patients in poorer group do better than some patients in the better treatment
4. Within cell variance > measurement error = real individual differences among the outcomes of patients who receive the same treatment.

Execution

1. Attrition always compromises random assignment -- introduces self-selection (i.e., patient variables).
2. Double blinds are routinely broken through monitoring of side-effects.
3. End-point analyses entail the assumption of counter-to fact conditions (with little implication for practice).

Generalizability

1. Inclusion criteria are not easily identified by clinician (usually at least includes a diagnostic assessment that clinicians are not trained to do; willingness to sign an informed consent)
2. Exclusion criteria also not readily assessable (seldom told how these are assessed or why they have been chosen)
3. Patients do not get to select treatments (as in practice) -- agree to random assignment among specific treatments
4. Patients often have to be recruited through advertisements (and thus are not representative of any known population)
5. Treatments are manualized and closely monitored; therapists are trained to specific criteria and performance is regularly reassessed

Independent Variables

Some independent variables are difficult to assign --e.g., dosage
Some independent variables are unethical to assign -- e.g., childhood abuse
Some independent variables are impossible to assign -- e.g., chronicity, severity.

Howrad´s SOLUTION

Gather assessments of lots of plausible potential confounds (setting, treatment, therapist, and patient variables).
Stay clearly focused on the research question (in clinical research this question should be patient focused -- what is the best treatment for this (type of) patient?).

NB: An empirically "validated Treatment" is one that has been subjected to at least two clinical trials, since clinical trials rarely fail to get significant results or come to the conclusion that the targeted treatment "doesn't work."

Probably the most well know study of this kind is the NIMH Treatment of Depression Collaborative Research Program (TDCRP) that has been described in great detail by Elkin 1994 in the latest edition of the Handbook of Psychotherapy and behavior Change that set out to evaluate the valkue of two psychotheraypy conditions (Cognitive-Bahavioral Treatment (CBT) and Interpersonal Therapy (IPT) for the treatment of depression compared to a standard treatment with imipramine. Meanwhile this study has aroused many controversial discussions in the field on the validity of the findings (see Kächele 1996). I shall take up some remarks by Jacobson, a well know psychotherapy researcher,who has sharpened the debate regrading the usefulness of RCT for making decisions on clinical practice (Jacobsen 1995 in Kächele 1996):

> First, it seems clear to me that clinical trials have generally provided optimal tests of the efficacy of CBT and other psychotherapies, and undoubtedly overestimate the value of these treatments as practiced by your typical clinician who is not rigorously trained, monitored, or supervised during the course of a trial.
> Our research and that of others raises questions about the exportability of CBT as a treatment for depression into naturalistic settings, since competence seems to drift downward even among highly experienced therapists who were well-trained to a certain level of competence, unless supervision/calibration remains quite intensive throughout the trial. Any treatment which requires such intensive supervision, even with highly experienced and well-trained therapists, may not generalize well to typical practitioner settings where there is little training and certainly no supervision.
>
> Second, the results in clinical trials are further skewed by subject selection procedures designed to homogenize the sample and detract from its representativeness to clinical practice. As Ellen Frank and Tracie Shea pointed out at the meeting, this subject selection practice cuts both ways. On the one hand, the single most often used basis for exclusion in depression trials is that subjects are not sufficiently depressed: so subjects treated in clinical trials are in all likelihood MORE depressed than those typically seeking therapy from practitioners. On the other hand, the exclusion of bipolar I and bipolar II patients, patients with dual

> diagnoses, including current episodes of substance abuse, panic disorder, eating disorders, and the like limit generalizability. But these two opposing processes (selecting for real major depressive disorder while at the same time screening out diagnostic complexity) result in a sample of patients that in some ways are easier to treat, and in other ways are harder to treat.
>
> Third, even taking into account the highly select sample of patients selected for these trials, the efficacy of CBT, as well as other available pharmacological and psychosocial treatments, are exceedingly modest from the standpoint of clinical significance. Whether psychosocial treatments, including CBT, exceed rigorously constructed placebos is still very much open to question.

These comments were reiterated and reinforced by another prominent member of the US-scientific comunity, by Martin Seligman. Reporting on the findings of a consumer satisfaction study - the Consumer Reports - he concluded:

> The efficacy study is the wrong method for empirically validating psychotherapy as it is actually done, because it omits too many crucial elements of what is done in the field (Seligman 1995, S.966).

To add to these critics Kächele et al. (1997) have compared the amount of treatment provided within such RCT by analyzing the studies reported on by Grawe et al (1994):

<u>Cognitive-Behavioral Therapies</u>
429 studies mean 11,2 sessions
434 studies mean 7, 9 weeks

<u>Humanistic Therapies</u>
70 studies mean 16,1 sessions
76 studies mean 11, 6 weeks

<u>Psychodynamic Therapies</u>
82 studies mean 27,6 sessions
80 studies mean 30,7 weeks

These findings again raise the question of the external validity of the RCT studies assuming that psychotherapy as practised in routine conditions deviate grossly from these findings.

Phase IV field study *(effectiveness research)*
Checking for facts describing what is routine practice in providing treatment in terms of time (number of sessions or month in treatment) Kächele et al. (1997) could identify three samples.

Sample A: 1800 psychodynamic oriented treatments delivered at the outpatient service of the department of psychotherapy at the Ulm University.
Sample B: 302 client-centered psychotherapies across Germany
Sample C: 496 behavioral therapies from the outpatient service of the department of clinical psychology at the Ruhr University Bochum

	mean	median	variation	
Sample A		22	6 - 1200	
Sample B	69,2, s 39,1	61	8 - 275	
Sample C	30,7, s 13,12	24	4 - 95	

The range of these treatments is considerable; it is most among psychoanalytic therapies. The graphical representations of these findings illustrate that the length of time in treatment replicates the findings of the dosage-effect relationship of Howard et al. (1987). Many treatments are terminated rather soon, a smaller number takes considerable time.
Investigations of these phenomena have far-reaching clinical implications because they correct the clinician's illusion (Vessey et al 1993) that he or she is treating an representative sample of patients. Epidemiological studies of the incidence and prevalence of psychosomatic or neurotic illnesses, of bodily dysfunctions or emotional disturbances, give an estimate of the need for services; the investigation of therapeutic practices yields an estimate of available resources to meet those needs; and, studies of the patterns of service utilization identify the constituencies served by the delivery system (Howard et al.1992).

Detailed field research describing what happens in treatment in many naturalistic dimensions on multiple cases combined with sophisticated outcome measurement supports finding specific treatments for patient groups with particular disorders. For many years substantial naturalistic research on specific groups of patients remained rare as the institutions had to take care of a broad spectrum of patients, which only sometimes allowed for the formation of groups homogeneous with respect to one disease. It took some time in psychotherapy to realize that this situation was not principally different from somatic medicine, where the development of multi-center studies led to progress because the desired homogeneity allowed better conclusions.

The multi-center study on the psychodynamic treatment of eating disorders that has been initiated by the Center for Psychotherapy Research in Stuttgart in 1992 includes a wide range of inpatient modalities all over Germany; it also has established the logistics of implementing the study in a considerable number of European countries. The German multi-center study is heavily supported by non-university institutions and offers also clinical exchange programs; so we feel that this type of research commitment may well turn out to be a prototype for the new look in psychotherapy research (Kächele et al. 1992).

Phase V patient focused treatment research

As the question of differential indication is what interests the clinician (Kächele & Kordy 1997) this approach still does not answer the crucial question: does this particular treatment work for this patient ?
The main limitation of clinical trials is that they can only confirm the efficacy of treatments that have already been developed. They are not generative. Also, they do not speak to individual clients, as a more qualitative or single subject design would. At this point in the discussion of clinical research we have to ask: how does one study the impact of treatment on an individual client? Single subject designs? Hermaneutic texts? Discourse analysis?

References

Dahl, H., Kächele, H., & Thomä, H. (Eds) Psychoanalytic Process Research Strategies. Berlin Heidelberg New York London Paris Tokyo: Springer.1988

Grawe, K.: Zurück zur psychotherapeutischen Einzelfallforschung. Z Klin Psychol, 17 (1988b) 4

Greenberg, L.: Research on the process of change. Psychotherapy Research 1 (1991) 3

Greenberg L, Pinsof W (Eds) The Psychotherepeutic Process: A Research Handbook . Guilford Press, New York 1986.

Greenspoon J (1955) The reinforcing aspect of two spoken sounds on the frequency of two responses. American Journal of Psychology 68:409-416

Heckmann H, Hölzer M, Kächele H, Robben H (1987) Resistance and transference as two main constituents in an "experimental analogue" of free association. In: Huber W (Hrsg) Progress in Psychotherapy Research. Presses Universitaire de Louvain, Louvain-la -Neuve, S 582-593

Howard KI., Kopta, S. M., Krause, M. S., & Orlinsky, D. E.: The dose-effect relationship in psychotherapy. American Psychologist, 41 (1986) 159

Howard KI., Lueger RJ, Schank D.: The psychotherapy service delivery system. Psychotherapy Research, 2 (1992)

Kächele H (1992) Narration and observation in psychotherapy research. Reporting on a 20 year long journey from qualitative case reports to quantitative studies on the psychoanalytic process. *Psychotherapy Research 2:1-15*

Kächele H (Hrsg) (1996) Letters from the SSCPnet. Ulmer Textbank, Ulm

Kächele, H. u. Mitarbeiter (1992). Planungsforum "Psychodynamische Therapie von Eßstörungen". PPmP DiskJournal. 42(3:1)

Kächele H, Eckert J, Hillecke W (1997) Zeit für Psychotherapie. in Vorbereitung

Kächele, H., & Kordy, H. (1997). Outcome Research. In T. v. Uexküll (Ed.), Psychosomatic Medicine (pp. in prep). New York: Urban & Schwarzenberg.

Kiesler DJ (1971) Experimental designs in psychotherapy research. In: Bergin AE, Garfield SL (Hrsg) Handbook of Psychotherapy and Behavior Change. Wiley & Sons, 36-74

Newmann FL, Howard KI: Therapeutic effort, treatment outcome, and national health policy. Amer Psychol 41 (1986) 181

Roth A, Fonagy P (1996) what works for whom ? - A critical review of psychotherapy research. Guilford Press, New York - London

Sandell, R Blomberg J, Lazar A, Schubert J,Carlsson J & Broberg J (1997) Self-reported outcomes and their clinical significance after psychotherapy and psychoanalysis in ther Stockholm outcome and psychotherapy and psychoanalysis procet (STOPP). Geilo, 28th Annual Meeting of the Society for Psychotherapy Research

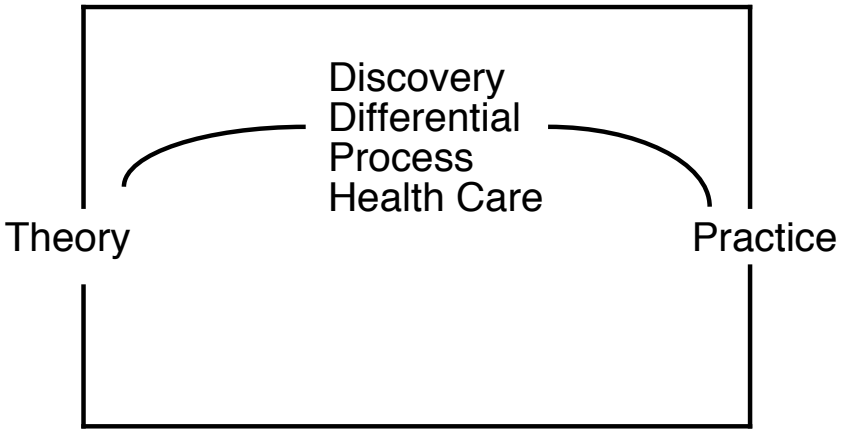
Seligman M (1994) What you can change and what you can´t. The complete guide to successful self-improvement. A Knopf, New York

Seligman MEP (1995) The effectiveness of psychotherapy. The *Consumer Reports* Study. Am Psychologist 50:965-974

Strupp HH (1960) Psychotherapists in Action: Explorations of the Therapist´s Contribution to the Treatment Process. Grune & Stratton, New York

Strupp HH, Hadley SW: A tripartite model of mental health and therapeutic outcome. Amer. Psychol. 32 (1977) 187

Vessey JT, Howard KI, Lueger RJ:. The clinician's illusion and the psychotherapy practice: an application of stochastic modelling. J Clin Con Psychol. (1992) (in press)


Discovery
Differential
Process
Health Care
Theory
Practice

Phases of Research

Phase I: descriptive study
(process research)

Phase II: experimental analogue study
(process-outcome research)

Phase III: clinical trial
(efficacy research)

Phase IV: field study
(effectiveness research)

Phase V: patient oriented treatment study *(efficiency research)*

> The efficacy study is the wrong method for empirically validating psychotherapy as it is actually done, because it omits too many crucial elements of what is done in the field
>
> (Seligman 1995, S.966).