

Department Psychologie

Klinische Psychologie und Psychotherapie
Abt. Psychoanalyse und psychodynamische Forschung
Prof. Dr. Wolfgang Mertens



LUDWIG-MAXIMILIANS-UNIVERSITÄT
Department Psychologie
Prof. Dr. Wolfgang Mertens
Leopoldstr. 13, 80802 München

Einige Überlegungen zur randomisierten klinischen Studie („RCT“)

Zur Evidenzbasierung und zur Postulierung der Evidenzgrade:

In der Forschungsmethodologie der Psychotherapieforschung hat in den letzten Jahren eine zunehmende Angleichung an verhaltensmedizinische und pharmakologische Standards der *evidence based medicine* (EBM) stattgefunden. Dies wird von vielen Funktionären im Gesundheitswesen begrüßt, da hierdurch endlich die gewünschte Vereinheitlichung und damit auch Vergleichbarkeit erzielt werden. Warum sollte in forschungsmethodologischer Hinsicht den verschiedenen Psychotherapieverfahren eine Sonderrolle eingeräumt werden, zumal die Verhaltenstherapeuten gegen diese Angleichung ohnehin keine Bedenken äußern? Der Wissenschaftliche Beirat hat sich ebenfalls dieser Tendenz angeschlossen.

In Analogie zur laboratoriumsexperimentellen Grundlagenforschung wird deshalb nun auch in der Psychotherapieforschung von interner und externer Validität gesprochen. Hohe interne Validität liegt immer dann vor, wenn das Ergebniskriterium (Erfolg der therapeutischen Behandlung) eindeutig auf die Treatmentvariable (therapeutische Behandlungsbedingung) zurückgeführt werden kann. Dies geschieht am besten durch die Anwendung von kontrollierten, randomisierten Studien.

Allerdings ergibt sich hierbei ein Dilemma: Je höher die interne Validität ist, die ein Forscher durch Ausschaltung bzw. Kontrolle aller Störvariablen erhalten kann, desto geringer fällt die externe Validität aus. D.h. die Ergebnisse können nur mit Schwierigkeiten auf Praxisbedingungen übertragen werden. Zwar müssen sich interne und externe Validität nicht völlig ausschließen, aber in der Regel sinkt doch die Generalisierbarkeit mit der Strenge der Anforderungen an die Sicherung der internen Validität.

Die Frage ist aber, ob die Hierarchisierung von interner und externer Validität, wie sie von der Evidenzbasierung gefordert wird, überhaupt sinnvoll ist: Lassen sich experimentelle Gütekriterien, die aus der Grundlagenforschung stammen, überhaupt auf die Praxis übertragen? Bei einer therapeutischen Behandlung handelt es sich a priori um ein komplexes Geschehen mit vielen Einflußvariablen, die noch dazu miteinander interagieren. Von daher wäre die bisherige Postulierung von Evidenzklassen (Studien mit Sicherung der internen Validität sind solchen, bei denen der externen Validität das Hauptaugenmerk gilt, bei der wissenschaftlichen Anerkennung von Psychotherapieverfahren vorzuziehen) von vornherein der Thematik der Psychotherapieforschung nicht angemessen.

Eng damit zusammenhängend, aber im Grunde bereits von nachgeordneter Bedeutung ist dann die Frage nach der Sinnhaftigkeit der Randomisierungsstrategie.

Durch Randomisierung soll v.a. der Einfluss patientenspezifischer Variablen in Psychotherapiewirkungs-Studien weitgehend *neutralisiert* werden. Denn der Zufall soll darüber entscheiden, welcher Patient zur Behandlungsgruppe oder zur Kontrollgruppe (Nichtbehandlung oder andere Behandlung) zugeteilt wird. Die Annahme lautet dabei, dass sich dadurch bestimmte Patientenmerkmale in beiden Gruppen gleich verteilt finden, sofern eine ausreichende Gruppengröße vorliegt.

Durch die Zufallszuweisung soll verhindert werden, dass sich in der Behandlungsgruppe vermehrt Patienten finden, die z.B. einen stärkeren Leidensdruck, ein besseres Reflexionsvermögen oder ein höheres Strukturniveau aufweisen. Es wäre dann nämlich nicht mehr entscheidbar, ob bzw. in welchem Ausmaß der Behandlungseffekt (die sog. abhängige Variable in Analogie zum Laboratoriumsexperiment) auf die Behandlungsvariable („experimentelles Treatment“ als unabhängige Variable) zurückzuführen ist und in welchem Ausmaß auf sog. Störvariablen oder Moderatorvariablen (die die zu untersuchende Kausalität zwischen abhängiger und unabhängiger Variable stören bzw. moderieren).

Da man realistischerweise immer von diesen Störvariablen ausgehen muss, kann die Interpretation, dass die zu prüfende therapeutische Behandlung zu einem besseren Ergebnis als eine andere Behandlung oder eine Nichtbehandlung geführt habe, sehr anfechtbar sein, solange nicht diesem Einwand mit Hilfe von Zufallszuweisung der Patienten zu einer Behandlungs- und einer Kontrollgruppe der Boden entzogen worden ist. Denn es sind ja immer auch sehr subtil wirkende moderierende Bedingungen denkbar: So könnte z.B. argumentiert werden, dass Patienten mit einem stärkeren Leidensdruck oder einem höherem Strukturniveau sich als Freiwillige für ein bestimmtes Untersuchungsdesign bzw. für ein bestimmtes Therapieverfahren zur Verfügung gestellt haben. Aussagen über die Wirksamkeit¹⁾ könnten somit jederzeit angezweifelt werden, was den mühseligen Aufwand einer Psychotherapiestudie mit einem einzigen Argument zunichte machen würde. Auf den ersten Blick führt deshalb kein Weg an einem randomisierten Kontroll-Gruppen-Design vorbei.

Zusätzliche Charakteristika der Randomisierten Kontrollgruppen-Strategie, wie einfache Verblindung (der Patient weiß nicht, ob ihm das wirkliche Medikament oder ein Placebo gegeben wird) und doppelte Verblindung (zusätzlich weiß auch der behandelnde Arzt bzw. der Versuchsleiter nicht, welches die Experimental- oder die Kontrollgruppe ist), wären nach Auffassung vieler empirischer Forscher nicht nur in der Medikamentenforschung, sondern auch in der Psychotherapieforschung wünschenswert, um Erwartungs- und Allegianzeffekte auszuschließen (dies war z.B. eines der Kritikpunkte von Grawe an der PAL-Studie von Rudolf), welche die interne Validität gefährden könnten. Blindstudien dieser Art sind aber aus leicht einsehbaren Gründen nicht durchzuführen.

Lediglich bei der Auswertung von Prozessstudien können die Ergebnisse einzelner Messzeitpunkte anhand von bestimmten Prozessmethoden (wie HUSS, ZBKT, SASB, RFS) für den Auswerter verblindet werden, sofern die Auswertung erst nach Abschluss aller Messzeitpunkte erfolgt. Der Auswerter weiß dann nicht, welchen Messzeitpunkten z.B. die einzelnen HUSS-Interviews zuzuordnen sind, und bestimmte Erwartungseffekte werden verhindert).

So bleibt für die Psychotherapieforschung sinnvollerweise nur die zufällige Zuteilung zur Experimental- und Kontrollgruppe übrig. Diese erscheint auf den ersten Blick – wenngleich auch aufwändig zu realisieren – als durchaus sinnvoll. Wie anders sollte sonst eine gültige Schlussfolgerung über den Effekt der Behandlung zustande kommen können? In dem Frankfurter Depressionsprojekt zum Vergleich der Wirksamkeit und Effizienz von KVT und

analytischer Psychotherapie von Leuzinger-Bohleber et al. (2005) kommt deshalb auch ein randomisiertes Kontrollgruppen-Design zur Anwendung.

Abgeschwächte Formen der Kontrolle finden sich im sog. *Quasi-Experiment*. Hierbei werden zwei (oder auch mehr Gruppen) miteinander verglichen, die nicht durch eine Zufallszuordnung zustandegekommen sind. Dennoch wird versucht, Stör- oder Moderatorvariablen konstant zu halten. Dies kann durch *Parallelisierung* oder durch *Auspartialisierung* geschehen. Bei der Parallelisierung werden im Vorhinein Behandlungs- und Kontrollgruppen hinsichtlich bestimmter Kriterien ähnlich zusammengestellt, z.B. hinsichtlich Alter, Geschlecht, bestimmter psychometrischer Kennwerte, die sich aufgrund eines Screenings ergeben haben. Bei der Auspartialisierung werden nachträglich mittels Kovarianz- oder Regressionsanalysen Gruppenunterschiede herausgerechnet. Wenn sich z.B. herausstellt, dass das Ausmaß der anfänglichen Symptombelastung Einfluss auf das Gesamtergebnis haben könnte, werden extreme Werte in beiden Gruppen auspartialisiert. Diese statistische Konstanthaltung ermöglicht aber auf jeden Fall nur eine schwächere Prüfung der Kausalhypothese, dass eine bestimmte Behandlungsform einer anderen Behandlungsform oder „treatment as usual“ überlegen ist.

Die Kritik an diesen beiden Methoden der Parallelisierung und der Auspartialisierung betrifft v.a. den Umstand, dass sie entweder im Vorhinein oder im Nachhinein immer nur eine sehr begrenzte Auswahl von Variablen konstant halten können; alle anderen möglicherweise bedeutsamen Faktoren werden dadurch nicht erfasst und könnten in der einen oder anderen Gruppe gehäuft auftreten, was einen sicheren Schluss erschwert.

Wurden bislang nur die Patientenmerkmale als mögliche moderierende Variablen betrachtet, so müssen nun auch noch die *Therapeutenmerkmale* berücksichtigt werden. Genau genommen müssen nämlich auch die behandelnden Therapeuten nach einer Zufallszuordnung auf die Patienten (bei den *Behandlungsbedingungen* ist dies im Fall von Therapievergleichsstudien wegen der unterschiedlichen Therapierichtungen nicht möglich) verteilt werden. Da sich aber in der Psychotherapie – anders als in der Medikamentenforschung - das Problem der Passung als ein wichtiger Wirkfaktor herausgestellt hat, muss diese Möglichkeit entfallen. Im Vergleich zum Laboratoriumsexperiment, in dem die Verabreichung der unabhängigen Variable immer nahezu gleich erfolgt (oder das Problem von vornherein entfällt, weil nur ein einziger Versuchsleiter eingesetzt wird), muss bei Psychotherapiewirkungs-Studien aber davon ausgegangen werden, dass wegen der Unmöglichkeit der Zufallszuordnung der Therapeuten die Unterschiedlichkeit von Persönlichkeit, Erfahrung, Manualtreue bzw. Adhärenz² einen erheblichen und nicht von vornherein per Randomisierung zu kontrollierenden Einfluss ausübt. Immerhin besteht aber die Möglichkeit der nachträglichen Auspartialisierung, wie z.B. in dem Stockholmer Psychoanalyse- und Psychotherapieprojekt (STOPP) von Sandell et al. (1999, 2001), was immerhin eine minimale Konstanthaltung und Vergleichbarkeit der Therapeutenvariable ermöglicht.

Welches sind nun die landläufigen Einwände gegen das RCT-Paradigma? Warum wird dieses seit geraumer Zeit bekämpft? So wird z.B. Buchholz in den Psycho-News-Letters nicht müde, in immer erneuten Ausführungen gegen den Randomisierungswahn herkömmlicher empirischer Psychotherapieforscher zu wettern. Aber die methodologisch versierte Psychoanalytikerin und Psychotherapieforscherin Marianne Leuzinger-Bohleber verzichtete in ihrer Depressionsstudie nicht zur Gänze auf ein RCT-Vorgehen, sondern entschied sich für ein randomisiertes *und* naturalistisches Vorgehen.

Im Folgenden werden in einer Gegenüberstellung die wichtigsten Charakteristika von randomisierten und kontrollierten Studien sowie die Argumente dafür und andererseits die Einwände gegen RCT als alleiniges oder höchstes Effizienzkriterium aufgelistet:

Charakteristika von und Argumente für „randomisierte und kontrollierte Studien“	Einwände gegen RCT als alleiniges oder höchstes Evidenzkriterium
randomisierte Behandlungszuordnung	<ul style="list-style-type: none"> • unter realen Bedingungen wählen Patienten in der Regel ihr Psychotherapieverfahren eigenständig aus. Die damit verküpften Erwartungen sind selbst behandlungsrelevant (Tschuschke, 2005) • die Passung zwischen Therapeut und Patient und zwischen Patient und Methode gilt als eine der wichtigsten Voraussetzungen für die Behandlung selbst (Kriz, 2005; Orlinsky & Howard, 1987) • wenn man die Entscheidungsautonomie des Patienten achtet, kommt die Behandlungszuordnung einer Entmündigung gleich (Legewie, 2000) • Selbst wenn diese sinnvoll wäre, ließe sie sich nur sehr bedingt methodisch tadellos durchführen. Bei der Vielzahl der möglichen und interagierenden Einflussfaktoren erscheint es fragwürdig, ob Randomisierung zu äquivalenten Gruppen führt, wenn die Stichprobengröße nicht ausreichend groß ist (Hsu, 1989)
Einrichtung von Nichtbehandlungs- und Placebogruppen („medizinisches Modell“)	<ul style="list-style-type: none"> • ethisch schwer oder überhaupt nicht vertretbar (Mundt & Backenstraß, 2001) • die pharmakologische Analogie ist mit psychotherapeutischen Belangen nicht vereinbar (Mundt und Backenstraß, 2001; Wampold, 2001); Placebokontrolle ist in diesem Bereich logisch widersinnig, da sie dazu dienen würde, den psychosozialen Anteil eines psychosozialen Verfahrens zu bestimmen (Meyer, 1996), die „Vergabe“ ist im Fall von Psychotherapie das Verum • deshalb wird auch ein eigenes Modell

	<p>für die Psychotherapieforschung gefordert, das seinem Gegenstand angemessen ist (Neukom, 2004; Leichsenring, 2004)</p>
<p>erforderlich sind eindeutige Diagnosen, und es darf keine Mulitmorbidität vorliegen v.a. wegen der störungsspezifischen Indikation</p>	<ul style="list-style-type: none"> • es handelt sich hierbei um nicht-repräsentative Untersuchungsgruppen (Tschuschke, 2005; Westen und Novotny, 2004; Silverman, 1996), denn die Mehrzahl der Patienten weist komplex bedingte Krankheitsbilder bzw. Multisymptomatiken auf (Henningsen und Rudolf, 2000; Kriz, 2005) • störungsspezifisches Vorgehen ist nur bei manchen Störungsbildern angebracht, bei vielen aber nicht (Westen und Novotny, 2004) • eine Einteilung in nosologisch eindeutig abgrenzbare Gruppen wie ICD- oder DSM-Klassifikationen ist fragwürdig (Beutler, 1998; Mundt und Backenstraß, 2001) • die Annahme, dass eine Technik für eine diagnostische Kategorie angewendet werden kann ohne die Berücksichtigung von lebensgeschichtlichen Zusammenhängen und Persönlichkeitsfaktoren ist zweifelhaft (Westen und Novotny, 2004), denn ähnlichen Oberflächensymptomatiken können unterschiedliche Ätiologien zugrunde liegen, die ein unterschiedliches Vorgehen erfordern (Sanderson, 1998; Westen und Novotny, 2004) • daraus ergibt sich zwangsläufig die Forderung nach naturalistischen Studien (Leichsenring, 2004; Tschuschke, 2005)
<p>theoretische Möglichkeit der Testung komorbider Gruppen im RCT-Design (Chambless und Hollon, 1998)</p>	<ul style="list-style-type: none"> • Praktisch nicht zu verwirklichen, da die im therapeutischen Alltag anzutreffenden Kombinationen die Möglichkeiten experimenteller Designs sprengen würde (Mundt und Backenstraß, 2001)
<p>standardisierte oder manualisierte Behandlung (Buchkremer und Klingberg, 2001)</p>	<ul style="list-style-type: none"> • Manualisierung bildet nicht die therapeutische Realität ab (Kriz, 2005; Smith, 1995)

	<ul style="list-style-type: none"> • allenfalls im Anfängerstadium handlungsrelevant (Zurhorst, 2003), später herrscht situationsgebundenes und ganzheitliches Erfahrungswissen mit starker gefühlsmäßiger Beteiligung vor (Dreyfus und Dreyfus, 1987) • die einseitige Bevorteilung solcher Verfahren, bei denen Manualisierung Bestandteil ihres Therapieverständnisses ist wie bei der KBT, führt zu verzerrenden Effekten bei Vergleichsstudien (Schmitt Freire, 2006) • psychotherapeutische Verfahren entziehen sich einer völligen Standardisierung aufgrund der subjektiven Bedeutungen, der individuellen Beziehungsgestaltungen und der spezifischen Lebensnarrationen (Smith, 1995)
festgelegte Stundenzahl	<ul style="list-style-type: none"> • unter naturalistischen Bedingungen erfolgt die Behandlung i.d.R. bis zum Erfolg (Legewie, 2000)
vorab festgelegte „messbare“ Erfolgskriterien	<ul style="list-style-type: none"> • Bedeutung nicht messbarer und nicht-vorhersehbarer bzw. bewusst intendierter Erfolge und Veränderungen (Legewie, 2000) • im Laufe der Behandlung und mit wachsender Einsicht können sich die Ziele ändern
gruppenstatistischer Wirksamkeitsnachweis	<ul style="list-style-type: none"> • fehlende Skalenqualität • irrelevant für den Einzelfall, da aggregierte Maße der zentralen Tendenz keine Aussagen darüber zulassen, welche Intervention für einen bestimmten Patienten geeignet ist (Krause, 1997; Westen & Novotny, 2004; Sandell 2006)
Trennung von Wirksamkeits- und Prozessforschung (Buchkremer und Klingberg, 2001)	<ul style="list-style-type: none"> • durch so genanntes Package-Testing (Westen und Novotny, 2004) bleibt unklar, <i>was</i> wirkt, daher sind keine weitergehenden Schlussfolgerungen bzgl. der Wirksamkeit möglich (Tschuschke, 2005)

	<ul style="list-style-type: none"> • daraus ergibt sich die Forderung nach Prozess-Ergebnis-Forschung (Tschuschke, 2005; Mertens, 2004)
<p>gesonderte Untersuchung spezifischer vs. unspezifischer Wirkfaktoren (Chambless und Hollon, 1998)</p>	<ul style="list-style-type: none"> • die Trennung zwischen Methode und Beziehung ist nicht für solche Therapieverfahren gültig, die in der Beziehungsgestaltung selbst den veränderungswirksamen Wirkfaktor sehen • in der Psychotherapieforschung werden interpersonelle Faktoren nicht nur als gemeinsame, sondern auch als spezifische Wirkfaktoren angesehen (Luborsky, 1984; Lambert und Bergin, 1994)
<p>nur RCT-Kriterien erfüllen die höchsten wissenschaftlichen Standards (Grawe, 2005)</p>	<ul style="list-style-type: none"> • impliziert ein Verständnis von Wissenschaftlichkeit, das von der Mehrheit der Psychotherapeuten und Psychotherapieforscher nicht mehr geteilt wird (Neukom, 2004), Wissenschaftlichkeit ist ein soziales Konstrukt (Kriz, 2004), das als solches hinterfragbar ist und verhandelbar sein muss • daraus ergibt sich die Forderung nach grundsätzlicher Gleichberechtigung verschiedener wissenschaftlicher Ansätze • Quantitative Forschung allein reicht nicht aus, um methodische Artefakte auszuschließen; wissenschaftlich fundierte Wirksamkeit sollte stattdessen mittels einer „multi-methodalen Konvergenz“ aufgezeigt werden (Fäh und Fischer, 1998; Fischer et al., 2007) • Daraus ergibt sich wiederum die Forderung nach Methodenpluralität (Mertens, 2004; Neukom, 2004)
<p>RCTs als alleinige Möglichkeit zur objektiven Beweissicherung (Borkovec und Castonguay, 1998)</p>	<ul style="list-style-type: none"> • die Grundannahme einer objektiven Wissensgenerierung wird im gegenwärtigen wissenschaftstheoretischen Diskurs als widerlegt angesehen (Mertens, 2005), methodologische Fragen können nicht in Absehung vom Inhalt d.h. der jeweiligen Fragestellung betrachtet werden, kontrollierte und naturalistische Studien

	<p>unterscheiden sich nicht hinsichtlich ihrer internen oder externen Validität, sondern untersuchen vielmehr unterschiedliche Gegenstände („Labortherapien“ und „Feldtherapien“) (Leichsenring, 2004)</p> <ul style="list-style-type: none"> • Wirksamkeit unter naturalistischen Bedingungen lässt sich nur unter naturalistischen Bedingungen nachweisen (Seligman, 1995; Leichsenring, 2004) • Level-1-entsprechende Niveaus (der Evidenzklassen-Hierarchie) sind auch für naturalistische Studien möglich (Leichsenring, 2004, Tschuschke, 2005)
Nur randomisierte Studien erlauben kausale Schlüsse	<ul style="list-style-type: none"> • prospektive, naturalistische Designs erlauben ebenfalls unter bestimmten Voraussetzungen kausale Schlussfolgerungen (Seligman, 1995; Leichsenring, 2004)

Anmerkungen:

1) Efficacy – effectiveness - efficiency

Mit „efficacy“ bezeichnen amerikanische Psychotherapieforscher die Wirksamkeit, wie sie sich unter Laborbedingungen einstellt, mit „effectiveness“ die Effektivität unter naturalistischen Bedingungen und mit „efficiency“ schließlich die Effizienz. Die Untersuchungen zur Effizienz berücksichtigen zusätzlich zur Wirksamkeit oder Effektivität auch den Kosten-Nutzen-Gesichtspunkt.

2) Adhärenz ist eine Bezeichnung für das Ausmaß, in dem sich ein Psychotherapeut an die manualisierten Verfahrensschritte einer bestimmten Psychotherapietechnik hält. Für die Zwecke einer evidenzbasierten Psychotherapie wird es von verschiedenen Psychotherapieforschern als notwendig erachtet, die therapeutischen Vorgehensweisen hinsichtlich detaillierter Verfahrensschritte und bestimmter Abfolgen in einem Manual festzulegen. Damit ist aber noch nicht gesichert, dass ein Psychotherapeut sich an diese Verfahrensschritte hält.

Die Adhärenz lässt sich entweder mittels Selbstbeurteilung am Ende einer Therapiestunde oder mit Hilfe eines Fremdratings anhand von audio- oder videographierten Therapiesitzungen bestimmen. Durch eine entsprechende Supervision kann die Adhärenz des Therapeuten verbessert werden. Die zugrunde liegende Prämisse dieses Forschungskonzepts besteht darin, dass die angewendete Verfahrenstechnik, die durch das Manual operationalisiert wird, eine bedeutsame Rolle für den Therapieerfolg spielt. Wenn dem aber so ist, sollte der betreffende Psychotherapeut auch tatsächlich z.B. kognitiv behaviorale Techniken anwenden und diese nicht mit psychoanalytischen oder gesprächstherapeutischen Elementen kontaminieren. Wie z.B. Jones und Pulos (1993) gezeigt haben, geschieht dies aber häufiger, als gemeinhin vermutet wird. Diese Psychotherapieforscher konnten

aufzeigen, dass diejenigen kognitiv behavioralen Therapeuten besser abschnitten, die Elemente der psychodynamischen Therapiekonzeption verwendeten, wie z.B. das Achten auf die Beziehung. Die strikte Einhaltung einer bestimmten Verfahrenstechnik wird aber von Kritikern als antitherapeutisch eingeschätzt, weil das spontane Eingehen auf einen individuellen Patienten und die Beachtung des jeweiligen Beziehungskontextes dadurch erschwert, wenn nicht sogar verunmöglicht werden. Die Bemühung um Adhärenz kann wie ein Prokrustesbett wirken, in dem intuitive, spontane, beziehungsregulierende und adaptive Elemente der Methodik zuliebe ausgeschaltet werden sollen; ein aus der Beschäftigung mit lebloser Materie hervorgegangenes Experimentallationale hat hierbei offensichtlich Pate gestanden.

3) Allegianz meint die Überzeugtheit eines Wissenschaftler/Psychotherapieforschers von seiner eigenen Theorie oder Behandlungsverfahrens, wodurch es zu verzerrenden Effekten kommen kann.

Literatur:

- Berns, U. (2004). Spezifische psychoanalytische Interventionen. *Forum der Psychoanalyse*, 20, 284 – 299.
- Beutler, L. E. (1998). Identifying empirically supported treatments: What if we didn't? *Journal of Consulting and Clinical Psychology*, 66, 113-120.
- Borkovec, T. D., & Castonguay L. G. (1998). What is the scientific meaning of empirically supported therapy? *Journal of Consulting and Clinical Psychology*, 66, 136-142.
- Buchkremer, G. & Klingberg, S. (2001). Was ist wissenschaftlich fundierte Psychotherapie? Zur Diskussion um Leitlinien für die Psychotherapieforschung. *Nervenarzt*, 72, 20–30.
- Chambless, D. L. & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, 66, 7-18.
- Fäh, M. & Fischer, G. (1998). Vom Sinn und Unsinn in der Psychotherapieforschung. Gießen: Psychosozial.
- Fischer, G., Barwinski R. & Eichenberg, Ch. (2007) Evidenzbasierte Psychotherapie? Wenn schon, dann richtig! Unveröffentlichtes Manuskript.
- Grawe, K. (2005). (Wie) kann Psychotherapie durch empirische Validierung wirksamer werden? *Psychotherapeutenjournal*, 1, 4-11.
- Henningsen, P. & Rudolf, G. (2000). Zur Bedeutung der Evidence-Based Medicine für die Psychotherapeutische Medizin. *Psychotherapie, Psychosomatik, medizinische Psychologie*, 50, 366-375.
- Hsu, L. (1989). Random sampling, randomization, and equivalence of contrasted groups in psychotherapy outcome research. *Journal of Consulting and Clinical Psychology*, 57, 131-137.
- Jones, E.E. & Pulos, S.M. (1993). Comparing the process in psychodynamic and cognitive-behavioral therapies. *Journal of Consulting and Clinical Psychology*, 61, 985 – 1015.
- Krause, R. (1997). Allgemeine psychoanalytische Krankheitslehre (Bd. 1). Stuttgart: Kohlhammer.
- Kriz, J. (2004). Methodologische Aspekte von „Wissenschaftlichkeit“ in der Psychotherapieforschung. *Psychotherapie und Sozialwissenschaft*, 6, 6-31.

- Kriz, J. (2005). Von den GRENZEN zu den PASSUNGEN. *Psychotherapeutenjournal*, 4, 12-20.
- Lambert, M. J. & Bergin, A. E. (1994). The effectiveness of psychotherapy. In: A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed.) (pp. 143 – 189). New York: Wiley.
- Legewie, H. (2000). Psychotherapie unter Rationalisierungsdruck. *Psychomed*, 2, 45-50.
- Leichsenring, F. (2004). Randomized controlled versus naturalistic studies: A new research agenda. *Bulletin of the Menninger Clinic*, 68, 137-151.
- Leuzinger-Bohleber, M., Hau, S. & Deserno, H. (2005). *Depression – Pluralismus in Praxis und Forschung*. Göttingen. Vandenhoeck & Ruprecht.
- Luborsky, L. (1984). *Principles of psychoanalytic psychotherapy: A manual of supportive-expressive therapy*. New York: Basic Books.
- Mertens, W. (2004) Plädoyer für eine theorie-, modell- und methodenplurale psychoanalytische Forschung – wider einen unfruchtbaren Dogmatismus theoretischer und methodischer Art. *Psychotherapie und Sozialwissenschaft*, 6, 48-66.
- Mertens, W. (2005). *Psychoanalyse. Grundlagen, Behandlungstechnik und Anwendung*. Stuttgart: Kohlhammer.
- Meyer, A. E. (1996). Grundzüge der Psychosomatik und ihrer Modelle. In: A. E. Meyer (Hrsg.).
- Mundt, C. & Backenstraß, M. (2001). Perspektiven der Psychotherapieforschung. *Nervenarzt*, 72, 11–19.
- Neukom, M. (2004). „Wissenschaftliche Psychotherapie? Die Ausgangslage der Debatte um die Kriterien der Wissenschaftlichkeit von Psychotherapien in Deutschland und in der Schweiz. *Psychotherapie und Sozialwissenschaft*, 6, 32-47.
- Orlinsky, D. E. & Howard, K. I. (1987) A generic model of psychotherapy. *Integrative eclectic psychotherapy*, 6, 6-27. (Dt.: Ein allgemeines Psychotherapiemodell. *Integrative Therapie*, 4, 281-308, 1988.)
- Sandell R., Blomberg, J., Lazar, A., Carlson, J., Broberg, J. & Schubert, J. (2001). Unterschiedliche Langzeitergebnisse von Psychoanalysen und Langzeitpsychotherapien: Aus der Forschung der Stockholmer Psychoanalyse- und Psychotherapienprojektes. *Psyche - Z Psychoanal*, 55, 277 – 310.
- Sandell, R., Blomberg, J. & a. Lazar, A. (1999). Wiederholte Langzeitkatamnesen von Langzeitpsychotherapien und Psychoanalysen. Erste Ergebnisse der „Stockholmer Outcome of Psychotherapy (STOP) Project. *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 45, 43-56.
- Sanderson, W. C. (1998). The case for evidence-based psychotherapy treatment guidelines. *American Journal of Psychotherapy*, 52, 382-387.
- Schmitt Freire E. (2006). Randomized controlled clinical trial in psychotherapy research: An epistemological controversy. *Journal of Humanistic Psychology*, 46, 323- 335
- Seligman, M. E. P. (1995). The effectiveness of psychotherapy: The Consumer Reports study. *American Psychologist*, 50, 965-974.

- Silverman, W. H. (1996). Cookbooks, manuals, and paint-by-numbers psychotherapy in the 90's. *Psychotherapy*, 33, 207-215.
- Smith, E. W. L. (1995). A passionate, rational response to the "manualization" of psychotherapy. *Psychotherapy Bulletin*, 30, 36-40.
- Tschuschke, V. (2005). Die Psychotherapie in Zeiten evidenzbasierter Medizin Fehlentwicklungen und Korrekturvorschläge. *Psychotherapeutenjournal*, 2, 106-115.
- Wampold, B. E. (2001). *The great psychotherapy debate: Models, methods and findings*. Mahwah, London: Lawrence Erlbaum Associates.
- Wampold, B.E. *The great psychotherapy debate. Models, method and findings*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Westen, D. & Novotny, C. M. (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin*, 130, 631–663.
- Zurhorst, G. (2003). Eminenz-basierte, Evidenz-basierte oder Ökologisch-basierte Psychotherapie? *Psychotherapeutenjournal*, 2, 97-104.