

Das Kuckucksei

Über das pharmakologische Modell in der Psychotherapie-Forschung¹

Dirk Revenstorf

Universität Tübingen

Vorbemerkung

Psychotherapie ist Bestandteil der Gesundheitsversorgung und damit ein Kostenfaktor, der Rechenschaft verlangt. Die empirisch fundierte Therapieforschung ist angetreten, die dazu nötigen Nachweise zu liefern. Da mit den Nachweisen auch Entscheidungen über Akkreditierung bestimmter und die Diskreditierung anderer Therapieverfahren gefällt werden, gewinnt Therapieforschung eine politische Dimension, bei der es um die Ausgabe knapper finanzieller Mittel für die Forschung, Ausbildung und die Erstattung von Behandlungskosten geht. Mindestens im Hinblick auf eine gerechte Verteilung dieser Ressourcen ist eine kritische Beleuchtung der Kriterien empirischer Evidenz in der Therapieforschung angebracht.

Einleitung

Ein traditionelles Problem in der psychologischen Forschung ist der auf Windelband (1904) zurück gehende Gegensatz zwischen nomothetischer und ideografischer Wissenschaft (Allport 1934). Akademische Psychologen sind daran interessiert, nomothetische Aussagen zu machen. Doch sind klinische Psychologen oft der Meinung, dass das, was Menschen im Leben bewegt und ihnen zum Problem wird, zu einem bisher nicht näher bestimmten Anteil sich nur individuell (d.h. ideografisch) verstehen und behandeln lässt. Damit wäre Psychotherapie grundsätzlich mit einer nomothetischen Beschreibung nur unvollständig erschließbar. Psychotherapeutisches Handeln ließe sich dann auch nicht systematisch aus einem theoretischen Zusammenhang deduzieren. Gegen einen einfachen Zusammenhang von Theorie und Praxis sprechen ohnehin einige bedeutsame Befunde der Therapieforschung. In der groß angelegten NIMH-Depressionsstudie (Elkin et al. 1994) beeinflussten Imipramin, Interpersonelle Therapie (IPT) und Kognitive Therapie alle drei unterschiedslos physiologische Reaktionen, dysfunktionale Gedanken und soziale Kompetenz und die jeweils

¹ Ich danke den Kollegen, Wilhelm Glaser, O.Bernd Scholz und Thomas Fydrich für die kritische Lektüre des Manuskripts und wertvolle Anregungen.

aus dem spezifischen physiologischen, kognitiven und sozialen Erklärungszusammenhang (Imber 1990) konstruierten Interventionen wirkten also nicht spezifisch, wie man annehmen sollte. Das spricht aber noch nicht gegen eine nomothetische Psychotherapieforschung, sondern für eine vernetzte Wirkungsweise von Psychotherapie, die auf nicht auf eine getrennte Wirkung physiologischer, kognitiver und sozialer Mechanismen und damit auf die Indikation einer bestimmten Methode für einen bestimmten Zweck reduziert werden kann (Abbildung 1).

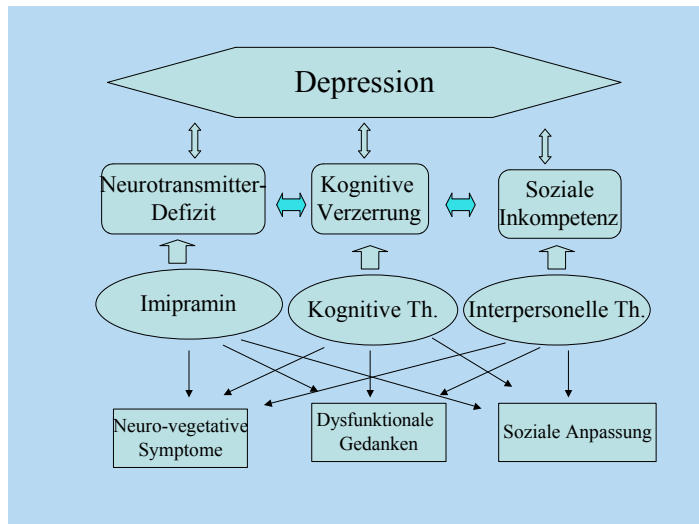


Abbildung 1: Multikausale Verknüpfung von Maßnahme und Wirkung (Messergebnissen) in der NIMH-Depressions-Studie (Elkin et al. 1994).

Aus der Schwierigkeit, Interventionen der Psychotherapie theoretisch herzuleiten, haben einige Therapeuten wie A. Lazarus (1979) die Konsequenz gezogen, sich zu einem technischen Eklektizismus zu bekennen oder wie M.H. Erickson programmatisch zu behaupten, jeder seiner Patienten (400 sind bei O'Hanlon & Hexum 1996 dokumentiert) erfordere eine eigene Theorie und einen neuen Ansatz (s. auch Ericksons Ges. Schriften (Rossi (Hrg.) 1995-98)).

Auf der anderen Seite bemüht sich die Klinische Psychologie um nomothetische Ätiologie von psychischen Problemen (ICD, DSM)². Und die Therapieforschung bemüht sich ebenfalls um einen nomothetischen Ansatz, indem sie Therapieverfahren differenziert - etwa in

² Man kann Klassifikationen für nicht nomothetisch halten, weil sie keine Kausalaussagen beinhalten; doch haben sie mit dem Anspruch auf Vollständigkeit und den deskriptiven Aussagen wie z.B. über Komorbidität ebenfalls gesetzesartigen Charakter.

behaviorale, kognitive, psychodynamische, systemische - und in ihrer Wirksamkeit nach überindividuell einheitlichen Maßstäben bewertet. Dabei orientiert sich die Forschung an Kriterien der Wissenschaftlichkeit, wie sie in Physik, Biologie und - speziell bezüglich der Therapieforschung - in der Medizin üblich sind. Auf diese Weise wird ein nach außen vertretbarer Standard an Objektivität gesucht, der Psychotherapie in ähnlicher Weise fassbar und respektabel macht wie die Medizin. Das entspricht dem Wunsch nach der Anerkennung als gleichrangiges Heilverfahren und nach Qualitätssicherung. Entsprechendes ist in der Legaldefinition (§1, Abs.3) des Psychotherapeuten-Gesetzes juristisch verankert, wonach Psychotherapie nur in der wissenschaftlich anerkannten Form ausgeübt werden kann.

Zum Nachweis der Wirksamkeit werden in der empirischen Psychotherapieforschung Kontroll-Gruppenexperimente herangezogen, wie sie in der pharmakologischen Forschung üblich sind, um ein Heilmittel zu überprüfen. Bei kritischer Beleuchtung werden schnell einige Grenzen der Möglichkeit deutlich, mit diesem Ansatz empirische Evidenz zu sichern. Dadurch wird aber auch der Anspruch, mithilfe derartiger Studien Therapieformen zu akkreditieren, infrage gestellt. Das Problem hat zwei Ebenen.

Auf der *inhaltlichen* Ebene ist zu bezweifeln, ob eine psychotherapeutische Behandlungstechnik wie ein Medikament wirkt, also durch spezifische Wirkmechanismen und ob sie u.U. für eine bestimmte Störung besonders indiziert ist – etwa so, wie Antibiotika gegen Bakterien aber nicht gegen Viren eine spezifische Wirkung entfalten. Dagegen sprechen z.B. die oben erwähnten Ergebnisse der Depressionsforschung. Oder ob Psychotherapie ein Verfahren ist, mit dem der psychische Apparat in ganzheitlicher Weise umgestimmt wird, weil durch eine für Patient und Therapeut überzeugende Behandlungstechnik ein Kontext hergestellt wird, in dem Selbstheilungskräfte mobilisiert werden. Die jeweils angewendete Technik sähe zwar spezifisch aus, würde aber nur eine von vielen Realisationsmöglichkeiten allgemeiner Therapiefaktoren darstellen, wie sie von Frank, Strupp oder Bandura formuliert wurden (vgl. Goldfried 1980). Psychotherapie wäre dann in seiner Wirkung mit einem Pharmakon nicht vergleichbar.

Auf der *methodologischen* Ebene erhebt sich die ganz andere Frage, ob die Gleichsetzung von Kriterien der Wissenschaftlichkeit für die Psychotherapie mit denen der Pharmakologie sinnvoll und ob das daraus entlehene Kontrollgruppenexperiment in der Psychotherapie dienlich ist. Nicht, dass auf wissenschaftliche Evaluation verzichtet werden könnte. Immerhin hat das bisherige Evaluationsmodell wichtige Erkenntnisse erbracht wie z.B. den Nachweis der generellen Wirksamkeit von Psychotherapie, wie auch die weit gehende Uniformität der Wirkung verschiedener Therapieformen (s. Fazit weiter unten). Wenn aber die bisherige Evaluationsstrategie manche für die Theorie und Praxis bedeutsamen Therapieformen systematisch nicht erfasst, weil diese nicht ins medizinische Modell passen und damit per Gesetz ihre Ausübung bedroht ist, ist das Evaluationsmodell zu überprüfen. Darauf soll hier näher eingegangen werden.

Rationalität als Grundlage der Wissenschaft

Die Naturwissenschaften stützen sich mit Erfolg auf Methoden der rationalen Analyse, wie

sich in der Physik oder Chemie durch ihre Vorhersagekraft bewährt haben. Dass deterministische Gesetzmäßigkeiten in der Physik schon bei der Thermodynamik in Probabilistik und bei Meteorologie in kurzfristige, vage Prognosen umschlagen, die jeder aus der Wettervorhersage kennt, wird oft übersehen. Die Psychologie ähnelt, wenn überhaupt ein Vergleich mit der Physik erlaubt ist, eher der Meteorologie als der Mechanik. Nicht nur dass sich der nomothetische Anspruch in der Psychotherapie auf Wahrscheinlichkeits-Aussagen, also Statistik reduziert. Es ist darüber hinaus zweckmäßig zu beleuchten wie weit der Begriff der aus Biologie und Medizin übernommenen Wissenschaftlichkeit trägt, um Überprüfungen und Vorhersagen über den Therapieerfolg zu versuchen, die dann u.a. auch für Legalisierung eines Therapieverfahrens maßgeblich sind.

Das derzeitige wissenschaftliche Selbstverständnis gründet sich auf ein Monopol der Rationalität, die einer bestimmten Stufe der Entwicklung des Bewusstseins entspricht und durch das Quadrupel (3-dimensionaler Raum, Unidirektionalität der Zeit, Kausalität und Subjekt-Objekt-Trennung) gekennzeichnet ist. Folgt man der bei E. Neumann (1995), J. Gebser (1986) und K. Wilber (1987) dargelegten anthropologischen These von einer Evolution des Bewusstseins, so gingen der Rationalität, die unser Denken seit Aristoteles oder spätestens seit der Renaissance und der Aufklärung beherrscht, mythologische und davor magische Bewusstseinsformen voraus. Sie wurden von der Rationalität überlagert, die wiederum ein Durchgangsstadium zu einem hypothetischen postrationalen Bewusstsein darstellt.

Jedes Stadium der Bewusstseinsentwicklung hat ein anderes Heilungsmodell hervorgebracht: Das magische Denken den Schamanismus, das mythologische Denken diverse quasireligiöse und religiöse Rituale (Orakel, Tempelschlaf, Handauflegen, Seelsorge) und das rationale Bewusstsein die naturwissenschaftlich begründeten Therapieansätze (s. Abbildung 2). Nun erscheint es äußerst zweckmäßig, die historisch gegebene rationale Bewusstseinslage als Rahmen zu nutzen, um Psychotherapie von Schamanismus und Orakel zu emanzipieren. Es spräche nichts dagegen, sich auf den Positivismus als Forschungsstrategie der Moderne auch für die Psychotherapie zu verlassen, wenn nicht dadurch etwas für das Verständnis des Menschen Bedeutsames verloren ginge. Und es besteht Grund anzunehmen, dass das so ist.

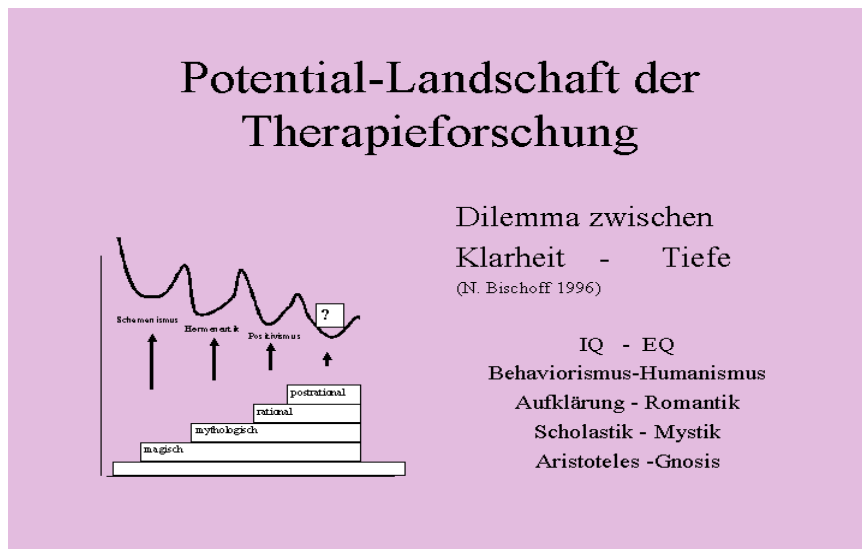


Abbildung 2: Rationalität als Stadium der Bewusstseinsentwicklung

Die Verpflichtung auf die Rationalität in den westlichen Kulturen ist nicht durchgängig angenommen worden. Seit Beginn des rationalen Zeitalters (d.h. in den letzten 2000 Jahren) bestand nach N. Bischoff (1999) immer eine Spaltung des Denkens in das, was Nietzsche apollinisch und dionysisch genannt hat. Offenbar war der Preis der zunehmenden Klarheit durch die Vernunft der Verlust dessen, was man als Tiefe des Verständnisses empfindet.

Es gab schon zu Zeiten der aristotelischen Akademie die Gnosis als Bemühung um Tiefe, zu Zeiten der Scholastik die Mystik und die Aufklärung provozierte als Antithese die Romantik. Ähnlich erwachsen aus dem Unbehagen mit dem verkürzten Menschenbild des positivistisch ausgerichteten Behaviorismus die humanistische Psychologie³ und entsprechende Therapieformen (Gesprächspsychotherapie, Gestalttherapie u.a.) und in den letzten Jahren kam es zu einer Rehabilitation des Gefühls in Form emotionaler Intelligenz als Gegenspieler zur rationalen Intelligenz (Goleman, 1995, Gershon 1998, Solomon 2000) und neuerdings wurde sogar der Begriff der spirituellen Intelligenz geprägt worden (Zohar & Marshall 2000). Während sich die Legalisierung der Psychotherapie auf die Klarheit stützt, könnte das, was hier vorläufig pauschal mit „Tiefe“ gekennzeichnet wird, für die Bestimmung der Befindlichkeit und der Behandlung psychischer Probleme Ausschlag gebend sein. Dann hätte Psychotherapieforschung einen breiteren Anspruch als die übrigen Naturwissenschaften, nämlich den, beides zu verbinden. Indem die Therapieforschung sich auf die rationale Analyse beschränkt, opfert sie der Klarheit die Tiefe.

Beispielsweise ist als Kritikpunkt an den kognitiv-behavioralen Therapieansätzen eine spezielle Form des Mangels an Tiefe zugunsten großer Klarheit auszumachen, nämlich die Auffassung, Therapie sei Übung, Habituation, Planung und Umstrukturierung von Sichtweisen. Damit wird vieles von dem, was Patienten leiden lässt und was sie in der Therapie thematisieren nicht adäquat erfasst⁴. Wenn ich etwa feststelle, vor zwanzig Jahren die falsche Berufswahl getroffen zu haben oder meine große Liebe nicht geheiratet zu haben oder wenn es darum geht, wie man mit einer lebensbedrohlichen somatischen Krankheitsdiagnose oder einem Verlusterlebnis umgeht, ist das als Konditionierungsprozess oder Problemlösung nicht angemessen zu bearbeiten.

Themen	Therapieziele
Hiatus zwischen möglich und wirklich Falsches Leben gelebt Zweifel in der Entscheidung Unvermeidbarkeit von Schuld Faustische Unrast Begrenztheit des Lebens Einsamkeit angesichts des Todes	<ul style="list-style-type: none"> ● Primat der Erfahrung ● Entwicklung der Person ● Rettung der Würde ● Liebesfähigkeit ● Suche nach Sinn

Tabelle 1: Krisenthemen und resultierende Therapieziele, die schwer operationalisierbar sind.

Die mit solchen Themen verbundenen Therapieziele der persönlichen Entwicklung, Sinnerfüllung und Wiederherstellung der Würde sind aber nicht nur theoretisch difizil; sie fallen außerdem leicht durch das Raster einer positivistischen Analyse, da diese Ziele nicht so einfach operationalisiert werden können wie die Anzahl der Migräneanfälle, Panik-Attacken oder Zwangshandlungen. Therapieverfahren, die sich diesen Aspekten der Heilung (Tabelle 1) zuwenden, können daher nur schwer mit dem positivistischen Ansatz überprüft werden, obwohl sie womöglich einen wichtigen Beitrag zur Psychotherapie leisten. Das positivistische Wissenschaftsverständnis erweist sich somit als ergänzungsbedürftig, wenn auch Therapieformen eine Chance haben sollen, überprüfbar zu sein, deren Fokus solche Ziele sind. Wünschenswert ist, dass das Verlassen des positivistischen Paradigmas mit einer dialektischen Überwindung von Klarheit und Tiefe verbunden wäre.

Die Überwindung der Verflachung von Psychologie kann nicht die Preisgabe der Rechenschaftspflicht (engl.: *Accountability*) der Psychotherapie bedeuten - insbesondere wenn mit Behandlungskosten zu rechnen ist, die von der Solidargemeinschaft der Versicherten getragen werden und mit Nebenwirkungen der Therapie, wie bei nicht fachgerechter Behandlung von posttraumatischer Belastung. Ein Negativ-Beispiel für den unüberprüften Umgang mit einem Therapieverfahren mit potentiell negativen Nebenwirkungen ist die

³ Gründung der Gesellschaft für Humanistischer Psychologie 1962 von Charlotte Bühler, Arthur Köstler, Carl Rogers, Abraham Maslow u.a.

Odyssee der Insulin- und Elektroschock-Behandlung für schwere Psychosen, die in ihrer Anwendung in den Psychiatrischen Krankenhäusern unabhängig von empirischen Belegen der Wirksamkeit Höhen und Tiefen der Beliebtheit erlebte (Curran 1959).

- 1949 wurden beide Verfahren als Revolution in der Behandlung für schwer depressive und schizophrene Patienten gefeiert (Curran, 1949)
- 1955 sagt derselbe Autor, dass es eigentlich nicht wirksam ist.
- Trotzdem wird 1959 das Verfahren in den meisten Einrichtungen noch verwendet; später wird es durch Psychopharmaka ersetzt.
- Auch wenn viele Psychiater zu der Auffassung kommen, dass diese Art der Therapie ineffektiv sei, ist sie in der Mehrheit der Psychiatrien nicht abgeschafft worden und wird es in vielen Kliniken zumindest bei schweren Fällen immer noch eingesetzt.

Daher ist die Forderung nach empirischer Evidenz für ein akkreditierbares Therapieverfahren unabweisbar. Und da die psychotherapeutische Heilkunst seit der Einführung der psychologischen Approbation analog der medizinischen Heilkunst gehandhabt wird, lag es ja auch nahe, entsprechend die dort üblichen Kriterien der Evidenz anzulegen. Soweit so gut. Nur muss geprüft werden, ob diese Kriterien passen.

Die Akkreditierung einer Therapieform gründet sich nach dem derzeitigen Verständnis auf drei Bestandteile: das Vorhandensein einer *Theorie*, einer *Verfahrensbeschreibung* und einer *Evaluation* der Ergebnisse (s. Wissenschaftlicher Beirat Psychotherapie (WBPT) oder auch Ehlert 2003, s. Abbildung 2). Die *theoretische* Begründung einer Therapieform wird problematisch, wenn die Verknüpfung von Veränderungsmechanismen mit Störungskategorien (Diagnostik) und Entstehungsmodellen (Ätiologie) nicht gelingt, was dann zum des technischen Eklektizismus führt. Tatsächlich fand sich ja in der schon genannten NIMH-Depressionsstudie (Elkin et al. 1994) dass diese Verknüpfung nicht einfach gelingt, d.h. eine Wirkung theoretisch nicht eindeutig zugeordnet werden konnte.

Die *Verfahrensbeschreibung* wird durch Manuale angestrebt, die ihre eigene Problematik haben, da sie meist von der Annahme spezifischer Wirkkomponenten ausgehen, die in den Manualen beschrieben werden – eine Annahme, die selbst der Überprüfung bedarf und eher fragwürdig ist. Denn in der Metaanalyse von 27 „Entblätterungs“-Studien (Wampold 2001, S.124) fand sich kein Hinweis darauf, dass das Weglassen einer wesentlichen spezifischen Komponente die Wirksamkeit des Therapie veränderte. Etwa untersuchten Jacobson et al. (1996) bei der Depressionsbehandlung die Komponenten Verhaltensaktivierung und kognitive

⁴ Interessanterweise würde ja nach einer Umfrage von A. Lazarus (1979) kein Verhaltenstherapeut zur

Therapie und fanden, dass das Weglassen jeweils einer der beiden Komponenten die Wirksamkeit erhöhte.

Auf die Probleme der theoretischen Begründung und der Verfahrensbeschreibung soll nicht näher eingegangen werden. Es sollen stattdessen einige Probleme der *Evaluation* betrachtet werden, die wiederum verschiedene Aspekte umfasst, nämlich Laborwirksamkeit (in kontrollierten Studien; engl.: efficacy), Wirksamkeit in der Praxis (in Kliniken und Ambulanzen; engl.: effectiveness) und die Wirtschaftlichkeit (engl.: efficiency). Da die Evaluation in der Praxis und die der Wirtschaftlichkeit gar nicht erst in Betracht gezogen werden, wenn nicht eine Wirksamkeit unter kontrollierten Bedingungen nachgewiesen wurde, stellt dies die entscheidende Schwelle für die Anerkennung der Wissenschaftlichkeit einer Therapieform dar.

Kritik des Kontrollgruppen-Designs

Für den Nachweis der Laborwirksamkeit wird in der Therapieforschung derzeit auf das Paradigma des randomisierten Kontrollgruppen-Designs zurückgegriffen, das in der Pharmakologie in Form des Doppelblindversuchs üblich ist. Das Kontrollgruppen-Design (engl.: Randomized controlled trial, RCT) gilt als "Goldstandard" in der Therapieforschung (s. Chambless & Hollon, 1998) und hat folgende Charakteristika

- 1) repräsentative Untersuchungs-Stichprobe
- 2) Stichprobengröße von mindestens 25 Klienten pro Design-Zelle
- 3) randomisierte Zuteilung zu Experimental- und Kontrollgruppe
- 4) Standardtherapie als Kontrollgruppe oder
- 5) Warteliste als Kontrollgruppe oder
- 6) nicht erkennbares Placebo als Kontrollbedingung
- 7) Therapeuten-blinde Anwendung von Placebo und Verum
- 8) Patienten-blinde Anwendung von Placebo und Verum
- 9) Auswerter-blinde Analyse der Daten
- 10) Hinreichend langer Katamnesezeitraum (6 Monate bis 1 Jahr)

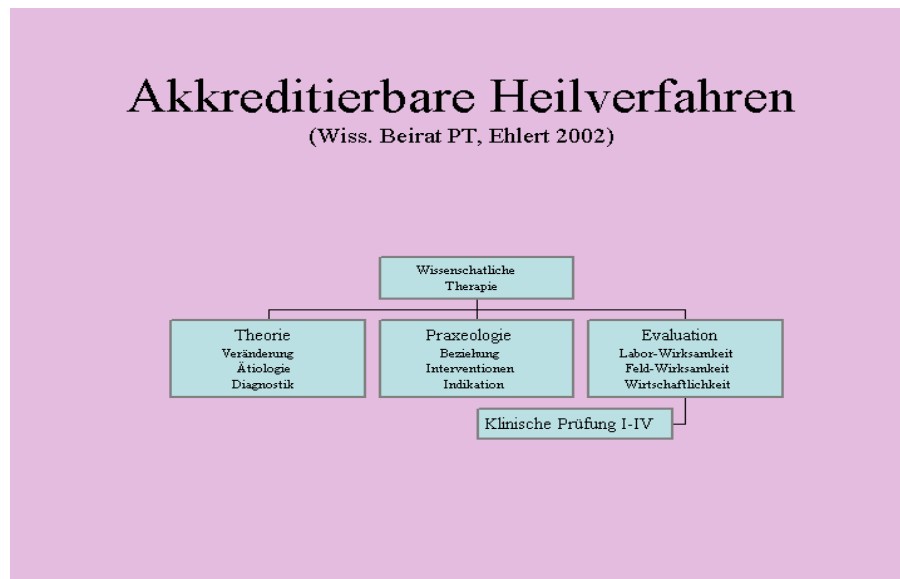


Abbildung 3: Komponenten der Wissenschaftlichkeit einer Therapieform (nach Ehlert 2003).

Notorische Irrepräsentativität

Selbstselektion: Wenn die Analyse der Daten den Therapieerfolg bestätigt - auf welche Grundgesamtheit ist das Ergebnis generalisierbar? In den Untersuchungen, die in Forschungseinrichtungen gemacht werden (i. Allg. Universitäts-Institute oder Kliniken) entscheiden sich Patienten, die Zugang zu dieser Einrichtung haben und sie gewählt haben, für oder gegen die Teilnahme an einer Studie. Oder die Teilnehmer werden durch Annoncen akquiriert. Meist sind es solche Patienten, die ihre Therapie nicht selbst zahlen können. Hier findet also immer eine Selbstselektion (engl.: patient preference) statt, indem nur Patienten an der jeweiligen Psychotherapie-Studie teilnehmen, die ihr Problem (a) psychisch attribuieren und (b) über begrenzte Geldmittel für eine Therapie verfügen und (c) zufällig im Einzugsbereich des Forschungsinstitutes wohnen und auf die betreffende Studie aufmerksam wurden. Die ersten beiden Gesichtspunkte spielen bei pharmakologischen Untersuchungen keine Rolle.

Komorbidität: Die vom Untersucher angelegten Ausschlusskriterien bringen eine weitere Einschränkung der Repräsentativität mit sich. Normalerweise wird nämlich in wissenschaftlichen Studien die seltene Sorte Patienten behandelt, die sich monosymptomatisch klassifizieren lässt – etwa als reine Suchtfälle, Agoraphobien, Depressive, Persönlichkeitsstörungen u.ä. Nicht nur existiert eine hohe Korrelation der Symptomstörungen (Achse 1 des DSM IV) mit bestimmten Persönlichkeitsstörungen (Achse 2, Fiedler 1998), sondern die Komorbidität innerhalb beider Achsen ist ebenfalls hoch. So weisen 70% der Fälle mit Depression Symptome von Angst, Zwang, oder PTSD auf und innerhalb der Angststörungen selbst trifft dies auch wieder zu (Davison & Neale 2002, S. 148, 317, 338). Innerhalb der Persönlichkeitsstörungen sieht es ähnlich aus: Borderline-Fälle etwa werden zu 55% als schizotypisch, zu 47% als antisozial, 57% als histrionisch diagnostiziert. Ähnliches trifft auf die dependente Struktur zu (ebenda S.456, 473, Comer 1995, S.635). Das heißt, die für einzelne Störungen entworfenen und geprüften Therapieprogramme sind für die allgemeine Klientel mit komplexen Komorbiditäten nur bedingt anwendbar. Und die Kombination verschiedener Verfahren für verschiedene Störungskombinationen

ist meist nicht systematisch (d.h. faktoriell) erforscht, um Interaktionseffekte aufzuklären. Im Gegensatz dazu werden häufig Breitbandprogramme angewendet, deren Komponenten nicht einzeln auf Wirksamkeit überprüft wurden.

Es ist daher fragwürdig, wie weit die Klassifikation von Patienten nach ICD oder ähnlichen Systemen angesichts der hohen Komorbidität trägt und ob nicht eine ganzheitliche und prozess-orientierte Betrachtungsweise wie sie in den humanistischen Therapieformen angestrebt wird, angemessener ist als eine störungsspezifische Therapiekonzeption. In jedem Fall sind Untersuchungs-Stichproben mit monosymptomatischen Fällen irrepräsentativ für die Grundgesamtheit der behandlungsbedürftigen Klienten, wie sie in der Praxis auftauchen und die Untersuchungsergebnisse stehen lediglich für Patienten, die sich für die jeweilige Untersuchung zur Verfügung stellen

*Zusammenfassend lässt sich für die Psychotherapie feststellen, dass die Repräsentativität einer Studienstichprobe für eine Grundgesamtheit von Patienten in der Alltagspraxis durch **Selbstselektion** und **monosymptomatische** Einschlusskriterien verhindert wird. Es bleibt daher unklar, für als wirksam erwiesene Therapieform hilfreich ist.*

Probleme von Kontrollgruppe und Zuteilung

Im Gegensatz zur pharmakologischen Forschung gibt es in der Psychotherapie kein glaubwürdiges Placebo. Placebothérapien sollen nur unspezifische und keine spezifischen Wirkungen entfalten. In der Medikamentenprüfung sind die unspezifischen Wirkungen psychologisch und die spezifischen Wirkungen physiologisch. Dort ist eine klare Trennung möglich. In der Psychotherapie sind Placebo und Verum beide psychologisch und daher schwer zu trennen. Zuweilen sind Versuche unternommen worden, etwas dem Placebo analoges in der Psychotherapie zu kreieren: „Behandlungen“, die der Patient dafür hält aber der Therapeut nicht (Aufmerksamkeitsplacebo, Diskussionsgruppen, unterstützende Beratung). Derartige Behandlungen sind zu wenig frei von therapeutischen Komponenten, um als Placebo gelten zu können und wirken wenig überzeugend, wenn sie als Realisierung nur einiger unspezifischer Therapiefaktoren intendiert sind⁵.

Daher wird auf andere Formen der Kontrollgruppe zurückgegriffen: Vergleichstherapie oder Warteliste. Die *Warteliste* als Kontrollgruppe hat ihre eigenen Schwächen: häufig fällt es schwer, ernsthaft Kranke ein halbes Jahr auf die Behandlung warten zu lassen, abgesehen davon, dass sie auch gar nicht warten werden, sondern das Experiment abbrechen. Das beeinträchtigt die Vergleichbarkeit der Gruppen und damit die Schlüssigkeit des Experiments. Etwa hatte in einer Studie zu Schlafstörungen nach der Randomisierung die Wartegruppe eine durchschnittliche Einschlafzeit von 70 min wie die Behandlungsgruppe. Zum Zeitpunkt der Katamnese nach 6 Monaten war die Wartegruppe auf solche Patienten geschrumpft, die eine durchschnittliche Einschlafzeit von 30 min hatten. Die schweren Fälle hatten eine andere Behandlung aufgesucht (Schulz & Weinmann 2001, Schlarb 2003).

⁵ Ein richtiges Psychotherapie-Placebo müsste u.a. enthalten: gleich gute Beziehung, gleiche Erwartungen und Präferenz des Patienten fürs Verfahren, gleiche Überzeugungskraft und Erwartung des Therapeuten – was allein schon durch dessen fehlende Blindheit ausgeschlossen ist.

Als Alternative zur Warteliste eine *Vergleichstherapie* als Kontrollgruppe eingesetzt werden, die entweder die bestmögliche Standardtherapie nach derzeitigem Stand der Kunst ist oder die eine bestimmte theoretische Position repräsentiert, um ein Entscheidungsexperiment durchzuführen – etwa hilft VT bei Migräne besser als Gesprächspsychotherapie? Ein solches Experiment ist aber nur valide, wenn der Therapeut von der Wirkung der von ihm durchgeführten Therapie überzeugt ist. Metaanalysen (Wampold 2002) haben herausgefunden, dass neben der Güte der Arbeitsbeziehung (engl.: *alliance*) die Überzeugung des Therapeuten von Richtigkeit seiner Intervention (engl.: *allegiance*) einer der bedeutsamsten Wirkkomponenten ist. Das erfordert bei komparativen Studien schulübergreifende Forscherteams, sodass die zu vergleichenden Therapieformen angemessen vermittelt werden. Das ist z.Z. eher nicht der Fall, da z.B. fast alle deutschen Universitätsinstitute, an denen Psychotherapieforschung betrieben wird, verhaltenstherapeutisch orientiert sind.

Für jedwede Art von Kontrollgruppen ist in der Psychotherapie weder die Unwissenheit des Therapeuten gegenüber dem "Medikament" (i.e. der jeweiligen Psychotherapieform) zu erreichen noch die Blindheit des Patienten gegenüber dem Verfahren wie es in der Pharmaforschung gefordert wird – zumal der Patient sich außerhalb der Therapie über das, was mit ihm geschieht informieren kann, um sich schlau zu machen. Außerdem erhebt sich die Frage, wie weit man die Autonomie des Patienten bei der Wahl der Therapieform übergehen kann. Patienten können erfreut oder erschreckt sein, wenn sie erfahren, dass ihnen in einer Therapiestudie per Randomisierung eine Reizüberflutung statt Entspannung zugeteilt oder vorenthalten wird. Aus diesem Grunde kämen eigentlich nur Cross-over Designs infrage, die allerdings eine unabhängige Evaluation der zu vergleichenden Verfahren erschweren.

Zusammenfassend lässt sich feststellen, dass Wartelisten wegen unzumutbarer Verzögerung der Behandlung bei hinreichender Therapiedauer und langer Katamnese kaum infrage kommen und außerdem zu Abbrüchen führen (engl.: attrition). Hinzu kommt, dass bei randomisierter Zuteilung die Autonomie des Patienten derart missachtet wird, dass Reaktanz nicht auszuschließen ist. Daher macht ein valides Therapieexperiment Vergleichstherapien im „Cross-over“-Design oder die Selbstzuweisung der Patienten mit nachträglicher Balancierung bezüglich möglicher Störvariablen erforderlich und die Zusammenarbeit von Forschern, die von den durchgeführten Therapieformen überzeugt sind.

Fallsstricke der Auswertung

Reliabilitäts-Validitätsdilemma: Die Auswertung einer Therapie erfordert die Objektivierung des Therapieerfolges. Dabei tritt das bekannte Reliabilitäts-Validitätsdilemma auf, dass nämlich ein sehr zuverlässiger psychologischer Test, wie ein IQ-Test, eine sehr gute Voraussage für das Abschneiden in einem anderen ähnlichen Test ermöglicht, aber relativ wenig für die Bewältigung im täglichen Leben. Befragt man letztere mithilfe von subjektiven Selbsteinschätzungen, erhält man zwar valide aber weniger zuverlässige Daten. Das betrifft auch die oben genannten Therapieziele (Tabelle 1). Ein Interview oder erste Eindrücke bei einer Begegnung, die im strengen Sinne überhaupt nicht reliabel sind, haben oft eine größere

Validität für den Therapeuten als ein Fragebogenergebnis. Diesem Dilemma kann man offenbar nicht entkommen. Aber es wird vielleicht erträglicher, wenn man akzeptiert, dass eine komplexe Datenmenge, wie sie ein Interview oder ein Personeneindruck liefert, ein sensibles Gefüge von Beobachtungen darstellt, in dem die fluktuierende Aufmerksamkeit des Betrachters jedes Mal einen etwas anderen Fokus findet und zwar zu validen aber nicht leicht replizierbaren (d.h. reliablen) Ergebnissen kommt.

Primat der Subjektivität: Die Widerständigkeit der Materie bei Operationalisierungsversuchen von subjektiver Erfahrung ist nicht etwa als Mangel der Psychologie zu verstehen. Auch in der Physik gilt das Primat der Subjektivität und jedem Messen ist eine subjektive Erfahrung vorausgegangen (Wärme, Schwerkraft, Wetter, usw. Abbildung 4). Nur fehlt es in der Psychologie oft an angemessenen korrelativen und kausalen Reduktionsmöglichkeiten (Searle 1993). Das ist aber kein Hinweis auf die Minderwertigkeit psychologischer Variablen, die auf subjektiver Erfahrung allein beruhen und möglicherweise valider sind als die, die auf Operationalisierungen reduziert wurden.

Physikalismus			
Reduktion	Physik	Medizin	Psychologie
Subj. Phänomen	Wärme	Schmerz	Irrat. Gedanke
Obj. Phänomen	Wasser verdampft	schreit ?	Depress. Ausdruck?
Korrelative Reduktion	Molekularbeweg.	C-Faser-Leitung?	Transmitterdefizit ?
Kausale Reduktion	Kinetische Energie	?	?
Neudefinition	Celsius°	?	?
Manipulation	Energiezufuhr	?	?

Abbildung 4: Möglichkeiten und Grenzen des Reduktionismus in der Physik, Medizin und Psychologie: In der Psychologie fehlen häufig die eindeutigen objektiven Phänomene, die Korrelation mit einer einfacheren Substrat-Ebene und auf kausale Mechanismen (nach Searle 1993).

Reaktivität: Ein weiteres Problem ist die Reaktivität des Messinstrumentes. Gut operationalisierbare Variable, die durch den Therapeuten einsehbar erhoben werden wie Nähe und Dauer des Kontakts mit dem phobischen Objekt oder introspektive Aussagen über die Befindlichkeit, unterliegen leicht der Verzerrung durch Gefälligkeit. So konnten Smith, Glass und Miller schon 1980 feststellen, dass jeglicher Vorsprung der behavioralen Therapie vor den verbalen verschwindet, wenn man die Reaktivität des Messinstrumentes berücksichtigt.

Klinische Relevanz: Die Bestätigung der Wirksamkeit einer Behandlung wird mit komplexer statistischer Methodik ermittelt: mit multivariaten Varianzanalysen, Post-hoc Prüfungen mit Bonferoni-Korrektur, Herauspriorisierung von Störvariablen, usw. Am Ende bleibt ein Signifikanzniveau übrig, das sehr begrenzte Aussagekraft hat. Die Statistik lehrt, dass wenn die Stichprobe groß genug ist auch kleinere Effekte signifikant werden und das kann bedeuten, dass durch das Verfahren womöglich kein einziger Patient geheilt wird. Wenn sich der Wert im Beck-Depressionsinventar im Durchschnitt um 10 Punkte von 30 auf 20 verringert, bleibt unbekannt, wie viele einen Wert unter 16 erreichen (geheilt sind). D.h. der Untersuchung fehlt trotz statistischer gesicherter Wirksamkeit die klinische Relevanz (Jacobson & Revenstorf 1986). Wichtig wäre die Frage, wie viele Leute denn nun wirklich gesund geworden sind (Abbildung 5).

Die nicht widerlegte Nullhypothese: Eine weitere bekannte Schwierigkeit statistischer Signifikanz-Aussagen stellt die unwiderlegte Nullhypothese dar. Zwar ist der α -Fehler bekannt (z.B. 5%), mit dem ein gefundener Unterschied als auch zufällig sein könnte. Aber es bleibt ja der β -Fehler unbekannt, mit dem ein auf dem gewählten Signifikanzniveau abgelehnter Unterschied auf einem wahren Unterschied beruht. Dem kann man bekanntlich durch Vergrößerung der Stichprobe oder Verringerung der Varianz entgegen treten. Wichtiger ist jedoch: eine nicht widerlegte Nullhypothese ist nicht der Beweis dafür, dass kein Unterschied zwischen unbehandelter Kontrollgruppe und der mit der Therapieform X behandelten Gruppe besteht, sondern ein Indikator dafür, dass der Beweis des Unterschieds noch aussteht.

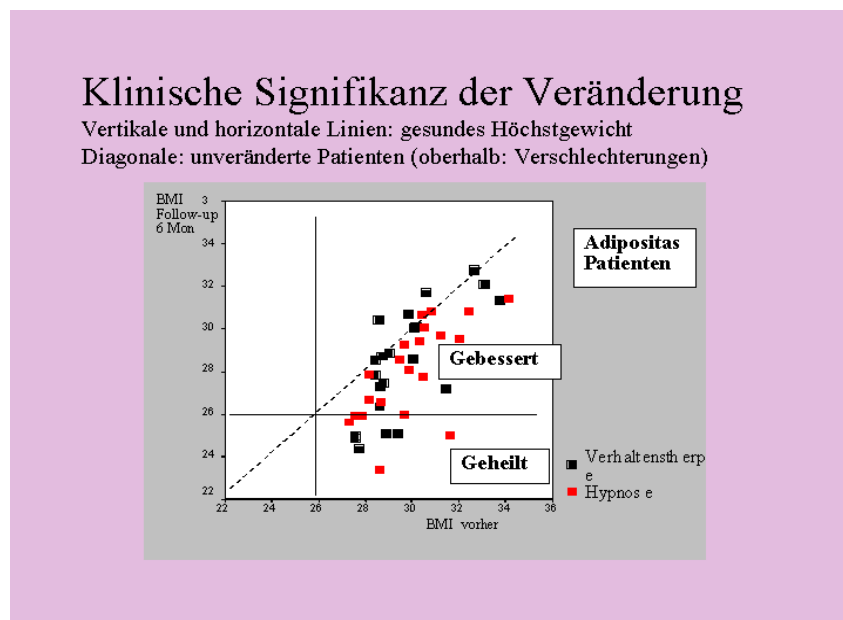


Abbildung 5: Darstellung der Patienten im Koordinatensystem von Prämessung (X-Achse) und Postmessung (Y-Achse) zur Verdeutlichung des Anteils der unveränderten (Diagonale), der verbesserten und geheilten Patienten (unterhalb der horizontalen Kriteriumslinie).

„File-drawer“-Problem: Damit im Zusammenhang ist das ebenfalls sattsam bekannte *„file-drawer“-Problem* zu sehen: Man kann nicht ausschließen, dass es zu jedem publizierten Ergebnis, das auf dem 5%-Niveau signifikant ist, mindestens 19 nicht publizierte Studien gleicher Art vorliegen, deren Ergebnisse nicht signifikant sind. Damit wäre die Zufallsquote von 5% bei dem publizierten Ergebnis trotz statistischer Signifikanz nicht überschritten. Das hat zwei mögliche Folgen: gefundene Ergebnisse bei anerkannten Therapieformen sind gar nicht wirklich wirksam, weil es zwanzig Mal so viele nicht signifikante gibt. Und umgekehrt: viele nicht signifikante Ergebnisse werden nicht publiziert, obwohl sie in einer Metaanalyse über alle Studien signifikante Resultate erbringen würden (was eher für die weniger intensiv beforschten Therapierichtungen zutreffen dürfte). Passt also eine Therapieform aus den oben genannten Gründen nicht ins aktuelle Wissenschaftsverständnis, dann werden mangels Forschung zu wenig Resultate produziert, als dass das nötige Signifikanzniveau öfter mal rein zufällig erreicht wird. Oder es werden nicht genügend insignifikante Ergebnisse veröffentlicht, um durch Aggregation Signifikanz zu erreichen.

Zusammenfassend lässt sich feststellen: Die statistisch abgesicherten Ergebnisse von Therapiestudien begrenzen sich oft auf die Betrachtung von Variablen, die gut operationalisierbar aber möglicherweise für das Erleben des Patienten irrelevant sind und die statistische Signifikanz sagt wenig über die klinische Bedeutsamkeit der Ergebnisse. Außerdem ist eine nicht widerlegte Nullhypothese kein Beweis der Unwirksamkeit und ein signifikantes Ergebnis ist bei massenhafter Forschungsaktivität und bevorzugter Publikation signifikanter Ergebnisse nicht aussagekräftig - genauso wenig wie wenige nicht signifikante Ergebnisse bei geringer Forschungsaktivität es sind. Das trifft auf die pharmakologische Forschung natürlich auch zu. Die Argumente sind alle seit langem bekannt (Bredenkamp 1972, Witte 1980), es scheint aber leicht in Vergessenheit zu geraten, wie dadurch nomothetische Aussagen in Psychologie relativiert werden.

Fazit

Die vorliegenden Metaanalysen von Kontrollgruppen-Studien attestieren der Psychotherapie einen sehr guten Wirkungsgrad, der durch zwei vielfach bestätigte Tatsachen gekennzeichnet ist. Erstens:

79-84% der Psychotherapie-Patienten (Effektstärken von 0,8 – 1,0) und
66% der Patienten in Placebo-Therapien (Effektstärke von 0,4)

geht es besser als dem Durchschnitt der unbehandelten Patienten. Und zweitens zeigen die Metaanalysen von Therapie-Vergleichen, dass Differenzen der Wirksamkeit zwischen den untersuchten Interventionsformen zufällig um Null verteilt sind (Wampold et al. 1997). Dass manchmal größere Differenzen nachgewiesen werden, liegt an der Verwendung reaktiver Masse (Smith, Glass & Miller 1980) oder daran, dass eine der verglichenen Psychotherapien inkompetent durchgeführt werden, weil der Therapeut von der einen Behandlung überzeugter ist als von der anderen (engl. als *allegiance*-Problem bekannt, vgl. Wampold, 2001).

Gibt es Unterschiede zwischen den Therapieformen ?

- H1: Unterschiede würden sich etwa wie die rote Kurve verteilen
- H0: Zufällige Unterschiede verteilen sich wie die schwarze Kurve

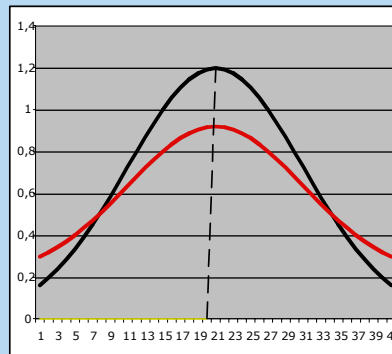


Abbildung 7: Abweichung von der Zufallsverteilung der Differenzen ist nicht signifikant: $\chi^2 = 241,2$ (df=276), d.h. Unterschiede zwischen Therapieformen in ihrer Wirksamkeit sind unbedeutend. (Wampold et.al. 1997).

Mit den Mitteln der traditionellen Therapieforschung lässt sich demnach zeigen, dass eine hohe Wirksamkeit der Psychotherapie vorliegt, die aber nicht spezifisch genannt werden kann. Darüber hinaus kann aber die traditionelle Therapieforschung methodisch als inadäquat betrachtet werden, sodass man zu dem paradoxen Ergebnis kommt: die mit den falschen Mitteln untersuchte Hypothese von der spezifischen Wirksamkeit der einzelnen Therapieverfahren erweist sich als falsch. Was würde wohl herauskommen, wenn man angemessene Untersuchungsmethoden verwendet? Und wie würden die aussehen?

Für die Weiterentwicklung der Psychotherapie werden heuristische Modelle benötigt, mithilfe derer die vorhandenen Erkenntnisse differenziert werden können. Die bisher gängige Heuristik in der Psychotherapieforschung ist die Auffassung von spezifischen und synergetischen Wirkungen der therapeutischen Maßnahmen (je mehr je besser). Der gegenwärtige empirische Status weist darauf hin, dass sehr unterschiedliche spezifische Interventionen gleich wirksam sind und vorläufig davon auszugehen ist, dass das Modell der allgemeinen Wirkfaktoren sensu Strupp oder Garfield oder das Kontextmodell sensu Frank genauso plausibel ist wie das der spezifischen Wirkungen. Beide Modelle ließen sich mit dem Kontrollgruppen-Design überprüfen, wenn man als Experimentalbedingungen neben spezifischen Wirkungen auch Kontexte oder allgemeine Therapiefaktoren wie therapeutische Bindung und Authentizität variiert. Dabei wären allerdings drei Dinge zu berücksichtigen:

- (a) Zufallsstichproben und randomisierte Gruppenzuweisung lassen sich nicht realisieren;

- (b) statistische Signifikanz sagt wenig über die klinische Signifikanz aus;
- (c) eine nicht widerlegte Nullhypothese beweist nicht, dass keine Unterschiede existieren und eine signifikanter Unterschied kann bei unbekannter Anzahl nichtsignifikanter (nicht publizierter) Ergebnisse immer noch zufällig sein (vgl. Tabelle 2).

Das *Repräsentativitätsproblem* (a) lässt sich lösen, wenn man den Anspruch der Generalisierung aufgibt. Wenn man die Aussage darauf beschränkt, dass die aktuelle Aufteilung der N Patienten im Experiment auf zwei (oder mehr) Behandlungsgruppen eine zufällige Realisation aller möglichen Aufteilungen dieser N Patienten auf die Behandlungsgruppen darstellt, lässt sich die Nullhypothese prüfen, dass eine andere Verteilung der Patienten bei ihnen nur zufällig davon abweichende Werte hervorgerufen hätte. Dazu kann man die Mittelwertsunterschiede zwischen den Gruppen mit den üblichen parametrischen Statistiken wie t-Test oder Varianzanalyse (bei kleinen Stichproben mit Fishers Randomisierungstest) überprüfen. Wird ein Unterschied signifikant, dann kann man ähnlich wie bei einem Einzelfalllexperiment sagen, dass ein Unterschied zwischen den Behandlungsmethoden für diese Gruppe von Patienten nachgewiesen wurde. Gewissermaßen eine Existenzaussage für eine untersuchte Gruppe nicht mehr für die Population, aus der sie eine Stichprobe sein soll (vgl. Glaser in Vorber.).

Das Problem der *klinischen Relevanz* (b) wird berücksichtigt, wenn man berichtet, wie viele Patienten geheilt wurden, was in Publikationen zunehmend mehr getan wird. Das *Signifikanzproblem* (c) wäre erträglich, wenn daran keine wirtschaftlichen Konsequenzen geknüpft sind wie die Einschränkung der Berufsausübung oder Erstattung der Behandlungskosten. So gewinnt das RCT-Design einen anderen Stellenwert: Es kann die Wirksamkeit einer Interventionsform vorschlagen, die in der Praxisevaluation (*Effectiveness*: Stufe VI der klinischen Prüfung) bestätigt werden müsste, Es kann nicht die Unwirksamkeit einer Methode nachweisen; und das Ergebnis gilt nur als Existenzaussage für die untersuchte Stichprobe. Man kann zwar davon ausgehen, dass unter ähnlichen Bedingungen ähnliches herauskommt – aber die Bedingungen und die Patienten in Instituten und der Praxis sind meist nicht ähnlich. Das RCT-Design kann daher als Forschungsinstrument dazu dienen, Anregungen für die Praxis zu geben. Da aber die Praxis aufgrund mangelnder Übertragbarkeit ganz anders aussehen kann als die Labortherapie, muss für die Akkreditierung von Psychotherapie die Praxis evaluiert werden. Das auch deshalb, um dem sonst unvermeidlichen Etikettenschwindel vorzubeugen, sich im Kassenantrag auf eine akkreditierte Therapieform zu beziehen und in der Praxis andere wirksame aber nicht akkreditierte Verfahren hinzu zu nehmen. Die Überprüfung von Wirtschaftlichkeit und Erfolg in kontrollierter Praxis und nicht die im RCT-Design gewonnene Information müssten daher darüber entscheiden, ob eine Behandlungsform als legitimiert gelten kann,.

Für die Forschungsförderung würde es bedeuten, dass die Praxis-Evaluation (Klinische Prüfung Stufe IV) gleichrangig mit der kontrollierten Therapie-Forschung behandelt wird und dass bei letzterer schulübergreifende Forscherteams und multizentrische Studien (wie etwa bei Elkin et al. 1989) die glaubwürdige Durchführung von Therapievergleichen garantieren, dass lange Behandlungszeiten (verg. Seligman 1995) und hinreichende Katamnesezeiträume vorgesehen werden, womit Wartelisten als unzumutbar entfallen. Auch die Randomisierung als Strategie müsste überdacht werden und gegebenenfalls durch Balancierung der Vergleichsgruppen bezüglich absehbarer Störvariablen oder durch cross-over-Designs ersetzt werden, und der Repräsentivitätsanspruch eingeschränkt wird. Die wissenschaftliche Anerkennung von Interventionstechniken sollte auf empirischer

Evidenz und zusätzlich einen vorläufigen Experten-Konsensus für schwer operationalisierbare Aspekte des Therapieprozesses beruhen. Die RCT-Forschung sollte nicht länger als Goldstandard gehätschelt werden.

Nachteile der RCT-Forschung	Vorteile der RCT-Forschung
Begrenzte Validität der Messung Irrepräsentative Stichproben Scheitern der Randomisierung Nicht berücksichtigte klinische Signifikanz Vernachlässigte Komorbidität Sterilisierung der Durchführung Nicht Umsetzbarkeit in die Praxis Unzumutbarkeit der Warteliste; Drop-out Problem Indifferenz der unwiderlegten Nullhypothese Filedrawer-Problem	Reliable Messung Anspruch auf (statistische) All-Aussagen über selbst seligierte Stichproben Eliminierte Störvariablen Statistische Signifikanz Störungsspezifität Standardisierung der Durchführung (Manual) Optimierung des Vorgehens

Tabelle 2: Vor- und Nachteile des Versuches RCT-Designs in der Psychotherapieforschung zu etablieren..

Das RCT-Design, das als Kuckucksei aus der Medikamentenforschung für die Therapieforschung übernommen wurde, kann hier nicht das leisten, was es verspricht. Denn eine medizinische Störung sieht im Labor so aus wie in der Praxis, die Anwendung der Medikation bleibt in beiden Situationen gleich und ein Verum unterscheidet sich äußerlich nicht vom Placebo. Das alles ist in der Psychotherapie nicht gegeben: Daran scheitert die Übertragung des RCT-Designs in der beschriebenen Art.

Aber die Kritik des medizinischen Modells der Psychotherapieforschung geht weiter. Während sich die Pharmaforschung auf die rationale Analyse begrenzen kann, reicht das für Erfassung und Heilung psychischer Probleme nicht aus. Die besinnungslose Reduktion auf das Messbare ist eine unzulässige Verkürzung der Kartografie menschlicher Probleme und die Rückbesinnung auf das Subjektive als den Ursprung von Erkenntnissen erscheint mir unvermeidbar.auf das Subjektive als den Ursprung von Erkenntnissen erscheint mir unvermeidbar.

Forderungen

- | | |
|--|---|
| <ul style="list-style-type: none"> • Forschungsförderung <ul style="list-style-type: none"> – Lange Behandlungszeiten – Hinreichende Katamnesezeiträume – Praxis-Evaluation (IV) – Schulübergreifende Forscherteams | <ul style="list-style-type: none"> • Wissenschaftliche Anerkennung <ul style="list-style-type: none"> – Empirische Evidenz – Experten-Konsens – Handlungsspielraum für die Praxis – Schulübergreifende Gremien |
|--|---|

Tabelle 3: Forderungen an die Therapieforschung.

Literatur

Bischof, N. (1999) Die Kraft der Mythen. Frankfurt, Fischer

Chambless, DL & Hollon, SD (1998) Defining empirical supported therapies. *Journal of Consulting and Clinical Psychology*, 66, 7-18

Comer, R (1995) *Klinische Psychologie*. Berlin: Spektrum

Curran D & Guttman, E (1949) *Psychological Medicine*. Edinburgh: Livingstone

Curran, D & Partridge, M (1963) *Psychological Medicine*. Edinburgh: Livingstone

Davison, GC, Neale JM & Hautzinger M (2002) *Klinische Psychologie*. Weinheim; Beltz

Ehlert, U (2003) *Kommentar zur Empfehlung des Wissenschaftlichen Beirats*

Elkin, I (1994) NIMH treatment of depression collaborative research program: Where we began and where we are. In A E Bergin & S L Garfield (Eds) *Handbook of psychotherapy and behavior change*. New York: Wiley

Elkin I, Shea M T, Watkins J T, Imber S D (1989) NIMH treatment of depression collaborative research program: general effectiveness of treatments. *Archives of General Psychiatry*, 46, 971-982

Fiedler, P (1998) *Persönlichkeitsstörungen*. Weinheim; Beltz

Gershon, MD (1998) *The second brain*. NY: Harper

Glaser, W (2004) *Nicht randomisierte Therapieexperimente (in Vorbereitung)*.

Goldfried, MR (1980) Towards the delineation of therapeutic change principles. *American Psychologist*, 35, 991-99

Goleman, D (1995) *Emotional Intelligence*. NY: Bantam

Imber S D, Elkin I, Watkins J T, Collins J E, Shea M T, Leber W R, Glass D R (1990) Modest specific effects among three treatments for depression. *Journal of Consulting and Clinical Psychology*, 58, 352-359

Jacobson N, Schmaling K B & Holzworth-Munroe A (1987) A component analysis of cognitive behavioral treatment of depression. *Journal of Counseling and Clinical Psychology*, 64, 295-304

Jacobson N S & Revenstorf, D (1988) Statistics for Assessing the Clinical Significance of Psychotherapy Techniques. *Issues, Problems and New Developments. Journal of Behavioral Assessment*, 10 (2), 133-145, 1988

- Lazarus, A (1979) Multimodale Verhaltenstherapie. Frankfurt: Fachbuchhandlung für Psychologie
- Margraf J. (1995) Verhaltenstherapie als Bestandteil der Psychotherapeutischen Versorgung. BdP-Gegenwart und Zukunft der Psychotherapie im Gesundheitswesen, Bonn: BdP
- Neumann, (1995) Ursprungsgeschichte des Bewußtseins. Frankfurt: Fischer
- O'Hanlon W H & Hexum A (1990) An uncommon casebook. The complete clinical work of Milton H. Erickson. New York: Norton
- Rossi, E.L. (1995-98) (Hrg.) Gesammelte Schriften von M.H. Erickson. Heidelberg: Auer
- Schlarb, A. (2004) Insomnie und Hypnose. Dissertation Psychol.Inst. Universität Tübingen
- Schulz, Chr & Weinmann U (2001) Hypnose bei Schlafstörungen. unveröffentlichte Diplomarbeit Psychol.Inst. Universität Tübingen
- Searle, JR (1993) Die Wiederentdeckung des Geistes. München: Artemis
- Smith M L, Glass, G & Miller T I (1980) Benefits of Psychotherapy. Baltimor: Johns Hopkins University Press
- Solomon, RC (2000) Gefühle und der Sinn des lebens. München: Zweitausendeins
- Wampold, B E (2001) The great Pschotherapy debate. London, Erlbaum
- Wampold B E, Mondin G W, Moody G, Stich F, Benson K J & Ahn H (1997) A metaanalysis of outcome studies. Comparing bona fide psychotherapies. Empirically „All must have prizes“. Psychological Bulletin, 122, 203-215
- Wilber, K (1987) Halbzeit der Evolution. Bern: Scherz
- WBT, Wissenschaftlicher Beirat Psychotherapie der Bundesrepublik (siehe homepage)
- Zohar D & Marshall I (2000) Spiritual Intelligence. New York: Bloomsberry