

158. Tschuschke V, Kächele H, Hölzer M (1994) Gibt es unterschiedlich ef-fektive Formen von Psychotherapie? Psychotherapeut 39: 281-297

## Gibt es unterschiedlich ef-fektive Formen von Psychotherapie?

Volker Tschuschke , Horst Kächele & Michael Hölzer

1. Einleitung
2. Ein historischer Überblick
  - 2.1. Eysenck und die Folgen: Zur Frage der Spontanremission und der generellen Wirksamkeit von Psychotherapie
  - 2.2. Zur Spezifität von psychotherapeutischen Wirkfaktoren
  - 2.3. Die Meta-Analyse als Methode zum Vergleich der Effektivität unterschiedlicher Psychotherapieformen
3. Die Berner Studie: Also doch unterschiedlich wirksame Psychotherapien?
4. Zurück zur Gegenwart: Zwei Studien über die Wirksamkeit analytischer Gruppentherapie
5. Zusammenfassung und Ausblick

## 1. Einleitung

Die Auseinandersetzungen um die Wirksamkeit von unterschiedlichen Formen von Psychotherapie haben in letzter Zeit wieder verstärkt Nahrung erhalten. Nachdem das Feld bereits früh durch die provokanten Thesen Eysenck's (1952) zur relativen Unwirksamkeit von psychotherapeutischen Behandlungen - im Vergleich zu den Remissionsraten bei "unbehandelten" Patienten - in eine letztlich produktive Unruhe gestürzt worden war (Grawe 1992), mehrten sich inzwischen Stimmen, die das von Luborsky 1975 formulierte 'Dodo bird-Verdict'<sup>1</sup> bezüglich vergleichbarer Effektstärken unterschiedlicher psychotherapeutischer Techniken zunehmend anzweifeln (Grawe 1992; Grawe et al. 1990a-d; Grawe et al. 1994; Shadish & Sweeney 1991; Shapiro & Shapiro 1982). Auch wird in diesem Zusammenhang immer wieder die Frage diskutiert, ob psychotherapeutische Wirkungen, ungeachtet der zugrundeliegenden Technik und eines präsumptiv zur Anwendung gelangenden therapeutischen Konzepts, lediglich auf 'gemeinsamen (kommunalen)' und weniger - oder überhaupt nicht - auf 'spezifischen Wirkfaktoren' beruhen (und deswegen - unabhängig davon welcher schultheoretischen Orientierung sie unterliegen - stets gleiche Effekte produzieren)<sup>2</sup>. Die Diskussion um mögliche Effektivitätsunterschiede verschiedener psychotherapeutischer Techniken erhält darüber hinaus zusätzliche Brisanz vor dem Hintergrund aktueller berufspoli-

---

<sup>1</sup>"all have won and all must have prizes"

<sup>2</sup>Dies war im übrigen die Ausgangsfrage des Vanderbilt I-Projekts in den späten 70er und frühen 80er Jahren, wo es um die Beantwortung der Frage ging, ob die 'unspezifischen' Beziehungsfaktoren entscheidend in der Psychotherapie seien - Psychotherapie also unspezifisch und somit in allen Formen von psychotherapeutischer Behandlung ein guter menschlicher Kontakt ausschlaggebend sei - oder ob es 'spezifische', auf der vom Therapeuten angewendeten Technik basierende, Wirkfaktoren gebe (Strupp u. Hadley 1979)? Die Vermutung einer eher unspezifischen Wirkung, weil professionelle und Laientherapeuten keine signifikanten Unterschiede in ihrer Effektivität zeigten, wurde später von Strupp weitestgehend revidiert, als die Versuchsanordnung in nicht ausreichendem Maße die Basis-Frage nach der 'Spezifität' beantworten konnte, da die Behandlungsdauer nicht ausreichte und College-Professoren als "Laientherapeuten" in kurzen Interventionen sicher gute Beziehungseffekte bei College-Studenten (als "Patienten" benutzt) erzielen konnten, was im Grunde im voraus zu vermuten war (Strupp 1993; Tschuschke 1990).

tischer Entwicklungen<sup>3</sup>). sowie angesichts knapper werdender öffentlich-sozialer Mittel und den damit verbundenen Anstrengungen politischer Entscheidungsträger in der psychotherapeutischen Versorgung anfallende Kosten umzuschichten. In Zeiten verstärkten Wettbewerbs bedeutet der Nachweis größerer Effektivität für die Vertreter des jeweiligen Verfahrens immer auch die Hoffnung auf Vorteile bei der Allocation verfügbarer Ressourcen.

Nach einem historischen Überblick über die wichtigsten Stationen der Effektivitätsdiskussion werden wir die Forschungslage zur differentiellen Wirksamkeit einiger psycho-therapeutischer Behandlungskonzepte und die diesbezüglichen aktuellen, teilweise leider auch polemischen, Stellungnahmen kritisch würdigen, um unter Rückgriff auf derzeit wissenschaftlich abgesichertes Wissen zu einer Versachlichung der Diskussion über vermeintlich nachweisbare Unterschiede psychotherapeutischen Handelns beizutragen. Wir werden zeigen, daß man seriöserweise derzeit nicht die Behauptung aufstellen kann, die eine oder andere Form von Psychotherapie sei generell effektiver als eine dritte. Um schon an dieser Stelle einem Mißverständnis vorzubeugen: Es mögen durchaus Unterschiede in verschiedenen Konzepten und Techniken im Hinblick auf die psychotherapeutische Veränderungspotenz bei eng umschreibbaren Störungen und Problemen gegeben sein. Die Annahme, daß 'spezifische' Wirkfaktoren in der einen oder anderen psychotherapeutischen Technik enthalten sind, ist angesichts der vorliegenden Ergebnisse der vergleichenden Psychotherapieforschung sehr plausibel. Nur, die bisherige Forschung konnte sie in der von verschiedenen Autoren derzeit unterstellten Ubiquität nicht nachweisen. Es steht außer Frage, dass weitere Forschung zu differentiellen Effekten eine *conditio sine qua non* für das Feld der Psychotherapie sein muß. Dies muß auch unser vorrangigstes Ziel bleiben, um in Übereinstimmung mit der von Grawe et al. 1994 formulierten Losung "von der Konfession zur Profession" fortzuschreiten. Aber angesichts einer derzeit noch völlig unzureichenden Faktenlage allgemeine Behauptungen über differentielle Wirkweisen unterschiedlicher Behandlungskonzepte aufzustellen, schafft mehr Verwirrung und Irreleitung als Klarheit und Sicherheit.

## **2. Ein historischer Überblick**

---

<sup>3</sup>siehe auch die Beratungen zum Psychotherapeutengesetz; Meyer et al. 1991

Empirische Forschung zur Wirksamkeit von Psychotherapie, d.h zu deren Ergebnissen und den diesen Ergebnissen zugrundeliegenden Therapieprozessen kann nach Shapiro (1990) rückblickend in drei Phasen unterteilt werden. Die erste Phase, beginnend mit einer 10-Jahres Katamnese des Berliner Psychoanalytischen Instituts (Fenichel, 1930), hatte ihren Höhepunkt in den Jahren 1950-1970. Sie suchte - vornehmlich zu Legitimationszwecken - Antworten auf die Frage, ob Psychotherapie überhaupt nützt. Die zweite Phase (1960-1980) fokussierte auf den Zusammenhang von Ergebnis und Prozeß: Welche therapeutischen Prozesse führen zu welchen Therapieergebnissen. Diese Form der Prozeß/Ergebnis Forschung ermöglichte auf der Basis korrelativ-statistischer Zusammenhänge Einblicke in die vielfältige Komplexität psychotherapeutischer Interventionen, der naturgemäß mit Hilfe vielfach stark reduktionistischer Ansätze schlecht beizukommen war. Die dritte, derzeit aktuelle Phase der empirischen Psychotherapieforschung trägt diesem Umstand Rechnung und intensiviert erneut die Untersuchung der prozessualen Vorgänge. Ihre leitende Frage lautet: "Welche mikroprozessualen Vorgänge konstituieren die klinisch konzipierten Makroprozesse?" (Kächele, 1992).

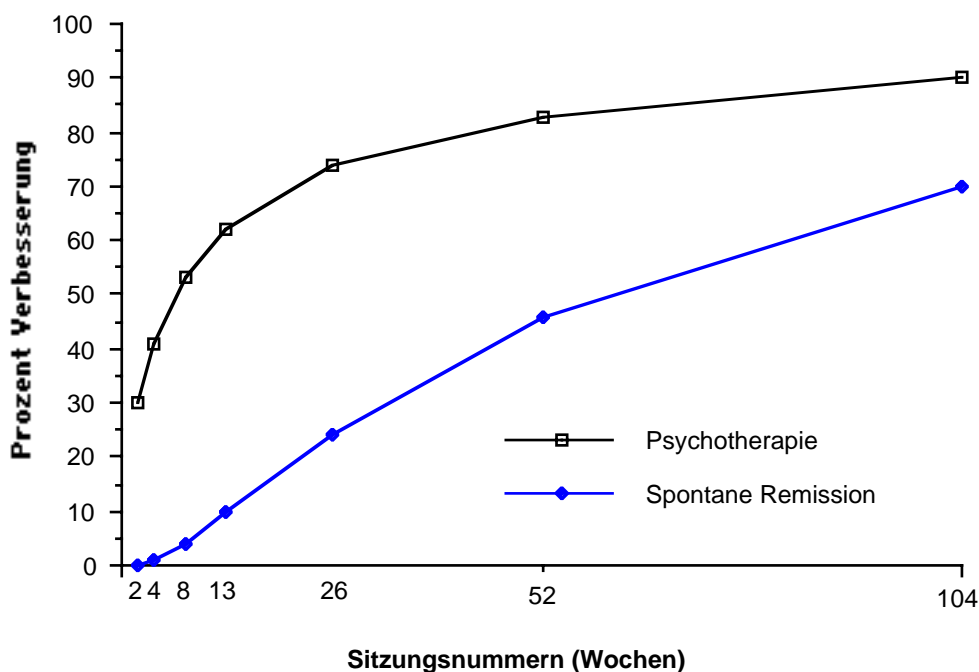
### **2.1.Eysenck und die Folgen: Zur Frage der Spontanremission und der generellen Wirksamkeit von Psychotherapie**

Eysenck's (1952) provokante These, daß Psychotherapie im Endeffekt nicht wirksamer sei als gar keine Therapie, markiert einen wichtigen Ausgangspunkt für die erste Forschungsphase. Sie basierte auf von Denker (1946) und Landis (1938) erhobenen Daten<sup>4</sup> sogenannter "psychoneurotischer" Patienten, die nicht spezifisch psychotherapiert, sondern lediglich "betreut" oder von Allgemein-ärzten behandelt worden waren. Er schloß aus den Verbesserungswerten von "unbehandelten" im Vergleich mit den "betreut/behandelten" Patienten, daß die sogenannten "spontanen Remissionsraten" unbehandelter Patienten keine Unterschiede zu Psychotherapie-Patienten erkennen ließen. Diese Schlußfolgerung ist in der Folgezeit heftigst kritisiert worden; u..a. erbrachten z.B. genaue Analysen der Originaldaten, daß Eysenck erhebliche Verfahrensfehler unterliefen, die seine Schlüsse in keiner Weise rechtfertigten (Dührssen 1962; Bergin 1971). In einer neuerlichen Überprüfung wurden Daten von 500 Patienten dieser ursprünglichen Stichprobe einer erst seit kurzen in der Psychotherapieforschung verwendeten statistischen Analyse unterzogen

---

<sup>4</sup>Bergin (1971) hat die Natur dieser "Daten" detailliert beschrieben.

und die Prozentzahlen der Verbesserungsrate pro Meßzeitpunkt mit Daten verglichen, die aus 2.431 aggregierten Patienten-Behandlungen gewonnen waren (Howard et al. 1986). Das Ergebnis ist frappierend eindeutig: Den unbehandelten Patienten der Eysenck'schen Stichprobe geht es erst nach zwei Jahren so gut wie Patienten aus psychotherapeutischen Behandlungen bereits nach ca. 15 Sitzungen (bei wöchentlich einer Sitzung, vgl. Abbildung 1). Oder anders ausgedrückt: Während nach ca. 26 Sitzungen Psychotherapie-Patienten im Schnitt symptomatische Verbesserungsrate um die 75% zeigen, weisen unbehandelte Patienten Raten von weniger als 24% Verbesserung auf (McNeilly u. Howard 1991).



**Abbildung 1:** Vergleich von Dosis-Effekt-Beziehungen für Psychotherapien und "Spontane Remissionen" (nach McNeilly u. Howard, 1991)

Auch wenn diese Ergebnisse nichts an Deutlichkeit zu wünschen übrig ließen, so operierten sie noch mit der Eysenckschen Grundannahme, daß "spontane Remission" irgendwann unbehandelten Personen den Leidensdruck letztlich vollständig nehmen werde. Daß dem keineswegs so ist, läßt sich mittlerweile anhand zahlreicher Studien belegen: Spontane Verbesserungen bei behandlungsbedürftigen psychischen Störungen treten ohne psychotherapeutische Interventionen offenbar in entscheidend geringerem Ausmaß auf, als dies von

Eysenck unterstellt wurde (Lambert et al. 1986; Landman & Dawes 1982; Shapiro & Shapiro 1982; Smith et al. 1980). Da Psychotherapiepatienten erst nach einer durchschnittlichen Symptombdauer von 7 Jahren (s. Meyer et al. 1991) fachpsychotherapeutische Hilfe in Anspruch nehmen, konnte Grawe (1992) mit freundlicher Ironie feststellen:

"Würden sich, wie Eysenck es behauptete, in einem Zeitraum von fünf Jahren bei über neunzig Prozent der Patienten die Beschwerden legen, so hätten wir es bei denjenigen Patienten, die ein Psychotherapeut schließlich zu sehen bekommt, mit dem prognostisch ungünstigen Rest von weniger als zehn Prozent zu tun, der sich als resistent gegen spontane Remissionen erwiesen hat. Bei dieser Restgruppe...wäre jede Besserungsrate über Null als Erfolg anzusehen. Besserungsraten in der Größenordnung von 44-64%, wie Eysenck sie aus seinen Zahlen für psychoanalytische und eklektische Psychotherapie errechnete, würden geradezu als fabelhaft gelten müssen und der Psychotherapie ein glänzendes Zeugnis ausstellen....Da die Effektstärken bei Smith et al. aus der Differenz zur jeweiligen Kontrollgruppe errechnet wurden, sind aus ihrem Wert eventuelle spontane Remissionen bereits herausgefiltert, so daß Effektstärken von diesem Ausmaß gewissermaßen als Netto-Effekt von Psychotherapie betrachtet werden können. Damit können Eysenck's generelle Zweifel an der Wirksamkeit von Psychotherapie als erledigt angesehen werden." (Grawe, 1992, S. 134 f)

Grawe selbst konnte inzwischen "in einer direkteren Abschätzung des Ausmaßes spontaner Remission bei Psychotherapiepatienten" (1992, S. 134) anhand der Daten aus 111 Studien mit nachgewiesenen unbehandelten Kontrollgruppen im Versuchsplan feststellen, daß die ermittelten wahren Effektstärken für Nicht-Behandlung sich deutlich von den Effekten psychotherapeutischer Interventionen unterscheiden. Für diese "spontanen Remissionseffekte" konnten nämlich Effekt-Stärken lediglich im Mittel um Null oder leicht darüber ermittelt werden<sup>5</sup>, während sich z.B. in der viel-zitierten Meta-Analyse von Smith et al. (1980), die auf der Grundlage von Daten aus 475 Psychotherapiestudien durchgeführt wurde, Effektstärken für Psychotherapie in Höhe von 0.85 fanden. Auch Luborsky et al (1988) berichteten durchschnittli-

---

<sup>5</sup>80 von 111 Maße der Verteilung streuen im Bereich von -0.16 bis 0.25

che Effektstärken über verschiedene Outcome-Studien und -Maße hinweg von ca. 0.80, also in vergleichbarer Größenordnung.

Wir sollten die Idee der Spontanremission endgültig begraben, wie dies Grawes Ausführungen eindrücklich nahelegen und verschiedene andere Arbeiten gleichsinnig unterstreichen (Bergin 1971; Schepank 1990). Wenn auch ganz allgemein festgestellt werden kann, daß nicht jeder Hilfesuchende von psychologischen bzw. psychotherapeutischen Behandlungen notwendigerweise profitiert, so gibt es heute dennoch keinen berechtigten Zweifel mehr an der generellen Wirksamkeit von Psychotherapie (Lambert et al. 1986).

Als Hinweis auf die Persistenz psychischer Probleme bei nicht-erfolgreicher Behandlung mögen die Ergebnisse eigener Untersuchungen dienen<sup>6</sup> (Tschuschke 1993, 1994).

Als Quintessenz der kreuzvalidierten Ergebnisse einer Untersuchung an einer eher kleinen Patientenstichprobe (N=16) konnte festgehalten werden, daß Patienten, die im "Hoch-Dosis-Therapie-Milieu" einer stationären Psychotherapie-Klinik nicht profitieren, sich auch im längeren Zeitraum nach der stationären Therapie ohne zusätzliche Therapie nicht verbessern trotz weiterer ambulanter psychodynamischer oder verhaltenstherapeutischer Anschlußtherapien (Tschuschke, 1994). Nur zwei Patienten erreichten in den 18 Monaten nach Beendigung der stationären Therapie durch anschließende analytische Einzelbehandlung einen klaren therapeutischen Zugeschwin. Bei diesen Patienten handelte es sich aber um bereits in der stationären Behandlung mäßig erfolgreiche Patienten, bei denen bereits während der stationären Behandlung entscheidende Veränderungen auf den Weg gebracht wurden, wie eine Analyse der Wirkfaktoren im Prozeß zeigte (Tschuschke & Dies, 1994).

Auch unter Verweis auf methodologische Schwächen meta-analytischer Studien (s. weiter unten), kann das generelle Ergebnis, daß sich psychologisch-psychotherapeutische Behandlungen ganz allgemein als effektiv erwiesen haben, "nicht wegerklärt" werden. Zahllose Untersuchungen haben unter

---

<sup>6</sup>Die erzielten günstigen therapeutischen Effekte bei 16 nach Kriterien des DSM-III-R schwer neurotischen, psychosomatischen und persönlichkeitsgestörten Patienten aufgrund von hochintensiver Langzeitbehandlung (fünf Gruppentherapie-Sitzungen pro Woche über sechs Monate) in stationären analytischen Gruppenpsychotherapien konnten über zwei katamnestische Nachuntersuchungen (nach 12 und 18 Monaten) kontrolliert werden wie auch die Persistenz therapeutischen Mißerfolges bei nicht erfolgreichen Patienten (Evaluierungen auf der Persönlichkeitsebene, der Symptom-ebene, persönlichen Therapiezielen, strukturellen Einschätzungen von behandelnden Therapeuten und - getrennt - unabhängigen Kliniker).

Verwendung von Kontrollgruppen-Versuchsplänen im Vergleich zu Behandlungsbedingungen therapeutische Effekte nachgewiesen, die auf die Überlegenheit psychotherapeutischer Interventionen verweisen.

"Noch immer gibt es da und dort Stimmen, die das Unternehmen Psychotherapie zur Gänze in Frage stellen. Die inzwischen verfügbare Befundlage verweist dieses Kritisieren in den Bereich des Ungebildetseins" (Kächele & Kordy 1992, S. 520).

## **2.2. Zur Spezifität von psychotherapeutischen Wirkfaktoren**

Die im ganzen gesehen ungeklärte Situation, ob nun verschiedene Behandlungsansätze gleiche Effekte hervorbringen oder nicht berührt die Frage, ob nicht doch allen psychotherapeutischen Ansätzen und Techniken eher gemeinsame Wirkfaktoren immanent sind? "Es besteht Konsens, dass grob gesehen- mit der Betonung auf "grob" - ganz verschiedene psychotherapeutische Interventionssysteme im kontrollierten Vergleich keine klaren Sieger und Verlierer erkennen lassen"(1990). Mit diesen Worten faßt A.E. Meyer die Ergebnisse zahlreicher Studien der vergleichenden Psychotherapieforschung zusammen. Er "beglückwünscht" Rosenzweig dafür daß, "er (a) dieses Ergebnis bereits 1936 erspürte, und weil er (b) dafür das salomonische Urteil des Dodo-bird aus "Alice in Wonderland" als surrealistische Kurzformel heranzog: "Everybody has won and all must have prizes". Ein wiederholt vorgetragenes Argument zur Begründung der scheinbar gleichartigen Effektivität von verschiedenen Psychotherapierichtungen fußt auf der Beobachtung sogenannter "unspezifischer", später von Meyer als "kommunale" Faktoren bezeichnet. In seiner Aufzählung kommunaler Faktoren findet sich u.a. die "Mobilisierung von Hoffnung...Anwendung einer Theorie, wie Heilung erzielt werden kann, eines oder mehrerer Verfahren, welche Heilung erzielen... das Angebot einer Helferbeziehung...eine Klärung oder Neudefinition von Problemen... die Suche nach konstruktiveren Problemlösungsmöglichkeiten" (1990). Ein entscheidendes Problem bei der Beforschung der Spezifität therapeutischer Wirkfaktoren benennen Lambert et al.:

"Clearly, the challenge to sort out unique from common factors is difficult, and it has not yet been accomplished despite the enthusiasm generated by exponents of therapeutic schools. Indeed, this separation could be infeasible be-



cause common factors may activate or mediate specific ones." (Lambert et al. 1986, S. 178).

Ob die Wirkung von Psychotherapie auf spezifische, d.h. auf der vom Therapeuten angewendeten Technik basierende Wirkfaktoren oder eher 'unspezifischen' Beziehungsfaktoren - Psychotherapie also unspezifisch und somit in allen Formen von psychotherapeutischer Behandlung ein guter menschlicher Kontakt ausschlaggebend sei - zurückzuführen ist, war auch die Ausgangsfrage des Vanderbilt I-Projekts in den späten 70er und frühen 80er Jahren (Strupp u. Hadley 1979). Die Vermutung einer eher unspezifischen Wirkung, weil professionelle und Laientherapeuten zunächst keine signifikanten Unterschiede in ihrer Effektivität zeigten, wurde später von Strupp weitestgehend revidiert, da die gewählte Versuchsanordnung in nicht ausreichendem Maße auf die Basis-Frage nach der 'Spezifität' bezogen war: Die Behandlungsdauer erwies sich als keineswegs ausreichend und im Grunde war von vorneherein zu vermuten gewesen, daß College-Professoren (die als "Laientherapeuten" bzw. als eine Art Placebo-Gruppe fungierten) bei College-Studenten (= die Patienten des Vanderbilt-1 Projekts) in kurzen Interventionen gute Beziehungseffekte erzielen würden (Strupp, 1993; Tschuschke 1990)<sup>7</sup>.

Auch wenn sich bei späteren Vergleichen von psychotherapeutischen Interventionen mit sogenannten "Placebo-Gruppen" klare Unterschiede zugunsten der Psychotherapie ergaben, macht der Ansatz der Placebo-Kontrolle in der Psychotherapieforschung eigentlich wenig Sinn, beruhen doch beide - Behandlung und Placebo - als menschliche Zuwendung auf psychologischen Mechanismen (Lambert et al. 1986).

"...Placebo-Psychotherapie ist ein hölzernes Eisen. Psychotherapie wird entweder als hilfreich erlebt und ist dann kein Schein, sondern Hilfe, oder sie ist unwirksam. Dann ist sie genau dieses. In der Pharmakologie lassen sich dagegen Scheinmedikamente herstellen, welche mit bloßen Sinnen nicht vom Verum zu unterscheiden sind." (Meyer 1990, S. 180).

---

<sup>7</sup>Auch Hans Strupp teilt unsere Auffassung, dass College Professoren diese psychotherapeutische Leistung nicht über längere Zeit für viele Klienten erbringen würden.

Bereits der Begriff "Placebo" wirft Probleme auf. Weitgehend wird unter "Placebo" ein unspezifischer Faktor verstanden, der in jeder Form menschlicher Begegnung enthalten sei, somit auch in der psychotherapeutischen. Die Ergebnisse der vergleichenden Effektivitäts-Forschung zeigen denn auch, daß Patienten in Placebo-Gruppen eindeutig mehr profitierten als Patienten in der Kontroll- bzw. Wartegruppe. Wie die Placebo-Forschung zeigt, muß von gemeinsamen bzw. unspezifischen Faktoren ausgegangen werden (wie z.B. Resonanz, Wärme, Interesse, Zuhören, Entspannen etc.), die Effekt-Stärken von immerhin 0.42 bis 0.58 generierten (Andrews & Harvey 1981; Landman & Dawes 1982; Prioleau et al. 1983; Smith et al. 1980). Es läßt sich aber auf jeden Fall festhalten, daß Psychotherapie offenbar mehr als nur menschliche Zuwendung zu bieten hat, denn die Effektstärken von Psychotherapie liegen im Durchschnitt höher als die von Placebo-Behandlungen wie Dawes (1983, zit.n. Meyer 1990) und Prioleau et al. (1983) belegten.

Die bisher weitestgehend vergebliche Suche nach spezifischen Wirkfaktoren bei einer spezifischen Psychotherapieform ist keinesfalls ein Argument gegen die Effektivität von Psychotherapie. Die nachgewiesenermaßen im Durchschnitt regelmäßig höheren Effekt-Stärken von Psychotherapie verweisen hingegen auf eine weitergehende therapeutische Potenz, die über die Placebo-Effekte hinausgehen und somit auf - außer den gemeinsamen Wirkfaktoren - zusätzliche Wirkkomponenten hindeuten, die vermutlich dem Bereich der spezifischen Wirkungen zuzuordnen sind, also in der spezifischen therapeutischen Interventionstechnik unter Zugrundelegung spezifischer therapeutischer Konzepte und Theorien gesucht werden müßten.

Meyer und seine Hamburger Arbeitsgruppe haben mit Hilfe eines 154 verschiedene Therapeuteninterventionen umfassenden Katalogs 50 Sitzungen von 12 Gesprächspsychotherapeuten und 13 psychoanalytisch orientierten Therapeuten untersucht. "153 von unseren 154 Kategorien waren kommunal. Die einzige Ausnahme betraf "Systematische Bearbeitung von Widerstand", welche nur bei Analytikern vorkam"...Psychoanalytiker und personzentrierte Therapeuten verwenden fast ausschließlich kommunale Elemente" (Meyer, 1990). Der Unterschied der beiden untersuchten Therapieformen ergab sich einaml aus der "Dosis", d.h. der relativen Häufigkeit bestimmter Interventionen, zum anderen aus der Korrelation dieser Häufigkeiten mit dem Ergebnis der Behandlung. Positiv korreliert mit der Effizienz der Behandlung waren

bei den Gesprächstherapeuten Interventionen wie "Gefühle erwähnen", "Verhalten zentral erwähnt", "Wünsche erwähnt", bei den Psychoanalytikern "systematische Annäherung an biographische Ursprünge" und "Durcharbeiten". Psychotherapeuten sind also alles andere als "placebologists" (Lambert et al. 1986).

Folgt man den Ergebnissen Meyers läßt sich folgern, daß Spezifität von Psychotherapie nicht über die Einzigartigkeit der jeweils realisierten Interventionen sondern vermutlich v.a. durch eine je spezifische Kombination von mehr oder weniger ubiquitären Interventionen hergestellt wird.

### **2.3. Die Metaanalyse als Methode zum Vergleich der Effektivität unterschiedlicher Psychotherapieformen**

Die Einführung der Technik der Meta-Analyse in die Psychotherapieforschung (Rosenthal 1983) läßt sich in ihrer Bedeutung nicht hoch genug veranschlagen. Auch wenn die Methodik sowohl heftigste Kritik als auch überwältigende Belobigung auf sich gezogen hat, ihr Erscheinen und ihre Ergebnisse versetzten das Feld jedenfalls (wieder) in jene Unruhe, die schon einmal zu erheblichen Fortschritten geführt hatte. Von einer - im positiven Sinne - "Revolution" in der Psychotherapieergebnisforschung bis hin zur Verdammung der Methode durch Eysenck als "an exercise in mega-silliness" reichen die Urteile. Mit der Methodik steht in den Augen ihrer Protagonisten ein wissenschaftlich objektives Instrumentarium zur Verfügung, da es gestatte, subjektive Bewertungen und Fehleinschätzungen der Wirkung psychologischer Interventionen zu überwinden und zu einer nüchternen Betrachtung real gegebener Fakten zu gelangen. Der Terminus "Meta-Analyse" ist neu, jedoch nicht das Prinzip der "Überblicks-Methodologie" ("survey methodology").

Ein so heterogenes Feld wie die Psychotherapie, mit unterschiedlichsten psychotherapeutischen Auffassungen und Techniken, mit seinen Angeboten und Praktiken in höchst verschiedenen Settings mit einem extrem unterschiedlichen Klientel generiert zwangsläufig ein maximal breites Spektrum an Ergebnissen bezüglich der erzielten Effekte. Da der Bereich des Meinbaren unendlich ist und die Diskussionen klinischer Fallbeschreibungen häufig genug beliebig und unverbindlich, stellte sich in der Psychotherapie von jeher das Problem einer nachvollziehbaren Objektivierung und Kumulierung von gesichertem Wissen (Lösel 1987). Die Meta-Analyse als statistische Prozedur zur

vergleichsweise objektiven Einschätzung bzw. als "Konzept der Sekundäranalyse"

"...versucht eine Synthese von empirischen Forschungsarbeiten durch Integration und Kumulation statistischer Kennwerte, die in den einzelnen Studien berichtet werden, wie z.B. p-Werte, Mittelwerte, und deren Differenzen, Varianzen, bzw. Standardabweichungen, Korrelationskoeffizienten, t-, F-, chi-Quadratwerte usw. Es wird also nicht, wie bei anderen Formen der Sekundäranalyse, auf die ursprünglichen Rohdaten zurückgegriffen." (Wittmann & Matt 1986, S. 20).

Im Kern stellt die meta-analytische Methodik objektive Kriterien bereit, die eine Zusammenfassung, Synthese und Integration empirischer Ergebnisse empirischer Studien erlaubt. Die Methodik selber trägt wiederum zu spezifischer formulierten Designs bei Untersuchungen bei. So erwies sich bei der Anwendung der Technik die Notwendigkeit einer präzisen Angabe therapeutischer Technik und Ziele, von mehr Vergleichbarkeit und Standardisierung einzelner Untersuchungs-Designs und - Methoden, um zu einer höheren Übereinstimmung bezüglich der Bewertungen der Qualität von Studien zu gelangen. Es gibt eine ganze Reihe von verschiedenen Techniken, die in ihrer Komplexität und Anwendbarkeit erheblich variieren und alle unter dem Terminus "Meta-Analyse-Technik" zusammengefaßt werden können.

Unterschiedliche meta-analytische Prozeduren werden verwendet. Man kann die Signifikanzen (p-Werte) als grundlegend betrachten (Rosenthal 1983) oder man verwendet Cohen's Effekt-Stärke-Berechnung<sup>8</sup>, oder man benützt zur Effektstärkeberechnung die Korrelationskoeffizienten @@@@ welche ????? (Fricke u. Treinies 1985; Lösel 1987; Wittmann u. Matt 1986). Natürlich gibt es zahlreiche methodisch-statistische Feinheiten über die man im Kreise der "Meta-Analyse Gourmets" ausgiebig diskutieren kann, wie z.B. die differentiellen Gewichtungen einzelner in die Analyse eingehender Parameter, um den unterschiedlichen Qualitäten der einzelnen in die Analyse eingehenden Studien besser gerecht werden zu können.

---

<sup>8</sup>d-Koeffizient: *Mittelwert Postmessung der Therapiegruppe minus Mittelwert Prämessung der Therapiegruppe bzw. Postmessung der Kontrollgruppe, dividiert durch die Streuung der Postmessung der Kontrollgruppe* dieses Verfahren wurde bei der Meta-Analyse von Glass und Mitarbeitern durchgeführt (Smith et al. 1980).

Meta-Analysen gestatten die Berechnungen von sogenannten "Effekt-Stärken" für den Erfolg der verschiedenen Behandlungsformen. Damit werden relative Zugewinne durch psychologisch/psychotherapeutische Maßnahmen meßbar und untereinander vergleichbar. Es handelt sich also um eine statistische "Krücke" für den Nachweis eigentlich qualitativer Effekte<sup>9</sup>.

Die vergleichende Therapieforschung mit Hilfe meta-analytischer Ansätze erbrachte in den siebziger Jahren im Endeffekt keine überzeugenden Befunde zur differentiellen Wirksamkeit verschiedener psycho-therapeutischer Methoden (Luborsky et al. 1975; Sloane et al. 1975). Auch die maßstabsetzende Arbeit von Smith, Glass und Miller (1980) rief zahlreiche Kritiker auf den Plan. Um die generellen und weitreichenden Schlußfolgerungen vor allem dieser Studie auf ihre Stichhaltigkeit zu überprüfen, wurden zahlreiche Re-Analysen durchgeführt, zum großen Teil mit einem fraktionierten Sample der ursprünglich in die Analyse eingegangenen Studien (Landman & Dawes 1982; Shapiro & Shapiro 1982). Die Smith et al.-Studie wurde in ihren prinzipiellen Schlußfolgerungen und in den ermittelten Effekt-Stärken sowohl bestätigt (Landman & Dawes 1982) als auch. modifiziert: bestätigt wurde das Effektstärkemaß psychotherapeutischer Intervention generell; zugleich wurde die Überlegenheit behavioraler gegenüber dynamisch/humanistischen Methoden (Shapiro & Shapiro 1982) in dieser Datenbasis demonstriert. Diese Re-Analysen wurden lt. Grawe im Lichte der Frage durchgeführt

"..ob es der Verhaltenstherapie gelingen würde, ihre vorweg in Anspruch genommene (Wolpe 1958; Eysenck & Rachmann 1968; Ullmann & Krasner 1969) Überlegenheit stichhaltig zu beweisen." (Grawe 1992, S. 142).

Deshalb richtete sich die Kritik an der Smith et al.-Studie (1980) u.a. auf die "Subjektivität der Ergebnis-Maße"<sup>10</sup>, sowie auf die willkürliche Auswahl von Patienten. Auch Wilson & Rachman (1983) bemängeln den Ausschluß vieler relevanter Verhaltenstherapie-Studien durch Smith et al.. Dagegen werden Shapiro & Shapiro (1982), von Wilson & Rachman für ihre im Vergleich zu Smith et al. veränderten Einschlußkriterien gelobt. Shapiro & Shapiro (1982)

---

<sup>9</sup>Weiterführende Details zur Technik der Meta-Analyse sollen an dieser Stelle nicht erörtert werden und können der einschlägigen Literatur entnommen werden (Fricke u. Treinies 1985; Rosenthal 1983).

<sup>10</sup> wegen deren hoher Korrelation mit der Effekt-Stärke

kommen denn auch zu anderen Schlußfolgerungen als Smith et al., nämlich, daß behaviorale Methoden mit einer durchschnittlichen Effekt-Stärke von 1.08 den dynamisch/humanistischen Methoden mit nur einer Effekt-Stärke von durchschnittlich 0.40 eindeutig überlegen seien.

Eine unseres Erachtens zentrale, klinisch bedeutsame Kritik an der Studie von Shapiro & Shapiro indes läßt sich analog derjenigen für die Smith et al.-Studie mit dem Hinweis auf die mangelnde klinische Repräsentativität der ausgewerteten Studien formulieren:

"As in the Smith et al. study, however, these authors (Shapiro & Shapiro 1982; die Verfasser) also attributed the larger effects sizes to strong biases in the behavioral and cognitive literature toward analogue studies, mild cases, and highly reactive criteria. They stated that the treatments and cases studied were unrepresentative of clinical practice but very representative of the simple experiments on these techniques that are frequently conducted in university settings." (Lambert et al. 1986, S. 166f.).

Gewichtig wiegt u.E. auch der Vorwurf an Shapiro & Shapiro, daß sie dynamische und humanistische Behandlungen in einen Topf warfen<sup>11</sup>, während die behavioralen Studien in zehn unterschiedliche Techniken aufgegliedert wurden<sup>12</sup>. Tatsächlich stellen die ursprünglichen Gruppierungs-Kriterien bereits die Weichen für spätere Meta-Analyse-Ergebnisse; vor allem stellen sie wegen ihrer Uneinheitlichkeit eine fundamentale Quelle späterer irrtümlicher Schlußfolgerungen dar.

"Shapiro & Shapiro grouped Gestalt with psychodynamic and humanistic therapies, but Smith & Glass (1977) had Gestalt among the behavioral therapies. Andrews & Harvey (1981) did their own restricting by separating client-centered from the psychodynamic and cognitive therapies, and placing it in a third superclass entitled "developmental". It is therefore impossible to compare the conclusions reached by these different meta-analyses." (Wilson & Rachman 1983, S. 60).

---

<sup>11</sup> darunter werden solche unterschiedlichen Ansätze gefaßt wie psychodynamische, klientenzentrierte, Encounter- und Gestalttherapien; trotzdem umfasst die von ihnen ausgewertete Stichprobe für diesen Bereich nur 16 Studien.

<sup>12</sup> insgesamt 172 Studien.

Auf diese Weise widerfährt weder den humanistischen noch den psychodynamischen Behandlungsverfahren Gerechtigkeit. Aus den Gegenüberstellungen von Meta-Analysen zweier "Therapie-Klassen" ("dynamic/humanistic" and "behavioral/cognitive" therapy classes") von Shapiro (1985) und dem von Grawe (1992) vorgenommenen Fazit ist jedenfalls aufgrund der genannten unzureichenden Klassifizierungen kein Rückschluß auf die Wirksamkeit bestimmter Therapieformen zu ziehen.

Damit ist ein besonders problematischer Punkt von Meta-Analysen angesprochen: Jeder Untersucher erhält aufgrund seiner Auswahlstrategie jeweils sein individuelles Sample an empirischen Effektivitäts-Studien, das mit anderen nicht mehr vergleichbar ist. Vor allem durch die Zusammenfassung von empirischen Studien zu Behandlungsklassen (vermeintlich homogenen therapeutischen Konzepten bzw. der Schulzugehörigkeit entsprechend) entstehen unübersehbare Fehlermöglichkeiten, denen allenfalls durch empirisch geprüfte Ähnlichkeiten im Behandlungsgeschehen begegnet werden kann (Lösel, 1987). Jeder Untersucher findet plausible Gründe, warum nur seine Zusammenstellung Berechtigung hat und die anderer Untersucher nicht. Und so kommt er/sie auch dazu, seine/ihre (manchmal sogar von ihm/ihr entwickelte, siehe Grawe et al., 1990) Form von Psychotherapie als die effektivste zu bezeichnen (Eckert, 1993). Hier eröffnet sich ein weites Betätigungsfeld für Apologeten der einzelnen therapeutischen Schulen, die Eier zu finden, die man finden will. Diese Kritik betrifft grundsätzlich jede Art von Review (Lösel, 1987); im Falle der vergleichenden Effektivitätsforschung innerhalb der Psychotherapie schaden aber falsche und einseitige Schlußfolgerungen erheblich. Was not täte, wären explizite und allgemein akzeptierte Klassifikationsschemata für die Durchführung meta-analytischer Studien sowie vor allem explizitere Beschreibungen und Kennzeichnungen theoretisch-konzeptueller Behandlungsverfahren in den Evaluationsstudien selbst.

Weitere fundierte Kritik bezieht sich auf die methodologisch sehr heterogenen Studien, die in Meta-Analysen eingehen (Shadish & Sweeney, 1991; Wilson & Rachman, 1983). Die Probleme der bisherigen meta-analytischen Studien lassen sich im Überblick zusammenfassen (u.a. Crits-Christoph et al. 1991; Lambert et al. 1986; Lösel 1987; Orlinsky u. Howard 1986; Shadish u. Sweeney 1991; Wilson u. Rachman 1983):

- 1) Unzureichende Daten- und mangelnde statistische Qualität der zugrundeliegenden empirischen Studien.
- 2) Selektivität der einbezogenen Studien.
- 3) konzeptuell unklare bzw. zweifelhafte therapeutische Aktivität, auch unter dem "Uniformitäts-" und "Inkommensurabilitätsproblem" (Apfel-Birnen-Problem) diskutiert;
- 4) Einbeziehung nicht klinischer Stichproben, häufig verwendete Klientel: Studenten mit leichteren Problemen und dann entsprechend nur kurzen Behandlungszeiten.
- 5) Die Kontrolle von Therapeuten-Kompetenz und -Erfahrung wurde fast vollständig in den Studien vernachlässigt, dabei gibt es zahllose Untersuchungen, die nachweisen, daß der Therapeuten-Faktor einer der entscheidendsten in der Psychotherapie überhaupt ist.
- 6) hoch reagible Ergebnis-Instrumente begünstigen behaviorale bzw. verhaltenstherapeutische Studien gegenüber anderen Behandlungsmethoden, da es in der Natur der Sache liegt, daß Verhaltenstherapeuten ihre Behandlung eher auf spezifische, sehr eng gefaßte Ziele "zuschneiden" ("tailored targets"). Vor allem auch spielen die eingesetzten Ergebnis-Maße eine kritische Rolle in differentieller Bewertung<sup>13</sup>.
- 7) nicht ausreichende Berücksichtigung tatsächlicher psychotherapeutischer Prozesse: Es ist zweifellos ein ganz erhebliches Handicap bisheriger Untersuchungen zur vergleichenden Effektivität, daß nicht abgesichert wurde, daß die untersuchten Therapien auch in der unterstellten Art und Weise erfolgten, also dem verwendeten Konzept entsprachen, denn dies war ja theoretisch jeweils die Voraussetzung für die später angestellten Therapievergleiche. Vergleich-

---

<sup>13</sup>Wie Shadish & Sweeney (1991) auch ausführen, beeinflußt dieser Aspekt behaviorale Therapie mehr als nicht-behaviorale. Wenn man annahme, daß Psychotherapie ein Null-Summen-Spiel sei, dann seien therapeutische Ergebnisse proportional zu Therapie-Inputs und die Anzahl an inputs sei endlich. Je mehr inputs spezifischen Zielen zugeordnet würden, desto weniger inputs können anderen Zielen zugeordnet werden. Entsprechend änderten sich die ersteren (spezifischen) mehr und die letztgenannten weniger.



chenden Effektivitäts-Studien sind von ihrer Anlage her nicht geeignet, Aussagen über psychotherapeutische Prozesse zu machen, insofern kann auch keine Aussage darüber gemacht werden, wie die am Ende ähnlich aussehenden Resultate zustande gekommen sind: Dies kann an gleichen Prozessen, es kann aber auch an unzureichenden Ergebnis-Analysen (Meta-Analysen) liegen. (Stichwort: manual-geleitete Therapien, prozessuale Psychotherapieforschung).

Meta-Analysen über Psychotherapien sagen aufgrund der genannten, insgesamt schwerwiegenden, Probleme lediglich etwas über den generellen Effekte von Psychotherapieformen aus. Sie sagen nichts über Moderator-Variablen, von denen es zahllose und gewichtige im therapeutischen Geschehen gibt. Auch sagen sie nichts aus über die Realisierungen von Therapie-Konzepten, also über den faktisch realisierten psychotherapeutischen Prozeß. Entscheidend ist auch hier die Erklärung: was, wann, warum, wodurch?

"Dodo birds are not very smart, so it is not clear why we would let them award prizes to begin with. That only happens in Wonderland. The dodo bird conclusion is an artifact of the dodo bird's failure to look for plausible mediators and moderators (as Luborsky et al. acknowledged in ending their article)." (Shadish & Sweeney 1991, S. 888).

Meta-Analysen sind trotz allem ein Fortschritt in der Psychotherapieforschung, weil sie uns zwingen, über die wirksamen Mechanismen und Prozesse psychotherapeutischer Arbeit nachzudenken, um zu präzisen Antworten zu kommen, die wiederum gezielt in Indikations- und Behandlungsentscheidungen eingesetzt werden zu können. Sie geben uns numerische Hilfen. Fortschritte in der Krebstherapie wären z.B. nicht möglich, wenn keine standardisierten Datenerhebungen international verwendet worden wären. Die Anwendung von Meta-Analysen in der Psychotherapieforschung benötigt vor allem einheitliche und allgemein akzeptierte Standards - neben der Verbesserung der Grundlagen, der Evaluationsstudien über Psychotherapien - , um nicht zu falschen Schlußfolgerungen zu gelangen. Es kann nichts besseres herauskommen als man zuvor hineingesteckt hat. Die Methode an sich ist nicht dafür zu kritisieren.

### **3. Die Berner Studie: Also doch unterschiedlich wirksame Psychotherapien?**

Seit seiner Habilitationsschrift an der Hamburger Psychiatrischen Universitätsklinik (Grawe, 1976), hat der Berner Lehrstuhlinhaber für Klinische Psychologie Grawe kontinuierlich und weithin vernehmbar die Diskussionen um differentielle Wirksamkeiten verschiedener Psychotherapieformen angeregt (Grawe 1986, 1989, Grawe et al. 1990) und mitgestaltet.

Vor allem in seinen jüngsten Beiträge, der Bestandsaufnahme zur "Psychotherapieforschung zu Beginn der neunziger Jahre" (Grawe 1992) wie auch dem Buch seiner Berner Arbeitsgruppe "Psychotherapie im Wandel - Von der Konfession zur Profession" (Grawe et al. 1994) wird deutlich, wem der Angriff gilt: allen anderen Psychotherapieformen als den behavioral-kognitiven, vor allem auch den psychodynamischen bzw. psychoanalytischen Behandlungsmethoden. In einem mehr als dreizehn Jahre währenden Mammut-Projekt analysierte das Berner Team um Grawe alle kontrollierten (und publizierten) Psychotherapie-Effektivitäts-Studien mit klinischen Populationen (insgesamt 897!) bis einschließlich 1983/84. Dies geschah mittels eines detaillierten Bewertungskataloges<sup>14</sup>, bei dem alle Studien nach acht Gütekriterien klassifiziert wurden (vgl. Tab. 1).

---

Hier etwa Tabelle 1

---

Die anhand des Bewertungskataloges erstellten Güteprofile (über die acht einzelnen Gütekriterien) dienten dem Berner Team als Grundlage für die Bewertung der empirische abgesicherten Wirkung der in den kontrollierten Therapie-Studien untersuchten Therapie-Methoden und damit für ihren generellen Anspruch, mit ihrer Monumental-Analyse das gesamte Gebiet der Psychotherapie mit ihren zahlreichen Konzepten, Methoden, Techniken und Effekten so gut beurteilen zu können, dass auch konkrete berufspolitisch brisante Empfehlungen formuliert wurden.

Die Schlußfolgerungen attestieren den behavioral-kognitiven Therapieformen eine überragend nachgewiesene Effektivität für den größten Teil des Spektrums psychischer Störungen; der Gesprächspsychotherapie wird weiterhin

---

<sup>14</sup> über 63 Seiten, ca. 1.000 Einzelmerkmale, nach denen die zugrundeliegenden Studien bewertet wurden

eine "überzeugend nachgewiesene Wirksamkeit" bescheinigt (S. 134), und mit deutlichen Abstrichen wird noch der psychoanalytisch orientierter Therapie und der psychoanalytischen Kurztherapie " ...eine einigermaßen gesicherte Wirksamkeit..." (S. 195) zugestanden. Dies gelte aber auch nur in den Anwendungsbereichen ambulanter Therapie von neurotischen und Persönlichkeitsstörungen sowie stationärer Therapie von schizophrenen Patienten. Alle anderen Psychotherapieformen haben noch keinen Preis verdient und haben lt. Grawe et al. keinen Anspruch darauf, in der kassenärztlichen Versorgung etabliert zu werden.

Die Schlußfolgerungen von Grawe et al. sind jedoch aus verschiedenen Gründen überzogen und vermutlich nicht haltbar. Zum einen orientierten sich die Urteile der Studie u.a. ganz offensichtlich an der Anzahl der durchgeführten empirischen Behandlungs-Evaluationen. Jedenfalls entspricht die Rangfolge der Bewertungen exakt der Häufigkeit der an diesen Behandlungsformen durchgeführten Studien: Verhaltenstherapie vor Gesprächstherapie und dann psychoanalytisch-psychodynamische Therapie, unter "Ferner liefern" erscheinen alle die Therapie-Techniken, bei denen wenige oder gar keine empirische Studien durchgeführt wurden (z.B. Analytische Therapie nach C.G. Jung, Bioenergetische Therapie, Daseinsanalyse, Familientherapie, Gestalttherapie, Individualtherapie, Katathymen Bilderleben, Musiktherapie, Autogenes Training, Psychodrama oder Transaktionsanalyse, um nur die bekanntesten Namen zu nennen). Dieses Faktum veranlaßte Eckert (1993) - ein namhafter Vertreter der Gesprächspsychotherapie - zu dem eher sarkastischen Kommentar:

"Viele sind gar nicht erst angetreten, drei haben gewonnen und zwei bekommen den Preis" (S.87). "Der Bewertungsmaßstab ist dann plötzlich nicht mehr die Effizienz des Verfahrens, sondern das Fehlen von Forschungen zur Effizienz." (Eckert 1993, S. 89).

In der Tat, in der Darstellung bzw. Kommentierung ihrer umfangreichen Ergebnisse werden von der Berner Gruppe zwei Aspekte ihrer Bewertung tendenziell konfundiert: Einerseits werden Studien zu bestimmten Psychotherapieverfahren mittels eigens erstellter Güteprofile beurteilt. Zum anderen werden Ergebnisse zur Wirksamkeit der entsprechenden Verfahren aus diesen Studien in Vergleichstabellen aufgelistet und kommentiert. Obwohl die Einführung von Güteprofilen zweifellos eine wichtige Neuerung in der Beur-

teilung (und hoffentlich auch in der Planung) von Psychotherapiestudien sein wird, bleibt fragwürdig, warum von den Autoren auch solche Studien munter interpretiert werden, deren Güteprofil nicht einmal ihren eigenen Standards entspricht. Wenn die Güteprofile stimmen, ist es doch ganz offensichtlich, daß solche Studien nichts beweisen. Hier ist die Chance zu einer valideren Stichprobe von Psychotherapiestudien aus auf Anhieb nicht ersichtlichen Gründen nicht wahrgenommen worden. Im Gegenteil das Verbleiben dieser (laut Güteprofil eher schlechten) Studien führt in der Folge zu einer "Entwertung" der damit untersuchten Verfahren.

Es wäre nichts gegen das Argument zu sagen, daß die

"...in ihrer Wirksamkeit eher nicht bestätigten Therapiemethoden.....aus der Versorgung kranker Menschen ausgeschlossen werden (sollten), solange sich nicht durch eine genügende Anzahl weiterer Untersuchungen eine andere Befundlage ergeben hat." (Grawe 1992, S. 138).

Es müßte aber dann auch ernsthaft die Forderung nach "Effizienz"-Prüfung erhoben werden (Überprüfung von Methoden im klinischen Feld), im Unterschied zur Überprüfung von "Effektivität" (kontrollierte Wirkungen im experimentellen Bereich; Kächele u. Kordy 1992), ein klinisch relevanter Unterschied, der bei Grawe et al. nicht diskutiert wird. Auf jeden Fall ist mit den o.g. Ausführungen von Grawe et al. noch nichts über die den genannten Ansätzen immanente psychotherapeutische Potenz ausgesagt. Daß sie noch nicht oder nicht ausreichend untersucht sind, besagt keinesfalls, daß sie einer solchen Untersuchung nicht zugänglich sind bzw. daß sie diesen Nachweis in Zukunft nicht zu erbringen imstande wären.

Ein deutliches Beispiel für die "Halbwert-Zeit" der Berner Aussagen liefern neuere Forschungsergebnisse zur vergleichenden Effektivität zwischen Verhaltenstherapie und Autogenem Training (AT) in einem vom BMFT-geförderten potenten Forschungsprogramm (Gieler u. Detig-Kohler, 1994). Die von der Grawe-Gruppe vergleichsweise abgebürstete Methode des Autogenen Trainings konnte in der wissenschaftlich erstklassigen Untersuchung den Vergleich mit der Verhaltenstherapie blendend bestehen.

Dieses Beispiel verdeutlicht, daß sich die wissenschaftliche Befundlage täglich ändern kann, daß sorgfältigere Studien in der Lage sind, inhärente psychotherapeutische Potenzen einzelner Verfahren valide herauszuarbeiten, und daß wir vor allem erst am Beginn einer größeren Periode von Ergebnis-

Aggregationen stehen, die weitreichende Schlußfolgerungen bereits zum heutigen Zeitpunkt obsolet erscheinen lassen.

Zudem lassen sich - teilweise vernichtende - Verurteilungen einzelner psychotherapeutischer Richtungen durch das Berner Projekt auch methodisch nicht nachvollziehen. Zunächst ist auch hier wieder der entscheidende Punkt (s.o. zur Kritik an der Meta-Analyse), welche Studien in eine Überblicksanalyse Eingang finden. Zwar behaupten Grawe et al., alle empirischen Kontrollstudien, die bis Ende 1983/Anfang 1984 weltweit greifbar waren, miteinzubeziehen. Die Reduktion der bis dahin veröffentlichten ca. 3.500 Studien auf nur 897, die dem Berner Kriterium "klinisch ausreichend" genügen konnten, läßt eine von der Berner Forschungsgruppe unabhängig stattfindende Replikation des Auswahlprozesses wünschenswert erscheinen. Vor allem die Gewichtungen, wie sie anhand des Beurteilungskataloges dann vorgenommen wurden, lassen sich grundsätzlich infrage stellen (Eckert 1993). Im Unterschied zu den international durchgeführten Meta-Analysen zur vergleichenden Effektivität verschiedener psychotherapeutischer Behandlungen berechneten Grawe et al. keine neuen Effektstärken. Grawe et al. (1994) stellen stattdessen sogenannte "Güteprofile" der für die einzelnen Therapierichtungen untersuchten Studien dar, die auf den o.g. acht Gütekriterien fußen. Studien, deren Güteprofile eine klare oder sehr überwiegende Tendenz in Richtung positive z-Werte aufweisen (besser als der Durchschnitt aller 897 zugrundeliegenden Studien), werden anschließend grundsätzlich positiv kommentiert. Angesichts der Brisanz der intendierten allgemeinen Aussagen des Berner Projekts erscheint es kaum ausreichend und wissenschaftlichen Kriterien nicht standzuhalten, wenn Grawe et al. lediglich darauf hinweisen, daß die Beurteiler-Gruppe zur Einschätzung der Güteprofile eine reliable Beurteilung in allen Kriterien - nebst der Einschußkriterien in die Übersicht - vorgenommen habe. Der Grad der Übereinstimmung ist mehr als relevant und hier sollte ein professioneller Psychotherapieforscher wie Grawe nicht nur an die Bereitschaft seiner Leser ihm zu glauben appellieren. Reliabilitäts-Koeffizienten, wie sie mittlerweile sogar für die Transkription von Psychotherapieprotokollen gefordert werden (Mergenthaler & Stinson 1990.), werden jedenfalls schmerzlich vermißt: Ob die in Bern trainierte Gruppe von Studenten, denen die Beurteilung der Studien oblag, auch außerhalb des "Berner Milieus" zu einem ähnlichen Ergebnis

gekommen wäre? Was würde passieren, wenn die Auswahl dieser Studien in gesprächstherapeutischen oder psycho-analytischen Zentren überprüft würde?

Die von der Berner Forschergruppe verwendeten acht Güteprofile, die wir weiter oben in der Tab. 1 dargestellt haben, stellen das Instrument für den größten Teil der monumentalen Werkes der Berner Forschergruppe ein (vgl. das Grawe et al.-Buch). Sie dienen als die Grundlage bei der Beurteilung der verglichenen Therapie-Verfahren. Wenn man nun diese von der Berner Gruppe erstellten Güteprofile der drei meist-untersuchten Therapierichtungen "behavioral-kognitive Therapieformen", "Gesprächs-psychotherapie" und "psychoanalytisch orientierte und analytische Kurztherapien" einander direkt gegenüberstellt, so zeigen sich scheinbar Vorteile speziell der VT-Therapieformen "Systematische Desensibilisierung" (S. 269), 'Reizkonfrontation' (S. 354) sowie der kognitiven Therapie-Varianten (S. 402 - 478) bzw. der 'Breitspektrum-Verhaltenstherapie'. Zusammengefaßt präsentieren Grawe et al. auf S. 79 die durchschnittlichen Güteprofile aller von ihnen verwendeten Güteprofile jeder Studie in vier Therapie-Bereichen: Qualitativ rangiert 'Breitspektrum-VT' vor 'Kognitiver Therapie nach Beck', dann 'Gesprächspsychotherapie' und dann 'Psychoanalytisch orientierte Therapie'.

Dieses in das Gutachten zum Psychotherapeutengesetz eingegangene "Mapping" der Wissenschaftlichkeit der großen Psychotherapieformen ist relativ und wie mittlerweile mehrfach erwähnt Produkt einer bestimmten Stichprobenbildung. Auch bei nur oberflächlicher Betrachtung stutzt der Leser allerdings über Inkonsistenzen: So z.B wird erwähnt, daß 12 Studien zur psychoanalytisch orientierten Psychotherapie existieren, nur 9 werden tatsächlich mit Güteprofilen aufgeführt (-3). In vier weiteren der verbleibenden Studien werden schizophrene Patienten oder Psychotiker therapiert (-4), in zwei Fällen davon chronische Schizophrenien; in einer weiteren Studie wird eine nicht psychotische Gruppe gegen eine chronisch-psychotische bzw. gegen eine akut psychotische Gruppe getestet (-1). Grawe et al führen diese Studien (die zudem das schlechteste Güteprofil dieser Gruppe zeigen) kommentarlos so an als ob tatsächlich chronische Schizophrenie regelhaft mit Hilfe psychoanalytischer Verfahren therapiert würden. U. E. ist es ein überaus fragwürdiges Vorgehen, wenn Besonderheiten der amerikanischen stationären Psychiatrie argumentativ für die ambulante Versorgung in der BRD herangezogen werden. Erinnern wir uns: Interne Validität (s. Abbildung 1) wurde von Grawe et al.

negativ gewichtet, wenn "das Ergebnis zugunsten einer in der Anlage der Untersuchung begünstigten Methode ausfiel" (S.75). Grawe kontra Grawe?! Interne Validität kann angesichts der Verzerrung in den untersuchten Patientenpopulationen seine eigene Untersuchung jedenfalls nicht beanspruchen! Vor diesem Hintergrund dürften nur 4 der aufgeführten Studien wirklich in den Überblick einbezogen und interpretiert werden. Beschränkt man sich (für die Untersuchung der Güteprofile) auf eine unserer Kenntnis nach in der ambulanten kassenärztlichen Versorgung (und gerade auf die wird in den Schlußfolgerungen von Grawe et al. abgehoben!) vorwiegend psychotherapeutisch versorgte Klientel, also auf Studien, in denen neurotische und psychosomatische Störungsbilder sowie Persönlichkeitsstörungen behandelt wurden, so ergibt sich ein differenzierteres Bild, das die Graweschen Schlußfolgerungen mehr als problematisch erscheinen läßt (wohlgemerkt: es wurden für diese Re-Analyse nur die von Grawe et al. selbst aufgeführten Güteprofile der zugrundegelegten Studien verwendet und exakt nachgerechnet) (vgl. Abb. 2).<sup>15</sup>.

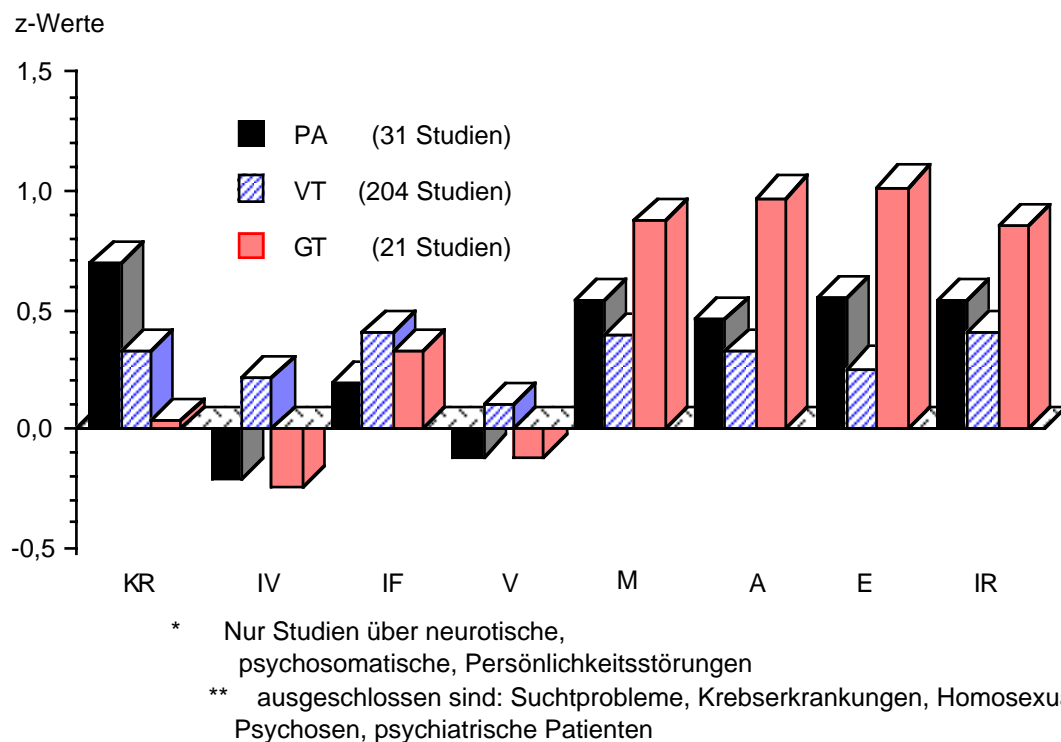
Es bleiben dann nach den genannten Kriterien folgende Studien. Die in der Abb. 2 zugrundeliegenden Studien umfassen für die PA 31 aus 37 Studien (nur "psychoanalytisch orientierte" oder "analytische Kurztherapie"), für die VT 204 aus 250 Studien (aufgenommen wurden die "typischen" VT-Techniken "Systematische Desensibilisierung", "Training sozialer Kompetenz", "Reizkontrolle", "Rational-Emotive Therapie", "Depressionstherapie nach Beck", "Depressionstherapie nach Lewinsohn" sowie "Breitspektrumverhaltenstherapie". Ausgeschlossen wurden z.B.: "Aversionstherapie", "Biofeedback", "Paradoxe Intention", die im übrigen alle ein extrem schlechtes Güteprofil bei Grawe et al. aufweisen, das weit unter z.B. denen der psychodynamischen Therapien liegt). Für die GT wurden nach den genannten Kriterien 21 aus 35 Studien ausgewählt. Bei den durchschnittlichen Güteprofil-Berechnungen je Therapieform wurde naturgemäß der relative Beitrag der einzelnen Therapie-Spezifität bei der Gewichtung mitberücksichtigt, also z.B., daß die Reizkonfrontation bei den VT-Studien in der Klinischen Relevanz mit

---

<sup>15</sup>Es ist ja im übrigen sehr verwunderlich, was die Berner Gruppe so alles an heterogenen Problemen und Erkrankungen zusammen mit Psychotherapie-Klientel in einen Topf geworfen hat, wo es doch gerade die Verhaltenstherapie ist, die heute den Ansatz betont, spezifisch zu therapieren (Baumann u. Perrez 1990).

0.55 den stärksten Effekt auf das Durchschnittsprofil hat (da 58 von 62 Studien berücksichtigt werden konnten), während z.B. nur 6 von 8 Studien zur Breitspektrumverhaltenstherapie mit einem durchschnittlichen Profilwert von 1,19 zu Buche stehen, also bei der Durchschnittsberechnung der zur Verfügung stehenden VT-Therapieformen diese letztere Form entsprechend weniger gewichtet werden muß.

Es bestehen signifikante Unterschiede im Mittelwertvergleich zwischen den Güte-kriterien der einzelnen Therapieformen in folgenden Fällen: Klinische Relevanz (PA > VT und GT, jeweils  $p < 0.05$ ); Interne Validität (VT > PA und GT, jeweils  $p < 0.01$ ); Güte u. Reichhaltigkeit der Auswertung (GT > VT,  $p < 0.01$ ) und Reichhaltigkeit der Ergebnisse (GT > VT,  $p < 0.01$ ).



**Abbildung 2:** Güteprofile von psychoanalytischen (PA), verhaltenstherapeutischen (VT) und gesprächspsychotherapeutischen (GT) Studien (nach Grawe et al., 1994)\* \*\*

(KR=Klinische Relevanz, IV=Interne Validität, IF=Güte der Information, V=Vorsicht bei der Interpretation, M=Reichhaltigkeit der Messung, A=Güte



und Reichhaltigkeit der Auswertung, E=Reichhaltigkeit der Ergebnisse, IR=Indikationsrelevanz)

Psychoanalytische Studien weisen eine signifikant höhere klinische Relevanz auf als GT- und VT-Studien, VT-Studien eine klar höhere Interne Validität als die der beiden anderen Verfahren und GT-Studien jeweils signifikant mehr Güte u. Reichhaltigkeit der Auswertung sowie mehr Reichhaltigkeit der Ergebnisse als VT-Studien. Alle anderen Unterschiede sind nicht signifikant. Der einzige Vorteil für die VT-Studien ergibt sich somit bei dem Gütekriterium Interne Validität und dementsprechend wohl in der Folge dann beim Kriterium "Vorsicht bei der Interpretation". Zugegeben: die Interne Validität bestimmt maßgeblich die Rückschluß-Möglichkeit auf die eingesetzte psychotherapeutische Interventionsmethode, d.h. ob die erzielten Wirkungen genau auf die Intervention und nichts anderes zurückgeführt werden kann. Aber hier scheinen alle drei Formen - absolut gesehen - nicht überzeugend zu sein. Dies ist ein berechtigter Vorwurf an die Produzenten der Studien. Die Therapiekonzepte selber sind damit jedenfalls nicht zu bewerten!

Bei den restlichen Gütekriterien M - IR muß es nachgerade erschrecken, wie groß die Unterschiede zwischen Gesprächspsychotherapie und Verhaltenstherapie - und tendentiell auch zwischen Gesprächspsychotherapie- und psychodynamischen Studien - im Durchschnitt sind. Die Gesprächstherapie-Studien schlagen bei der Reichhaltigkeit der Messung, der Auswertungsgüte, der Reichhaltigkeit der Ergebnisse und der Indikationsrelevanz alle anderen Therapie-Studien klar aus dem Feld. Andersherum gesagt: Verhaltenstherapeutische Studien schneiden bei typisch psychotherapeutischer Klientel relativ schlecht ab, und das bei ihrem leichten Vorsprung bei der Internen Validität! Das kann nur heißen: Trotz tendentiell besserer Rückführbarkeit der Untersuchungsergebnisse auf die manipulierte Variable "Behandlungsintervention" kommen tendentiell schlechtere Indikationsschlüsse, geringere Ergebnisreichhaltigkeit bei deutlich niedrigerer klinischer Relevanz heraus. Wie kann man da von einer Überlegenheit behavioral-kognitiver Forschungsdesigns ausgehen, wie Grawe et al. (1994) dies in euphorisch anmutenden Tönen an verschiedenen Stellen immer wieder anklingen lassen? Aus den Daten der Re-Analyse ist jedenfalls keine Überlegenheit der Verhaltenstherapie herauszuinterpretieren! Und dabei wurden die ganz deutlich schlechteren Profile von

"Aversionstherapie" (s.d. S. 390), "Biofeedback" (S. 370) und "Paradoxe Intention" der VT (S. 400) noch nicht einmal miteinbezogen.

Ein Vergleich der Studien, die die Berner Forscher-Gruppe selber als aussagekräftig ausgewählt hat und eine Beschränkung auf eine durchschnittlich zu erwartende ambulante Klientel führt zu Güteprofilen für die VT-Studien, die im Mittel deutlich hinter der Gesprächspsychotherapie und den psychoanalytisch orientierten Therapien liegen. Die Berner Schlußfolgerungsmentalität zugrundegelegt, ließe sich daraus nur der einzig möglicher Schluß ziehen, daß nämlich die Beforschung der Verhaltenstherapie im Psychotherapiebereich eindeutige Defizite gegenüber den beiden anderen Therapierichtungen aufweist. Falls diese Güteprofile für die Fundiertheit der einzelnen Therapieformen erhalten sollen, dann bliebe nur eine mögliche Schlußfolgerung: Verhaltenstherapie ist empirisch weniger gut fundiert als Gesprächspsychotherapie und psychodynamische Verfahren fundiert! So viel zur Stichprobenbildung.

Natürlich ist Grawe et al. (1994) in dem Punkte vollständig zuzustimmen, daß die analytische Langzeitbehandlung über mehrere hundert Stunden bisher praktisch nicht mit den Kriterien des randomisierten Kontrollgruppen-Designs untersucht worden ist. Als einzige Ausnahme sprechen Grawe et al. merkwürdig euphemistisch über die "Menninger-Studie" als "...eine der bedeutendsten bisher je durchgeführten Therapiestudie(n)..." (S. 173), die über Jahrzehnte hinweg die Therapieergebnisse von 42 schwergestörten Patienten beobachtete (Wallerstein 1986). Gar nicht mehr merkwürdig wirkt der Euphemismus, wenn man versteht, wozu er dienen soll: einer vernichtenden Kritik an der Langzeitpsychoanalyse. Wenn Grawe et al. schon die Ausnahme machen, im Falle der Langzeitpsychoanalyse auch naturalistische Studien zu diskutieren, warum bedienen sie sich dann der methodologisch so veralteten Menninger-Studie, anstatt sich mit den methodisch höherwertigen Studien des "Heidelberger Katamnese-Projektes" (Bräutigam et al. 1990, Kordy et al. 1983), der Berliner Therapie-Studie von Rudolf et al. (1991) oder den in der Übersichtsarbeit von Bachrach et al. (1991) erwähnten Evaluationen von Psychoanalysen jüngerer Datums auseinander-zusetzen (vgl. auch Kächele u. Kordy 1994)? Die Ausnahmen die im Bereich der Interpersonalen Therapie von Grawe et al. gemacht werden (der Einbezug von Studien nach 1983) weisen daraufhin, daß das spätere Datum dieser Untersuchungen nicht einziges Ausschlußkriterium gewesen sein können. Vielleicht hat hier "Erkenntnis und

Interesse" sich auf eine Art ergänzt, die der psychoanalytischen Langzeitbehandlung nicht gerade zugute kam und kommen sollte. Die Ergebnisse der Menninger-Studie sprechen in der Tat nicht für "heroische Indikationen" für jahrelangen Langzeitbehandlungen bei schwerst gestörten Patienten. Wir vertreten die Auffassung, daß die methodische Seite der Menninger-Studie diese aus heutiger Sicht "aussageunfähig" macht.

Dennoch: die Kritik an der mangelhaften empirischen Untermauerung psychoanalytischer Behandlungskonzepte bleibt voll bestehen. Dieses Manko muß sich die Psychoanalyse selber ans Revers heften. Es darf nicht sein, daß eine solch klinisch etablierte Methode so wenig über ihre wahren Wirkmechanismen und Effekte bei speziellen Krankheitsbildern dokumentieren kann. Das Argument ist schlicht nicht haltbar, daß die Effekte und Prozesse psychoanalytischer Therapie nicht meßbar seien, dazu sind viel zu viele der im Setting auftretenden Variablen beobachtbar und damit einschätzbar. Die Mystifizierung psychoanalytischer Prozesse als unbeobachtbar und bei oder durch Beforschung zerstörbar dient u.E. eher Abwehrgründen und läßt sich vor einem weltweit akzeptierten Wissenschaftsverständnis nicht länger tabuisieren. Psychoanalyse geriete so ins gesellschaftliche Abseits. Hier ist also die Kritik der Grawe-Gruppe berechtigt und wir stimmen ihr prinzipiell zu. Wir sind mit Grawe der Meinung, daß hier deutliche Defizite auszugleichen sind, im wissenschaftlichen und theoretischen Erkenntnisinteresse sowie im Interesse unserer Patienten.

"Eine Psychotherapierichtung, die weiterhin (mit Freud) glaubt, daß jeder Praktiker eo ipso auch Forscher sei, muß aus gruppendynamischen Gründen alle "fremdдоктroyierten" Forschungsergebnisse ablehnen - es sei denn, sie bestätigen zufällig die eigene Person. Grawe sieht diesen Mißstand scharf und zutreffend. Ich kann mich da nur genieren und es für mich besser machen." (Hoffmann 1992, S. 163).

Die von Grawe et al. (1994) direkt vorgenommenen Wirkungsvergleiche (im Kapitel 4.9) der drei bekanntesten - und wie zuvor erwähnt wurde, auch bei Grawe et al. (1994) am besten abschneidenden - Therapieformen (VT vs. PA, VT vs. GT, GT vs. PA in Kapitel 4.9) entbehren hinsichtlich der methodischen Durchführung wie hinsichtlich der dann vollzogenen Schlußfolgerungen allerdings jeglicher Originalität, insbesondere da die Autoren die zum

Zwecke des Vergleichs von ihnen verwendete Methode der Ergebnisübersicht auf Seite 46/47 an Luborsky et al (1975) kritisieren.

"Obwohl dieser erste Versuch der Quantifizierung gewiß einen Fortschritt darstellte, so ließ die Differenziertheit der Auswertung in dieser Untersuchung doch sehr zu wünschen übrig. Wenn man das Ergebnis einer Vergleichsuntersuchung zwischen zwei Therapieformen, in der die Wirkungen der untersuchten Therapien u.U. durch eine breit angelegte Messbatterie zu erfassen versucht wurden, zu einem einzigen "-", "-" oder "Unentschieden" zusammenfaßt, dann kann diese grobe Kategorisierung nur in Ausnahmefällen das tatsächliche Ergebnisbild dieser Untersuchung angemessen repräsentieren, denn dieses ist in der Regel nicht nur viel reichhaltiger und komplizierter, sondern auch interessanter. Die Notwendigkeit, von den Besonderheiten der einzelnen Untersuchung in einem gewissen Ausmaß zu abstrahieren, ist jeder Übersichtsarbeit immanent. Aber Luborsky et al. sind nach unserem Dafürhalten mit ihrer Kondensierung der Untersuchungsergebnisse einer Therapievergleichsstudie auf eine von drei möglichen Kategorien doch einige Schritte weiter auf dem Weg der Vereinfachung gegangen, als es der Natur des Untersuchungsgegenstandes guttut." (S.46/47).

Auf Seite 655f führen Grawe et al. eine Methode des Therapievergleiches mit eben den von ihnen zuvor kritisierten Reduktionen auf " " und "-" ein, mit der dann Verhaltenstherapie, Gesprächspsychotherapie sowie Psychoanalytische Verfahren "direkt" miteinander verglichen werden. Als "besser" (+) oder "schlechter" (-) wurden alle Zahlenwerte in allen verwendeten Veränderungsmaßen nach Therapie-beginn interpretiert, ungeachtet jeglichen Signifikanz-Niveaus, wobei im Prinzip jeder unterschiedliche Wert interpretiert wurde, und sei es auch nur ein Punkt auf einer 100er-Skala. "Beim Vergleich zwischen analytischer Therapie und Verhaltenstherapie ist es auf den ersten Blick offensichtlich, dass die Mehrzahl der 487 Vergleiche eindeutig zugunsten der Verhaltenstherapie ausfällt." (S.662). Die Schlußfolgerung, *"Verhaltenstherapie ist im Durchschnitt hochsignifikant wirksamer als psychoanalytische Therapie"* fällt den Autoren dann nicht mehr schwer. (Kursiv im Original). Wir denken, daß Grawe et al. an dieser Stelle ihrerseits "einige Schritte weiter auf dem Weg der Vereinfachung gegangen" sind "als es der Natur des Untersuchungsgegenstandes guttut". Grawe contra Grawe?

Es soll in diesem Zusammenhang nicht unerwähnt bleiben, daß alle in den weiter oben erwähnten Meta-Analysen verarbeiteten Psychotherapie-Wirkksamkeits-Studien, wie auch die von der Berner Gruppe untersuchten Therapie-Studien, ältere Arbeiten einer Forschungsperiode dar, die nach der eingangs erwähnten Einteilung von Shapiro den Phasen 1 oder 2 zugeordnet werden müssen. Bezugnehmend auf die im "Forschungsgutachten zu Fragen eines Psychotherapeutengesetzes" (Meyer et al. 1991) berichtete "Ergebnis- und Befundlage", kritisiert Eckert:

"...liegen zwischen Durchführung und Veröffentlichung von größeren Therapiestudien etwa 10 Jahre. Die im Gutachten berichteten "Ergebnislagen" basieren danach auf einer Therapiepraxis, wie sie vor 17 und mehr Jahren üblich war. Die ältesten berücksichtigten Studien erfassen therapeutisches Handeln, wie es vor 60 Jahren praktiziert wurde." (Eckert 1993, S. 89).

Forschung in der Psychotherapie entwickelte sich erst allmählich innerhalb der letzten 35 Jahre. Die Hinwendung zur detaillierten Einzelfallforschung und überhaupt zur Prozeß-Ergebnis-Forschung erleben wir erst seit den letzten 10 - 15 Jahren. Die Qualität der Methoden und Forschungsdesigns hat in diesem Zeitraum in erheblichem Maße zugenommen, damit auch dokumentierend, daß die bis ca. 1980 erfolgten Psychotherapie-Studien in nur sehr einzelner Weise, generell aber nicht ausreichend, Therapieprozesse und diesbezügliche Effekte abgebildet haben. Alle Meta-Analysen, die z.B. Grawe (1992) ins Feld führt (s.o.), sind noch dieser "Frühzeit der Psychotherapieforschung" (Meyer, 1990) zuzurechnen. Wie sollten wir also auf solch einer Basis zu der Auffassung gelangen, wir wüßten, welche Therapieformen was und in welcher Zeit bewirken? Psychotherapieforschung ist derzeit erst auf dem Wege, tatsächliche differentielle Wirkfaktoren und spezifische Wirkmechanismen zu entdecken. Sie war in der Vergangenheit dazu noch gar nicht in der Lage. Es wäre für alle Beteiligten fatal, das Rennen bereits vor dem Startschuß für beendet zu erklären!

Noch problematischer wird es, zu einer sachlich berechtigten Abklärung von therapeutischen Potenzen verschiedener psychotherapeutischer Behandlungsformen zu gelangen, wenn man im Vergleichstest bereits bei der Einbeziehung von Basis-Studien willkürliche Kriterien (s.o.) walten läßt und z.B. verfügbare Information zurückhält, weil sie nicht ins gewünschte Bild paßt.

Grawe et al. (1994) kommentieren z.B. die höchst wichtigen und interessanten Befunde der wohl größten systematischen Therapievergleichs-Studie, die jemals vorgenommen wurde, in diesem Zusammenhang sehr unvollständig, obwohl weitergehende Informationen vorlagen.

Es handelt sich bei dieser Studie um die sogenannte "NIMH"-Studie ("National Institute of Mental Health Treatment of Depression Collaborative Research Program"), bei der 162 depressive Patienten ambulant an verschiedenen Zentren in den USA hinsichtlich verschiedener Behandlungen über 16 Wochen miteinander verglichen wurden: 1) Kognitiv-behaviorale Therapie nach Beck, 2) Interpersonale Therapie, 3) medikamentöse Behandlung mit Imipramin und klinischer Betreuung in Form von einem halbstündigem Gespräch pro Woche mit dem betreuenden Psychiater sowie 4) Placebo-Medikament plus klinischer Betreuung wie zuvor. Grawe et al. (1994) schreiben u.a. sehr verkürzt (S. 524 f.), daß die Imipramin-Gruppe deutlich am schlechtesten abgeschnitten habe und ziehen es vor über neueste Erkenntnisse dieser Studie eher Vermutungen anzustellen als zu zitieren, obwohl sie die 1992 erschienenen Ergebnisse einer Follow-up Untersuchung (Shea et al. 1992) in ihr 1994er Werk durchaus hätten einbeziehen können. Aber die Ergebnisse sind wohl nicht so gut, wie man sich das gewünscht hätte.

Es kommt nämlich kein substantieller Unterschied zwischen kognitiv/behavioraler Therapie, Interpersonaler Therapie sowie medikamentöser Therapie oder Placebo-Therapie heraus! Tendentiell stärkste Effekte hatten sogar die medikamentöse Behandlung sowie die Interpersonale Therapie bei der Verhinderung von Rückfällen von Patienten ohne Medikation. Die Autoren schließen daraus, daß eine Kurztherapie bei einem solchen Patienten-Klientel eben nicht ausreicht, um eine volle Genesung zu erzielen, besonders nicht bei schwerer depressiven Patienten (die Rückfallquoten waren nach 18 Monaten zwischen 33% und 50%). Leichter depressive Patienten sprachen besser auf Psychotherapie an (kognitiv/behaviorale und Interpersonale Therapie).

Ein wenig zufriedenstellendes Ergebnis berichten Holon et al. (1992) über eine sehr ähnlich gelagerte Kurztherapie-Studie gleichfalls mit nicht-psychotisch depressiven Patienten (ebenfalls kognitive Therapie im Vergleich mit medikamentöser Therapie mit Imipramin, allerdings ohne Placebo-Kontrolle).

Die Effektstärken waren sehr niedrig nach 12-wöchiger Behandlung: 0.59 bzw. 0.66 und die Rückfallquote sehr hoch.

Die Ergebnisse beider Studien unterstützen in keiner Weise die von Grawe et al. (1994) bzw. von anderen VT-orientierten Forschern (Shapiro 1985, zit. n. Grawe et al.) gemachten Schlußfolgerungen. Die Collaborative Study konnte die von Grawe et al. vehement vertretenen Effekte kürzerer psychotherapeutischer Interventionen und insbesondere die der kognitiv-behavioralen Therapie gerade für die schwerer depressiven Patienten nicht bestätigen. Es gibt, so scheint es, doch Belege dafür, daß kürzere Therapien nicht "...gleich gut oder besser..." (S. 696) sein müssen als längere Behandlungen.

Als Fazit läßt sich für das Berner Monumental-Projekt festhalten: Auch die Arbeiten der Berner Forscher um Grawe erbrachten nicht die klärenden Hinweise auf differentielle Wirksamkeiten unterschiedlicher Psychotherapieformen, trotz des immensen Arbeitsaufwandes und der aufgestellten Behauptungen. Zu viele Probleme überlagern die Schlußfolgerungen, die von Grawe et al (1994) gezogen werden; zu viele Voraussetzungen gehen undiskutiert in die Auswertung ein.

- 1) Konstruktion und Arbeitsweise der Gütekriterien, die den Güteprofilen zugrundeliegen, sind zweifelhaft, z.B. überlappen sich einige Kriterien fast vollständig
- 2) die dargelegten Güteprofile, verändern sich radikal legt man ein an der tatsächlichen kassenärztlichen Versorgung orientierte Stichprobe von Studien zugrunde
- 3) Die Güteprofile haben keine Konsequenz im Sinne eines Ausschlusses schlechter Studien aus der Untersuchung. Statt dessen wird auf eine Stichprobenbildung abgehoben, deren Reliabilität behauptet aber nicht nachgewiesen wird. Schlußfolgerungen basieren damit auf nachweislich auf der Auswertung schlechter Studien!
- 4) In der Argumentation werden Güteprofile und Effektivitätsmerkmale konfundiert

- 5) die in die Übersichten aufgenommenen Studien sind den Psychotherapieforschungsphasen 1 und 2 zuzuordnen und damit in ihrer Aussagefähigkeit relativiert.
- 6) es fanden keine separaten Berechnungen über die Zusammenhänge zwischen Erfahrung der Therapeuten und Behandlungseffekten statt, obwohl die Tabellen des Grawe et al.-Buches diese Informationen weitgehend liefern und bei Mitberücksichtigung dieser Informationsquelle signifikant positive korrelative Beziehungen zwischen der Qualität der Güteprofile und dem Level der Therapeuten-Erfahrung bestehen (gilt für alle VT- und PA-Studien, die wir bei unserer Zusammenstellung berücksichtigt haben, nicht für die GT-Studien).
- 7) die in die Überblicke eingehenden Studien wurden nicht auf Konzeptrealisierungen getestet (was natürlich den Produzenten der Studien anzulasten ist und nicht Grawe et al.).
- 8) es gibt keine ausreichende (oder gar keine) Kontrolle von Langzeiteffekten von Therapien bei den meisten in die Vergleiche eingehenden Studien.
- 9) Was ist eigentlich "Breitspektrumverhaltenstherapie"? Dieser völlig neu auftauchende Terminus wird von Grawe et al. in die Diskussion (und Therapievergleiche) geworfen, wobei die Abgrenzungen zu anderen VT-Verfahren und -Vorgehensweisen völlig im Unklaren verbleiben (S. 502f.). Der Eindruck läßt sich nicht vermeiden, daß hier willkürlich einzelne Studien gebündelt werden, um sie als die neueste, wirkungsvollste Methode der VT mit jeweils bestem Güteprofil in der Argumentation zu gebrauchen

Wenn auch nur einige dieser Aspekte in der Diskussion um die Berner Studie berücksichtigen würden, erscheinen die meisten Schlußfolgerungen der Autoren nicht mehr ohne weiteres nachvollziehbar.



#### **4. Zurück zur Gegenwart:**

##### **Zwei Studien über die Wirksamkeit analytischer Gruppentherapie.**

Als Hinweis auf die Möglichkeiten einer neuen präziseren Prozeß-Ergebnis-Forschungen (Stichwort "multiple Einzelfallforschung") zitieren wir hier stellvertretend zwei Psychotherapie-Studien, die die genannten Gütekriterien leicht erfüllen, ja sogar die Profile der besonders hervorgehobenen Studien um Sloane et al. (1975) bzw. Grawe (1976) (beide zit. nach Grawe et al. 1994) erreichen dürften. Diese Studien lassen sich zur sogenannten "Phase IV"-Forschung zählen (Linden, 1987), bei der tatsächliches therapeutisches Handeln in der alltäglichen klinischen Praxis zu den Effekten dieser Behandlungen in Beziehung gesetzt werden.

Beide Studien untersuchten Patienten aus stationären analytischen Langzeitgruppen-psychotherapien (Strauß & Burgmeier-Lohse 1994; Tschuschke 1993). Außer Kontrollgruppen wurden alle Design-Anforderungen an hochwertige Therapie-Studien erfüllt, von einer umfangreichen Outcome-Batterie, die Ebenen wie Selbsteinschätzung, Therapeuten-Beurteilung und unabhängige Kliniker einbezog, Längsschnittprozeß-Untersuchungen mit zahlreichen komplexen Selbst- und Fremdratings, Herstellung von Prozeß-Ergebnis-Beziehungen bis zu katamnestischen Nachuntersuchungen von 12 und 18 Monaten. Die interne Validität der Studien war weitestgehend gesichert, die Reichhaltigkeit der klinisch und theoretisch brauchbaren Ergebnisse und der Indikationshinweise sind klar ersichtlich, die klinische Relevanz gleichfalls evident. Beide Studien sind weitgehend vergleichbar hinsichtlich des Schweregrades der psychischen Probleme ihrer Klientel, weiterhin bezüglich der Behandlungsdauer und -frequenz - also der Zeit/Dosis-Beziehung - und des Nachuntersuchungszeitraums sowie hinsichtlich der Forschungsdesigns mit teilweise sogar identischen Meßinstrumenten.

Es zeigten sich für beide Studien über die Messungen zu unterschiedlichen katamnestischen Nachuntersuchungszeitpunkten im Prä - Followup-Vergleich, über das jeweils gesamte Sample beider Studien hinweg (also inklusive der wenig verbesserten Patienten), Effekt-Stärken, die die von Grawe (1992) oder Shapiro (1985, zit. nach Grawe et al 1994) zitierten durchschnittlichen Effektstärken behavioral/kognitiver Therapien bei weitem überschreiten. Die durchschnittliche Effektstärke der "Kieler Gruppenpsycho-

therapiestudie" (Strauß & Burgmeier-Lohse 1994) ( $n_1 = 31$  Patienten) betrug zur Katamnese (Prä vs. Katamnese) 1.51, die der "Stuttgarter Studie" (Tschuschke 1993) ( $n_2 = 16$ ) betrug zur Katamnese (Prä vs. Katamnese) 1.63 über jeweils alle eingesetzten Erfolgsmaße hinweg.

Die Rückführung der erzielten Therapie-Effekte wurde im Prozeß genauestens kontrolliert. D.h., die Effekte stationärer psychoanalytischer Gruppenpsychotherapie bei schwer gestörtem, teilweise erheblich chronifiziertem neurotischen, psychosomatischem und persönlichkeitsgestörtem Klientel beiderlei Geschlechts konnten eindeutig auf die Wirkungen der analytischen Gruppenpsychotherapie zurückgeführt werden. Die fast identischen Effektstärken dieser völlig unabhängig voneinander durchgeführten Studien sprechen nicht für die vorschnellen Schlüsse von Grawe et al. (1994) bzw. Shapiro (1985). Sie demonstrieren allerdings, was detaillierte und damit zuverlässigere Forschung zutage fördern kann.

## **5. Zusammenfassung und Ausblick**

Es bleibt an dieser Stelle zusammenfassend festzuhalten, daß der Effekt psychologisch-psychotherapeutischer Behandlungen im Schnitt nach den genannten umfangreichen Analysen bei einer Effektstärke von 0.85 (Smith et al. 1980), 0.86 (Andrews & Harvey 1981), 1,03 (Shapiro & Shapiro 1982) bzw. 1.08 (Landman & Dawes 1982) deutlich höher als bei Nicht-Behandlung bzw. bei anderen Formen von Behandlung liegt. Die durchschnittliche Effektstärke von dreizehn Meta-Analysen rangiert bei 0.81, wobei die Streuung mit 0.24 relativ gering ist (Lambert et al. 1986, S. 160). Zum Vergleich: die Effektstärken für andere Maßnahmen liegen deutlich niedriger - für ein neunmonatiges Lesetraining in der Grundschule muß z.B. ein Effekt von ca. 0.67 veranschlagt werden, die Wirkungen von Antidepressiva streuen zwischen 0.40 und 0.81 (Lambert et al. 1986).

"Thus, the effect sizes produced through the application of psychotherapies are typically as large or larger than those produced by a variety of methods typically employed during medical and educational interventions." (Lambert et al. 1986, S. 159).

Lehnen wir uns zurück. Warum erbrachte die Psychotherapieforschung bisher so ähnliche Ergebnisse? Ist den "Meta-Meta-Analysen" und der Berner Mo-

numental-Analyse zu glauben, die behaupten, es gebe leichte bzw. sogar konsistente Vorteile behavioral/kognitiver Therapieansätze z.B. dynamisch/humanistischen Ansätzen gegenüber? Kein Thema jedenfalls mehr muß die Frage sein, ob Psychotherapie überhaupt günstige Effekte aufweist, die z.B. "spontanen Remissionen" überlegen sind. Erstens sind "spontane Remissionen" in viel geringerem Ausmaß zu erwarten, als uns dies Eysenck (1952) seinerzeit glauben machen wollte (Grawe 1992). Das Konzept der "Spontanremission" bei ernsten psychischen Problemen gehört endgültig begraben. Zweitens weist Psychotherapie nachweislich deutlich bessere Effekte auf als unter Placebo-Bedingungen einheitlich festgestellt wurde (Andrews & Harvey 1981; Landman & Dawes 1982; Smith et al. 1980). Psychotherapie beinhaltet mehr als "lediglich" menschliche Zuwendung ("psychotherapists are more than placebologists", Lambert, et al. 1986, S. 163).

Es bleibt die Frage nach den differentiellen Möglichkeiten unterschiedlicher psychotherapeutischer Behandlungsansätze. Die seit ca. 15 Jahren innerhalb der Psychotherapieforschung praktizierte Technik der Meta-Analyse wie auch die aufwendigen Untersuchungen der Berner Forschergruppe um Grawe jedenfalls erbrachten - bei aller Euphorie für die erhofften Auswege aus der Misere der Unüberschaubarkeit, Heterogenität und Widersprüchlichkeit in der herkömmlichen Literatur - nicht die klare Ergebnis-Lage, die man sich von diesen aufwendigen Studien erhofft hatte. Es muß verwirren, wenn Untersucher an fast identischen Stichproben Meta-Analysen durchführen und zu unterschiedlichen Schlüssen kommen (Andrews u. Harvey 1981; Berman et al. 1985; Landman u. Dawes 1982; Luborsky et al. 1975; Shapiro u. Shapiro 1982; Smith et al. 1980). Interessant sind die noch gar nicht so alten Überlegungen von Grawe et al. (1990), daß es sich bei den in den Meta-Analysen ermittelten Therapieeffekten häufig nicht um gefundenen Fakten, sondern um vom Untersucher "hergestellte Produkte" handelt, weil abhängig nicht nur von den Therapie-Techniken, sondern auch von Meß- und Auswertungsmethoden. Dem wäre - besonders auch im Hinblick auf die neueren Grawe et al.-Ausführungen (1994) - nichts hinzuzufügen: Grawe contra Grawe.

Psychotherapieforschung ist noch nicht weit genug gediehen, um konsistent spezifische Effekte unterschiedlicher therapeutischer Techniken für alle Störungsbilder nachgewiesen zu haben. Systematische Therapieforschung gibt es erst seit ca. 40 Jahren. Es besteht überhaupt keine Veranlassung dazu,

davon auszugehen, daß systematische, wissenschaftliche Arbeit auf diesem Sektor in ihrem Aufwand und ihren Kosten, verglichen mit Forschungen auf anderen Gebieten, geringer ausfallen sollte. Schlußfolgerungen im Hinblick auf differentielle Effektivität einzelner psychotherapeutischer Behandlungsansätze sind zu diesem Zeitpunkt nicht möglich. Daran ändert leider auch die aufwendige Berner Therapieevaluationsstudie nichts, wie gezeigt werden konnte. Grawe selbst wird sich wohl noch an seine eigene Kritik an vorschnellen Schlüssen bezüglich einer vermeintlich nachgewiesenen Überlegenheit verhaltenstherapeutischer Therapie-Techniken erinnern: die maßgeschneiderten Effektnachweise bei der VT seien dafür verantwortlich, daß größere Effektstärken bei der VT als bei der GT oder der PA z.B. resultieren (Grawe 1986).

"Aber wer sich gefreut hat, hat sich zu früh gefreut. Das zitierte Ergebnis (VT schneide besser ab als GT und PA, die Verf.) steht nämlich auf wackeligen Füßen, und ich werde es gleich umstoßen. Es ist einfach ein Artefakt der metaanalytischen Methodik." (Grawe 1986, S. 522).

Der einzig sinnvolle Ausweg aus einer Vielfalt an widersprüchlichen Aussagen meta-analytischer Untersuchungen in der Psychotherapie liegt wohl nur in akribischer, sehr aufwendiger psychotherapeutischer Prozeß-Ergebnis-Forschung (Dahl et al. 1988; Greenberg & Pinsof 1986; Rice & Greenberg 1984; Tschuschke & Czogalik 1990). Psychotherapeutische Prozesse sind komplex - Forschung auf diesem Gebiet benötigt entsprechende Sorgfalt und langen Atem. Voreilige, und völlig ungerechtfertigte, Schlußfolgerungen bewirken nur falsche Hoffnungen und unangebrachte Aufregung; darüber hinaus verbreiten sie sich schnell in eine aufnahmewillige Öffentlichkeit hinein, wo sie zu unrealistischen und falschen Erwartungen führen und Schaden anrichten können. Der Paradigma-Wechsel in der Psychotherapieforschung mit der Hinwendung zu aufwendigen und detaillierten - sowohl kontrollierten aber ebenso wichtigen naturalistischen Prozeß-Ergebnisstudien (und multiplen Einzelfallanalysen) mit wesentlich sorgsameren Designs wurde von der Grawe-Arbeitsgruppe nicht ausreichend berücksichtigt. Alleine aus einer sehr präzisen Prozeßbewertung kann in Zukunft Erkenntnis gewonnen werden, welche therapeutischen Konzepte und Techniken für die wirklichen Veränderungen unter psychotherapeutischem Einfluß verantwortlich gemacht werden können.

Was bleibt? Es dürfte - trotz aller Kritik - nicht zu pathetisch sein, wenn wir die Behauptung aufstellen, daß die "monumentale Analyse" der Berner Therapieforscher so epochal für das Gebiet der Psychotherapie sein wird wie seinerzeit die Arbeit von Eysenck (1952). Damals war das Feld in eine "produktive Unruhe" gestürzt worden mit der Behauptung, Psychotherapie sei nicht wirksam. Paradoxerweise führte gerade die Eysencksche Provokation zum Nachweis der Wirksamkeit von Psychotherapie (McNeilly u. Howard 1991). Warum sollte die "Grawesche Provokation" nicht gleichfalls zur Aktivierung von angemessener Forschung führen, die die spezifischen Wirkfaktoren von psychotherapeutischen Behandlungskonzepten und -techniken nachweisen, sofern sie existent sind?

Also: Es bleiben nicht mehr viele Gründe, auf Dodobirds zu hören. Aber sich derzeit von süßen Klängen anderer Art becirren zu lassen besteht genauso wenig Veranlassung. Wir sollten nicht den großen Versprechungen erliegen ("Von der Konfession zur Profession"), wenn zu diesem Zweck unsere Bereitschaft zum Glauben stärker strapaziert wird als unsere kritische Vernunft. Der Psychotherapie und denen, die sich von ihr helfen lassen wollen, nutzen weder Wonderland noch Wunscherfüllungs-Denken. Wir alle brauchen Fakten, die stimmen. Die haben wir noch nicht. Nur präzise und aufwendige Forschungen werden sie erbringen.

## Literatur

- Bachrach HM, Galatzer-Levy R, Skolnikoff A, Waldron S (1991) On the efficacy of psychoanalysis. *J Amer Psychoanal Assoc* 39:871-916
- Bräutigam W, Rad M von, Engel K (1990) Wirkfaktoren psychoanalytischer Therapien aus der Sicht des Heidelberger Katamnese-Projektes. In: Lang H (Hrsg.) *Wirkfaktoren in der Psychotherapie*. Springer, Berlin, Heidelberg, New York, London, Paris, Tokyo
- Crits-Christoph P, Baranackie K, Kurcias J et al. (1991) Meta-analysis of therapist effects in psychotherapy outcome studies. *Psychother Res* 1:81-91
- Baumann U, Perrez M (Hrsg.) (1990) *Lehrbuch Klinische Psychologie*, Bd. 1. Huber, Bern, Stuttgart, Toronto
- Bergin AE (1971) The evaluation of therapeutic outcomes. In: Bergin AE u. Garfield SL (Eds.) *Handbook of Psychotherapy and Behavior Change*, 1st ed., John Wiley & Sons, New York, Chichester, Brisbane, Toronto, Singapore, S. 139-189
- Dahl H, Kächele H, Thomä H (Eds.) (1988) *Psychoanalytic Process Research Strategies*. Springer, Berlin, Heidelberg, New York, London, Paris, Tokyo
- Dührssen A, Jorswieck E (1962) Zur Korrektur von Eysenck's  
Berichterstattung über psychoanalytische Behandlungsergebnisse. *Acta Psychotherap* 10: 329-342
- Eckert J (1993) Zur Begutachtung der psychotherapeutischen Verfahren im "Forschungsgutachten" zum Psychotherapeutengesetz: Viele sind gar nicht erst angetreten, drei haben gewonnen und zwei bekommen den Preis. *Psychother Forum* 1:87-91
- Eysenck HJ (1952) The effects of psychotherapy: An evaluation. *J Cons Psychol* 16:319-324
- Fricke R, Treinies G (1985) *Einführung in die Metaanalyse*. Huber, Bern
- Gieler U, Detig-Kohler C (1994) Nähe und Distanz bei Hautkrankheiten. *Der Psychotherapeut* 39 (im Druck)
- Grawe K (1976) *Differentielle Psychotherapie I*. Hans Huber, Bern
- Grawe K (1986) Die Effekte der Psychotherapie. In: Amelang M (Hrsg.) *Bericht zum 35. Kongreß der Deutschen Gesellschaft für Psychologie (DGfP) in Heidelberg, Band II*, Hogrefe, Göttingen, S. 515-534
- Grawe K (1989) Von der psychotherapeutischen Outcome-Forschung zur differentiellen Prozessanalyse. *Zeitschrift für Klinische Psychologie* 18: 23-34
- Grawe K (1992) Psychotherapieforschung zu Beginn der neunziger Jahre. *Psychol Rundsch* 43:132-162
- Grawe K, Caspar F, Ambuehl H (1990) Differentielle Psychotherapieforschung: Vier Therapieformen im Vergleich. *Z klin Psychol* 19: 292-376
- Grawe K, Caspar F, Ambühl H (1990a-d) Die Berner Therapievergleichsstudie: Fragestellung und Versuchsplan. *Z Klin Psychol* XIX (ganzes Heft)
- Grawe K, Donati R, Bernauer F (1994) *Psychotherapie im Wandel. Von der Konfession zur Profession*. Hogrefe, Göttingen, Toronto, Zürich

- Greenberg LS, Pinsof WM (Eds.) (1986) *The Psychotherapeutic Process. A Research Handbook*. Guilford Press, New York, London
- Henry WP, Strupp HH, Butler SF, Schacht TE, Binder JL (1993a) Effects of training in time-limited dynamic psychotherapy: Changes in therapist behavior. *J Consult Clin Psychol* 61:434-440
- Henry WP, Schacht TE, Strupp HH, Butler SF, Binder JL (1993b) Effects of training in time-limited dynamic psychotherapy: Mediators of therapists' responses to training. *J Consult Clin Psychol* 61:441-447
- Holon SD, DeRubeis RJ, Evans MD, Wiemer MJ, Garvey MJ, Grove WM, Tuason VB (1992) Cognitive therapy and pharmacotherapy for depression. Single and in combination. *Arch Gen Psychiatry* 49:774-781
- Howard KI, Kopta S, Krause M, Orlinsky D (1986) The dose-effect relationship in psychotherapy. *Am Psychol* 41:159-164
- Kächele H (1992a) Psychoanalytische Therapieforschung 1930-1990. *Psyche*, 46/3, 259 - 285
- Kächele H (1992b) Die Persönlichkeit des Therapeuten und ihr Beitrag zum Behandlungsprozeß. *Z Psychosom Med Psychoanal* 38:227-239
- Kächele H, Kordy H (1992) Psychotherapieforschung und therapeutische Versorgung. *Der Nervenarzt* 63:517-526
- Kächele H, Kordy H (1994) Effektivität und Effizienz von hochfrequenten Langzeittherapien (Unveröff. Manuskript)
- Kordy H, Rad M von, Senf W (1983) Success and failure in psychotherapy: Hypotheses and results from the Heidelberg follow-up project. *Psychother Psychosom* 40:211-227
- Lambert M, Shapiro D, Bergin A (1986) The effectiveness of psychotherapy. In: Garfield S, Bergin A (Eds.) *Handbook of Psychotherapy and Behavior Change*. John Wiley & Sons, New York, Chichester, Brisbane, Toronto, Singapore, S. 157-211
- Landman J, Dawes R (1982) Psychotherapy outcome - Smith and Glass' conclusions stand up under scrutiny. *Am Psychol* 37:504-516
- Linden M (1987) *Phase-IV-Forschung. Antidepressiva in der Nervenarztpraxis*. Springer, Berlin
- Lösel F (1987) Methodik und Problematik von Meta-Analysen - Mit Beispielen der Psychotherapieforschung. *Gruppendyn* 18:323-343
- Luborsky L, Singer B, Luborsky L (1975) Comparative studies of psychotherapy: Is it true that "Everybody has won and all must have prizes"? *Arch Gen Psychiatry* 32: 995-1008
- Luborsky L, Crits-Christoph P, Mintz J, Auerbach A (1988) *Who Will Benefit From Psychotherapy?* Basic Books, New York
- Luborsky L, Singer B, Luborsky L (1975) Comparative studies of psychotherapies: Is it true that "everyone has won and all must have prizes"? *Arch Gen Psychiatry* 32:995-1008
- McNeilly C, Howard KI (1991) The effects of psychotherapy: A reevaluation based on dosage. *Psychother Res* 1:74-78
- Mergenthaler E, Stinson CH (1990) Transcriptions transcription standards. *Psychotherapy Research* 2: 125-142

- Meyer AE, Richter R, Grawe K, Graf von der Schulenburg J-M, Schulte B (1991) Forschungsgutachten zu Fragen eines Psychotherapeutengesetzes, Universitäts-Krankenhaus Hamburg-Eppendorf
- Meyer AE (1990) Wodurch wirkt Psychotherapie? In: Lang H (Hrsg.): Wirkfaktoren der Psychotherapie, Springer, Berlin, Heidelberg, New York, London, Paris, Tokyo, S. 179-189
- Meyer, AE (1990) Kommunale Faktoren in der Psychotherapie als Erklärung für nicht grob unterschiedliche Ergebnisse - Ein Mythos mehr in der Psychotherapieforschung? *Psychother. Psychosom. med. Psychol.* 40, 152-157
- Orlinsky DE, Howard KI (1986) Process and outcome in psychotherapy. In: Garfield SL, Bergin AE (Eds.) *Handbook of Psychotherapy and Behavior Change*. John Wiley & Sons, New York, Chichester, Brisbane, Toronto, Singapore, S. 311-381
- Prioleau L, Murdock M, Brody N (1983) An analysis of psychotherapy versus placebo studies. *Behav Brain Sci* 6:275-310
- Rice LN, Greenberg LS (Eds.) (1984) *Patterns of Change. Intensive Analysis of Psychotherapy Process*. Guilford Press, New York, London
- Rosenthal R (1983) Assessing the statistical and social importance of the effects of psychotherapy. *J Consult Clin Psychol* 51:4-13
- Rudolf G, Grande T, Porsch U (1991) *Die therapeutische Arbeitsbeziehung. Untersuchungen zum Zustandekommen, Verlauf und Ergebnis analytischer Psychotherapie*. Springer, Berlin, Heidelberg, New York, London, Paris, Tokyo, Barcelona
- Schepank H (Hrsg.) (1990) *Verläufe - seelische Gesundheit und psychogene Erkrankungen heute*, Springer Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, Barcelona
- Shadish W, Sweeney R (1991) Mediators and moderators in meta-analysis: There's a reason we don't let Dodo Birds tell us which psychotherapies should have prizes. *J Consult Clin Psychol* 59:883-893
- Shapiro DA, (1990) Lessons in history: three generations of therapy research. Paper delivered at the Annual Meeting of the Society of Psychotherapy Research, Wintergreen, USA
- Shapiro DA, Shapiro D (1982) Meta-analysis of comparative therapy outcome studies: A replication and refinement. *Psychol Bull* 92:581-604
- Shea T, Elkin I, Imber SD, Sotsky SM, Watkins JT, Collins JF, Pilkonis PA, Beckham E, Glass DR, Dolan RT, Parloff MB (1992) Course of depressive symptoms over follow-up. Findings from the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Arch Gen Psychiat* 49:782-787
- Sloane R, Staples F, Cristol A, Yorkston N, Whipple K (1975) *Psychotherapy versus Behavior Therapy*. Harvard University Press, Cambridge
- Smith ML, Glass GV, Miller TI (1980) *The Benefits of Psychotherapy*. Johns Hopkins University Press, Baltimore
- Strauß B, Burgmeier-Lohse M (1994) *Evaluation einer stationären Langzeitgruppenpsychotherapie - Ein Beitrag zur differentiellen Psychotherapieforschung im stationären Feld. Gruppenpsychother Gruppendyn (im Druck)*



- Strupp HH, Hadley SW (1979) Specific vs. nonspecific factors in psychotherapy. A controlled study of outcome. *Arch Gen Psychiat* 36:1125-1136
- Strupp HH, Binder JL (1991) Kurzpsychotherapie. Klett-Cotta, Stuttgart
- Strupp HH (1993) The Vanderbilt Psychotherapy Studies: Synopsis. *J Consult Clin Psychol* 61:431-433
- Tschuschke V (1990) Spezifische und/oder unspezifische Wirkfaktoren - ein Problem der Einzel- oder auch der Gruppenpsychotherapie? In: Tschuschke V, Czogalik D (Hrsg.) (1990) Psychotherapie - Welche Effekte verändern? Zur Frage der Wirkmechanismen therapeutischer Prozesse. Springer Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, Barcelona
- Tschuschke V (1993) Wirkfaktoren stationärer Gruppenpsychotherapie - Prozeß-Ergebnis-Relationen. Vandenhoeck & Ruprecht, Göttingen
- Tschuschke V (1994) Hinweise zur Dauerhaftigkeit therapeutischer Effekte und "Nicht-Effekte". Unveröff. Manuskript
- Tschuschke V, Czogalik D (Hrsg.) (1990) Psychotherapie - Welche Effekte verändern? Zur Frage der Wirkmechanismen therapeutischer Prozesse. Springer Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, Barcelona
- Tschuschke V, Dies RR (1994) Intensive analysis of therapeutic factors and outcome in long-term inpatient groups. *Intern J Group Psychother* 44 (2): 185-208
- Wallerstein RS (1986) Forty-two Lives In Treatment - A Study of Psychoanalysis and Psychotherapy. Guilford Press, New York
- Wilson G, Rachman S (1983) Meta-analysis and the evaluation of psychotherapy outcome: Limitations and liabilities. *J Consult Clin Psychol* 51:54-64
- Wittmann W, Matt G (1986) Meta-Analyse als Integration von Forschungsergebnissen am Beispiel deutschsprachiger Arbeiten zur Effektivität von Psychotherapie. *Psychol Rundsch* 7:20-40