

Perspektiven zur "Wissenschaftlichkeit" von Psychotherapie

Prof. Dr. Jürgen Kriz

Text für wissenschaftlich Interessierte, welche sich auch kritisch mit Fragen der aktuellen evidenzbasierten Strömungen auseinandersetzen möchte (Anm. R. Mischol / [psyaspect](#)).

Vorbemerkungen

Zu Beginn eines neuen Jahrtausends sind wir mit der Gefahr konfrontiert, unter dem Bann(er) angeblich "wissenschaftlicher Standards" ins Zeitalter der Glaubenskämpfe zurückkatapultiert zu werden, aus dessen dogmatischen Fesseln uns einst eine wissenschaftsgeleitete Aufklärung herausgeführt hat. Inspiriert von dem Buchtitel "Psychotherapeutische Perspektiven am Beginn des 21. Jahrhunderts" mag es daher erlaubt sein, zu Beginn dieses Beitrags einen Wunsch für die Zukunft zu äußern: Aus der Perspektive meines Themas halte ich es für das wünschenswerteste und dringendste Anliegen, daß die Mystifikation und der Mißbrauch des Attributes "wissenschaftlich" eingedämmt werden.

Nach gängigem Verständnis besteht nämlich eine der wichtigsten Aufgaben von Wissenschaft darin, Denk- und Handlungsräume von Menschen zu erweitern. Dazu ist es dienlich, möglichst umfangreiche Erkenntnisse über die Welt zu sammeln - oder auch aktiv zu generieren - und in Verstehenszusammenhänge einzubetten. Wesentlich für ein wissenschaftliches Vorgehen ist dabei, Alternativen in Fragestellung und Herangehensweisen nicht zu vermeiden, sondern diese in offensiver Weise möglichst zu erweitern und (dann!) kritisch zu diskutieren. Dazu gehört auch, daß die eigenen stillschweigenden Vorannahmen und damit die methodologischen und methodischen Voraussetzungen reflektiert und diskutiert werden. Wissenschaft sollte sich als Anwalt für eine Pluralität vieler unterschiedlicher Perspektiven verstehen, zumindest ideologischen Verengungen trotzen und einen möglichst großen Teil der "Welt" durch eine Vielzahl an Fragestellungen sowie Methoden für deren Untersuchung für die Allgemeinheit erschließen.

Solche eher globalen Standards für Wissenschaft sind unter Wissenschaftstheoretikern seit langem weitgehend unumstritten. Und im Allgemeinen wird die aufklärerische Vielfalt an Perspektiven und Zugangsweisen als ein Fortschritt gegenüber den dogmatischen und ideologischen Beschränkungen des Mittelalters gesehen und begrüßt. Umso beunruhigender ist es, wenn an der Schwelle zu einem neuen Jahrtausend unter Berufung auf "Wissenschaftlichkeit" versucht wird, im Bereich der Psychotherapie die Vielfalt der Perspektiven und Zugangsweisen und mit ihnen den Raum an Denk- und Forschungsmöglichkeiten drastisch zu beschneiden und - wenn auch nur innerhalb der Grenzen des Nationalstaates Deutschland - eine ganze Gesellschaft in ein bestimmtes Glaubens- und Wertekorsett hinsichtlich psychotherapeutischer Behandlungsmöglichkeiten zu zwängen.

Dies klingt wie eine ungeheuerliche Behauptung - und es bedarf daher zumindest des Verweises auf einige Fakten, um den Realitätsgehalt dieser Aussage überprüfbar zu machen.

1. Die Gleichschaltung der psychotherapeutischen Arbeits-, Forschungs- und Denkmöglichkeiten im Namen der Wissenschaft

Das 1983 in Deutsch erschienene zweibändige "Handbuch der Psychotherapie" (amer. Original 1981) des Professors für klinische Psychologie und Trägers mehrerer wissenschaftlicher Preise, Raymond J. Corsini weist rund 70 "Therapieansätze" (so das Vorwort) aus. Ein Blick in den Anzeigenteil einer Zeitschrift wie "Psychologie heute" zeigt, daß dies sicher nur die - von Corsini als nennenswert akzeptierte - "Spitze eines Eisberges" ist, zu dem vermutlich mehrere hundert Bezeichnungen gehören, die den Anspruch erheben, ein "Therapieverfahren" zu sein.

Mir schiene auf der Basis des derzeitigen Forschungsstandes nicht akzeptabel, auch nur den größeren Teil dieser Ansätze als "wissenschaftlich begründet" zu bezeichnen, da sie z.B. im wissenschaftlichen Diskurs - selbst bei einem sehr weit gefaßten Verständnis - kaum präsent sind und u. a. auch die Qualität der Ausbildung meist keineswegs hinreichend belegt ist. Auf der anderen Seite bin ich sicher, daß ein rein auf wissenschaftlicher Basis und ohne berufs- und standespolitische Interessen geführter Diskurs aller deutschen Fachwissenschaftler für den Bereich von Psychotherapie konstatieren würde, daß weit mehr als die gegenwärtig zugelassenen beiden "Richtlinienverfahren" als wissenschaftlich fundiert und anerkannt zu gelten haben.

Dafür spricht, daß auch der 1982 von Bastine, Fiedler, Grawe, Schmidtchen und Sommer - allesamt Universitätprofessoren mit klinisch-therapeutischer Fachkompetenz - herausgegebene Band "Grundbegriffe der Psychotherapie" mindestens 28 Therapieverfahren ausweist. Selbst wenn man einige dieser Verfahren für unsere Betrachtungsperspektive zu Oberkategorien zusammenfassen müßte, bleiben immer noch deutlich mehr als ein Dutzend übrig. Und das, obwohl die Herausgeber im Vorwort betonen: "Wir wenden uns (...) gegen den Wildwuchs (an) Therapien.." und das Buch "gründet sich wissenschaftstheoretisch auf ein empirisch-wissenschaftliches Grundverständnis ... Aufgrund dieses Wissenschaftsverständnisses werden im Buch Konzepte bevorzugt, die sich in der Theorienbildung einer empirischen Überprüfung stellen." Ebenso weist meine eigene Zusammenstellung "Grundkonzepte der Psychotherapie" (Erstauflage 1985) - das an vielen deutschen Universitäten Prüfungslektüre für Diplompsychologen ist - rund ein Dutzend Therapieansätze aus (die ebenfalls teilweise noch zu berufspolitisch begründeten Oberkategorien zusammengefaßt werden können). Das gleiche gilt für das vierbändige Werk "Psychotherapeutische Verfahren" des Tübinger Professors Dirk Revenstorf Anfang der 80er Jahre.

Die ungewöhnlichen Verweise auf Amtstitel und Prüfungskontexte dienen der Betonung der Verankerung im und Anerkennung durch das Wissenschaftssystem. Doch im Spätherbst 1999 wurde den beiden ersten (und in Fachkreisen unumstritten aussichtsreichsten) Antragstellern nach dem neuen Psychotherapeutengesetz die notwendige "Anerkennung der Wissenschaftlichkeit" bestritten - worauf konsequenterweise andere ihre Antragsvorbereitungen erst einmal einstellten. Erschreckend ist, daß dies durch einen sog. "Wissenschaftlichen Beirat" bewirkt wurde - aufgrund von Kriterien für "Wissenschaftlichkeit", die dieser Beirat selbst erlassen hatte. Es kann hier nicht der Ort sein, auf Detailprobleme dieser Entscheidung einzugehen und etwa zu fragen, ob es überhaupt Aufgabe des "Wissenschaftliche Beirat" war, eigene Kriterien für die Anerkennung von "Wissenschaftlichkeit" zu entwickeln, statt die Anerkennung durch andere (eben durch die Wissenschaft) zu untersuchen; und ob es statthaft oder zumindest schicklich war, die im Frühjahr 1999 selbst beschlossenen Kriterien ein halbes Jahr später und *nach* Vorliegen der Anträge (die natürlich auf die ursprünglichen Kriterien zugeschnitten waren) per Tischvorlage unmittelbar vor der

Abstimmung wesentlich zu ändern. Mit diesen und weiteren "merkwürdigen" Aspekten werden sich noch Juristen, Politiker (und später vielleicht einmal: Wissenschafts-Historiker) zu beschäftigen haben. Wichtiger ist es, kurz die Konsequenzen dieser Entscheidung zu beleuchten. In der Annahme, daß Wissenschaftler in der Lage sein müßten, die Konsequenzen ihrer Entscheidungen mit ins Kalkül zu ziehen, muß davon ausgegangen werden, daß diese Konsequenzen entweder intendiert oder zumindest fahrlässig billigend in Kauf genommen wurden:

Da "Psychotherapie" nach dem Gesetz nur Anwendung "wissenschaftlich anerkannter Verfahren" beinhaltet, dürften damit Psychotherapeuten keine anderen Verfahren als die beiden Richtlinienverfahren ausüben (faktisch sogar nur Verhaltenstherapie, da in Fachkreisen bezweifelt wird, daß die tiefenpsychologischen Verfahren diese Kriterien auch nur annähernd selbst erfüllen, da sie auch für diese allzu eng und inadäquat sind - was z. B. die Diskussion zwischen Tschuschke, Kächele und Hölzer auf der einen und Grawe auf der anderen Seite im *Psychotherapeut* 1994/95 deutlich macht). Was aber nicht ausgeübt werden darf, kann auch nicht (oder nur unter extrem erschwerten Bedingungen) praktisch erprobt und beforscht werden. Mit gutem Grund könnte man dann auch die Berechtigung der Lehre an den Universitäten problematisieren (zumal auch hier immer mehr vordergründige "Effektivität" als zentrales Kriterium zur Beschneidung von Pluralität verwendet wird). Das Spektrum der Ansätze wäre schlagartig auf *einen* Ansatz (oder ggf. zwei) reduziert und diese radikale Reduktion für Jahrzehnte zementiert. Wer Therapie anbietet und wer Therapie nachfragt, wäre (von sehr erschwerten Ausnahmefällen abgesehen) auf diese eine Richtung festgelegt. Und sehr schnell würden dann auch die nicht mehr gebrauchten, gelehrten und erforschten alternativen Zugänge, Menschenbilder, Konzepte, Denkweisen und Methoden aus der lebendigen Wissenschaft und der Gesellschaft verschwinden. Dies darf wohl durchaus als eine Kulturrevolution im Bereich der Psychotherapie (und der damit verbundenen Lebenskonzepte, Menschenbilder, Wertvorstellungen usw.) bezeichnet werden.

Auch wenn ich hoffnungsvoll sicher bin, daß vor allem wach gewordene Politiker und Verwaltungsjuristen diesem Alptraum bald ein Ende setzen werden, kann ich nicht glücklich darüber sein, weil dies zeigt, daß Selbstregulationsmechanismen der Wissenschaft in diesem Sektor derart versagt haben. Es ist erschreckend, wenn wissenschaftliche Vielfalt, Pluralität der Perspektiven und unterschiedliche Forschungszugänge - kurz: der Kern dessen, was viele für das Wesen von Wissenschaft ansehen - nur *gegen* eine massive Reduktion im Namen von "Wissenschaftlichkeit" und letztlich allein mit Hilfe von Verwaltungsjuristen und Politikern aufrechterhalten bzw. wiederhergestellt werden können.

Wie kann es dazu kommen, daß ein "Wissenschaftlicher Beirat" Kriterien für "Wissenschaftlichkeit" erläßt, die sich derart verheerend auch auf Wissenschaft auswirken würden? Immerhin ist beispielsweise das nach Meinung dieses Beirates nicht "wissenschaftlich anerkannte" Verfahren "Gesprächspsychotherapie" (GT) von Rogers selbstverständlich nicht nur in den eben angeführten und allen weiteren Lehrbüchern über unterschiedliche Psychotherapieverfahren vertreten, sondern es wird an über zwei Dutzend deutschen Universitäten regelmäßig gelehrt und ist Bestandteil von wissenschaftlichen Prüfungen. Zahlreiche wissenschaftliche Kongresse fanden statt und es gibt wohl kaum eine wissenschaftliche Zeitschrift im Bereich der Psychotherapie, die im vergangenen Jahrzehnt nicht vielfach auf Konzepte und Forschungsergebnisse der GT Bezug genommen hätte. Selbst die viel zitierte umfangreiche Meta-Analyse von Grawe, Donati und Bernauer (1994) - der bei aller Kritik niemals vorgeworfen wurde, ein Verfahren *zu gut* bewertet zu haben - kam zu dem Schluß: "Vergleicht man die Ergebnistabellen für die Prä-Post-Vergleiche und die Kontrollgruppen-Vergleiche

mit den analogen Tabellen zu den meisten anderen Therapieformen, dann muß man der Gesprächspsychotherapie eine sehr überzeugend nachgewiesene Wirksamkeit bescheinigen" (S. 134) und: "Es führt kein Weg an der Schlußfolgerung vorbei: Gesprächspsychotherapie ist nachweislich ein sehr wirksames Verfahren für ein weites Spektrum an Störungen" (S. 135).

Im Frühjahr 1998 unterzeichneten denn auch binnen weniger Wochen 80 "Deutsche Universitätsprofessorinnen und -professoren im Bereich Psychotherapie / Klinische Psychologie / medizinische Psychologie" einen Aufruf, in dem u. a. nochmals betont wurde: "Die Gesprächspsychotherapie gehört sowohl international als auch in Deutschland seit Jahrzehnten zu den praktizierten und bewährten Verfahren. Tausende von Patienten wurden mit Gesprächspsychotherapie erfolgreich ambulant bzw. stationär behandelt. In zahlreichen Lehrbüchern der Psychotherapie / Klinischen Psychologie wird dieses Verfahren als wissenschaftlich ausgewiesen und als effektiv dokumentiert. An vielen deutschen Universitäten gehört die Gesprächspsychotherapie sowohl zur Forschung als auch zur Lehre und somit zum Prüfungsstoff u. a. im Hauptdiplom in Psychologie...".

Nun sollte man meinen, daß eine deutlichere Dokumentation der "wissenschaftlichen Anerkennung" kaum denkbar ist, als wenn dies explizit von der ganz überwiegenden Mehrheit der fachlich kompetenten Wissenschaftler an Universitäten betont und zudem durch die Präsenz im wissenschaftlichen Diskurs in Fachbüchern und -zeitschriften und letztlich noch in Forschung und Meta-Analysen unterstrichen wird.

Wenn trotzdem der "Wissenschaftliche Beirat" mehrheitlich zu dem Schluß kam, der Gesprächspsychotherapie diese Bestätigung als "wissenschaftlich anerkanntes psychotherapeutisches Verfahren" nicht zu erteilen, wurde hier offensichtlich das Attribut "wissenschaftlich" in einer Weise definiert, die keinen Bezug zur Wissenschaftlergemeinschaft und deren (zumindest deutlich mehrheitlicher) wissenschaftlicher Bewertung herstellt. Da man einen "direkten Draht zum Weltgeist" oder ähnliche überirdische Wahrheitsquellen und -kriterien ausschließen muß und kann, ist dies nur erklärlich, indem man sich auf "streng objektive" (und von der Intersubjektivität der scientific community unabhängige) Kriterien beruft.

Meines Erachtens ist dies ein extremes Beispiel für die Mystifikation von bestimmten methodischen Zugängen und ein fast magischer Glauben an die Bedeutung statistischer Zahlen, den ich bereits vor Jahrzehnten mit dem Verweis glossiert habe, daß manche Wissenschaftler an der Tür zum Rechenraum mit dem Mantel scheinbar auch ihren inhaltlichen Sachverstand abgeben (Kriz 1975, 1981). Trotz vehementer Kritik seitens der Methodiker - von der Signifikanztest-Kontroverse in amerikanischen Zeitschriften in den sechziger Jahren, über den Beitrag des Statistikers L. Guttman mit dem bezeichnenden Titel "What is not what in statistics", bis hin zu dem markanten Aufsatz Kromreys (1999) "von den Problemen anwendungsorientierter Sozialforschung und den Gefahren methodischer Halbbildung"- wird vielfach den Outputs formaler Prozeduren, wie Signifikanzsternen oder statistischen Parametern, eine Bedeutung eingeräumt, die diesen nicht zukommt. Es ist hier sicher nicht der Ort, Grundkurse kritischer Statistik zu reproduzieren (u. a. Kriz 1973, Kriz & Lisch 1988). Aber für diesen Kontext mag es sinnvoll sein, zumindest exemplarisch am Beispiel der sog. "Effektstärke" die Probleme von Parameter-Gläubigkeit aufzuzeigen, da Effektstärke durchaus eine wesentliche Rolle in dem vom Beirat verordneten "Wissenschaftlichkeits"-Standard spielt.

2. Exkurs: Effektstärke versus Effektivität von Psychotherapie

Es gibt zahlreiche wissenschaftliche Debatten um inhaltliche Fragen darüber, wie die Effektivität von Psychotherapie zu konzeptionalisieren, zu operationalisieren und zu messen sei. Im Gegensatz zu den klar operational definierten Größen etwa der Physik haben unterschiedliche Therapierichtungen entsprechend ihren unterschiedlichen Zielvorstellungen nicht nur unterschiedliche inhaltliche Konzepte von "Effektivität", sondern selbst unterschiedliche Studien ein und derselben Richtung verwenden je nach Detailfrage(n) recht verschiedene Kriterien und damit Operationalisierungen. Hinzu kommt die grundsätzliche Frage, ob "effektive" Veränderungen unter streng kontrollierten Laborbedingungen mit ihren Anforderungen an "reine" Störungen und "reine" (also meist manualisierte) Behandlung valide Aussagen erlaubt über "effektive" Veränderungen unter üblichen Bedingungen - d. h. bei Menschen, die oft nicht nur eine Störung haben und die von Therapeuten behandelt werden, die nicht "manual-rein" vorgehen.

Das zentrale Problem bei einer Diskussion über therapeutische "Effektivitäts"-Maße, nämlich daß "Äpfel" nicht nur mit "Birnen" sondern eher mit "Armbanduhren" verglichen werden, läßt sich strukturbedingt nicht beseitigen: Es besteht eben kein Konsens dahingehend, was Effektivität von Therapie bedeutet - und mit einem Konsens ist auch nicht zu rechnen, weil die Heterogenität hinsichtlich der Vorgehensweisen und Maße nur die Heterogenität der Menschenbilder, Lebensziele, Erfahrungen, subkulturellen Gewohnheiten und Vorlieben in der Bevölkerung widerspiegelt.

Gleichwohl entspricht es dem Zeitgeist, sich nicht so sehr offen über die unterschiedlichen Lebensziele und Werte in der Gesellschaft auseinanderzusetzen, sondern nach "Effektivität" zu rufen. Und wenn man denn schon nicht weiß, wie die unterschiedlichen Vorstellungen von Effektivität unter einen Hut zu bringen sind, dann zumindest von "der Wissenschaft" ein genaues Maß für "Effektivität" zu fordern - in der trügerischen Hoffnung, damit sei das inhaltliche Problem wenn schon nicht gelöst, dann doch zumindest vom Tisch. Zu Recht kann man darauf hoffen, daß die Heterogenität der Wissenschaft tatsächlich immer wieder Vorschläge für scheinbar einfache "Lösungen" eigentlich unlösbarer Probleme hervorbringen wird.

Dem offenbaren Bedürfnis, die Eigenwerte von unterschiedlichen Zugangsweisen zur Welt auch im Bereich von Psychotherapien nicht einfach akzeptieren zu wollen, sondern doch "Äpfel" mit "Armbanduhren" irgendwie vergleichen zu können, entspringt die "Effektstärke". Sie hat sich in Meta-Analysen durchgesetzt, damit die unterschiedlichen Erfolgsmaße aus unterschiedliche Studien mit unterschiedlichen Fragestellungen an unterschiedlichen Patienten mit unterschiedlichen Therapiemethoden überhaupt formal zusammengefaßt und dann z. B. Gruppen miteinander verglichen werden können.

Die Effektstärke, ES, ist üblicherweise definiert und operationalisiert als (vgl. Reinecker 1996)

$$ES = (M_e - M_k) / S_k$$

wobei:

M_e = Mittelwert von T der Experimental(= Behandlungs-)gruppe i.d. Post-Messung

Mk = Mittelwert von T der Kontrollgruppe i.d. Post-Messung

Sk = Standardabweichung von T der Kontrollgruppe i.d. Post-Messung

T = Test-/Meßvariable (z.B. Klin. Test)

Die inhaltlichen Probleme - d.h. der "Preis", für den ein solcher Vergleich überhaupt möglich ist - sind nach dem oben Gesagten beträchtlich und betreffen hier insbesondere die genauere Spezifikation des "T". Doch sollen in diesem Abschnitt zunächst solche *inhaltlichen* Aspekte völlig außer acht gelassen werden (einige werden später noch kurz aufgegriffen). Vielmehr soll hier gezeigt werden, daß auch *formal* ES nur bedingt etwas über Effektivität aussagen kann - d. h. auch dann, wenn das "T" überhaupt nicht problematisiert wird.

Der Grund liegt darin, daß aus der obigen Formel hervorgeht, daß dann ES groß ist:

i) wenn $Me - Mk$ groß ist - das ist weitgehend wohl auch das, was man sich unter "Effektivität" vorstellt (sofern eben auch T dem hinreichend entspreche)

ii) wenn Sk klein ist - das aber hat fast nichts mit Effektivität zu tun, sondern hängt von Besonderheiten der Gruppe in Relation zu T ab

Insgesamt "mißt" ES somit eher statistische Diskriminierung zwischen zwei Gruppen, die das, was man sich inhaltlich klinisch unter "Effektivität" vorstellt, aber eben nur bedingt wiedergibt - ja ggf. sogar ins Gegenteil verkehren kann.

Hierzu ein Demonstrationsbeispiel mit einfachen, runden Zahlen.

Annahme: Man habe eine klinisch relevante Testgröße T, diese liege zwischen 0 (völlig störungsfrei) und 100 (schwerst gestört). Der Mittelwert von T in der Normalbevölkerung sei 40, die Standardabweichung 15.

Fall A: homogene, störungsspezifische Gruppe:

M (der störungsspezifischen Grundgesamtheit) sei 60 (wo auch der Postwert der Kontrollgruppe bleibe), s sei 5 (und bleibe es für die Postwerte der Kontrollgruppe);

die Therapie erbringe eine mittlere Verbesserung um 10 Punkte, d. h. $Me = 50$

Es gilt dann:

$$ES = (Me - Mk)/Sk = (50 - 60)/5 = 2,0 - \text{eine hohe Effektstärke !}$$

Fall B: heterogene, "Mischgruppe":

Wieder sei $M = 60$ (also im Mittel genau so gestört wie oben), s sei 15 (das spiegelt realistisch die Heterogenität wieder)

die Therapie erbringe eine mittlere Verbesserung um 20 Punkte, d. h. $Me = 40$

Es gilt dann:

$$ES = (Me - Mk)/Sk = (40 - 60)/15 = 1,3 - \text{eine deutlich geringere Effektstärke !}$$

De facto aber liegt in diesem Beispiel bei Fall B das M der Behandlungsgruppe nach der Therapie genau auf dem der Normalbevölkerung. Die Behandlungsgruppe ist somit nach der Therapie als "völlig normal" anzusehen.

In Fall A hingegen, trotz der um 50% höheren Effektstärke, liegt die Behandlungsgruppe 10 Punkte über (=schlechter) der Normalbevölkerung; Daraus folgt (sofern s gleich geblieben ist): über 97,5% der Behandlungsgruppe liegen hier über dem Mittel der Normalbevölkerung, sind also gestörter.

Fall A hat somit die deutlich höhere Effektstärke, Fall B aber die deutlich höhere Effektivität der Therapie !

Die inhaltlichen Zusammenhänge, in denen in der Literatur über "Effektstärke" gesprochen wird, lassen zweifeln, daß diese formalen Zusammenhänge bzw. Unterschiede hinreichend bekannt sind und daß bewußt ist, daß mit "Effektstärke" lediglich eine spezifische statistische Diskriminationsfähigkeit zwischen Gruppen gemeint ist. Diese ist im Fall A unbestritten höher - doch was haben die Therapeuten oder die Patienten davon ? Vielmehr wird meist so getan, als messe die Effektstärke tatsächlich und unbesehen einen therapeutischen Effekt. Es zeigt sich darüber hinaus, daß es falsch wäre, Effekte in heterogenen, störungs-unspezifischen Gruppen als unbedeutsamer oder weniger aussagekräftig abzutun: Zumindest in formal-quantitativer Hinsicht ist eher das Gegenteil richtig (da störungs-unspezifische Gruppen typischerweise eher eine größere als eine kleinere Varianz gegenüber homogenen Gruppen aufweisen).

Es zeigt sich somit auch an diesem Beispiel die bekannte Tatsache, daß die Genauigkeit einer Messung noch nichts über deren Sinnhaftigkeit aussagt. Allerdings sind Versuche, inhaltliche und gesellschaftliche Wertfragen als Fragen wissenschaftlicher Messung zu kaschieren, keineswegs neu. In einem anderen Beitrag über "Vermessene Qualität" (Kriz 1996) habe ich kürzlich nochmals in Erinnerung gebracht, wie sich beispielsweise Mitte des 19. Jahrhunderts zahlreiche - auch heute noch in Lehrbüchern als "renommierte Gelehrte" ausgewiesene - Forscher an dem Programm beteiligten, die damals gängigen rassistischen Ansichten "wissenschaftlich" zu unterstützen bzw. zu „beweisen“. Diese Beweise erfolgten auf der Basis von scheinbar harten objektiven Daten aus dem Bereich der aufblühenden naturwissenschaftlich orientierten Humanwissenschaften und unter Verwendung akribischer Meßmethoden. So veröffentlichte beispielsweise S. G. Morton mehrere Bände von exakten Meßwerten und Tabellen über Hirnvolumina. Vor dem Hintergrund der allgemein gängigen Annahme, daß die Hirngröße unmittelbaren Aufschluß über den Entwicklungsstand bzw. die angeborenen geistigen Fähigkeiten gebe, konnte er so wissenschaftlich präzise belegen, daß Weiße ganz oben, Indianer in der Mitte und Schwarze ganz unten in der Hierarchie stünden (bei Weißen übrigens Teutonen und Angelsachsen oben, Juden in der Mitte und Hindus unten). Ähnlich argumentierte der französische Chirurg und Anthropologe Broca (dem zu Ehren heute noch das motorische Sprachzentrum im Vorderhirn Broca-Zentrum benannt wird): Mittels verschiedener Schädel- und Körpermessungen und daraus konstruierter Indizes versuchte auch er, die angeborene Dummheit "minderwertiger Rassen" zu belegen: Für eine Stichprobe von 60 Weißen und 35 Schwarzen ergab sich dabei eine durchschnittliche Schädellänge hinter dem Hinterhauptsloch von 100,385 Millimetern für Weiße und 100,875 Millimetern für Schwarze, aber eine Schädellänge vor dem Hinterhauptsloch von 90,736 zu 100,404 Millimetern (man beachte das bemerkenswerte Verhältnis von inhaltlicher Relevanz und Meß-"Exaktheit"). Daraus schloß er 1872: „Es ist daher unbestreitbar..., daß der Körperbau des Negers in dieser Hinsicht wie in vielen anderen dem des Affen angenähert ist.“ (zit. n. Gould 1983). Ein anderes „Kriterium“ - das Verhältnis des

Unterarmknochens zum Oberarmknochen (!) - hatte er zuvor aufgegeben, weil die "richtige" Reihenfolge sich dabei nicht ergab. Es ist einsichtig, daß auch für die Verhältnisse innerhalb der weißen Rasse beliebte Vorurteile mit ebensolcher wissenschaftlichen Autorität vertreten und weitgehend durch „objektive Meßdaten“ belegt wurden.

Ist diese historische Verknüpfung von ideologischer Intention und "wissenschaftlich" präziser Quantifizierung wirklich so weit von der aktuellen Situation entfernt, wo nahezu beliebige "Kriterien" gewählt werden (z.B. Nachweis in "4 von 8" Störungsbereichen mit mindestens "3 störungsspezifischen" Studien, wobei bei "8" störungs-unspezifischen nur noch jeweils 2 erforderlich sind, etc.), um dann durch Abzählen eine scheinbare Objektivität vorzugeben, ohne daß die völlige Subjektivität der "Kriterien" auch nur irgendeinen Anlaß zu Selbstzweifeln gibt, während alle anderen in der Scientific Community anerkannten Daten und Belege für "Wissenschaftlichkeit" und Effektivität bewußt ignoriert werden?

3. Wie stellt man die Effektivität eines Psychotherapie-Verfahrens denn nun "richtig" fest?

Es ist offenbar nicht leicht, auf den ersten Blick die Tücken der beiden häufig hörbaren Fragen zu durchschauen - nämlich (a): "Wirkt eine bestimmte Psychotherapie überhaupt (und wenn ja: wie effektiv)?" und (b): "Wirkt Vorgehensweise A besser als Vorgehensweise B?" Die letztere Frage ist typisch für sog. Therapievergleichsstudien, bei der ersteren geht es um Therapieerfolg bzw. Evaluation (Bewertung einer Maßnahme). Beide Fragen klingen nämlich ganz "einfach und unschuldig" und scheinen üblichen naturwissenschaftlichen Fragestellungen ziemlich ähnlich zu sein - etwa: "Bewegt sich ein Objekt überhaupt (und wenn ja: mit welcher Geschwindigkeit)?", bzw.: "Bewegt sich Objekt A schneller als Objekt B?"

Doch der Ähnlichkeits-Schein trügt gewaltig: Während nämlich "Bewegung", "Geschwindigkeit" und "Geschwindigkeitsdifferenz" in der Physik für alle verbindlich operational (d.h. unmittelbar auf Beobachtungs-Handlungen bezogen) definiert sind, gilt dies für keinen der o. a. Begriffe in Bezug auf Psychotherapie.

Denn ein kleiner Blick in die Debatte klinischer Psychologie zeigt, daß keineswegs klar ist, wie man "Therapie-Erfolg" überhaupt definieren soll (bevor man sich überlegt, das Definierte dann zu messen). Hinsichtlich der inhaltlichen Ziele, was überhaupt erstrebenswert ist, herrscht bekanntlich keineswegs Konsens. So hat sich z. B. für das große Spektrum sog. "chronischer Krankheiten" - angefangen von schwerer Schizophrenie, Hirnschädigungen, Asthma, Diabetes, etc. - oder andere nicht (mehr) heilbare Krankheiten wie z. B. Krebs das Konzept der "Bewältigung" durchgesetzt, bei dem eine "Ausheilung" der Primärsymptome nicht als sinnvoller Therapie-Erfolg angestrebt werden kann, sondern eben eine Erhöhung der "Lebensqualität" und ein "erträglicher Umgang" mit der Krankheit. Doch wie sind diese - ebenfalls wohl nicht von jedem gleich verstandenen - Begriffe zu definieren?

Offenbar besteht zwischen der Frage, wie ein bestimmter Effekt möglichst wissenschaftlich präzise gemessen werden kann, und der Frage, welche Effekte aus wissenschaftlicher Sicht überhaupt zu erwarten sind und was sie bedeuten, ein großer Unterschied. Solange man beispielsweise davon ausgeht, daß sich die Sonne um die Erde dreht, lassen sich viele wissenschaftlich präzise Registrier- und Meßverfahren ersinnen, um den Weg der Sonne zu beschreiben und sogar um möglichst gute Vorhersagen zu machen. Es sei beachtet, daß selbst akzeptierte Effekte nicht unbedingt etwas über wissenschaftliche Begründbarkeit aussagen: So sind die Effekte von Akupunktur inzwischen mit

wissenschaftlichen Methoden belegt und weitgehend akzeptiert - aber es fehlt (meines Wissens) eine wissenschaftlich fundierte Theorie, die Akupunktur ins abendländische Paradigma einzureihen imstande wäre. Die Frage nach der Wissenschaftlichkeit berührt somit auch zentrale Grundlagen- und Grundsatzfragen.

Selbst dann, wenn man hinsichtlich der Konzept-Definitionen weit mehr Einigkeit hätte, als faktisch vorherrscht: Wie werden diese Konzepte nun konkret operationalisiert und gemessen, d. h. an welchen konkreten empirisch erfaßbaren Größen wird das Konzept denn festgemacht? Werden - wie meist üblich - Tests und Fragebögen eingesetzt, muß (mit Devereux 1992, S. 27) gefragt werden: Zu welchem Universum des Diskurses gehören überhaupt die erhobenen Daten? Sind Fragebogen-Daten eher als Meinungen oder klinische Diagnosen zu werten? Therapie-Forscher setzen "selbstverständlich" das Letztere voraus - worin und wann ist dies begründet?

Es sei beachtet, daß alle diese Fragen zwar erörtert, aber nicht geklärt werden können. Denn man kann hinsichtlich dieser Fragen zwar Entscheidungen herbeiführen, aber mit guten Gründen hätte man sich meist natürlich auch anders entscheiden können. Oft werden allerdings nicht einmal die getroffenen Entscheidungen als solche dokumentiert, sondern es wird so getan, als wären die Definition des Erfolgs, dann dessen Operationalisierung und dann die Messung inhaltlich so selbstverständlich, daß man nun eine formale Methode einfach "anwenden" könne.

Nachdem man diese Entscheidungen (explizit oder implizit) getroffen hat, taucht die nächste Gruppe an Fragen auf, über die entschieden werden muß (und auch diese Entscheidungen sind wieder keineswegs "selbstverständlich"): An welchen Patienten werden die Messungen überhaupt durchgeführt? Nimmt man an, daß die Therapiemethode A bei allen Patienten gleich wirkt? - Eher wohl nicht! Doch wonach wählt man aus? Reicht es, eine bestimmte "störungsspezifische" Gruppe" zu wählen? Und warum wird diese abgegrenzt - es ist nämlich wiederum keineswegs "selbstverständlich", daß unsere klinischen Klassifikationen, auf der Ebene von Krankheitsverläufen gewonnen, auch die optimalen Klassifikationen für Therapieverläufe sind. Geht es eher um möglichst homogene Symptome oder um eine homogene Schwere von Beeinträchtigungen oder um einen homogenen Verlauf der bisherigen Krankengeschichten oder um homogenen "sozialen Support" (der erheblich mit dem Therapie-Verlauf interagieren dürfte) oder...oder..? Auch dies sind nur vordergründig methodische Fragen, denn dahinter stehen unterschiedliche inhaltlich-theoretische Vorstellungen, für die man so oder so Partei ergreift, wenn man sich für das eine oder das andere entscheidet - und wieder ist eine Entscheidung etwas anderes als die algorithmisierte Durchführung einer Methode.

Welche Therapeuten von Richtung A läßt man nun die zu beforschenden Therapien durchführen? Eine möglichst repräsentative Auswahl? (Woher gewinnt man die Basisinformation dafür?) Oder nimmt man möglichst erfahrene Therapeuten? (Was heißt das - und: Wie realistisch ist es anzunehmen, daß erfahrene Therapeuten genau "Methode A" in "Reinkultur" vertreten?). Wählt man (was wohl realistisch ist) jene, die man gerade bekommen kann: Sind das dann nicht eher jene, die vielleicht unter Patientenmangel leiden (und vielleicht überzufällig Anfänger sind oder andere Eigenschaften aufweisen, die sie nicht gerade zu den Aushängeschildern für diese Therapieform machen)?

Wenn diese Fragen nun alle entschieden sind: Was bleibt, als kleine "Zwischenbilanz", eigentlich noch von der Ausgangs-Frage übrig, (z.B.) "ob und wie effektiv Therapiemethode A wirkt?" Was wir nach den bisherigen Entscheidungen erwarten könnten, wäre eine Information darüber, wie

Patienten der Gruppe X, nach den Kriterien C, D und E... zusammengestellt, hinsichtlich Antworten und Beobachtungen der Variablen F, G und H... bei Therapeuten mit den Eigenschaften I, J und K... sich verändern. Beantwortet das noch unsere obige Frage - bzw. ist diese Information noch als angemessene Basis anzusehen, um Aussagen über "die Therapieform A" zu machen?

Daß diese Probleme allzu wenig diskutiert werden, mag wohl daran liegen, daß als Metapher (oder gar Ideal) der komparativen Psychotherapie-Forschung immer noch die Pharma-Forschung angesehen wird, bei der eine Substanz A zu jeder Zeit und an jedem Ort nach einer bestimmten, algorithmusartigen Prozedur hergestellt werden kann und dann auch in derselben Weise wirken sollte. Dabei ist diese Vorstellung schon in der Pharma-Forschung eine Fiktion, da das Resultat der "Einwirkung" von A auch vom System abhängt und höchstens für isolierte, möglichst monoklonale Zellen im Reagenzglas ein zutreffendes Modell sein mag. Daher sind selbst in der Pharmaforschung ergänzende Feldstudien eine wichtige Gepflogenheit. Bei realen Patienten, so weiß auch jeder gute Arzt, muß erprobt werden, ob und wie im Detail und in welcher Dosierung A wirkt oder ob nicht besser B verabreicht werden sollte (wobei diese Rangordnung der "Effektivität" zwischen A und B beim nächsten Patienten trotz gleicher Symptome genau umgekehrt sein kann). Daher müssen die Effekte meist über größere Gruppen gemittelt werden (mit allen Problemen, die nicht-lineare, rückgekoppelte Systeme aufwerfen - vgl. Kriz 1997). Vernachlässigt man auch noch Aspekte wie z.B. zirkadiane Schwankungen (d.h. die stark unterschiedliche Ansprechbarkeit eines Organismus auf Wirkstoffe in Abhängigkeit vom Tageszyklus) und zahlreiche andere Faktoren, so kann man unter allen diesen Einschränkungen im Pharma-Bereich von der Wirkungsweise von A sprechen.

Was kann es aber bedeuten, die "Wirkung von Psychotherapie-Methode A" zu untersuchen? Wenn man A im Kontext der experimentellen Psychologie auf wenige, klar umrissene Handlungen reduziert oder A über exakte Manuale definiert, die dann genau befolgt werden, mag dies mit den o. a. Einschränkungen ähnlich wie im Pharma-Bereich gelten. Dies setzt allerdings voraus, daß der Experimentator möglichst wenig Entscheidungsfreiheit hat, von den vorher festgelegten Bedingungen abzuweichen. Zwar könnten prinzipiell auch Entscheidungsbäume vorgegeben werden; aber erstens reichen dafür unsere inhaltlich-theoretischen Kenntnisse nicht aus, zweitens würde das Design viel zu kompliziert, varianzanalytisch aufgefüchert und damit massenstatistisch kaum mehr untersuchbar werden (s. u.).

Der Gegenstandsbereich, über den Aussagen zu treffen wären, bezieht sich aber zum größten Teil nicht auf Situationen, die durch solche experimentellen Labordesigns abgebildet werden. "Psychotherapie-Richtung A" wird in der Realität üblicherweise durch Therapeuten repräsentiert, die ihre Therapie "A" von unterschiedlichen Ausgangsbedingungen (Lebens- und Bildungs(ver)läufen), in unterschiedlichen Ausbildungsstätten und mit einer erheblichen Varianz an theoretischen und praktischen Vorstellungen erlernt haben und sich (spätestens!) danach (hoffentlich!) auch um andere Konzepte und Vorgehensweisen von Therapie gekümmert haben - kurz: ihre therapeutische Kompetenz und Vorgehensweise verändert haben. Was also kann "Therapie A" mehr bedeuten als eine sehr heterogene Menge an Therapeuten, die mit heterogenen Kompetenzen, Vorstellungen und Vorgehensweisen sich einem therapeutischen Ansatz "A" zuordnen (genauer: dem meist zusätzlich ebenfalls heterogenen Ideen-Konglomerat unter der Bezeichnung "A"). Manuale für bestimmte Ansätze sind möglich und werden auch verwendet - aber man kann nicht das Leben in ein Labor sperren, nur weil die Forschung dort etwas artefaktärmer verläuft. Und selbst bei der Verkümmern von Psychotherapie zu einer sklavisch-beschränkten Manualvollstreckung am Patienten wird es "Therapieform A" nie so rein geben können wie die "Pharma-Substanz A". Eine äquivalente

Metapher wäre daher vielmehr die folgende: Ein Pharma-Konzern schickt viele Menschen mit recht unterschiedlichen Vorstellungen über die Wirkung von Heilkräutern in unterschiedlichen Regionen mit dem Auftrag, dort nach heilenden Kräutern zu suchen, diese (nach weit auslegbaren Vorschriften) zu verarbeiten und letztlich und vorgefertigte Fläschchen mit der Aufschrift "A" abzufüllen. Selbst wenn man nun einige dieser Fläschchen mit der Aufschrift "A" (und im Detail sehr unterschiedlichem Inhalt) mit der gesamten ausgeklügelten Methodik der Pharma-Forschung untersuchen würde: Jeder Konzern müßte zu Recht einen Prozeß fürchten, wenn er nun Werbung für die wissenschaftlich geprüfte Wirkung des Medikamentes "A" behaupten würde. Ebenso aber ist die Formulierung: "Therapieform A bewirkt..." faktischer Unsinn, bzw. Etikettenschwindel.

Dieses Problem verschärft sich noch, wenn eine Therapieform aus einem sehr großen Spektrum an überaus unterschiedlichen Einzelvorgehensweisen oder -verfahren besteht, was z. B. für "die Verhaltenstherapie" gilt (vgl. Margraf 1996, Linden & Hautzinger 1996): Systematische Desensibilisierung, Aversionstherapien, Biofeedbacktherapie, Reizkonfrontation, Stimuluskontrolle, Token-Programm, Selbstinstruktionstraining, Training sozialer Kompetenz, Rational Emotive Therapie (um nur einige zu nennen) entstammen nicht nur recht heterogenen theoretischen Erklärungskontexten von Störung und Veränderung, sondern dürften wohl kaum alle kompetent von einem einzelnen Verhaltenstherapeuten beherrscht werden (und manualisierte Vorgehensweisen haben nochmals andere Gewichtungen). Dazu sind spezifische Vorgehensweisen für spezifische Störungsbilder untersucht und für wirksam befunden worden. Doch obwohl dies forschungslogisch und für die weitere Entwicklung auch aus meiner Sicht durchaus positiv zu bewerten ist, hat dies für die hier diskutierte "Wissenschaftlichkeits"-Debatte eine gravierende Schwäche: Wenn ein Verhaltenstherapeut einen Patienten mit Störungsbild "S" nicht mit jener Vorgehensweise behandelt, für die experimentell der Wirksamkeitsnachweis erbracht wurde (weil er eben nicht alle ausführen kann - oder sogar guten Grund hat, bei diesem Patienten anders vorzugehen), über welche globale Wirksamkeit "der Verhaltenstherapie" wird dann eigentlich noch "als wissenschaftlich gesichert" gesprochen? Wählen andererseits Therapeuten *immer* jene, in einer Untersuchung belegte, Vorgehensweise, so kann man fragen, wieso aus einer massenstatistischen Wahrscheinlichkeitsaussage für deren Überlegenheit dann eine faktisch 100%- zu 0%-Entscheidung gegen andere verhaltenstherapeutische Vorgehensweisen folgt. Ein Dilemma, das sich nicht lösen läßt. Auch dieses Argument zeigt, daß die Globalfrage zur "Wissenschaftlichkeit" einer Therapieform seriöserweise nicht beantwortet werden kann, weil nicht klar ist, worauf sich die praktisch überhaupt beziehen soll.

Dabei hätten wir nun noch fast jene Fragen vergessen, um die in der Literatur über Psychotherapie-Forschung gemeinhin so viel Aufhebens gemacht wird: Soll man einen Vergleich "vorher"/"nachher" (d.h. ein pre-post-design) durchführen oder aber eine Verlaufsstudie (panelanalyse), bei der zu vielen Zeitpunkten die Information erhoben wird? Wie soll eine Kontrollgruppe zusammengesetzt sein, die den Therapie-Erfolg gegen Zufallseinflüsse bzw. Spontanremission sichert (d.h. gegenüber der Tatsache, daß sich auch bei einem Teil der unbehandelten Patienten nach einer Zeit "Besserung" einstellt)? Wie groß soll die Stichprobe an Patienten sein usw.? Doch die in diesem Absatz skizzierten Fragen sind vergleichsweise zu den vorher aufgeworfenen trivial, da es sich hier tatsächlich eher (wenn auch nicht vollständig) um statistisch-methodische Fragen handelt, bei denen die Kriterien daher (vergleichsweise) klar sind und deren adäquate Beantwortung zum Handwerkszeug des Forschers gehören sollte.

Weit schwieriger und unklarer ist es da schon wieder, statt die Wirkung *einer* Therapieform A untersuchen zu wollen die Frage nach einem Therapie-Vergleich zu stellen. Denn nun müssen nicht nur sämtliche oben aufgeführten Entscheidungen (mit den dahinter liegenden theoretischen Vorstellungen und Problemen) getroffen werden. Sondern zusätzlich ist natürlich zu entscheiden, hinsichtlich welcher Aspekte (bzw. welches Aspekte-Raumes) ein solcher Vergleich überhaupt stattfinden könnte. Erinnern wir uns: Die obigen Entscheidungen führten sinnvollerweise zu einer, für A ganz spezifischen Auswahlkombination aller dieser Entscheidungsaspekte. Wenn dieselbe Entscheidungssequenz analog für Therapieform B durchgeführt wird, bleiben dann überhaupt noch irgendwo Überschneidungen, um über einen Vergleich etwas aussagen zu können?

Nun, wer nur statistische Modelle im Kopf hat und von klinischer Psychologie nichts versteht, könnte auf die Idee kommen, alles per Zufall auszuwählen - d. h. Patienten werden per Zufall auf zufällig ausgewählte Therapeuten der Schulen A und B aufgeteilt und die Veränderungen werden mit zufällig gezogenen Kategorien, Tests etc. bestimmt. Abgesehen davon, daß diese Vorgehensweise fernab jeder praktischen Realisierung liegen würden, gäbe es zwei Möglichkeiten: Entweder man hat ein vollständiges Design, in dem alle Kategorienkombinationen vertreten sind, oder man führt tatsächlich nur eine Zufallsauswahl durch. Im ersten Fall muß man nur ein Dutzend der oben aufgeworfenen Entscheidungsfragen (die keineswegs Anspruch auf Vollständigkeit erheben!) mit je 5 Abstufungen bzw. Alternativen berücksichtigen, um die Absurdität zu erkennen: Es ergeben sich 512, das sind über 244 Millionen, Kombinationen. Man müßte also eine "Stichprobe" vom Umfang mehrerer europäischen Staaten erheben, um pro Kombination auch nur jeweils einen einzigen Menschen zu untersuchen. Im zweiten Fall wären die Untersuchungs- (und Kontroll-)gruppen zwar klein - jedoch was sagt es aus, wenn Therapieform A bei anderen "Störungen" etc. hinsichtlich anderer Kriterien "besser" abschneidet als Therapieform B? Dies wäre der Aussage äquivalent, daß Geigenunterricht effektiver ist als Klavierunterricht - nachdem man einige Kinder, die ein Instrument lernen wollten, zufällig auf verschiedene Lehrer an unterschiedlichen Geigen- und an Klavierschulen aufgeteilt hätte und nach einem Jahr die Eltern, mit ihren unterschiedlichen Vorstellungen darüber, was überhaupt der Sinn des Musikunterrichtes sein könnte, diese "Erfolge" beurteilen ließe. (Und auch hier wird, nebenbei bemerkt, das "Ergebnis" zwar "objektiver", aber nicht sinnvoller, wenn man mit großem Meßaufwand die Geschwindigkeit des Fingerbeugens, die Abweichungen von der Idealfrequenz jedes gespielten Tones etc. genau erhebt und eine aufwendige Statistik darüber stülpt - obwohl beide genannten Variablen sicher auch etwas mit dem Erfolg zu tun haben).

Diese eher anekdotische Schilderung einiger Probleme und Schwierigkeiten beim direkten Vergleich nur zweier Therapie-Ansätze hat wohl deutlich werden lassen, welche immensen theoretischen und konzeptionellen Defizite aufzuarbeiten wären, bevor die obigen Fragen einigermaßen kompetent entschieden werden können und man "loslegen" könnte, herumzurechnen. Es sollte übrigens auf der Hand liegen, daß es da nicht unbedingt leichter ist, Meta-Analysen durchzuführen, d.h. Analysen, in die "Ergebnisse" zahlreicher anderer Studien eingehen, um zu einer Gesamtaussage zu kommen. Das Wort "Ergebnisse" wurde deshalb in Anführungszeichen gesetzt, weil eben jedes "Ergebnis" einer solchen Einzelstudie implizit oder (selten) explizit die gesamte Entscheidungssequenz über alle oben aufgeworfenen Fragen (und weitere) enthält.

Hier stellen sich zusätzlich Fragen wie: Welche Studien werden einbezogen? Wie begründet sich die Auswahl? Wie wird mit unterschiedlichen Stichproben, mit unterschiedlicher Anzahl von Erfolgskriterien umgegangen (d.h. ist eine Studie an 50 Personen genau so aussagekräftig wie eine an 1000, und ist eine Studie mit nur einem klaren Effekt nur 1/10 so wichtig zu nehmen wie eine andere

mit 10 nachgewiesenen Effekten)? Wie werden Abhängigkeiten berücksichtigt (sind 10 fast gleiche Kriterien 10 mal so wichtig wie nur eines in einer anderen Studie)?

Wir wollen es bei diesen Fragen bewenden lassen. Es zeigt sich, daß die Frage nach der Wirksamkeit oder gar nach dem Vergleich keineswegs einfach beantwortet werden kann. Zudem gibt es neben den aufgezeigten Ebenen noch viel grundlegendere - z.B. welche Modelle von "Ursache-Wirkung" wir unseren Überlegungen zugrunde legen. Wenn wir ernst nehmen, was uns die moderne naturwissenschaftlich fundierte Systemtheorie lehrt, dann sind in komplex vernetzten Systemen diskontinuierliche Verläufe zu erwarten - und Psychotherapieforscher reden auch gern davon, daß die unterschiedlichen Faktoren, die den Erfolg einer Therapie beeinflussen, miteinander vernetzt sind. Daraus würde aber folgen, daß z. B. keine Gruppenvergleiche derart durchgeführt werden dürften, daß die Veränderungen auf Datenniveau aggregiert (d.h. zusammengefaßt - etwa über Mittelwertbildung) werden. Je nach individuellem Ausgangspunkt (der zu erheben wäre) kann nämlich ggf. eine "große Ursache" dann keine Wirkung, eine "kleine Ursache" eine große Wirkung entfalten. Auch solche Fragen müßten natürlich vorher explizit geklärt werden, bevor man nach irgendwelchen "statistischen Verfahren" etwas "berechnet" (und damit natürlich unwissentlich in dieser Frage eine bestimmte Stellung bezieht - meist jene, daß man nicht von diskontinuierlichen Verläufen und damit, entgegen den verbalen Beteuerungen, auch nicht von vernetzten Prozessen ausgeht). Solange wir in der Erforschung und wissenschaftlichen Debatte, was überhaupt unter Psychotherapie-Effekten zu verstehen ist, soweit am Anfang stehen, kann die großangelegte und wissenschaftspolitisch brisante "Sammlung" von Effekten eben nur Effekt-Hascherei sein.

4. Alternativen zu "vermessenen" Standards/Entscheidungen

Wird damit dann alles beliebig - d. h. können wir keine Aussagen zur Wissenschaftlichkeit und Wirksamkeit von Psychotherapieverfahren machen? Dies stände nur zu befürchten, wenn Wissenschaftler meinen, "im Rechenraum" bleiben zu müssen (um das Bild oben nochmals zu verwenden), statt diesen schleunigst zu verlassen, um sich die inhaltliche Kompetenz wieder anzuziehen.

Wenn wir unter "wissenschaftlich" beispielsweise auch verstehen,

- mit wesentlichen Theorien eines Faches in Einklang zu stehen,
- von der scientific community akzeptiert zu werden,
- bestimmte fachspezifische aber auch ansatzspezifische Kriterien zu erfüllen,

dann brauchen wir uns nicht allein auf sog. "objektive" Daten - mit allen aufgezeigten Problemen - berufen. Vielmehr gibt es auch Therapieansätze, die seit vielen Jahren klinisch gut erprobt sind - hier müßten wir nur die entsprechenden Klinikchefs befragen. Die Tatsache, daß mehrere zehntausend Patienten mit einem Ansatz in den letzten Jahrzehnten erfolgreich und mit Abrechnung durch die Kassen behandelt wurden (was z. B. für die vom "Beirat" nicht "wissenschaftlich anerkannte" Gesprächstherapie gilt), ist zwar nicht unbedingt ein Beweis für "Wissenschaftlichkeit", aber doch auch nicht irrelevant. Immerhin könnten Wissenschaftler diese Fakten systematischer erheben und damit noch stärker zum Gegenstand des wissenschaftlichen Diskurses machen. Die Effektivität einer manualisierten Anwendung ist schließlich auch noch kein Beleg für die "Wissenschaftlichkeit" dieses Verfahrens selbst, sondern eben nur der Nachweis der (Labor-)Effektivität mit

wissenschaftlichen *Methoden*. Und zu den wissenschaftlichen Methoden kann genau so gut auch die empirische Erhebung der Heilerfolge außerhalb des Labors gehören.

Es war ein mutiger Schritt von hoher Brisanz, als 1995 Martin E. P. Seligman mit einer groß angelegten Studie in den USA mit der von ihm geleiteten Consumer-Report-Studie den Aspekt der Labor-Wirksamkeit (efficacy) durch den Aspekt der klinischen Brauchbarkeit und Bewährtheit (effectiveness) bereicherte, indem das breite Spektrum realer therapeutischer Praxis untersucht wurde. Wie von vielen zuvor vermutet, zeigten diese Ergebnisse über die Wirksamkeit von Psychotherapie in der realen Praxis ein deutlich anderes Bild als die efficacy-Studien: Die Consumer-Report-Daten belegten für längere Behandlungen durchaus einen höheren Besserungsgrad als für kurze; Psychotherapie wurde durch Medikamente (in der Regel) nicht verbessert; keine Therapieform konnte allgemein oder mit Hinblick auf bestimmte Störungsbilder eine höhere Wirkung als andere nachweisen; Psychologen, Psychiater und Sozialarbeiter waren gleich effektiv - aber deutlich besser als Eheberater und (nicht fachtherapeutische) Ärzte. Besonders bedeutsam ist, daß Seligman einen vehementen Artikel für die Bedeutsamkeit der effectiveness-Studien schrieb, worin er offen bekannte, daß er zuvor selbst die efficacy-Kriterien für den einzigen "Gold-Standard" hielt, die 1994 nochmals durch die APA (American Psychological Association) formuliert worden waren. Nach seiner Kritik dieser Kriterien als zu einseitig (Seligman 1995) wurde er 1997 zum Präsidenten der APA gewählt (Amtszeit 1998) - was insofern bedeutsam ist, als in der gegenwärtigen deutschen Debatte mit Verweis auf die Bedeutung der APA fast ausschließlich die efficacy-Kriterien von 1994 referiert werden. Die späteren Entwicklungen in der APA aber, die nicht so gut ins Bild einer rigorosen efficacy-Ideologie passen, werden dabei kaum erwähnt.

Eine solche Untersuchung zur Gewinnung umfangreicher Daten über die Wirksamkeit von Psychotherapie im Spektrum alltäglicher Praxis aus der Sicht der Betroffenen sollte auch in Deutschland nachvollzogen werden. Zumindest theoretisch ist denkbar, daß diese Daten flankierend um exemplarische Erhebungen von Kassendaten ergänzt werden, um Kenntnisse über die Wirksamkeit auf weitere Aspekte auszudehnen - etwa: Medikamentenverbrauch, Krankenhaus-, Kur- und Fehlzeiten, Kosten für (nicht-psychotherapeutische) Arztbesuche etc. jeweils vor und nach den Psychotherapien. Ich weiß nicht, ob und wie schnell wir an die nötigen Daten kommen - der Datenschutz wird hierzulande ja sehr ernst genommen, wofür wir eben auch Preise zu zahlen haben.

Eine weitere Entscheidungsgrundlage wäre, ein fundiertes Delphi-Verfahren in Gang zu setzen, bei dem eine Erhebung bei den Fachprofessoren (denn diese definieren letztlich, was Wissenschaft ist!) erfolgt, deren Ergebnisse in mehreren Schritten rückzumelden und zu revidieren sind. Dabei müssen dann z. B. die Abweichungen von den Modalwerten der Rückmeldungen in den nächsten Schritten von jenen sorgsam begründet werden, die bei dieser Abweichung bleiben wollen. Durch ein solches Vorgehen würden wir Wissenschaftler dokumentieren, daß wir unsere Entscheidungskompetenz und -befugnis ernst nehmen und diese nicht allein an mechanisierte Auswahlprozeduren delegieren, die nicht selten aufgrund wissenschaftstheoretisch überholter Methoden mit ansatz-spezifischem Bias behaftet sind.

Das würde allerdings bedeuten, daß klinische Professoren schon jetzt die Verantwortung für das, was sich als Wissenschaft darstellt, ernst nehmen und sich nicht allein von statistischen Maßzahlen aus artifiziellen Untersuchungen leiten lassen. Ich bin ganz sicher: Wenn wir unseren eigenen Sachverstand, unsere Lehrbücher, das, was wir unterrichten und die Vielfalt der mit unterschiedlichen Methoden erbrachten Ergebnisse wirklich ernst nehmen, kann ein breiter Konsens

gefunden werden, in dem sich dann weit mehr als nur die beiden Richtlinienverfahren als wissenschaftlich fundiert und klinisch erprobt herausstellen. Es ist vielleicht auch nicht verboten, die Erfahrungen aus Nachbarländern mit einzubeziehen und nicht der Fiktion zu unterliegen, an den nationalstaatlichen Grenzen Deutschlands hörten der Sachverstand und die Wissenschaft auf. Die Ignoranz gegenüber Erfahrungen und Ergebnissen anderer Wissenschaftler als derjenigen des derzeitigen deutschen Mainstreams wäre jedenfalls kein Beleg für "Wissenschaftlichkeit".

Langfristig sollten wir als Wissenschaftler allerdings auch das Ziel (weiter) verfolgen, durch Ausarbeitung einer möglichst umfassenden Psychotherapie-Theorie unter Einbeziehung all dessen, was unsere Wissenschaft vom Erleben und Verhalten zu bieten hat (und einige Nachbarwissenschaften), einerseits bei möglichst vielen der praktisch vorfindlichen Therapiemethoden klar wissenschaftlich rekonstruieren und begründen zu können, warum wann was wie wirkt bzw. warum (und wann etc.) etwas nicht so wirkt, wie behauptet. Bis wir diesem Ziel aber einen bedeutsamen Schritt nähergekommen sind, wäre es gegenwärtig redlich, zu akzeptieren, daß therapeutische Handlungskompetenz sich historisch in heterogenen Schulen entwickelt hat - bei denen viele Gründer und spätere Mitarbeiter allerdings keineswegs, wie oft abfällig in der öffentlichen Polemik suggeriert wird, idiosynkratisch-sektenhafte Eigenbrötler mit rational nicht begründbaren Vorlieben waren, sondern durchaus gut ins Wissenschaftssystem integrierte Forscher, oft mit Professorenrang, deren Vorstellungen klar und explizit am wissenschaftlichen Fachdiskurs teilhatten und teilhaben und das heutige Spektrum wissenschaftlicher Kenntnisse der klinischen Psychologie wesentlich bereichert haben.

5. Pluralität - ein Kennzeichen von Wissenschaft

Nun wird in der Debatte um die Wissenschaftlichkeit und um eine fortschreitende weitere Verwissenschaftlichung von Psychotherapie oft ins Feld geführt, daß die Vielfalt der Therapie-Richtungen eher als ein Manko zu werten sei, und es ein wichtiges Ziel wissenschaftlicher Bemühungen sein sollte, zu einer einheitlichen Psychotherapie zu kommen. Was ist davon zu halten?

Meines Erachtens wird bei dieser Forderung der Aspekt einer rekonstruierenden Beschreibung (und damit teilweise Erklärung) von Zusammenhängen mit dem Aspekt einer präskriptiven Handlungsanweisung vermengt. Diese Vermengung ist leider gerade im Bereich der Psychotherapie nicht untypisch - etwa wenn in der Gesprächspsychotherapie die drei sog. Basisvariablen (Akzeptanz, Empathie, Kongruenz), die als wichtige beschreibende Aspekte durchaus fruchtbar sind, in der Therapeutenausbildung trainiert werden. Dies ist nämlich so fragwürdig wie ein Training der Wangenmuskulatur, nur weil man empirisch beobachten kann, daß Menschen, die sich gerade freuen und glücklich sind, die Mundwinkel nach oben ziehen. Und ein Training, um möglichst "echt" zu wirken, steht in geradezu paradoxalem Gegensatz dazu, "echt" zu sein. Entsprechend teile ich zwar das Bestreben, zu einer möglichst einheitlichen, umfassenden (rekonstruktiven) Theorie therapeutischen Geschehens zu kommen - aber als praktische Tätigkeit kann ich mir eine Zukunft der Psychotherapie nur als ein Spektrum unterschiedlicher Ansätze vorstellen, das die Heterogenität menschlichen Lebens widerspiegelt (s.u.). Und selbst die Vorstellung einer einheitlichen Theorie muß nochmals relativiert und differenziert werden - besonders wenn man grundsätzliche Fragen von "Wissenschaftlichkeit" im Auge hat.

Während man nämlich im vergangenen Jahrhundert noch vielfach daran glaubte, Wissenschaft könnte den Anteil letztlich "wahrer" (und kulturunabhängiger) Erkenntnisbausteine ständig vergrößern, woraus sich dann nach und nach die eine große, wahre Welttheorie zusammenfügen ließe, ist die überwiegende Mehrheit der Wissenschaftstheoretiker heute deutlich bescheidener geworden. Zwar dient in der "Mutter" aller Naturwissenschaften, der Physik, eine einheitliche Welttheorie vielen immer noch als prinzipielle Leitidee. Aber selbst in der Physik - zumindest im Verständnis ihrer führenden Vertreter - hat man längst die Idee aufgegeben, eine solche Theorie könne zu einer objektiven, vom Menschen unabhängigen Ordnung dieser Welt vordringen. Vielmehr ist selbst dort akzeptiert, daß immer nur spezifische Antworten auf spezifische Fragen möglich sind. Beispielsweise wird man auf "Teilchen"-Fragen andere Antworten erhalten als auf "Wellen"-Fragen, so daß allein schon wegen der Fragen-Relativität Wissenschaft nicht von Kultur getrennt werden kann. Dies hat der Physik-Nobelpreisträger Werner Heisenberg bereits 1955 wie folgt klar herausgestellt: "Wenn von einem Naturbild der exakten Naturwissenschaften in unserer Zeit gesprochen werden kann, so handelt es sich eigentlich nicht mehr um ein Bild der Natur, sondern um ein Bild unserer Beziehung zur Natur." Die Physik hat sich mit dieser Bescheidenheit sehr gut abgefunden. Niemand käme auf den Gedanken, den Physikern deswegen Unexaktheit oder gar Unwissenschaftlichkeit vorzuwerfen, weil sie nicht klären wollen und können, ob die Welt nun "wirklich" aus Wellen oder aus Teilchen besteht.

De facto ist die Entwicklung selbst in den "harten" Naturwissenschaften und trotz der Sehnsucht nach einer Einheitstheorie gerade entgegengesetzt dazu verlaufen, wie es die Gegner therapeutischer Vielfalt "im Dienste der Wissenschaft" gern hätten: Die historisch frühen Wissenschaftsbereiche Physik und Chemie haben sich im Laufe der wissenschaftlichen Entwicklung nämlich nicht etwa vereint (auch wenn z.B. unter manchen Perspektiven chemische Vorgänge durch physikalische Erkenntnisse besser verstanden werden). Vielmehr entwickelten sich immer mehr eigene Disziplinen: physikalische Chemie, Biologie, Biochemie, Molekularbiologie usw. Man könnte hier geradezu von "Schulbildung" sprechen. Solche Physiker, die sich als Anhänger einer Einheitswissenschaft verstehen, könnten dies zu Recht bemängeln und beispielsweise anderen Wissenschafts"schulen" mangelnde Professionalität vorwerfen, indem sie argumentieren, es sei reine "Konfession", daß diese Biochemie oder Molekularbiologie betrieben, wo man doch besser alles auf Physik zurückführen solle. Ein solches Argument träfe sogar durchaus den Kern des Problems: Es sind nämlich in der Tat meist persönliche Vorlieben, Fähigkeiten, Interessen, Lebenswege etc., die jemand zum Biochemiker oder aber zum Molekularbiologen werden lassen - und weniger rational-quantitative Analysen der Frage, welche der "Schulen" ihre Probleme effektiver oder wissenschaftlicher löst. Daher ist für Naturwissenschaftler die Frage auch relativ uninteressant, ob Biochemie bessere Erklärungen eines bestimmten pathologischen somatischen Prozesses als Molekularbiologie liefert (und Entsprechendes gilt sogar für weitere Unterdisziplinen): Man würdigt dort eher die unterschiedlichen Zugänge, versucht ggf. sich von den Erkenntnissen für die eigene Arbeit befruchten zu lassen und pflegt ansonsten maximale Koexistenz.

Verlassen wir nun den Phänomenbereich mikroskopisch-materieller Prozesse und wenden uns dem Phänomenbereich von solchen makroskopischen psychischen und interaktiven Prozessen (einschließlich deren somatischen Korrelaten) zu, welche vor allem Übergänge von unauffällig zu pathologisch und umgekehrt im Fokus haben - kurz: dem Phänomenbereich von Psychotherapie. Auch hier haben sich in diesem Jahrhundert unterschiedliche Schulen gebildet. Die Gründe dafür sind zunächst im Kern dieselben wie für die eben skizzierte Differenzierung in Wissenschaftsdisziplinen: Es gibt unterschiedliche Vorlieben, Fähigkeiten, Interessen, Lebenswege, Menschenbilder etc., die

jemand zum Psychoanalytiker oder aber zum Gestalttherapeuten werden ließen. Die Gesamtsituation ist aber im Bereich der Psychotherapie komplexer: Während es für ein Lichtquant gleich gültig ist, ob es im Rahmen der Wellen- oder der Teilchenphysik untersucht wird und während es für eine pathologische Körperzelle gleichgültig ist, ob sie biochemisch oder molekularbiologisch behandelt wird, sind den Patienten ihre Behandlungen nicht gleichgültig und damit auch deren gesamter Kontext einschließlich der Ergebnisse nicht gleich gültig. Im Bereich der Psychotherapie kommt somit nochmals ein mindestens ebenso großes Spektrum an unterschiedlichen Vorlieben, Fähigkeiten, Interessen, Lebenswegen, Menschenbilder etc. seitens der Behandelten hinzu. Diese ggf. zu ignorieren (und Patienten analog zu Gegenständen der Physik oder zu einem Zellhaufen zu sehen), dokumentiert vielleicht *ein* bestimmtes Menschenbild von Behandlern, löst aber dieses Heterogenitätsproblem nicht. (Und in Ignoranz sollte zudem nicht das herausragende Zeichen von "Wissenschaftlichkeit" gesehen werden.)

Ich vermag somit in der Heterogenität psychotherapeutischer Ansätze prinzipiell nichts Verwerfliches zu sehen, da sie nur eine Widerspiegelung der Heterogenität von Lebens"weisen" sowohl seitens der Behandler als auch der Behandelten darstellt. Die Konfession, sich vorzugsweise für Verhaltenskontingenzen oder aber für Sinndeutungen des Lebens zu interessieren, sagt ebenso wie die Konfession, sich vorzugsweise für biochemische oder für molekularbiologische Abläufe zu interessieren, wenig über Professionalität aus: Den jeweiligen Interessen kann man sowohl höchst dilettantisch oder gar unseriös wie auch sehr professionell nachgehen. Hier eine bipolare Beziehung zwischen Konfession und Profession konstruieren zu wollen, mag zwar eine gewisse Suggestivkraft entfalten, kann aber letztlich nicht überzeugen.

Die sinnvolle Heterogenität der Psychotherapieansätze macht zudem auch deutlich, warum die Frage nach der "effektivsten" Therapie ins Leere geht, denn was überhaupt erstrebenswert ist (das dann ggf. effektiv zu erreichen wäre), ist untrennbar mit den unterschiedlichen "Lebensweisen" der Menschen verbunden. So besteht kein Zweifel, daß Mais eine den Boden sehr effektiv ausnutzende Nahrungspflanze ist; ebenso ist eine Flachlandschaft wie der amerikanische Südwesten zweifelsfrei effektiver zu bearbeiten als eine Berg- und Seenlandschaft mit vielen engen Tälern. Trotzdem käme wohl niemand auf die Idee, die Schweizer Berge abzutragen, mit dem Erdreich die Seen aufzufüllen, um so eine Flachlandschaft herzustellen, auf der dann statt der blühenden Vielfalt an Kulturpflanzen eine Mais-Monokultur angelegt werden kann - obwohl diese fraglos effektiv für Großmaschinen wäre: Es geht aber eben nicht nur um Fragen der Effektivität - nach welchen "wissenschaftlichen Kriterien" auch immer vermessen -, sondern auch um die Grundfrage darüber, wie wir überhaupt leben wollen. Und ich bin nicht sicher, ob nicht selbst jene Wissenschaftler, die derzeit mit aller Macht die blühende Vielfalt der Therapieansätze "plattmachen" wollen, um möglichst nur noch einen einzigen "effektiven" Ansatz für alle zu verordnen, sehr bald in dieser geistigen Monokultur etwas vermissen würden.

Fragen der Lebensart stehen interessanterweise sogar im Zentrum jener Metapher, die meist fälschlich verwendet wird, um aus der Vielfalt mit Hilfe eines künstlich inszenierten Konkurrenzkampfes doch noch eine Monokultur zumachen. Verwiesen wird auf einen Wettlauf, der Lewis Carolls "Alice im Wunderland" entstammt, und bei dem etwa ein Dutzend Tiere beteiligt sind. Zitiert wird in der Diskussion dabei meist die Entscheidung des "Dodo-Bird" (dt.: Brachvogel), mit der dieser die Frage: "Aber wer ist Sieger?" entscheidet, nämlich: "Alle sind Sieger und jeder muß einen Preis bekommen" (vgl. Luborsky et al. 1975). Gebraucht wird diese Metapher häufig dafür, um die Absurdität der Aussage "alle sind Sieger" in Bezug auf die Therapierichtungen zu demonstrieren.

Schließlich leuchtet es in unserer auf Konkurrenz ausgerichteten Gesellschaft sofort jedem ein, daß keineswegs alle Sieger sein können. Zumal wenn dann die auch bei Caroll unmittelbar folgende Frage in den Fokus gerät: "Aber wer soll die Preise stiften?" rief ein ganzer Stimmenchor zurück." - Spätestens hier wird deutlich, daß es bei begrenzten Preisen (bzw. Krankenkassen-Budgets) eben darum zu gehen hat, daß manche einen Preis erhalten und andere nicht. Die Aussage des Dodo darf daher nicht hingenommen, sondern muß als Mangel wissenschaftlicher Differenzierungsfähigkeit gedeutet werden. Ohnedies erscheint die gesamte Organisation des Wettlaufs geradezu ein Alptraum an Inkompetenz für jede an objektiven Kriterien und Standards orientierte "wissenschaftliche" Psychotherapieforschung zu sein. So heißt es im Text: "Er legte zuerst die Rennbahn fest, eine Art Kreis ('auf die genaue Form kommt es nicht an' sagte er), und die Mitspieler mußten sich irgendwo auf der Bahn aufstellen, wie es sich gerade traf. Es gab kein 'Eins-zwei-drei-los!', sondern jeder begann zu laufen, wann er wollte..." (Caroll 1963, S.29). Hier scheint genau jener Dilettantismus zu herrschen, den nur präzise Effektivitätsmaße und genaue Vermessungen (z.B. der Rennbahn) eindämmen können.

Doch der Schein trügt. Verglichen mit manchen "wissenschaftlichen" Abhandlungen zur komparativen Wirksamkeit ist das Vorgehen des "Dodo-Bird" von tiefer Weisheit erfüllt. Der Kontext des Rennens ist nämlich der, daß die Tiere in einen Teich gefallen und tropfnaß und verdrossen waren: "Das Wichtigste war natürlich, wieder trocken zu werden" heißt es bei Caroll. Angesichts dieses Kontextes steht das Rennen plötzlich in völlig anderem Licht: Es ist nun wirklich egal, wie genau die Form einer Kreisbahn erreicht wird; gleichgültig ist auch, von wo die Tiere losrennen und wer wann durch welches Ziel geht. Wichtig ist es vielmehr, das gemeinsame Ziel zu erreichen: nämlich trocken zu werden. Und da kein Tier mit "wissenschaftlichen", aber völlig irrelevanten Kriterien über Bahnform, exakte Zeit- und Wegmessung und dergleichen das Geschehen behinderte (was vermutlich die Erkrankung etlicher Tiere zur Folge gehabt hätte), erreichten alle dieses Ziel, jeder hatte gewonnen und daher durchaus auch einen Preis verdient!

Mich hat immer schon gewundert, warum die Weisheit des "Dodo-Bird" in der Psychotherapie-Debatte so geflissentlich übersehen wird (wobei die Frage außer Acht bleiben soll, ob zumindest die Kontexte der Psychotherapie-Verfahren sorgfältiger rezipiert wurden als der Kontext des Rennens). Sind die unterschiedlichen Psychotherapie-Ansätze nicht eher mit den unterschiedlichen Tieren von "Alice im Wunderland" vergleichbar als mit Formel-1-Wagen, bei deren Rennen ein Sieger hinreichend objektiv ermittelt werden kann? Wie Carolls Tiere weisen auch die Therapierichtungen eine große individuelle Unterschiedlichkeit auf; und sie sind wie diese historisch zu sehr verschiedenen Zeitpunkten und von recht unterschiedlichen Standpunkten aus gestartet. Im Gegensatz zu Formel-1-Wagen lag der tiefere Sinn ihres Entstehens und ihrer Weiterentwicklung nicht darin, als Sieger aus irgendeinem Rennen hervorzugehen, sondern der Vielfalt menschlichen Lebens und dessen Beeinträchtigungen besser gerecht zu werden, als jede einzelne Richtung es könnte. Und das gemeinsame Ziel der unterschiedlichen Menschen, die zu Therapeuten aus unterschiedlichen Therapierichtungen gehen, ist es, "trocken" zu werden, d.h. eine ihren spezifischen Lebensvorstellungen und -umständen entsprechende Weise zu finden, mit der sie weniger leidvoll diese je unterschiedlichen Lebenswege meistern können.

Darüber hinaus sollte auch nicht unerwähnt bleiben, daß "Psychotherapie" ein weitaus breiteres Spektrum an Antworten bietet als das, worauf sich der Fokus in der gegenwärtigen Diskussion immer mehr verengt. Man mag nur an Aspekte wie Persönlichkeitsentwicklung, Individuation, Entängstigung u.a. denken, zu denen unterschiedliche Psychotherapieansätze durchaus Effektives beizutragen

hätten, deren Perspektive aber weiter reicht als das, was Krankenkassen und die dahinter stehende Sozialgemeinschaft sinnvollerweise bezahlen sollten. Denn es ist klar, daß nicht für alle Leistungen die Sozialgemeinschaft aufkommen kann und soll (unabhängig von der Frage, wie viele Milliarden die Sozialgemeinschaft z. B. zur Entsorgung des mit privatwirtschaftlichen Gewinnen verbundenen Atommülls aufwendet). Doch die wichtige und relevante Frage: "Wer soll die Preise stiften" führt auf ein ganz anderes Diskussionsfeld als die Frage nach der Wissenschaftlichkeit oder der Wirksamkeit von Psychotherapie - und die Vermengung dieser Fragen gehört zu den vielen Unredlichkeiten, mit denen die gegenwärtige Debatte belastet ist.

6. Quintessenz

Bei einer angemessenen Würdigung der hier zusammengetragenen Aspekte müssen wir uns von der Fiktion verabschieden, wir könnten und sollten in nächster Zeit auf wissenschaftlich redliche Weise entscheiden, welches Therapieverfahren das beste oder effektivste ist (bzw. dies für eine allzu kleine Anzahl von Therapieverfahren entscheiden). Die Hauptgründe sind, nochmals zusammengefaßt:

n es gibt nicht "das Therapieverfahren X", das sich wie eine Pharmasubstanz rein herstellen und in saubere, leere Flaschen (= un"voreingenommene" Therapeuten-Kandidaten) abfüllen ließe. Selbst bei hochgradig manualisierter Therapie (die ohnedies niemals Paradigma für alle Ansätze werden kann, ohne deren Wesen umzubringen) führen diese kompetente Menschen aus, die ihre je unterschiedlichen Begabungen und Teilkompetenzen auch nutzen wollen.

n Patienten sind - selbst nach reinen Störungsgruppen zusammengestellt - nur sehr artefiziell als "Störungsträger" bzw. "Manifestation einer Störung" anzusehen. Nicht nur, daß Klassifikationen und Ätiologien zwei völlig verschiedene Konzeptualisierungen sind; selbst eine "reine" Depression vom Typ 300.40 DSM III-R verläuft anders als z. B. eine Virusinfektion, ein Beinbruch oder ein Asthma (und selbst dort weiß ein guter Arzt um die Grenzen manualisierter Vorgehensweisen und wie stark andere, nicht so leicht meßbare und im Detail vorschreibbare ärztliche Kompetenzen gefragt sind). Weder lassen sich somit die für eine weitgehende Manualisierung notwendigen Regelverläufe vorhersagen (ohne die aber das Konzept einer "Therapieform X" wieder zerfließt) noch wird dies der Heterogenität an Lebensweisen und -bildern der Menschen, deren Beeinträchtigung 300.40 zuzuordnen wäre, auch nur annähernd gerecht.

n Es gibt derzeit keine Forschungsergebnisse, die seriös etwas über ein "Therapieverfahren" in der geforderten Globalität aussagen. Es liegen - gottlob - zahlreiche Ergebnisse zu zahlreichen Kontingenzen vor, die für Therapie als Wissenschaft von großem Wert sind und zur Qualität der Therapeutenausbildung viel beitragen können. Es sei auch nicht bestritten, daß unterschiedliche Richtungen in unterschiedlichem Ausmaß - vor allem aber auch: in unterschiedlicher Art und Weise! - Ergebnisse vorgelegt haben. Diese lassen sich aber nur sehr bedingt - und schon gar nicht reduziert auf einzelne Parameter - miteinander vergleichen. Insbesondere geben sie aber nur Auskunft über bestimmte Aspekte, die mit jeweils einer Richtung X zu tun haben - nicht über die "Richtung X" als ganze (was immer damit gemeint sein soll).

n Die Anzahl und die Art der Forschungsergebnisse sind keineswegs irrelevant - aber sie sind kein Gradmesser für "Wissenschaftlichkeit" oder gar "Effektivität" oder dergleichen. Die ganz überwiegende Zahl der Studien wurde nicht "von Richtungen" oder "Schulen" durchgeführt, sondern von Forschungseinrichtungen (i. W. Universitäten), für welche die Sozialgemeinschaft viele Milliarden

an Mitteln bereitgestellt hat. Die Ergebnisstrukturen spiegeln primär diese Wissenschaftsstrukturen wider und hängen mit den üblichen Karrieremustern - und diese wiederum mit den Publikations- und Zitiermustern - zusammen. Diese Strukturen - und nicht etwa die *inhaltlichen* Themen - begünstigten bestimmte methodische Zugänge.

Es gibt somit derzeit keine rational wissenschaftlich begründete Möglichkeit, die Unternehmung des "Dodo-Birds" doch noch in einen Wettlauf umzufunktionieren, damit sich möglichst wenige Tiere die Preise teilen und die anderen in ihrer Eigenart und Funktion als wertlos und überflüssig deklarieren können. Es gibt insbesondere keine wissenschaftliche Basis, um über die Ziellinie zu entscheiden, die alle Tiere zu passieren haben. Solange keine monoklonen Behandler an monoklonen Störungsträgern in reiner Form Manuale vollstrecken, sondern unsere gesellschaftliche Pluralität sich auch in der Heterogenität unterschiedlicher Menschenbilder mit unterschiedlichen Lebensweisen und -zielen ausdrückt, solange wird Bedarf an unterschiedlichen Therapierichtungen sein.

Natürlich kann und darf über die kontroversen Vorstellungen immer wieder gesellschaftlich diskutiert werden. Und hier kann auch die Wissenschaft mit ihren Forschungsergebnissen zur Debatte beitragen. Die Wissenschaft kann und sollte sich selbst aber nicht zum Richter in diesen Wertfragen aufspielen. Und sie sollte dies schon gar nicht in unredlicher, verdeckter Form tun, indem inhaltliche und gesellschaftliche Wert-Fragen als "methodische" Fragen und Probleme umdefiniert werden. Hier erleben wir leider in der Tat einen zunehmenden Mißbrauch von "Wissenschaft" und "wissenschaftlicher Methodik", nämlich als Waffe, um pluralistische Wertdebatten zu unterlaufen und möglichst für eine bestimmte Partei zu entscheiden. Seriöse Erörterungen der obigen Fragen führen eben in jeweils überaus komplexe Zusammenhänge und münden in das angeführte heterogene Geflecht aus unterschiedlichen Lebensweisen, Zielen, Werten und Vorstellungen, das sich nicht durch Berufung auf eine oder wenige Zahlen homogenisieren und damit gleichschalten läßt. Genau das wird aber z. B. mit Operationalisierungen und Messungen in diesem Bereich allzu oft versucht: Statt den gesellschaftlichen Diskurs zwischen den unterschiedlichen Gruppen über deren Wertvorstellungen zu führen, wird unter dem Deckmantel von scheinbar neutraler "Methodik" und "Wissenschaftlichkeit" Partei im Wertestreit ergriffen. Wenn dieser kognitive Taschenspielertrick gelingt - und ein Blick auf viele gegenwärtige gesellschaftliche Werte-Auseinandersetzungen zeigt leider, *daß* er nur allzu häufig und allzu gut gelingt -, braucht man dem Gegner im Disput keine inhaltlichen Argumente mehr zu liefern und sich selbst der Diskussion auszusetzen, sondern kann sich hinter einer "korrekten Methodik" und dem Prädikat "wissenschaftlich" verschanzen. So lassen sich dann Menschen mit anderen Wertvorstellungen als "unwissenschaftlich" oder "methodisch unzulänglich" diskreditieren und der Gesellschaft letztlich die eigenen Wertvorstellungen möglichst unangefochten aufdrücken.

Die Akzeptierung der Heterogenität der Therapieansätze bedeutet keineswegs, daß jeder Therapeut tun kann und sollte, was ihm beliebt, oder daß Wissenschaftler einem bunten Treiben von Therapieansätzen nur zuzusehen hätten. Vielmehr sind die Wissenschaftler gerade mit ihrer inhaltlichen Kompetenz gefragt, zu entscheiden, welche Therapieverfahren auf dem jeweiligen Stand des Wissens bevorzugt (d. h. von der Sozialgemeinschaft finanziert) zum Einsatz kommen. Diese inhaltliche Kompetenz liegt allerdings vor allem nicht im Rechnen und Zählen, sondern in der Kenntnis der Bücher und Fachliteratur, der entsprechenden Diskurse und in der Qualifikation, vorgelegtes Material bewerten zu können, wie oben ausgeführt wurde. Allerdings müßte viel weniger über "Verfahren", sondern zumindest zugleich über Ausbildungswege und -institutionen diskutiert werden - denn ein guter Therapeut ergibt sich nicht durch Glaubensbekenntnis zu irgendeinem

Verfahren, sondern durch eine gute Ausbildung (ein Aspekt, dem m. E. viel zu wenig Rechnung getragen wurde). Und selbstverständlich müssen und dürfen auch die Erkenntnisse aus der Prozeßforschung oder aus der Erforschung der realen Wirksamkeit (wie im Consumer-Report) mit einbezogen werden.

Realistischerweise wird auch eine solche "wissenschaftliche Anerkennung" nur bei einem Dutzend oder weniger Ansätzen (je nach Zusammenfassung in Clustern - und mit jeweils qualifizierten Ausbildungsgängen) konsensfähig sein. Das mag für zahlreiche andere ggf. ungerecht erscheinen, die nur eine sehr erschwerte Chance haben, sich dann noch als eigenständiges Verfahren zu etablieren. Aber zunächst einmal würde das Spektrum an zugelassenen Richtungen weit besser der vielfach betonten Pluralität unserer Gesellschaft gerecht werden. Darüber hinaus würde damit der allergrößte Teil faktischer psychotherapeutischer Behandlungen der letzten Jahrzehnte erfaßt sein. Und letztlich hätten auch andere Richtungen in dieser Heterogenität gute Chancen, sich durch Anbindung an vorhandene zu qualifizieren.

In einer so auch zum Nutzen der Patienten beruhigten therapeutischen Kulturlandschaft könnten sich Wissenschaftler endlich wieder an ihre Hausaufgaben machen: nämlich die Vielfalt möglicher Kontingenzen zwischen therapeutischen Vorgehensweisen und Lebensschicksalen (mit ihren ggf. mehr oder minder typischen Verläufen) leidender Menschen weiter zu erforschen, die Ergebnisse zu systematisieren und auf dieser Basis die theoretische Diskussion voranzutreiben. Hier gibt es unglaublich viel zu tun, und hier sind Wissenschaftler gefragt. Zu Preisrichtern in umfunktionierten Wettkämpfen, die letztlich sogar die Forschungs- und Denkmöglichkeiten beschränken und daher zutiefst antiwissenschaftlich sind, sollten sich seriöse Wissenschaftler möglichst nicht hochstilisieren lassen.

Literatur

Caroll, L. (1963). Alice im Wunderland. Frankfurt: Insel-Verlag.

Corsini, R.J. (Hg.) (1983). Handbuch der Psychotherapie. 2 Bde. Weinheim: Beltz

Bastine, R., Fiedler, P. A., Grawe, K., Schmidtchen, S., Sommer, G. (Hrsg) (1982). Grundbegriffe der Psychotherapie. Weinheim: Edition Psychologie

Devereux, G. (1992). Angst und Methode in den Verhaltenswissenschaften. Frankfurt: Suhrkamp

Gould, S, J. (1983). Der falsch vermessene Mensch. Basel: Birkhäuser

Grawe, K. (1995). Welchen Sinn hat Psychotherapieforschung? Eine Erwiderung auf Tschuschke, Kächele & Hölzer (1994)*Psychotherapeut*, 40, 281-297

Grawe K., Bernauer, F., Donati, R. (1994). Psychotherapie im Wandel. Göttingen: Hogrefe

Guttman, L. (1977). What is not what in Statistics. *The Statistician* 26: 81-107

Kriz, J. (1973). Statistik in den Sozialwissenschaften. Einführung und kritische Diskussion. Reinbek: Rowohlt (ab 5. Aufl. Opladen: Westdeutscher Verlag)

- Kriz, J. (1975). Datenverarbeitung für Sozialwissenschaftler. Einführung in Grundlagen, Programmierung und Anwendung. Reinbek: Rowohlt (rororo-studium Bd. 45)
- Kriz, J. (1981). Methodenkritik empirischer Sozialforschung. Eine Problemanalyse sozialwissenschaftlicher Forschungspraxis. Stuttgart: Teubner
- Kriz, J. (1985). Grundkonzepte der Psychotherapie. München: Urban & Schwarzenberg (ab 4. Aufl. 1993: Beltz/PVU)
- Kriz, J. (1996). Grundfragen der Forschungs- und Wissenschaftsmethodik. In: Hutterer-Krisch, R., Kriz, J., Parfy, E., Margreiter, U., Schmetterer, W., Schwendtner, G.: Psychotherapie als Wissenschaft - Fragen der Ethik, Bd. 5 der "Serie Psychotherapie" (Hrsg.: G. Sonneck), Wien: Facultas, 15-160
- Kriz, J. (1997): Systemtheorie. Eine Einführung für Psychotherapeuten, Psychologen und Mediziner. Wien: Facultas (ab 3. Aufl. 1999: UTB)
- Kriz, J. & Lisch, R. (1988). Methodenlexikon; für Mediziner, Psychologen, Soziologen. München: PVU
- Kromrey, H. (1999). Von den Problemen anwendungsorientierter Sozialforschung und den Gefahren methodischer Halbbildung. Sozialwissenschaften und Berufspraxis, 22: 58-77
- Linden, M. & Hautzinger, M. (Hrsg) (1996). Verhaltenstherapie.3 Heidelberg: Springer
- Luborsky, L., Singer, B., Luborsky, E. (1975). Comparative Studies of Psychotherapies: Is it True that "everyone has won and all must have prizes"? Archives of General Psychiatry, 32, 995-1008
- Margraf, J. (Hrsg) (1996). Lehrbuch der Verhaltenstherapie. Heidelberg: Springer
- Reinecker, H. (1996). Therapieforschung. S. 31-47 In: Margraf, J. (Hrsg). Lehrbuch der Verhaltenstherapie. Heidelberg: Springer
- Revenstorf, D. (1982). Psychotherapeutische Verfahren. 4 Bde. Stuttgart: Kohlhammer
- Seligman, M.E.P. (1995). The Effectiveness of Psychotherapy. Amer. Psychologist, 50, 965-974
- Tschuschke, V., Kächele, H., Hölzer, M. (1994). Gibt es unterschiedlich effektive Formen von Psychotherapie? Psychotherapeut, 29, 281-291
- Tschuschke, V., Hölzer, M., Kächele, H., (1995). Ach, du liebe "Güte". Eine Einladung statt einer Erwiderung. Psychotherapeut, 30, 304-308

Erschienen in: Hermer, Matthias (Hrsg): Psychotherapeutische Perspektiven am Beginn des 21. Jahrhunderts. Tübingen: DGVT-Verlag, 2000, 43-66