

**Stellungnahme der DGPT  
zum Entwurf der Neufassung der Verfahrensregeln zur Beurteilung der wissenschaftlichen Anerkennung von Methoden und Verfahren der Psychotherapie des Wissenschaftlichen Beirates Psychotherapie („Methodenpapier“)**

erstellt von **Prof. Dr. Wolfgang Mertens**, Ludwig-Maximilians-Universität München, Department Psychologie, Institut für Klinische Psychologie/Psychoanalyse.

I. Vorbemerkungen

Die detaillierte Ausarbeitung von Verfahrensregeln durch den Wissenschaftlichen Beirat Psychotherapie (WBP) für die wissenschaftliche Anerkennung psychotherapeutischer Verfahren und Methoden wird grundsätzlich begrüßt. Die Transparenz, die durch die Operationalisierungen ermöglicht wird, erlaubt nunmehr eine differenzierte Einschätzung darüber, inwieweit nicht nur experimentalpsychologische Prüfkriterien, sondern auch übergeordnete Kriterien, wie Versorgungsrealität und -relevanz, Einbeziehung des Fachwissens der Praktiker, Belange der Prozessforschung, Impact für die zukünftige Psychotherapieforschung, berücksichtigt werden.

Eine immer wiederkehrende Kritik an der ersten Fassung aus dem Jahr 1999 lautete dahingehend, dass einseitig nach neopositivistischen Überzeugungen hinsichtlich experimenteller Prüfkriterien vorgegangen worden sei, die eine unterschiedliche Auffassung von psychotherapeutischen Wirkfaktoren aufgrund anders gearteter methodologischer und methodischer Vorstellungen haben. Es verwundert nicht, dass diese Kritik gern als überflüssige „Sophistifizierung“ abgetan wird. Tatsächlich sind diese Überlegungen aber für die vorliegende Einschätzung konstitutiv; mit ihrer Hilfe wird es vor allem möglich, auch die zentralen Punkte der Versorgungsrealität, der Einbeziehung des Fachwissens praktizierender Psychotherapeuten sowie die Beurteilung des Anreizes für die zukünftige Psychotherapieforschung einzuschätzen. Methodologische Überlegungen bilden also aus moderner wissenschaftstheoretischer Sicht nicht irgendeinen idiosynkratischen Appendix, sondern sie sind grundlegend für eine umfassende Gegenstandsbetrachtung.

Somit eröffnet die Neufassung der Verfahrensregeln die Möglichkeit, einige basale methodologische Probleme der empirischen Überprüfung von Therapien anhand des vorliegenden Entwurfs einzuschätzen und anzusprechen.

Von grundsätzlicher Wichtigkeit erscheint nach wie vor die Frage, ob die vom Wissenschaftlichen Beirat ausgearbeiteten Kriterien zur Beurteilung der Qualität von Psychotherapiestudien der Unterschiedlichkeit der Verfahren gerecht werden können, ob somit die verfahrensspezifische Fachkundigkeit ausreichend Berücksichtigung findet, ferner, ob die Belange der Versorgungsforschung ausreichend berücksichtigt werden.

Dabei ist es prinzipiell denkbar, dass trotz unterschiedlicher therapeutischer Vorgehensweisen Diagnostik, Krankheitskategorien sowie Zielkriterien mit identischen Methoden erfasst werden können. Dennoch sollte diese Annahme nicht als selbstverständlich vorausgesetzt werden, sondern Ergebnis einer kritischen Diskussion sein. Ansonsten entsteht die Gefahr, dass mit einem sehr grobmaschigen Kamm bestehende Unterschiede nivelliert und ein ex cathedra formulierter, ausschließlich laborexperimentellen Kriterien genügender Outcome zur obersten Richtschnur erklärt wird, der aber für die Versorgungsbedingungen zu wenig Bedeutsamkeit hat. Da wir aber weit davon entfernt sind, von einer Einheitspsychotherapie ausgehen zu können, muss die existierende Pluralität in einem so komplexen Gebiet wie dem der Veränderung von Erleben und Verhalten akzeptiert und der Tatsache Rechnung getragen werden, dass Patienten, die Psychotherapie in Anspruch nehmen, meistens eine Multimorbidität und eine Polysymptomatik aufweisen.

Tatsächlich spricht manches dafür, dass sich die unterschiedlichen therapeutischen Schulen auch im diagnostischen Vorgehen, in der Art, psychische und körperliche Leidenszustände zu konzeptualisieren sowie in den Zielkriterien unterscheiden. Wäre dies nicht der Fall, würde sich die Unterschiedlichkeit tatsächlich nur auf bestimmte „Techniken“ beziehen. Tatsächlich gibt es aber Unterschiede in den Menschenbildannahmen sowie in methodologischen und methodischen Auffassungen.

So kommt es z.B. im Unterschied zu einer Beobachtungssprache, die funktionale Aspekte des Verhaltens anhand bewusst zugänglicher Selbsteinschätzung abzubilden

versucht und einen stärker pragmatischen Umgang mit Daten ermöglicht, bei psychoanalytisch begründeten Verfahren mehr auf die Feinzeleierung von Daten an, d.h. auf die vielen bewussten und unbewussten Bedeutungsnuancen, die eine Handlung im Erleben einer Person hat. Der psychoanalytische Gegenstand der inneren Welt wird eher als fluktuierend, multipel determiniert, mit fließenden Übergängen, autopoetisch konzipiert (vgl. Schüle 1999) und kommt daher modernen nichtlinearen dynamischen Konzeptualisierungen von innerer und äußerer Natur sehr viel näher. Die „Übersetzung“ dieser eher parallel kodierten in eine serielle „Top-down“ Sprache der Selbsteinschätzungsfragebögen, Symptomskalen und Interviewauswertungen impliziert einen mehr oder weniger großen Bedeutungsverlust, abgesehen von den Selbsttäuschungsmöglichkeiten der Befragten. Mit der Praktikabilität der Datenerhebung geht somit ein „epistemischer Chauvinismus“ (Kuhl 2007) einher: Der rationalen Verwörterbarkeit von subjektiven Erfahrungen wird das ständige Verflochtensein des analytischen sequenziellen Denkens mit subsymbolischen und ganzheitlich affektiven Erfahrungen geopfert. Das wird zwar aus pragmatischer Sicht eher für nebensächlich erklärt; tatsächlich könnte die Vernachlässigung dieser methodologischen und methodischen Überlegungen aber den Grund für die vielen widersprüchlichen Befunde im Bereich der Psychotherapieforschung bilden. Mit anderen Worten wird deshalb auch dafür plädiert, sich bei den Überlegungen für eine gegenstandsadäquate Beurteilung von Prüfkriterien nicht ausschließlich an den Kriterien einer evidenzbasierten Medizin und pharmakologischer Forschungsdesigns zu orientieren, um damit die „Droge Psychotherapie“ einem szientistischen Alltagsdenken schmackhafter zu machen, sondern stattdessen dem bereits eingeschlagenen Weg einer qualitativ hochwertigen Wirksamkeitsforschung in psychotherapeutischen Prozessen weiterhin zu folgen. Dazu gehören auch qualitative Methoden der Psychotherapieforschung.

Buchholz (1999) z.B. hat auf die Unterscheidung von empirischer Forschung und professioneller Praxis hingewiesen, die in den Sozialwissenschaften bereits eine längere Tradition hat. Wenn empirische Forschung lediglich das intersubjektiv Auszählbare in einem veralteten naturwissenschaftlichen Verständnis von Objektivität als wissenschaftliches Faktum gelten lässt, sind die empirischen Befunde eher als „dünn“ zu bezeichnen. Notwendig erscheint es deshalb, auch das Professionswissen der Praktiker einzubeziehen, um die methodologischen Einseitigkeiten früherer empirischer Forschung in der Psychotherapie zu vermeiden und den Weg zu einer gehaltvolleren und angemesseneren Forschungspraxis in der psychodynamischen Psychotherapieforschung zu gehen. Empirisch ist keineswegs nur das, was sich mit Zahlen belegen lässt, sondern auch all das, was mit geschulten Sinnen und professioneller Kompetenz als qualitative Erfahrung wahrgenommen, reflektiert, interpretiert und kontextualisiert

wird. Natürlich lassen sich auf diese Weise gewonnene Daten in einem weiteren Schritt auch numerisch abbilden und statistisch verrechnen.

Im Unterschied zu der Doktrin einer positivistischen Einheitswissenschaft verfügt jede Disziplin über ihre eigenen Methoden, welche die „Empirie“ auf je eigene Weise konstituieren (vgl. Hampe 2001 2004). Deshalb müssen letztlich alle Methoden auch im Verbund mit Praktikern auf ihre Adäquanz hin überprüft werden. Aus diesem Grund sollten zukünftig auch die Praktiker vermehrt darin geschult werden, sich über die Angemessenheit einer off-line-Forschung Gedanken zu machen (vgl. Kernberg 2007). Und entsprechend müssen auch die Ausbildungsinstitute ihren Nachwuchs schulen.

Dennoch nimmt in den letzten Jahren das Interesse praktizierender Psychotherapeuten an Psychotherapieforschung erfreulicherweise zu. Ein wichtiger Grund für das gestiegene Interesse an psychodynamischer Wirksamkeitsforschung ist sicherlich auch, dass in den letzten 10 bis 15 Jahren sehr differenzierte psychodynamische Einschätzverfahren entstanden sind, die sowohl für die Status- und Verlaufsdiagnostik als auch für die interaktive und mikroprozessuale Verlaufsforschung anspruchsvolle Instrumente bereitstellen (siehe z.B. Bänninger-Huber, Benecke, Benjamin, Bucci, Fonagy, Kächele, Krause, Luborski und Kächele, Jones, Merten, Rudolf, Weiss und Sampson, u.a.), die auch die meisten Praktiker psychodynamischer Verfahren überzeugen.

Wenn es gelingt, die bisherige Tradition anspruchsvoller Fallgeschichten, die das Übertragungs-/Gegenübertragungs- und Beziehungs-Geschehen verdichtet und fokussiert abbilden (z.B. Leuzinger-Bohleber 1995), mit einer multimethodischen off-line-Forschung zu verbinden, können für die zukünftige Psychotherapieverlaufs- und -ergebnisforschung innovative Maßstäbe gesetzt werden. Vereinzelt ist dies bereits ausgeführt worden (z.B. Benecke und Krause 2005, A. Buchheim, Kächele et al. 2006, Deserno et al. 1998, Leuzinger-Bohleber und Stuhr 1997, Rudolf 2001, Rudolf, Grande und Oberbracht 2000, Rudolf, Grande und Dilg 2001). Vor allem kann dann auch davon ausgegangen werden, dass diese gegenstandsadäquatere Vorgehensweise ungleich stärker Praktiker anspricht als ein klassisch positivistisches Kontroll-Paradigma. Deren Mitarbeit ist aber wiederum für die zukünftige Psychotherapieforschung und auch für Teilbereiche der Grundlagenforschung unerlässlich.

Zu bedenken ist hierbei auch, dass der neopositivistische Normenkatalog, der für viele methodische Evaluationen noch maßgeblich ist, bei zeitgenössischen Wissenschaftsphilosophen schon seit geraumer Zeit wenn nicht generell in Frage gestellt, so doch hinsichtlich der Erreichbarkeit relativiert worden ist. Es wäre deshalb bedauernd, wenn der klinische Reichtum des herkömmlichen Fallberichts mit der detaillierten Schilderung der Übertragungs- und Gegenübertragungsprozesse gänzlich paper-and-pencil-Einschätzungen des Outcomes weichen sollte, nur weil diese Verfahren in einem herkömmlichen Wissenschaftsverständnis als maximal reliabel gelten. Die Gültigkeit der Schlussfolgerungen darf deshalb nicht einem überholten objektivistischen Verständnis zum Opfer fallen, das – wie mehrfach betont – seit vielen Jahren selbst in den Naturwissenschaften als veraltetes Ideal gilt. Denn die methodische Zurichtung einer komplexen zwischenmenschlichen Erfahrungswirklichkeit anhand von Daten aus Fragebögen und Befundskalen, die szientistisch dann als die eigentliche „Empirie“ verstanden werden, obwohl sie eine theoretisch präparierte Selektion darstellen, übersieht, wie sehr diese angebliche Objektivierung eine theoretisch konstruierte Empirie darstellt. Diese bleibt nicht weniger interpretationsbedürftig als eine epistemische Haltung, die versucht, sich im möglichst vorurteilsfreien Zuhören mit allen Sinnen der Wirklichkeit im Gegenüber und der Beziehung anzunähern und diese Eindrücke auf systematische Weise festzuhalten. Die Alternative kann deshalb heutzutage nicht mehr lauten: objektive Empirie versus spekulative Intuition, sondern eine Forschungskultur, in der theoretisch konstruierte Methodik im Verbund mit (alltags-) theoretisch vermittelter und methodisch eingesetzter Intuition zum Tragen kommt.

Deshalb wäre eine Methodik des „case study plus“ wünschenswert, d.h. ein Vorgehen, bei dem herkömmliche Fallberichte mit einer angemessenen und differenzierten offline-Forschung ergänzt werden. Zugleich wird dafür plädiert, dass die zukünftige Forschung innerhalb psychoanalytisch begründeter Verfahren vermehrt stattfinden und hochrangige Standards erfüllen sollte.

II. Die folgende Einschätzung erfolgt auf diesem skizzierten Hintergrund:

1. ein einheitswissenschaftliches Vorgehen kann der tatsächlichen Unterschiedlichkeit der aus verschiedenen Kulturen und Denktraditionen stammenden psychotherapeutischen Verfahren nicht gerecht werden
2. die psychoanalytisch begründeten Verfahren, sind aus den dargelegten Gründen aber dennoch einen Kompromiss mit bisherigen sog. objektiven Methoden eingegangen, die aber eher unterkomplex sind
3. diese lassen sich aber mit innovativen naturalistischen Forschungs-Designs durchaus verbinden
4. dennoch müssen einige Fragen an die Vereinbarkeit kritisch betrachtet werden, wenn der Kompromiss zu weitgehend erscheint und dem Innovativen zu wenig Raum gegeben wird.

*Zu Punkt 1. Wissenschaftliche Anerkennung (Seite 3 unten)*

Wirksamkeitsforschung und Prozessforschung werden als zwei unterschiedliche Fragestellungen betrachtet, weil sich aus der Erklärung der Wirkungsweise eines Verfahrens nicht zwingend auf seine Wirksamkeit schließen lässt.

Bei der Erforschung der Wirkungsweise kann aber nicht von der Wirksamkeit derselben abstrahiert werden, es sei denn, es bliebe bei einer rein theoretischen Begründung, ohne dass sich das Modell in seiner Instantiierung bewähren muss. Aus der Tatsache der Unterschiedlichkeit sollte deshalb nicht auf die Nachrangigkeit der Erforschung der Wirkungsweise geschlossen werden. Denn diese gehört zu den primär anstehenden Aufgaben der Psychotherapieforschung, nachdem die Wirkung von Psychotherapie als belegt gilt, und die Unterschiede zwischen den einzelnen Schulen in Meta-Analysen immer wieder als gering und nicht substantiell nachgewiesen worden sind (auch wenn dies mit den bislang eingesetzten und aus unserer Sicht immer noch unbefriedigenden Methoden zu tun haben könnte).

Wenngleich von den Autoren sicherlich nicht intendiert, könnte die Formulierung, dass „die Erklärung der Wirkungsweise ... ein zusätzlicher Aspekt“ (S. 4 oben) ist, die Interpretation nahe legen, dass einem eher theorielosen, pragmatistischen Verständnis

der Vorrang eingeräumt wird, nach dem Motto, „Hauptsache, ein therapeutisches Verfahren wirkt, egal wie“.

Weitere Anmerkung:

Da die Wirksamkeitsforschung den Beweis dafür erbracht hat, dass Psychotherapien wirken, sollte der erneute, zukünftige Forschungsaufwand primär der Frage gelten, ob in neuen, methodisch verfeinerten Studien die Wirkungsweise genauer als bislang erforscht werden kann. Unter Umständen ließen sich sogar aufgrund differenzierterer Methoden als bislang verfahrensspezifische Unterschiede für bestimmte Leidenszustände bei bestimmten Patienten und bestimmten Psychotherapeuten-Patienten-Paaren ermitteln (z.B. Ablon und Jones 1998; Blatt und Shahar 2004).

Auch wenn dies auf den ersten Blick eher für die Psychotherapieforschung wichtig zu sein scheint als für den Nachweis der Wirksamkeit neuer Verfahren, könnten durch die Formulierungen des WBP dennoch bereits die erforderlichen Weichen für eine zukünftige anspruchsvollere Forschung zum Zwecke des Nachweises der Wirksamkeit therapiespezifischer Verfahren gestellt werden.

Dabei wird der Reiz dieser Forschung sicherlich darin bestehen, in entsprechenden Forschungsdesigns den Surplus an Wirkung (verfahrensspezifische Wirkung minus dem Einfluss der kommunalen Faktoren) nachzuweisen. Dass dies nicht mit Hilfe zwar reliabler, aber zu einfacher Selbsteinschätzungsfragebögen (wie z.B. der SCL-90) geschehen kann, sondern nur mit ungleich aufwändigeren Verfahren, verweist auf die Anstrengungen, die zukünftig auf diesem Gebiet zu erbringen sind. Schon an dieser Stelle ist jedoch auch darauf hinzuweisen, dass zukünftig den maßgeblichen Psychotherapieverfahren im Vorfeld gleiche Chancen zur finanziellen und organisatorischen Verwirklichung derartiger Vorhaben einzuräumen sind. Diese (berufs- und hochschulpolitischen) Voraussetzungen sind bislang aber keineswegs gegeben (siehe z.B. Bucci 2005), ja verändern sich gegenwärtig sogar rapide in Richtung einer einseitigen Bevorzugung neurowissenschaftlicher Verfahren.

*Zur experimentell nachgewiesenen Wirksamkeit (S. 4 Mitte)*

Erfreulicherweise kommen die Verfasser des WBP zunächst zu der Einschätzung, dass „aus der experimentell nachgewiesenen Wirksamkeit nicht zwingend auf die Wirksamkeit unter Praxisbedingungen geschlossen werden“ kann. Mit dieser

Einschränkung erfolgt eine stärkere Gewichtung der externen Validität und eine kleine Abweichung gegenüber der Empfehlung des Gemeinsamen Bundesausschusses, die eine eindeutige Präferenz für die RCT-Studien ausgesprochen hat.

Sie ziehen daraus allerdings nicht den Schluss, dass der naturalistischen Überprüfung von vornherein der Primat einzuräumen ist, sondern bleiben dennoch ganz am klassischen laborexperimentellen Vorgehen orientiert, in dem zunächst die interne Validität zu maximieren ist und dann in einem zweiten Schritt auch die externe Validität bestimmt werden kann (vgl. Campbell und Stanley 1963), wobei sich in der Regel ein Zielkonflikt zwischen interner und externer Validität ergibt, d.h. je höher die interne Gültigkeit ist, desto geringer wird die externe Gültigkeit. Auf die Künstlichkeit und Lebensferne dieses naturwissenschaftlich-positivistischen Ideals der Maximierung der internen Validität für humanwissenschaftliche Sachverhalte ist in der psychologischen Forschung seit vielen Jahren hingewiesen worden. Die in den letzten Jahren erfolgte intersubjektive Wende in den Human- und Sozialwissenschaften verschärft die bisherige Kritik einer unbesehenen Übernahme naturwissenschaftlich-positivistischer Standards in die humanwissenschaftliche Forschung sogar noch um einiges. Therapien sind soziale Unternehmungen, in der eine Vielzahl dyadenspezifischer Variablen miteinander interagieren; ein randomisiertes Kontrollgruppen-Design, bei dem die klassische Experimentalforschung Pate stand, weist zu starke artifizielle Bedingungen auf (vgl. Kriz 2004, Leichsenring 2005). Üblicherweise wird eine Randomisierung angewandt, um akzessorische Effekte wie Placebowirkungen, suggestive Wirkungen durch die Überzeugtheit eines Arztes von einem Medikament auszuschließen. Es geht also darum, die Wirkung des Verum – also die reine Substanzwirkung zu isolieren. In den psychodynamischen Psychotherapieverfahren ist aber gerade das, was über die Randomisierung ausgeschlossen werden soll, nämlich die Wirkung der persönlichen Beziehung, das Verum. Die Zufallszuweisung von Patienten zu Behandlungs- und Kontrollgruppen widerspricht allen Erkenntnissen von der Wichtigkeit einer freien Therapeutenwahl, bei der zumeist schon in den ersten Minuten eine mehr oder weniger gute Passung zwischen Patienten und Therapeuten zustande kommt, die als eine wichtige prognostische Variable gilt. Die sog. Verblindung sowie die Kontrolle von Einflussgrößen, die experimentalpsychologisch als störend eingeschätzt werden, sind nur bei experimentellen Treatments von sehr kurzer Dauer überhaupt sinnvoll. Aber selbst wenn – wie bei Pharmastudien – nur ein Medikament verabreicht wird, ist die individuelle Bedeutungszuschreibung (siehe Placebo-Studien) von ausschlaggebender

Bedeutung, wie auch die Erforschung der Demand-Charakteristiken des Experiments erbracht hat (siehe z.B. Bungard 1980). Zudem sind wohl nur bei kurzen Therapien die Einrichtung einer Wartekontrollgruppe sowie eine Zufallszuweisung, bei dem die individuelle Auswahl von Therapeuten seitens des Patienten und die subjektive Indikationsstellung des Therapeuten (i.S.v. Dantlgraber, 1989) vernachlässigt werden können, ethisch überhaupt vertretbar.

Problematisch ist des Weiteren, dass bei der Annahme einer Homogenisierung von Experimental- und Kontrollgruppe anhand von diagnostischen Klassifikationen übersehen wird, dass vor allem Patienten in längeren Therapien zumeist Komorbiditäten aufweisen. Eine rein phänomenologische Klassifikation in ICD- oder DSM-Kriterien klammert zudem die interne Vielgestaltigkeit bestimmter Krankheitsbilder aus. Auch die Überzeugung, dass die experimentelle Variable immer unter den gleichen Bedingungen erfolgen kann, selbst wenn eine strikte Manualisierung eingehalten werden könnte, muss als illusionär bezeichnet werden. (Dies gilt übrigens auch für pharmakologische Wirksamkeitsstudien, da auch hier immer von einer nicht erfassbaren Interaktion von Verabreichungsmodalität und Suggestions-/Placebo-/Compliance-Effekten ausgegangen werden muss.)

Das Bestreben, Ursächlichkeit unter reinen Laborbedingungen nachzuweisen, muss somit zwangsläufig an der Komplexität humanwissenschaftlicher Sachverhalte scheitern. Dass nur mit Hilfe des Nachweises der internen Validität ein objektiver Ursache-Wirkungs-Zusammenhang in der Psychotherapieforschung nachgewiesen werden könne, stellt deshalb leider einen weiteren Mythos der an Mythen bislang so reichen Geschichte der Psychotherapieforschung dar.

Deswegen wäre, wenn nicht sogar ein Verzicht, so doch eine Neubewertung bezüglich der Rangordnung in „(1.)“ und „(2.)“ (s. S. 5 oben) erforderlich: An erster Stelle sollten bei Wirksamkeitsnachweisen und in der Psychotherapieforschung Untersuchungen unter naturalistischen Bedingungen stehen, deren methodische Qualitäten durchaus eine Rangreihe bilden können (vgl. Leichsenring 2004, der verschiedene Evidenzstufen ausgearbeitet hat). Die hoch künstliche, an die Untersuchung der psychologischen Grundlagenforschung und der Pharmaforschung angelehnte „Labor-Efficacy“ sollte nur als letzte aller möglichen Untersuchungsbedingungen herangezogen werden, etwa wenn es darum geht, vereinzelt Aspekte bestimmter Interventionen mikroanalytisch bestimmen zu wollen. Dabei sollten aber die Einschränkungen dieses Vorgehens im

Auge behalten werden, um nicht zu falschen Schlüssen zu kommen und sich fälschlicherweise auf der absolut sicheren Seite zu wähnen.

Ebenso sollte auch der Kostenaspekt randomisiert-kontrollierter Studien berücksichtigt werden, der nichtuniversitär verankerte Therapieverfahren besonders belasten würde.

Wenn aus politischen Gründen der RCT-Ansatz dennoch unverzichtbar erscheint (um die vorgebliche „Härte“ dieser Methode gegenüber einer szientistischen Öffentlichkeit zu demonstrieren), dann wäre allenfalls denkbar, dass in einem „gemischten Design“ randomisierte Gruppenzuweisungen und präferierte Gruppenzuweisungen gebildet werden. Ferner wären experimentelle Studien für Verfahren von kurzer Dauer, wie z.B. fokalthérapeutische Maßnahmen, Desensibilisierungstechniken, EMDR u.ä., denkbar, ohne aber zu gutgläubig von der Überlegenheit dieses Vorgehens von vornherein auszugehen. Ansonsten sollten aber die Vorteile der naturalistischen Methodik – wobei es mittlerweile auch genügend kritische Stimmen zu einer objektivistischen pharmakologischen Wirksamkeitsforschung aus dem naturwissenschaftlichen Lager gibt - selbstbewusst und offensiv vertreten werden.

Diese Gewichtung tritt im Kriterienkatalog (S. 29 oben) v.a. unter „Studiendesign“: 8. „Gruppenzuweisung“, dergestalt auf, dass unter 1) eine angemessene Randomisierung zur Sicherstellung bekannter und unbekannter prognostisch relevanter Merkmale genannt wird.

*zu 2. „Beurteilung einzelner Studien zur Wirksamkeit von Psychotherapiemethoden, 2.1. Feststellung der untersuchten Methode“ (S. 13 untere Hälfte)*

„Die psychotherapeutische Vorgehensweise ist in der Regel durch ein Manual oder manual-ähnliche Behandlungsrichtlinien konkretisiert.“

Diese Bestimmung taucht im Kriterienkatalog (S. 28) unter „Intervention“: 3.

„Operationale Definition der Interventionen“ und „6. Manualtreue, Treatment integrity“ wieder auf.

Grundsätzlich ist es natürlich richtig, die Verfahrensschritte der therapiespezifischen Richtungen und Module zu bestimmen. Allerdings birgt die Diktion des „Manuals“ und der „Manualtreue“ das Missverständnis in sich, als würde einer hochgradig

reglementierten Vorgehensweise das Wort geredet, was zusammen mit der oben kritisierten „Labor-Efficacy“ wiederum an ein eher künstliches Vorgehen erinnert, das mit dem naturalistischen Vorgehen nur wenig Ähnlichkeit aufweist. Die Existenz von Manualen könnte den Irrglauben entstehen lassen, dass dieses bereits die Behandlung ausmache, was zur unrealistischen Selbstüberschätzung von Novizen führt. (Häufig wurden in der Vergangenheit manualisierte Kurztherapien in experimentellen Settings von universitären Hilfskräften durchgeführt, so als bräuchte man zur Durchführung von Therapien keine Erfahrung und keine persönlichen Qualifikationen.)

Allerdings zeigen Untersuchungen, dass es durchaus Sinn machen kann, die (flexible) Einhaltung von therapiespezifischen Verfahrensschritten zu überprüfen. So konnten z.B. Ablon und Jones (1998) mit Hilfe der Methode des Psychotherapieprozess-Q-Sort aufzeigen, dass Kognitive Verhaltenstherapeuten immer dann besser abschnitten, wenn sie entgegen ihrem eigenen Selbstverständnis psychodynamische Therapieelemente, wie z.B. die Berücksichtigung der Beziehung, einfließen ließen. Und spätestens seit der Wallerstein-Studie weiß man, dass Psychoanalytiker oftmals entgegen ihrem eigenen Selbstverständnis auch psychotherapeutisch stützend arbeiten.

Für einen wissenschaftlichen Vergleich verschiedener Therapieverfahren scheint es deshalb wichtig zu sein, dass die betreffenden Therapeuten sich tatsächlich an die behandlungstheoretischen Konzepte ihrer Richtung halten. Aus diesen und anderen Gründen wurden einige Jahre lang manualisierte Therapieinterventionen forschungsmäßig favorisiert. Bislang steht aber der Nachweis aus, dass manualisiert durchgeführte Therapien im Ergebnis wirklich besser sind (Beutler et al. 2004).

Inwieweit allerdings eine „Spezifizierung der Bedingungen für die einzelnen Interventionen“ (Kriterienkatalog (S. 28), „Intervention“: 3. „Operationale Definition der Kontrollbedingungen“, davon die Operationalisierung 1) als Sprachregelung sinnvoll ist, bedarf sicherlich einer intensiven Diskussion. In der Praxis zeigt sich auf jeden Fall, dass nur sehr globale Verfahrensschritte manualisiert werden können und dass sich dyadenspezifisch viele Variationsmöglichkeiten ergeben (vgl. Freuds Metapher der psychoanalytischen Psychotherapie als Schachspiel, von dem sich nur die ersten Züge fixieren lassen).

*Zum überprüften Indikationsbereich (S. 13 unten)*

Für die Diagnostik werden die Diagnostiksysteme ICD und DSM angegeben, wobei hinzugefügt wird: „Es können jedoch ergänzend auch Informationen, welche die untersuchte Patientenstichprobe einer Studie zusätzlich kennzeichnen, festgehalten werden (S. 13f.).

Die Angabe „z.B.“ lässt also offen, ob auch andere Diagnostiksysteme mit entsprechendem Interview (wie z.B. das OPD-Interview der OPD, ab 2006: OPD-2) bei der diagnostischen Zuordnung von Patienten zu den verschiedenen Anwendungsbereichen eingesetzt werden können.

Die unter psychodynamischen Gesichtspunkten erfolgte Konstruktion der OPD ermöglicht unter anderem vor allem eine Berücksichtigung der strukturellen Vulnerabilität des jeweiligen Patienten, dessen Diagnose dadurch eine erhebliche Differenzierung erfährt, was v.a. für die Prognose des Behandlungserfolgs ausschlaggebend sein kann.

Die mangelnde Spezifizierung, welche ergänzenden Informationen neben ICD oder DSM noch hinzugezogen werden können, scheint eine gewisse Beliebigkeit zu suggerieren. Sind aber strukturierte klinische bzw. voll standardisierte Interviews wie z.B. SKID oder DIPS (wie unter A2: Diagnosestellung) mit dem OPD-Interview vergleichbar?

Dem WBP muss aber zugestanden werden, dass eine Normierung dieser zusätzlichen Einschätzungen schwierig ist, denn wäre dies möglich, gäbe es vermutlich auch keine verschiedenen Psychotherapieverfahren mehr.

*Zu B. Kriterien zur Bewertung der internen Validität, Punkt 2. Erhebung der spezifizierten Einschluss- und Ausschlusskriterien mittels valider Methoden ( S. 28 oben)*

In vielen herkömmlichen Studien wurden Patienten wegen ihrer Komorbidität ausgesondert (in der Untersuchung von Westen und Morrison waren dies etwa zwei

Drittel). Diese aus der Sicht des Randomized Clinical Trial (RCT) folgerichtige Absicht, homogene Untersuchungsgruppen zu erhalten, ist aber äußerst kritisch zu betrachten, weil sie die Generalisierbarkeit derartiger Studien wiederum stark einschränkt. Denn unter naturalistischen Bedingungen erfüllt die überwiegende Anzahl der Patienten die Kriterien für mehrere diagnostische Kriterien. Hier wird anhand der Operationalisierungskriterien zudem ein Ausmaß an diagnostischer Genauigkeit für die Identifizierung der Komorbidität festgelegt, das im Feld nicht zu erfüllen ist (So verschwinden z.B. Angstsymptome nach wenigen Therapiesitzungen, depressive Verstimmungen tauchen auf, neue charakterliche Haltungen verstärken sich, ein körperliches Symptom verschlimmert sich, strukturelle Beeinträchtigungen scheinen intensiver ausgeprägt usw.).

*Zu B. Kriterien zur Bewertung der internen Validität, Punkt 11 Follow-up-Messung ( S. 29)*

Hier heißt es unter Operationalisierung von Punkt 11): „zeitlich störungsangemessene Katamnese mit hoher Ausschöpfung“. Die Angemessenheit wird nicht weiter spezifiziert. Zudem impliziert der Ausdruck „störungsangemessen“, als sei daran gedacht, die Zeiträume zwischen Behandlungsende und Durchführung einer Katamnese von einer spezifischen psychischen Krankheit abhängig zu machen, was angesichts der Komplexität der meisten Krankheitsbilder zu einem ungerechtfertigten Vorgehen führen könnte. Katamnesenzeiträume sollten vielmehr von der zurückliegenden Behandlungsdauer festgesetzt werden.

Bislang ist der Nachweis der Wirksamkeit von Therapien fast ausschließlich für Kurzzeittherapien erfolgt, wobei das katamnestische Intervall in der Regel eher kurz war. Über längere Zeiträume existieren nur ganz wenige Untersuchungen. Deshalb ist die behauptete Wirksamkeit vieler Kurztherapien eher als fraglich zu bezeichnen. Es könnte sein, dass die Nachhaltigkeit von Langzeittherapien jedoch um einiges größer ist als die Effekte von Kurzzeittherapien (z.B. Keller et al., 1997). Auf jeden Fall besteht hier ein weiterer Forschungsbedarf.

Bedeutsam ist deshalb auf jeden Fall die Länge der Zeit zwischen Behandlungsende und Katamnese. Die tatsächliche Wirksamkeit eines Verfahrens lässt sich vermutlich erst im Rahmen von langfristig angesetzten Katamnesen (zwei bis drei Jahre nach

Behandlungsende, wenn nicht sogar noch länger) nachweisen (vgl. Sandell 1997). Deswegen wären genau betrachtet mehrere katamnestiche Messzeitpunkte erforderlich. (So sind in der Rudolf'schen Projektstudie Psychoanalytische Langzeittherapie zwei katamnestiche Messzeitpunkte angegeben.)

Soll diese Katamnesenforschung wiederum nicht ausschließlich nur am Output interessiert sein, wäre auch für die Zeit zwischen Behandlungsabschluss und Katamnesenzeitpunkt eine qualitative Untersuchung darüber erforderlich, wie ehemalige Patienten die gewonnenen Fähigkeiten und Einsichten in Handlungen und Erlebnisregulierungen umsetzen, als maligne erkannte Beziehungskonstellationen vermeiden oder unterbrechen u.a.m. und wie förderlich oder abträglich sie hierbei ihre Umwelt erleben.

Der hierbei gegebene enorme, v.a. auch mit qualitativen Methoden vorzunehmende Forschungsaufwand (vom Therapiebeginn bis zur abschließenden Auswertung der Katamnese ist mit einer Dauer von mindestens sechs bis acht Jahren zu rechnen) verdeutlicht noch einmal die enormen Forschungsbemühungen, die auf einem derart komplexen Feld wie der Psychotherapieforschung zu erbringen sind. Auch in der Öffentlichkeit muss deshalb das Bewusstsein dafür geschärft werden, dass für anspruchsvolle Untersuchungen psychotherapeutischer Versorgung mindestens ebenso viele finanzielle Mittel bereitgestellt werden müssen, wie für Untersuchungen medikamentöser Therapien. Auf diese Rahmenbedingungen geht der WBP allerdings an keiner Stelle ein.

### *Fazit:*

**Der Katalog zur Beurteilung der Studienqualität von Psychotherapiestudien (allgemeine methodische Qualität, interne Validität, externe Validität, Ergebnisse einer Studie) ist eine sehr differenzierte Ausarbeitung methodischer und fachwissenschaftlicher Kriterien. Die methodischen Überlegungen haben allerdings eindeutig das Übergewicht gegenüber fachwissenschaftlichen Angemessenheitsüberlegungen. Insofern kann man auch von einer Dominanz klassischer naturwissenschaftlicher Kriterien sprechen, wie sie lange Zeit in der medizinischen und psychologischen Forschung vertreten worden sind. Die Kritik an dieser Art der Forschung ist bekanntlich, dass sie zwar objektive, reliable und intern valide Ergebnisse hervorgebracht hat, diese aber zugleich auch wegen der**

vielen stillschweigend übergangenen Voraussetzungen und Implikationen sehr interpretationsbedürftig waren. Im vorliegenden Fall ist zu befürchten, dass angesichts der geforderten Gewichtung von einer weitgehenden Praxisirrelevanz ausgegangen werden muss. Standards der Pharmaforschung dürfen nicht unbesehen übernommen werden. Gerade im humanwissenschaftlichen Bereich können die vielen unerkannten Bedeutungszuschreibungen nicht per methodischem Kunstgriff eliminiert werden, sondern müssen hinsichtlich ihrer idiosynkratischen Einflüsse sehr viel präziser studiert werden, was unter anderem auch für sorgfältige Einzelfallanalysen spricht. Auch lässt sich der Verdacht nicht ganz von der Hand weisen, dass die propagierte Methodologie von vornherein kürzere Therapien bevorzugt.

Aus diesem Grund müsste zumindest der Einwand diskutiert werden, warum trotz vorliegender Kritik bekannter (psychodynamischer) Psychotherapieforscher in diesem Entwurf deren Einwände immer noch zu wenig berücksichtigt worden sind. Warum ist innovativen Vorschlägen, z.B. zur stärkeren Differenzierung naturalistischer Studien, so wenig Raum gegeben worden?

Wäre es nicht eine einmalige Chance, mit Hilfe des Beurteilungsinstruments des Wissenschaftlichen Beirats Psychotherapie innovative Ergebnis- und Prozessforschung bei den verschiedenen Psychotherapieverfahren anzuregen, ein noch stärkeres Interesse bei praktizierenden Psychotherapeuten zu wecken und auf diese Weise auch Psychotherapieforschung zum curricularen Bestandteil der psychotherapeutischen Weiterbildung zu machen, statt ein einseitiges Wissenschafts-, Forschungs- und Behandlungsverständnis zu fördern, das die psychotherapeutische Praxis nur unzureichend erfasst?

Die oberste Priorität sollte die Versorgungsrealität haben; deshalb müssen methodologische und methodische Überlegungen und Vorgehensweisen noch einmal gründlich reflektiert werden.

## Literatur:

- Ablon, J.S. & Jones, E.E. (1998). How expert clinicians' prototypes of an ideal treatment correlated with outcome in psychodynamic and cognitive behavioral therapy. *Psychotherapy Research*, 8, 71 – 83.
- Benecke, C. & Krause, R. (2005). Facial-affective relationship: Offers of patients with panic disorders. *Psychotherapy Research*, 15, 178 – 187.
- Beutler, L.E., Malik, M., Alimohamed, S., Harwood, T.M., Talebi, H., Noble, S. & Wong, E. (2004). Therapist variables. In M.J. Lambert (Ed.) *Bergin and Garfield's Handbook of psychotherapy and behavior change* (pp. 226 – 306). New York, 5. Aufl.
- Blatt, S.J. & Shahar, G. (2004). Psychoanalysis – with whom, for what, and how? Comparisons with psychotherapy. *Journal of the American Psychoanalytic Association*, 52, 393 – 447.
- Bucci, W. (2005). Process research. In: E. S. Person, A. M. Cooper & G.O. Gabbard (Eds.), *Textbook of psychoanalysis* (pp. 317 – 333). Washington, DC: American Psychiatric Publishing, Inc.
- Buchholz, M. B. (1999). *Psychotherapie als Profession*. Gießen: Psychosozial.
- Bungard, W. (Hg.) (1980). *Die "gute" Versuchsperson denkt nicht - Artefakte in der Sozialpsychologie – München: Urban und Schwarzenberg*.
- Campbell, D.T. & Stanley, J.C. (1963). Experimental and quasi-experimental designs for research on teaching (pp. 171 – 246). In N.L.Gage (Hg.), *Handbook of research on teaching*. Chicago.
- Dantlgraber, J. (1989). Bemerkungen zur subjektiven Indikation für Psychoanalyse. *Psyche – Z. Psychoanal*, 36, 191 – 225.
- Deserno, H. et al. (1998). „Wiederholen“ der Übertragung? *Psychotherapie, Psychosomatik, medizinische Psychologie*, 48, 287 - 297.
- Hampe, M. (2001). Theorie, Erfahrung, Therapie. Anmerkungen zur philosophischen Beurteilung psychoanalytischer Prozesse. Kommentar zu Leuzinger-Bohleber et al. und Sandell et al. *Psyche – Z Psychoanal.*, 55, 328 – 337.
- Hampe, M. (2004). Pluralität der Wissenschaften und Einheit der Vernunft – Einige philosophische Anmerkungen zur Psychoanalyse. In M. Leuzinger-Bohleber, H. Deserno und S. Hau (Hg.), *Psychoanalyse als Profession und Wissenschaft. Die psychoanalytische Methode in Zeiten wissenschaftlicher Pluralität* (S. 17 – 32). Stuttgart: Kohlhammer.
- Kächele, H. & Thomä, H. (1993). Psychoanalytic process research: Methods and achievements. *Journal of the American Psychoanalytic Association*, 41, (Suppl.), 109 – 29.
- Kächele, H. , Albani, C., Buchheim, A. et al. (2006). The german specimen case, Amalia X: Empirical studies. *International Journal of Psychoanalysis*, 87, 809 – 826.
- Keller, W., Dilg, R., Westhoff, G., Rohner, R. & Studt, H.-H. (1997). Zur Wirksamkeit ambulanter jungianischer Psychoanalysen und Psychotherapien – eine katamnestiche Studie. In M. Leuzinger-Bohleber & U. Stuhr (Hg.), *Psychoanalysen im Rückblick*. Gießen: Psychosozial-Verlag.
- Kernberg, O.F. (2007). The coming changes in psychoanalytic education: Part II. *International Journal of Psychoanalysis*, 88, 183 – 202.
- Kriz, J. (2004). Methodologische Aspekte von „Wissenschaftlichkeit“ in der Psychotherapieforschung. *Psychotherapie und Sozialwissenschaft. Zeitschrift für Qualitative Forschung*, 6, 6 – 31.

- Kuhl, J. (2007) Psychologie des Selbstseins. In J. Kuhl & A. Luckner, Freies Selbstsein. Authentizität und Regression. Philosophie und Psychologie im Dialog (S. 49 -81). Göttingen: Vandenhoeck & Ruprecht.
- Leichsenring, F. (2004). Randomized controlled vs. naturalistic studies. A new research agenda. *Bulletin of the Menninger Clinic*, 68, 137 – 151.
- Leichsenring, F. (2005). Are psychodynamic and psychoanalytic therapies effective? A review of empirical data. *International Journal of Psychoanalysis*, 86, 841 – 868.
- Leichsenring, U. und Rieger, B. (2004). Psychotherapeutische Behandlungsverfahren auf dem Prüfstand der Evidence Based Medicine (EBM). Randomisierte kontrollierte Studien vs. Naturalistische Studien – Gibt es nur einen Goldstandard? *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 50, 203 – 217.
- Leuzinger-Bohleber, M. (1995). Die Einzelfallstudie als psychoanalytisches Forschungsinstrument. *Psyche – Z Psychoanal*, 49, 434 – 481.
- Leuzinger-Bohleber, M. & Stuhr, U. (Eds.) (1997). *Psychoanalysen im Rückblick. Methoden, Ergebnisse und Perspektiven der neueren Katamnese Forschung*. Gießen: Psychosozial.
- Rudolf, G. (2001). Wie können strukturelle Veränderungen in analytischen Langzeitherapien empirisch erfasst werden? In W. Bohleber & S. Drews (Hg.), *Die Gegenwart der Psychoanalyse – die Psychoanalyse der Gegenwart* (S. 546 – 566). Stuttgart: Klett-Cotta.
- Rudolf, G., Grande, T. & Oberbracht, C. (2000). Die Heidelberger Umstrukturierungsskala. Ein Modell der Veränderung in psychoanalytischen Therapien und seine Operationalisierung in einer Schätzskaala. *Psychotherapeut*, 45, 237 – 246.
- Rudolf, G., Grande, T. & Dilg, R. (2001). Strukturelle Veränderungen in psychoanalytischen Behandlungen – Zur Praxisstudie analytischer Langzeittherapien (PAL). In U. Stuhr, M. Leuzinger-Bohleber & M. Beutel (Hg.), *Langzeittherapie. Perspektiven für Therapeuten und Wissenschaftler* (S. 238 – 259). Stuttgart: Kohlhammer.
- Sandell, R. (1997). Langzeitwirkung von Psychotherapie und Psychoanalyse. In M. Leuzinger-Bohleber & U. Stuhr (Hg.), *Psychoanalysen im Rückblick*. Gießen: Psychosozial-Verlag.
- Wallerstein, R.S. (2005). Outcome research. In E.S. Person, A.M. Cooper & G.O. Gabbard (Eds.). *Textbook of psychoanalysis* (pp. 301 – 315). Washington, DC: American Psychiatric Publishing.
- Westen, D., Nowotny, C.M., Thompson-Brenner, H. (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin*, 130, 631 – 663.