



Deutsche Gesellschaft für **Psychoanalyse,**
Psychotherapie,
Psychosomatik und Tiefenpsychologie e.V.

PSYCHO-NEWS-LETTER NR. 45

EIN KLEINER LITERATURRUNDFLUG

im Auftrag des Vorstands der DGPT
Verfaßt von
Michael B. Buchholz
email: buchholz.mbb@t-online.de

Mitte Juli 2006

WIRKT DER THERAPEUT ODER DIE METHODE ?

Ja, es ist wahr, dass wir in diesem Jahr den 150. Geburtstag Freuds mit beachtlicher medialer Aufmerksamkeit feiern konnten. Wir feiern eine Person. Nein, es ist nicht wahr, dass die Psychoanalyse nun auf einer Siegesstraße eingemündet wäre. Die Rolle der von den (be)handelnden Personen distanzierteren Methode ist unklar. Wer hat den seit April 2005 einsetzenden Boom pro Psychoanalyse und pro Freud verstanden, wer kann ihn erklären? Ist es wahr, dass in der Redaktion eines Hamburger Nachrichtenmagazins zwei antipsychoanalytische Feuilleton-Redakteure gingen und durch zwei „pro“ ersetzt wurden? Warum dann aber zogen ZEIT und GEO schon so bald nach? Und die andern Blätter im Wald raschelten im lauen Winde unisono mit, warum? Woher plötzlich die vielen recht sachlich informierenden Sendungen bei SAT1 und in anderen Kanälen? Ist Psychoanalyse als Methode einfach ein vermarktungsfähiges sujet „unter anderen“ im medienbeherrschten taumelnden Kommunikationsstillstand geworden?

Übersehen wir bei allem feierfreudigen Taumel nicht, dass es auch durchaus herausfordernde Bücher gibt wie die von **Todd Dufresne** (2003) mit dem Fatwah-Titel „Killing Freud“. Darin Kapitel wie „How Psychoanalysis Brings Out the Worst in Everyone“. Der Autor will erklärtermaßen provozieren: „...the seductive bad manners of psychoanalytic exchange are not merely a principle of good clinical practice, but has itself become a veritable culture, if not a worldview...“ Das ist wohl richtig und so preist der Autor sein Werk:

“This work is ultimately intended as a joyous, Nietzschean-inspired *celebration* for intellectuals and therapists; a celebration that is long overdue and happily announced”.

Ja, es ist wahr, dass die Psychoanalyse und ihre Kultur Erfolge feiern kann. Das geht soweit, dass sie an ihrem Erfolg zu scheitern droht. Worte wie Ödipus oder Verdrängung sind Standardrepertoire, Begriffe wie „narzisstisch“ oder „borderline“ erweisen sich als etwas

mehr von der Diskurskonjunktur abhängig, sind aber meist prima alltagstauglich geworden; der Sinn für sprachliche Artikulation bei zunehmender gesellschaftlicher Erhitzung verdampft. Wo **Freud** sich wunderte, dass seine Krankengeschichten wie Novellen zu lesen seien und **Schnitzler** antwortete, er schreibe Diagnosen, haben wir es mit einer alltäglich gewordenen Praxis dieses Staunens zu tun. **Molière** noch machte sich über „eingebildete Kranke“ lustig. Unsere ausgebildeten Kranken kommen heute mit schon fertigen Diagnosen - ADHS, Trauma, Asperger, Persönlichkeitsstörung, Tourette – zu uns und fordern jene Behandlung, über deren Tauglichkeit oder Untauglichkeit sie aus Tagesgazetten fertig gebackene Urteile mitbringen. Sie *erzählen* zu Beginn weniger Geschichten als dass sie sich mit perfekter Psycho-Terminologie *darstellen*. Und sie fordern, von der Behandlung unberührt als Person, in ihrem Lebensstil verbleiben zu können, weiter in monströse Filme zu gehen, endlos Parties zu feiern oder sich mit six-packs und anderem vollzudröhnen. Durchgesetzt hat sich die Vorstellung, dass eine Methode einem Problem begegnet, die Personen davon aber unberührt bleiben könnten, ja sogar sollten. Daß in dieser Ab-Spaltung der Person von ihrem Problem das allerernsteste Problem versteckt sein könnte, entzieht sich mehr und mehr der Einsicht. Im Wettstreit um „die bessere Methode“ der „Problemlösung“ scheinen auch Psychoanalytiker, die es besser wissen müssten, dafür zu erblinden. Psychoanalytische Begrifflichkeiten dienen auch als Marker der Differenzen zwischen den Kulturen. Sie können im Kampf der Kulturen – der in den Köpfen entflammte, als das Wort dazu so schön florierende – waffenfähig verwendet werden. Dafür ein kleines Beispiel. In dem wunderbar zarten, zutiefst melancholischen Roman „Atlas für verschollene Liebende“ beschreibt **Nadeem Aslam** das schwierige Leben eines schon älteren Pakistani in London, der sich seit Jahrzehnten um Integration seiner Landsleute bemüht; er verzweifelt, weil ihm dies nicht einmal bei und mit seiner islamisch-konservativ bleibenden Ehefrau Kaukab gelingt. Seine Frau hängt abgöttisch an ihrem Sohn:

“Aber dann fällt ihr ein, was ihr eine Frau beim Einkauf über die neuesten westlichen Theorien bezüglich der Verbindung zwischen Mutter und Sohn erzählt hat: ‚Sie sagen, alle Mütter wollen insgeheim mit ihren Söhnen ins Bett gehen. Ein Haufen Männer in Anzügen und Krawatten habe gestern Abend im Fernsehen darüber geredet‘. Kaukab war angewidert gewesen, hatte sich bei dieser Vorstellung geekelt: ‚So etwas sagten in Pakistan vulgäre Straßenhändler und Fischverkäuferinnen in Basaren, aber hier in England sprachen *gebildete* Menschen im *Fernsehen* darüber. Auf solche Art über die Liebe einer Mutter zu reden! Diese unmoralische und dekadente Kultur hatte es darauf abgesehen, alles, was an der menschlichen Existenz rein und erhaben war, in den Schmutz zu ziehen. Tatsächlich kontrollierten Mütter jeden Morgen als Erstes, ob der Penis ihres männlichen Babys steif wurde, aber dabei handelten sie aus Sorge um das Kind und seine ordnungsgemäße Entwicklung; und wenn der Junge ungefähr dreizehn war, prüften sie seine Bettwäsche und Schlafanzüge auf Anzeigen männlicher Samenergüsse, aber genauso behielt eine Mutter auch die Entwicklung ihrer Tochter im Auge.“

Selbstverständlich! Aus Sorge! Aus Liebe! Prüfen! Auf Reinheit! Wer würde „ordnungsgemäße Entwicklung“ nicht wollen? Und Mütter betreiben morgendlich solcherart methodische Qualitätssicherung aus edelsten Motiven. Das ist zarter und verzweifelter Spott bei Nadeem Aslam, das ist bitter, weil die Verzerrungen in den Köpfen von Personen so nicht zu korrigieren sind. Könnte Psychoanalyse zu Zeiten unter das Regime eines Karikaturenstreits geraten? Eine mögliche Antwort auf die Frage nach dem rasenden Stillstand bei qualitätsgesichertem Fortschritt findet sich wieder bei **Zadie Smith** in ihrem hier schon mehrfach gepriesenen Roman „White Teeth“:

“In fact, nothing moves. Nothing changes. They are running at a standstill. Zeno’s Paradox. But what was Zeno’s *deal* here (everybody’s got a deal), what was his *angle*? There is a body of opinion that argues his paradoxes are part of a more general *spiritual* programme. To

- (a) first establish multiplicity, the *Many*, as an illusion, and

(b) thus prove reality as seamless, flowing whole. A single, indivisible *One*.

Ja, es ist wahr, dass mit diesen Paradoxien sich auch die Psychoanalyse herumschlagen muß, mit ihrer *einen* Wahrheit, die ja immer nur szientifisch als aus Methode bestehend gedacht wird. Weil wir den Wert der Methode allein meinen verteidigen zu müssen, denken wir uns die Person des Therapeuten immer nur additiv dazu; denn der Therapeut, so die offizielle Lehre, ist ja lehranalytisch purifiziert. Personen zählen nicht, sie kommen in diesem Typ von wissenschaftlichem Diskurs nicht vor.

Jedoch: Eine an sich wertvolle Methode kann durch eine Person, wenn deren Kompetenz gegen Null geht, insgesamt wertlos werden; umgekehrt gilt das gleiche. Das Verhältnis von Person und Methode ist deshalb multiplikativ, nicht additiv zu denken! Und wenn ein solches komplexes Verhältnis noch mit fremden Kulturen multipliziert wird, finden wir skurrile Verschrägungen wie bei Kaukab. Mit der Kunst des Multiplizierens und der Mathematik, angewandt in empirischer Forschung also, kommt ein ganz anderer Faktor ins psychotherapeutische Spiel. Aber auch hier kommt nicht einfach etwas heraus, Ergebnisse sind nicht einfach.

Dies als kleine Vorwarnung, denn in diesem PNL kann ich etwas kompliziertere psychotherapiewissenschaftliche Verrenkungen nicht vermeiden. Studieren wir also geduldig, wie im letzten PNL versprochen, zunächst neue Nachrichten aus Schweden.

STOPP – THERAPEUTEN UND IHRE LEHRANALYSEN

Im Mai-Heft 3 (2006) von „Psychotherapy Research“ publizieren **Rolf Sandell, Jan Carlsson, Johan Schubert, Johan Grand, Anna Lazar** und **Jeannette Broberg** neue Befunde aus dem schwedischen STOPP-Projekt, über dessen Design und Anlage ich schon mehrfach berichtet habe, zuletzt im letzten PNL. Die letzten gesundheitsökonomischen Nachrichten waren irritierend, wenn nicht sogar deprimierend, was die Verfasser der Studie selbst deutlich einbekannten. Diesmal ist die Lage und ihre Befunde nicht viel besser; es geht um die Rolle der Lehrtherapien bzw. Lehranalysen und um die Frage, welchen Effekt diese auf das Befinden von Patienten haben.

Diese Frage ist knifflig empirisch zu untersuchen, berufspolitisch brisant und ideologisch nicht keimfrei. Es gibt Positionen, die eine Lehranalyse als unnötig ansehen, es sei denn für persönliche Probleme des Therapeuten und auf der anderen Seite des Spektrums finden sich die Festlegungen der internationalen

Verbände, die da manches vorschreiben bezüglich der Dauer und der Frequenz. Erfahrene und wohlangesehene Psychoanalytiker haben sich vehement gegen die Lehranalyse als Pflicht ausgesprochen, deren schädliche Wirkungen im Kontext der institutionalisierten Machtausübung vielfach angesprochen und erreicht, dass wohl überall ein „non-reporting“-System eingerichtet wurde: der Lehranalytiker kann nicht verpflichtet werden, über die „Fortschritte“ seines Analysanden zu berichten, ja er *darf* es nicht einmal mehr – das war ja zu manchen Zeiten anders. Empirisch beforscht allerdings wurde die Frage nach deren Wirksamkeit, nach Ausbildung von therapeutischer Kompetenz wenig; meist zentrierten die Debatten um „Identität“. Welche Effekte Lehrbehandlungen haben, gemessen an Patientenbefunden – dazu gibt es, wie die Autoren der Studie mitteilen, weltweit nicht mehr als 6-8 Studien mit

inkonsistenten Ergebnissen. Die Frage ist also offen.

In der hier mitgeteilten Berechnung der schwedischen Daten gingen die Autoren so vor, dass sie den patientenseitigen Outcome zu 9 Meßzeitpunkten („stage scale“) erhoben – von einem Punkt vor Behandlungsbeginn, zu mehreren Zeitpunkten während der Behandlung bis hin zu mehreren Katamnesezeitpunkten. Nur das GSI (General Symptom Index) wird hier nun in die Berechnung einbezogen.

167 Therapeuten wurden danach klassifiziert, welche Arten von Patienten sie haben:

“one of the classes of therapists (class 4) had patients starting from a very high level of symptom distress whereas the patients of another class of therapists (class 5) were well beneath the mean in the normal group already when they started treatment. The other three – and larger – classes started their therapies around the level of the caseness criterion, two (Classes 1 and 3) somewhat above, and one (Class 2) slightly below.

The patients in three of the therapist classes improved at a low to moderate rate across the stages, whereas those in the largest class (Class 1) showed considerable improvement. Especially interesting is Class 3, with 16% of the therapists, whose patients actually had a positive slope. This means they deteriorated from a rather high level of distress to begin with (0.91) to even more distressed levels as their treatments progressed. Although this deterioration was not statistically significant ($z = 1.41$; $p < .20$, two-tailed), we focused our further analyses on this particular class” (S. 310)

Die Patienten der Klasse 3 (16% der Therapeuten umfassend) verschlechterten sich zunächst leicht, verbesserten sich dann, um sich anschließend weiter zu verschlechtern, ihre Werte im GSI stiegen leicht an – aber statistisch nicht signifikante Veränderungen! Nach einigen statistischen Prozeduren stellt sich dann heraus,

“that the therapists in Class 3 were more than three times as likely to have been in

psychoanalysis as in psychotherapy for their training therapies” (S. 310)

Auch dieser Befund ist noch nicht signifikant, aber er wird es, wenn man die Dauer der Lehranalyse in die statistische Analyse einfließt. Die Autoren interpretieren das so, dass Therapeuten mit langer eigener *Lehranalyse*, die also neutrale, abstinenten und v.a. schweigsame Lehranalytiker hatten, in ihren *Psychotherapien* (wenn sie selbst im Sitzen behandeln), schlechter werden, nicht aber unbedingt, wenn sie selbst Analysen durchführen.

Aber was zeigt sich eigentlich nun?

Zugegeben, ich werde nicht recht schlau hinsichtlich der Darstellung der Befunde. Wenn man den Einfluß der Länge von Lehrbehandlungen auf Patienten-Outcome untersuchen will, hätte es dann nicht nahegelegen, Therapeutenklassen zu bilden, die genau das abbilden (also Klassen von Therapeuten mit langen, mittleren und relativ kurzen Lehranalysen bzw.- -therapien) und *erst dann*, innerhalb dieser Klassen danach einzuteilen, ob sie schwere oder weniger schwer gestörte Patienten behandeln? So teilen die Autoren eine Menge komplizierter statistischer Umrechnungen mit, die zur Folge haben, dass die Klasse-3-Therapeuten manchmal als Klasse 4 erscheinen und dann wieder nicht und sie müssen in ihrer Diskussion mitteilen:

“Thus, the relation between training therapy duration and patient outcome seems to have been curvilinear”.

Das besagt, ein *linearer* Zusammenhang ist gerade nicht nachzuweisen! Länge der Lehranalyse und Alter hatten wenig gemeinsame Varianzanteile – 5%; das wirkt auf den ersten Blick erklärungsbedürftig. Je länger eine Lehranalyse dauert, umso älter wird man ja. Eher habe jedoch die Erfahrung der Therapeuten eine Rolle gespielt. Die Autoren kritisieren ihren nicht eindeutigen Befund, indem sie alternative Erklärungen

diskutieren: Die Auswahl der Therapeuten war nicht repräsentativ für die schwedische Therapeutenpopulation. Sie schlagen als „most direct interpretation“ vor:

“Thus, therapists’ training therapy, when it has been a long psychoanalysis, is not very productive for their patients in psychotherapy”. (S. 312)

Es *könnte* also sein, dass lange Lehranalysen weniger hilfreiche therapeutische Kompetenzen hervorbringen, wenn Therapeuten mit solchen Erfahrungen dann Psychotherapien machen, die sie vermutlich *wie* Analysen – mit langem Schweigen etc. – durchführen. Zum Schluß heißt es:

“Whereas the current results may support such doubts, they do not show that training therapy as such is unnecessary or even counterproductive for psychoanalytically oriented therapists. Also, one should certainly respect the opinions of many clinicians that their training therapies have been important and valuable to them both personally and professionally ... It is difficult to see how therapists-to-be would otherwise learn how a person might feel being a patient, how experienced therapists ‘do it’, and how theoretical concepts manifest themselves” (S. 314)

Fraglos ist es gut, wenn der Einfluß von langen Lehranalysen empirisch untersucht wird; so etwas wirkt immer antidogmatisch und klärend. Daß so ausgebildete Therapeuten dann vielleicht mehr schweigen, als ihren ihnen gegenüberstehenden Patienten gut tut, kann man sich vorstellen – aber was ist hier

eigentlich genau untersucht worden? Weniger die statistischen Prozeduren und Umwandlungen als die eigentliche Untersuchungsstrategie hinterlassen eine Menge Fragen: Warum die bereits besprochene Klassenbildung nach Patientenschweregraden, warum nicht die unabhängige Variable genau definieren und dann die abhängige Variable, den Outcome bei den Patienten klassifizieren? Warum wird nicht mitgeteilt, wie viele Therapeuten mit langen Analyse in den anderen Klassen gefunden wurden? Und welche Ergebnisse dort anzutreffen sind? Die schwachen Veränderungen, dazu kurvilinear-schwankend, sind nur teilweise signifikant – wie soll man das verstehen? Warum nicht auch die anderen erhobenen Outcome-Werte außer dem GSI heranziehen?

Diese Fragen bleiben offen und werden offenbar auch von den Autoren als Schwäche möglicher Eindeutigkeit gewertet; so jedenfalls fasse ich den Rückgriff auf die „opinions of many clinicians“ im Schlusssatz auf. Eine solche Bemerkung freut zwar das klinische Herz. Aber sie ist deshalb verwunderlich, weil hier ja klinisches Denken empirisch überprüft werden sollte und man nun genau auf das zurückgreift, was überprüft werden sollte; auch das zeigt die Unsicherheit der Befunde. Vielleicht also, dass die korrigierende Tendenz, die man von empirischer Forschung erwartet, diesmal weniger beizutragen hatte.

DIE NIMH TDCRP-STUDIE UND VERWIRRENDE AUSWERTUNGSPROBLEME

Ein kluger Kopf, so erinnere ich aus meinem ersten Semester, hatte einmal gesagt, am Anfang müsse ein jeder Student das EGPF, das „*Erste Gesetz Psychologischer Forschung*“ lernen. Es lautet:

ERST NACH ABSCHLUß EINER
UNTERSUCHUNG WEIß MAN, WIE MAN SIE

SINNVOLLERWEISE HÄTTE DURCHFÜHREN MÜSSEN. Das kommt einem, ich weiß auch nicht so recht warum, irgendwie in den Sinn und verlässt einen nicht so ganz, wenn man sich mit den nächsten, durchaus interessanten Studien beschäftigt.

NIMH TDCRP ist ein Kürzel für ein Forschungsprogramm über „Treatment of

Depression Collaborative Research Programm“, das am „National Institute of Mental Health“ unter Federführung von **Irene Elkin** u.a. durchgeführt wurde; erste Berichte darüber erschienen bereits Mitte der 1990er Jahre. Verglichen wurden vier „treatment conditions“: CBT (cognitive behavioral), IPT (interpersonal psychotherapy), IMI-CM (imipramine hydrochloride plus clinical management) und PLA-CM (placebo + clinical management) an 250 Fällen. Diese waren ihren Behandlungsbedingungen nach einem screening-Verfahren durch Zufall zugewiesen worden. Es handelt sich also nicht um eine naturalistische Studie, sondern um eine, die strengen experimentellen Bedingungen genügen will. Unter diesen Bedingungen hatte die Studie überraschenderweise erbracht, dass sämtliche Behandlungen annähernd gleich gute Ergebnisse in der Behandlung der Depression zeigen; sie waren gemeinsam der Placebo-Bedingung überlegen. Die Daten dieser Studie werden hier nun genutzt, um anderen Fragestellungen nachzugehen.

Therapeut oder Methode? Einerseits ...

In der von **Irene Elkin, Lydia Falconnier, Zoran Martinovich** und **Colleen Mahoney** berichteten Studie über „Therapist effects“ (März-Heft 2/2006 von „Psychotherapy Research“) werden allerdings nur CBT und IPT miteinander verglichen hinsichtlich der Frage, ob „Methode“ oder „Person“ (des Therapeuten) einen erkennbar differenzierbaren Einfluß haben. Bei den Behandlungen unter Medikamentengabe meinte man wohl, so etwas ausschließen zu können und untersuchte jetzt nur

- 119 Patienten
- davon 59 in CBT und 60 in IPT.
- 17 Therapeuten waren beteiligt: 8 CBT-Therapeuten und 9 IPT-Therapeuten.
- Behandlungsdauer 16 Wochen,

- beide Verfahren hochgradig *manualisiert*.

Ein interessantes kleines Nebenergebnis: nur 23% der IPT-Patienten, aber 32% der CBT-Patienten beendeten vorzeitig ihre Behandlung. So ganz gleich, wie früher mitgeteilt, können die Behandlungen also nicht gewirkt haben, wenn man diese dropouts mit ins Kalkül nimmt...

Zu verschiedenen Zeitpunkten (vor, während und nach der Behandlung) wurden Messungen durchgeführt; hier kommt es v.a. auf die Befunde mit der Hamilton-Rating-Scale-for-Depression (HRSD) an.

Diskutiert wird, dass es mehrere Persönlichkeitszüge solcher Patienten gibt, die eine gute Kooperation mit dem Therapeuten bei der Symptomerleichterung erschweren:

- Persönlichkeitsstörungen,
- Feindseligkeit,
- Perfektionismus und
- zu hohe Erwartungen an die Therapie.

Gerade bei Patienten mit diesen Charakteristika hätten sich also Differenzen zeigen müssen, die auf unterschiedliche persönliche Kompetenzen der Therapeuten zurückzuführen wären, auch wenn die Therapeuten ansonsten in gleicher Weise das Manual anwenden. Denn der Umgang mit diesen schwierigen Patientenmerkmalen, so der empirische Gedanke, gehe auf die *allgemeine* therapeutische Kompetenz zurück, nicht auf die spezifische Wirkung der im Manual angewiesenen Behandlungsstrategien. Hinsichtlich ihrer allgemeinen Kompetenzen aber unterscheiden sich Therapeuten. Will man die Wirkung des Manuals an sich kalkulieren, muß man deshalb berücksichtigen, wie schwer gestört die Patienten eines Therapeuten sind. Damit hat man also drei Klassen von Variablen: die Methode (IPT oder CBT), den Störungsgrad der Patienten, die allgemeine Kompetenz des Therapeuten.

Zum Zuge kommt deshalb bei der statistischen Auswertung eine hierarchisch lineare Modellierung (HLM), wobei diese drei Komponenten auf verschiedenen Ebenen berücksichtigt werden; die Schwere der Störung der Patienten kommt also neben den methodischen und therapeutischen Faktoren hinzu.

“In these analyses, treatment outcomes are described not as being specific to some particular session or time but rather as a rate of change in an outcome variable per week, using a three-level HLM, with weeks in therapy (or “time”) at Level 1, patients at Level 2, and therapists at Level 3. Treatments were also entered at Level 3. Slope variation (adjusted for baseline dependent variable) was parsed into two components: therapist variance and patient variance.” (S. 149)

Man muß sich etwas in diese statistischen Feinheiten hineindenken, weil das ganze März-Heft von „Psychotherapy Research“ dieser Frage gewidmet ist: *macht der Therapeut die Therapie oder wirkt die Methode (unabhängig von der Person)?* Und dazu gibt es gleich ein paar aufregende **Knaller und Knüller**.

Doch zunächst zu den Befunden von Elkin et al. Wenn man also auf mehreren Levels analysiert, kann man sich das am besten als einen Würfel vorstellen, dessen Kanten aus „Therapeut“, „Patient“ und „Methode“ bestehen; dann gibt es noch für jeden Messzeitpunkt einen solchen Würfel. Die Behandlungsform (CBT oder IPT) ist eine Kovariable der „Kante“ des Therapeuten. Je nach Auswertungsrichtung – von welcher Seite des Würfels aus man zu analysieren beginnt – bekommt man verschiedene Verteilungen; beginnt man bei der Therapeutenseite heißt der Befund:

“there were no overall significant effects resulting from the therapist”.

Und noch weiter:

“most of the variation in rates of change in this model lies at the patient level rather than the therapist level”. (S. 150)

Der Varianzanteil des Therapeuten betrage nicht mehr als 0-4%!! Allerdings – einige Ausreißer, jene Therapeuten mit besonders schlechten oder besonders guten patientenseitigen Ergebnissen, wurden aus der Analyse ausgeschlossen! Die Daten von deren Patienten überraschenderweise jedoch nicht!

Und dann heißt es zusammenfassend:

“There was no indication of significant therapist effects in the current analyses despite the use of a more efficient model (because all available data may be included) and a conception of treatment progress that focuses on rates of change, adjusted for baseline expectations, instead of focusing on status at a somewhat arbitrary time point. Although there was not much reliable variance at the therapist level, there was substantial variability within therapists at the patient level. Because it is likely that therapists have differential levels of effectiveness with different types of patients, we also looked at the interactions of therapist with patient initial severity and with several indexes of patient difficulty. There were virtually no significant interactions.” (S. 151)

Die gleichen Therapeuten haben also durchaus bei verschiedenen Patienten variierende Ergebnisse, aber insgesamt ist der Therapeuteneffekt gering – so die Aussage, die mich wieder wundert! Denn es stellen sich auch hier wieder eine Menge Fragen.

So gibt es keine Analysebefunde, die die Methoden miteinander vergliche! (Außer man erinnert sich, dass Elkin et al. schon 1995 darüber berichtet hatten, dass IPT leicht, ohne Signifikanz, überlegen war – nur: gilt das auch noch bei der hier angewandten statistischen Analysetechnik??) Hier jedenfalls nur die Mitteilung einer leicht höheren Abbruchquote bei CBT! Zudem lässt ein hoher Grad von manualisierter Therapieform *natürlich* ein solches Ergebnis erwarten, weil die Manualisierung ja genau den Effekt hat, dass die Person des Therapeuten keine Rolle spielt, sondern eben die Anwendung der Methode

qua Manual. Die Einhaltung des Manuals ist durch Stichproben an Videos sorgfältig überwacht und somit gesichert worden, dass die Therapeuten tatsächlich methodenkonform vorgehen. Und – diejenigen, die den Befund hätten irritieren können, nämlich die (positiven/negativen) Ausreißer wurden aus der Statistik eliminiert, während die Daten über ihre Patienten in der Analyse mit berechnet wurden. Wo genau ist da die Grenze gezogen worden?

Und natürlich werden nun in der „discussion“-Sektion des Beitrags frühere Befunde anderer Autoren zu positiven Effekten der Person von Therapeuten ziemlich kritisch durchleuchtet daraufhin, dass die Mitteilung von Therapeuteneffekten auf falschen Berechnungen, unangemessenen Modellen, zu kleinen Stichproben etc. beruht habe, hier also neue Diskussionen und Forschungen notwendig würden.

Aber andererseits ...

Diese Irritation in der Lektüre der Befunde wird nun noch gesteigert, wenn der direkt nachfolgende Artikel von **Dong-Min Kim, Bruce E. Wampold** und **Daniel M. Bolt** mittels anderer Auswertungsstrategien des gleichen Datensatzes (!) der NIMH TDCRP-Studie also, zu genau gegenteiligen Schlüssen gelangt. Diese Autoren beginnen ihren Beitrag genau dort, wo Elkin et al. geendet haben: mit der Referierung derjenigen Literatur, die bislang gezeigt habe, dass es Therapeuteneffekte gebe und sie zitieren teils die gleichen Autoren, die von Elkin et al. gerade statistisch demontiert worden waren, nun als Kronzeugen - *pro* Therapeuteneffekt.

Der von den Autoren benutzte einleitende Vergleich ist erhellend: als man den Ackerbau wissenschaftlich zum Zweck höherer Ausbeute zu standardisieren begann, hatte man völlig die regionalen Bedingungen und die persönlichen Gewohnheiten der Farmer selbst ignoriert

und gelangte folglich zu unterschiedlichen Befunden trotz gleicher „Methoden“; es ist demnach nicht nur wichtig, was man ermittelt, sondern auch „what was left out“. Elkin wird aus dem Jahre 1999 zitiert, worin diese Autorin den von **Kiesler** kritisierten „Einheitsmythos“ zustimmend erwähnt, wonach es eben ein Mythos sei, dass alle Therapeuten einer Schule tatsächlich dieselbe Therapie machten. Will man das experimentell kontrollieren, muß man berücksichtigen, dass Therapeuten ihrerseits bestimmte Überzeugungen haben („therapist allegiance“) und sich Erfahrungen und Theorien verbunden fühlen. Lässt man sie dann andere Verfahren ausüben, um den Erfordernissen kontrollierter Studien („crossed design“) gerecht zu werden, hat man immerhin wenigstens ihren *allgemeinen* level therapeutischer Kompetenz unter experimenteller Kontrolle. Lässt man sie hingegen im „nested design“ nach ihren Präferenzen arbeiten, hat man zwei Variablen – allgemeine Kompetenz und Methodenspezifität – unentwirrbar vermischt, dafür aber die „therapist allegiance“ eindeutig erfasst.

Zu diesem einen Design-Dilemma kommt ein weiteres hinzu: Ob man Therapeuten als singuläre statistische Elemente, nämlich ohne klaren Verweis auf eine größere Stichprobe ansieht oder als „random factor“ spielt auch eine gewichtige Rolle. Sieht man sie als zufällig ausgewählte Repräsentanten einer größeren Stichprobe, kann man natürlich erheblich weitere Schlüsse ziehen, während man ansonsten Aussagen auf die Studienteilnehmer begrenzen muß. Dies Problem ist natürlich von extremer Bedeutung, wenn man die Frage „Therapeut oder Methode?“ untersucht. Sind die Therapeuten dann Repräsentanten ihrer Methode oder ihrer Persönlichkeit? Wie man das auffasst, hat weitgehende Folgen für die Auswertungsprozeduren.

Unter diesen Vorüberlegungen wird nun das Ziel dieser Studie formuliert:

“The purpose of the current study was to use analytic strategies that (a) take into account the nesting of patients within therapists, (b) are able to consider therapists as a random factor in various ways, (c) use estimation procedures that are robust for unequal sample sizes within therapists and between treatments, and (d) model treatment and therapist effects simultaneously.” (S. 162 f.)

Auch hier kommt eine hierarchische lineare Modellierung (HLM) als statistische Modellbildung zur Anwendung. Es werden nun verschiedene samples gebildet, deren detaillierte Nachzeichnung ich mir hier untersage; das Ergebnis ist, dass bis zu 12% der Varianz des Outcome durch den Therapeuten verantwortet werden. Die Variation hängt teils von den Pretests der Patienten ab; ob es sich also um leichtere oder schwerere Formen der Depression handelt. Bei letzteren spielt der Therapeut eine erheblich größere Rolle. Die Autoren fassen zusammen:

“Several multilevel analyses of the NIMH TDCRP data revealed sizable therapist effects, ranging from 1% to 12% depending on the outcome variables and the model adopted. Overall, a simple mean of all the estimates were about 8%. Given that the estimate of the population treatment effect (i.e. the difference between CBT and IPT) in this analysis was zero, it seems clear that therapists were an important source of variability in these data. Although many caveats are discussed next, the results suggest that, with regard to outcomes, therapists are more important than treatment.” (S. 167)

Jetzt ist der Knall da, denn diese Aussage steht der von Elkin et al. genau entgegen gerichtet und der Kontrast wird noch gesteigert:

“It appears from the various analyses that variability among therapists is greater as the initial severity of the patients increases” (S. 168)

Das wird von den Autoren des ganzen Heftes von „Psychotherapy Research“ durchaus als scharfer Kontrast gesehen und ist es fraglos auch gerade deshalb, weil eben innerhalb des Designs die manualgerechte Anwendung der jeweiligen Methode so genau überwacht worden war. Trotzdem lassen sich hier nun Befunde demonstrieren, die die Rolle des Therapeuten deutlich herausstreichen und das hatten Elkin et al. genau anders gesehen; sie machten ja die Variation unter den Patienten für die Differenzen innerhalb der gleichen Klassen von Therapeuten verantwortlich. Kim et al. halten mit Kritik an Manualisierungen nicht hinter dem Berg, indem sie formulieren, wenn man die Therapeuteneffekte eben durch Manuale ausschalte und ignoriere, würden die Methodeneffekte überschätzt. Dennoch lassen sie sich also in dieser manualisierten Studie nachweisen, jedenfalls mit den von Kim et al. verwandten Methoden.

Nun, hier handelt es sich um keine Kleinigkeit bei solchen Differenzen. Das betont auch die Herausgeberin dieses Heftes, **Clara E. Hill**, in ihrer Einführung. Das Heft enthält weitere Beiträge, die nun nicht eigene Studien vorstellen, sondern die aufgeworfenen methodischen und inhaltlichen Fragen zu diskutieren beginnen, teils auf sehr elaboriertem Niveau. Clara Hill empfiehlt:

“I encourage readers to read all six articles and consider all the issues involved and come up their own conclusions.”

Versuchen wir, dieser Aufforderung Folge zu leisten.

WHAT KIND OF SCIENCE IS THIS?

So fragt ein erster Kommentator in dieser statistisch-methodischen Kontroverse, **Stephen Soldz**

von der Boston Graduate School of Psychoanalysis und weist auf einige wichtige Punkte hin. Beide Studien verwandten unterschiedliche Outcome-

Maße, aber auch unterschiedliche statistische Modellierungen – nun, das wussten wir schon und man sollte meinen, da es Korrelationen zwischen verschiedenen Outcome-Maßen gibt, sollten diese in irgendeinem Zusammenhang stehen. Er entwickelt folgende, für die Entscheidung der Frage, etwas luftige, nicht falsche, aber zu allgemein bleibende Argumentation:

“These, and other contemporary advanced statistical techniques, are capable of leading to quite wrong answers if not used carefully and thoughtfully. Because these analyses are complex and their putput is somewhat foreign, it can also be harder to tell whether the analysis has gone awry. The point here is not to discourage use of these techniques, because they constitute a decided increase in our analytic capabilities and allow us to address questions that we previously could hardly imagine. However, their use requires care, because we are not in Kansas anymore. In many cases, as with the studies in this Special Section, appropriate use of them will require that a member of the research team be a statistical or a researcher who has become familiar contemporary statistical thought and practice. As we move toward the multilevel era, the days of research as a cottage industry conducted by a single researcher, with perhaps the assistance of a graduate student who took an extra statistics course, may be numbered” (S. 174f.)

Das wird in den Ohren von Elkin et al. und Kim et al. nicht ganz so freundlich ankommen, denn diese Leute sind versierte Statistiker und gewohnt, große Tabellen auf dem Bildschirm zu lesen. Die indirekte Empfehlung, einen guten Statistiker anzustellen, hat einen merkwürdigen sound. Soldz verstrickt sich in Widersprüche. Während er einleitend feststellt, Stichprobendifferenzen seien für die Ergebnisdifferenzen nicht verantwortlich, meint er nun, die Stichprobe sei für die Anwendung solcherart elaborierter statistischer Prozeduren immer noch nicht groß genug; eben das könnten versierte Statistiker einem beibringen. Seine Eingangsfrage:

„What kind of science is this?“ beantwortet er am Schluß mit dem Prädikat: “a healthy science”. Das ist fairerweise auch so zu sehen; denn hier müssen Differenzen ausgetragen werden und dazu sind wissenschaftliche Journale da. Im Heft wird der Debatte reichlich Raum gegeben und sehr solide argumentiert. Dennoch gibt es in der Bewertung der „science“ selbst da bekanntlich Differenzen. Zur Eröffnung des Jahrtausends, also knapp 6 Jahre vorher, stellte **Marvin Goldfried**, damaliger Vorsitzender der Society of Psychotherapy Research, deren Journal wir gerade studieren, fest, die empirische Psychotherapieforschung sei in einem „infant state“. Gut also, 6 Jahre später wird das Kind schon gewachsen sein. Es hat fraglos mehrere (durchaus kluge und sensible) Köpfe. Zeigt sich die Wichtigkeit der es erziehenden Personen in Gestalt der Autoren und Studienleiter?

Paul Crits-Christoph und **Robert Gallop** beteiligen sich ebenfalls an der Diskussion um die Klärung der Befunddifferenzen und führen diese auf unterschiedliches hierarchisches lineares Modellieren zurück, auf die statistische Prozedur also und stellen fest, dass es bisher keinen Literaturüberblick über Therapeuteneffekte in der empirischen Literatur gegeben habe. Und sie zitieren Elkin et al. (2005), wo diese Forschungsgruppe feststellte, dass die Differenzen zwischen Therapeuten *am besten in naturalistischen Studien* untersucht würde – man erinnere sich: die Patienten wurden durch Zufall zugewiesen, die Therapie war manualisiert und ihre Dauer vorgegeben. Das ist das Gegenteil von „naturalistisch“. Hier wird also zugestanden, dass es das Design der ganzen Studie war, das möglicherweise das Ergebnis hervorbrachte, nämlich Eliminierung der Therapeuteneffekte! Sie könnten, interessanterweise von Elkin selbst zugestanden, in naturalistischen Designs überhaupt erst sichtbar werden. Elkin et al. begründen in einem sehr knappen, einseitigen Diskussionsbeitrag

die Wahl ihrer linearen Modellierung; Bruce Wampold und Daniel Bolt weisen in einem etwas längeren Beitrag darauf hin, dass es in der Frage von Therapeuteneffekten „clever ways to make them (and everything else) disappear“ gebe. Diese „clever ways“ sind im einzelnen:

Wenn man die Patientenvariation vergrößert, wird auch der *Fehler Typ I* erhöht: Je größer die statistische Aufblähung der Patientenvariation, umso weniger kann man zwischen „wahren“ Differenzen und Zufallsvariationen unterscheiden. Das leuchtet unmittelbar ein, weil man dann natürlich auch die Therapeuteneffekte unsichtbar macht.

Weiter seien entscheidend die Werte bei Beendigung der Therapie, denn Longitudinalanalysen haben das Potential, die Patientenvariation zu erhöhen – klar, denn dann spielen unkontrollierbare Effekte eine immer größere Rolle. Damit aber werden die Therapeuteneffekte eliminiert bzw. von Zufallsvariationen nicht mehr sicher unterscheidbar. In ihren eigenen Berechnungen sind diese Autoren genauso verfahren, haben den allgemeinen 8%-Wert ermittelt und schätzen ihn gerade deshalb als bedeutsam ein, weil andere statistische Überlegungen ihn eher unwahrscheinlich machen.

Wampold und Bolt kritisieren nun auch den Ausschluß der therapeutischen „Ausreißer“ aus der statistischen Analyse und kritisieren zurecht die gleichzeitige Beibehaltung der Patientendaten in der Stichprobe, wie Elkin et al. es getan haben und landen schließlich einen „Knüller“, den man erst entdecken muß, denn er ist im Zitat eines Zitats versteckt:

“It should be noted that Elkin et al. admitted finding evidence of therapist effects. In their analyses, the proportion of variance attributable to therapists ranged from 0% to 4.5%, admittedly small, but larger than any other factor in their analysis (except, of course, for patient variation). As well, in terms of recovery, “the range was wide, with 0% to 63% of patients

‘recovering’ with different therapists” (p. 151). Moreover, the correlation of recovery rates for therapists and percentage of patients completing treatment was .73. Elkin concludes that “one therapist ...did very poorly (out of four patients, retaining only one, who did not recover) and two did very well (retaining over 80% of their patients, with over 60% of their patients ‘recovering’” (p. 151), result that was not explainable by caseloads” (S. 186)

Wampold und Bolt zitieren also Befunde aus der „gegnerischen“ - eben auch hier referierten - Studie von Elkin et al., die der Generalaussage der Studie selbst zu widersprechen scheinen und sie folgern noch einen Satz weiter:

“That specially selected, trained, supervised, and monitored therapists providing manualized, empirically supported treatments demonstrated this degree of variability suggests that therapists effects are indeed important”.

Trotz Manualisierung und Supervision, trotz Auswahl und Ausbildung dieser wohl trainierten Therapeuten gibt es also individuelle Effekte – eigentlich möchte man kaum glauben, dass es anders sein kann.

Löst man sich von der NIMH-Debatte und liest nun in diesem interessanten Heft noch ein bisschen weiter, stößt man auf die Arbeit von **Jacques P. Barber, Robert Gallop et al.** über die Rolle der therapeutenseitigen Überzeugungen bei der Behandlung von Drogenabhängigen. Diese Autoren bringen, ganz unabhängig von der NIMH-Studie über die Behandlung von Depressionen, noch eine weitere Möglichkeit ins Spiel.

Sie stellen bei 95 Patienten fest, dass es kurvilineare Beziehungen zwischen manualisierter Behandlung und Outcome gab, aber diese von der therapeutischen Kooperation („Alliance“) abhängen! Ist die Kooperation zwischen Patient und Therapeut gut, spielt es weniger eine Rolle, ob sich der Berater an das Manual hält oder nicht; das ist aber bei solchen Patienten von größerer Bedeutung, mit denen die Qualität der Zusammenarbeit

weniger gut ist. Wenn also ein Interaktionsmaß angewandt wird, nämlich die Einschätzung der Kooperation, dann kommen die *allgemeinen* Kompetenzen eines Therapeuten im Umgang mit einem schwierigen Menschen deutlicher zur Geltung als der besondere Effekt einer Manualdifferenz. Wer gut mit seinem Patienten kann, darf auch mal ein bisschen vom Manual abweichen, ohne dass die

Outcome-Werte gleich sinken, so könnte man vielleicht zusammenfassen.

Die therapeutische Allianz könnte also die Rolle einer Moderatorvariablen beanspruchen. Vielleicht klärt sich das Düstere der verwirrenden Statistiken und Methodenprobleme so auf einleuchtende Weise, aber das müssen wir natürlich weiteren Studien überlassen.

GESUNDHEITSKOSTEN, NOCH EINMAL...

Eine deutsche Studie von **Susanne Kraft**, **Bernd Puschner**, **Michael J. Lambert** und **Hans Kordy**, ebenfalls im gleichen Heft von „Psychotherapy Research“, untersucht an 176 Teilnehmern einer naturalistischen Studie, wie sich Gesundheitskosten im allgemeinen und Krankenhaustage im besonderen unter Psychotherapie verändern. Über einen Zeitraum von 4 Jahren wurden die Gesundheitskosten und die Krankenhaustage der Teilnehmer aus den Kassenakten erfasst; das Ergebnis ist erfreulich: wenn körperliche Beschwerden durch Psychotherapie abnahmen, kam es auch zu einer Abnahme der Gesundheitskosten. Alter und die Zahl der therapeutischen Sitzungen korrelierten freilich leicht negativ – hier verweise ich auf meinen PNL vom Februar 2006 über das Alter.

Auf diesem Weg haben wir jetzt eine mehrfach irritierende Befundlage ermittelt. Diese letzten gesundheitsökonomischen Befunde scheinen denen von Sandell et al., über die ich im PNL-44 berichtet habe, zu widersprechen. Dort war ja sogar von einer leichten Steigerung der Gesundheitskosten bei lange analysierten Patienten die Rede gewesen, hier haben wir nun eine Studie, die die gesundheitsökonomischen Kosten psychodynamischer Behandlungen gerade wiederum sinken sieht. Das könnte für die Vermutung sprechen, dass hier doch besondere schwedische Umstände eine

größere Rolle spielen, denn in dieser Hinsicht hatte die schwedische Studie sowieso vielen anderen internationalen Befunden widersprochen.

Die Frage nach den Therapeuteneffekten im Unterschied zur Methode könnte sich unter Hinweis auf das Arbeitsbündnis klären, obwohl man auch hier natürlich weitere Befunde abzuwarten hat. Doch auch sie würden natürlich nichts, rein gar nichts, für den einzelnen Fall klären. Das Argument von **Siegfried Zepf**, über das ich im März-PNL berichtet hatte, wonach es völlig unklar bleibt, ob eine in der Studie A untersuchte therapeutische Methode mit der von mir oder jemand anderem angewandten therapeutischen Methode gleichen Namens irgendetwas zu tun habe, bleibt triftig. Wenn der Therapeut eine Rolle spielt, *muß* unklar bleiben, was die Worte „gleiche Methode“ bedeuten könnten! Und nur wenn der Therapeut *keine* Rolle spielen würde, kämen die empirischen Gruppendesigns, über deren Probleme hier einiges zu erfahren war, überhaupt sinnvoll zum Einsatz! Das bedeutet nicht weniger als dass mit diesen Untersuchungen, wendet man sie nur reflexiv an, die medizinanaloge Forschungsstrategie selbst auf dem empirischen Prüfstand steht! In der Medizin soll es die Pille sein und nicht die „Droge Arzt“, deren Wirkung nachgewiesen werden muss, soll ein Medikament zugelassen werden. Aber

wenn es die „Droge Arzt“ ist, der eine entscheidende Rolle zugestanden werden muß, wird die medizinale Forschungsstrategie einer Revision unterzogen werden müssen.

Die Frage, ob wir es mit „infant science“ oder „healthy science“ zu tun haben, ist deshalb nicht ohne Bedeutung. Kann man überhaupt solcherart instrumentell Wirkungen in der Psychotherapie messen? Brauchen wir eine neue Debatte um viel weiterreichende Fragen aus den philosophischen, anthropologischen und an Sinn orientierten Dimensionen der Psychotherapie? Und wie könnte so etwas in den berufspolitischen Verengungen weiterhelfen?

Und schließlich bleibt die Frage, ob die „Big-Business“-Vermutung von Soldz, wir bräuchten laborierte Statistike(r)n sich so unbedingt halten lässt? Können wir uns psychotherapeutische Forschung nur noch denken, indem wir sie laborierten Spezialisten und deren Teams überlassen, die jene Fragen beantworten, die ihnen von professionellen Praktikern beantwortet

werden? Oder handelt es sich hier um einen Prozeß des „outsourcing“ therapeutischer Kompetenzen, wie er in anderen Bereichen längst gang und gäbe ist? Knüpfen wir vielleicht unsere Neugier und unser Interesse, unsere Erfahrung und klinische Befriedigung zu sehr an Antworten von anderen? Könnte es auch Wege geben, Forschungen selbst zu initiieren? Lassen sich therapeutische und wissenschaftliche Kompetenzen so vereinbaren, dass regionale Teams ihre eigenen Fragen mit lokal begrenztem Anspruch, aber in psychotherapiewissenschaftlicher Solidität beantworten können?

Es gibt Trends innerhalb der psychotherapiewissenschaftlichen Forschung, die in diese ganz andere Richtung eines neuen Junktims weisen, ein Junktim, das „Heilen und forschen“ nicht mehr in der Person, sondern in jeder Institution zu verknüpfen versucht. Werfen wir dazu einen Blick nach England. Hier gilt:

SMALL IS BEAUTIFUL

Ein neuer Band „What is Psychotherapeutic Research“ (2006 bei Karnac erschienen), herausgegeben von **Del Loewenthal** und **David Winter**, visiert diese Dimension eines erweiterten und zugleich regional begrenzten Junktims an. Die Herausgeber stellen in ihrer Einleitung etwas fest, was zum „outsourcing“ gehört, dass sich ein neuer Typus von Psychotherapieforscher herausgebildet habe, „with a particular interest in psychotherapeutic research who may or may not have been practitioners“. Auch in England wird Psychotherapie auf einem Postgraduierten Level gelehrt, aber die meisten haben unklare Vorbehalte gegenüber Forschung und gegen Zusammenarbeit mit Universitäten. Die Gefahr, dass daraus eine Marginalisierung

der Psychotherapie insgesamt entstehen könnte, wird schon lange deutlich gesehen. 1993 wurde deshalb ein United Kingdom Council for Psychotherapy (UKCP) ins Leben gerufen und 2002 die erste Forschungstagung abgehalten, deren Beiträge hier nun dokumentiert werden – mit Beiträgen renommierter Autoren aus den USA und aus dem übrigen Europa. Das Ziel wird einfach genannt:

“It is hoped that what is presented will inspire current practitioners and those in training to develop themselves, and hence their practice, through research”.

Die Verbindung von Selbst-Entwicklung und Förderung der Praxis ist durchaus nicht zufällig gleichrangig nebeneinander gestellt. In fünf Teilen werden Forschungsfragen von der ersten Idee bis

hin zu durchführbaren Projekten „mit Bordmitteln“ beschrieben und anschaulich gemacht, was sich gewinnen lässt. Die einzelnen Methoden werden vorgestellt, gezeigt, wie der therapeutische Prozeß, aber auch systematisch begleitende Outcome-Evaluation mit minimalen Mitteln betrieben werden kann, und schließlich wird ein letzter Abschnitt unter die Überschrift gestellt „Researching the Therapist and the Therapeutic Context“ – wir sind also mitten im Thema der Rolle von Therapeut und Methode.

Mark Aveline, von 2003 bis 2004 Präsident der Society for Psychotherapy Research eröffnet den Band mit einem grundlegenden Beitrag über Fragen der Psychotherapieforschung: „But how can the practitioner know that he or she is practising effectively and efficiently“ (S. 3), so artikuliert er gleich die Hauptfrage und schließt bald den Satz an: „Each therapy is a form of research“. Aber darin ist jeder Therapeut in seinem Blickwinkel naturgemäß beschränkt:

„However, even experienced therapists have a relatively small pool of experience from which to draw conclusions that they hope will improve their practice“. (S. 4)

Forschung müsse sich drei Charakteristika einfügen: Sie muß

- Hypothesen der empirischen Evidenz aussetzen,
- Ergebnisse öffentlich zugänglich machen und
- von Bereitschaft getragen sein, ggf. Theorien im Licht besserer Befunde aufzugeben.

Daraus entsteht ein „research cycle“, der mit dem Befund bei einzelnen Patienten beginnt, über ein Stadium der Hypothesenbildung zu qualitativen Pilot-Studien führe, aber dabei begleitet wird von „clinical practice“ und „clinical questions“ und schließlich zu Maßnahmen der Qualitätssicherung ebenso wie zu RCT-Studien führen kann. Einzelne Fragen, die mit qualitativen bzw.

quantitativen Mitteln untersucht werden können, werden aufgelistet; z.B. wie und ob sich die Narrative eines Patienten während des Prozesses verändern (qualitativ) oder ob verschiedene Arten der Verbesserung etwas mit der Dauer von Therapien zu tun haben (quantitativ). Die detaillierte Untersuchung einer einzelnen Fallgeschichte wird einbezogen, denn auch dies könne helfen, das eine wichtige Ziel zu erreichen:

“I look forward to improvements in methodology whereby quantitative researchers are more open to *contextualism* and *reflexivity* in their work and qualitative researchers find ways to enhance the *validity* and *generalizability* of their findings” (S. 21)

Dies bedeutet nicht nur eine Annäherung von qualitativen und quantitativen Forschungsbemühungen, sondern – wie der gesamte Rest des Buches verdeutlicht – auch eine andere Neu-Orientierung: Die Forderung nach „evidence based practice“ muß ergänzt werden durch „practice based evidence“.

Psychotherapiewissenschaftliche Forschung wird in allen Beiträgen als methodisch an sozialwissenschaftlich Standards orientiert ausgewiesen. Das bedeutet, Placebo-Wirkungen zu testen „is an example of the misapplication of drug trial methods“, wie es im nächsten Beitrag von **Rubaie** klar heißt. Psychotherapiewissenschaft muß sich mit vier methodischen Grunddimensionen befassen: sie bewegt sich zwischen

- den Polen der konstruktiven Bedeutungsgebung und der aufzählenden Verrechnung,
- der Generierung von Hypothesen und deren Bestätigung,
- den induktiven und deduktiven Schlussfolgerungen und
- zwischen dem Subjektiven und dem Objektiven.

Hinzu fügen könnte man dieser Auflistung noch, dass viele Beiträge eine zeitgeschichtliche Dimension sehen, die in der psychotherapiewissenschaftlichen

Forschung zu wenig berücksichtigt sei, etwa Fragen zum gesellschaftlichen Wandel, der auf Krankheitsauffassungen durchschlägt oder auf das Gender-Verhältnis, auf Rollenprobleme oder Themen der Intimität.

“Further, many distinguished natural scientists (e.g. Hayward, 1993; Maturana & Varela, 1987) concur that there is no universal methodology that can lead to the kind of certainty that many once hoped” (S. 35)

Es hat etwas Realistisches, wenn hier der Status der Ungewissheit, aus dem uns die Forschung nicht entläßt, als gleichsam nicht-anders-zu-erwartender bestätigt wird; und weil das so sei, weil eben die Hoffnungen auf Gewissheit sich nicht werden bestätigen lassen, darf man einzelne Projekte gewissermaßen pro domo, „im eigenen Haus“ durchführen – und sollte das auch. Dazu wird hier angeregt.

Der Beitrag von **Martin Milton** betont, „the notion of ‘evidence’ is not a straightforward, unambiguous, clear notion“, sondern selbst kontextabhängig. Man muß die Daten in einen Rahmen einfügen und nur in diesem sind sie interpretierbar. Er zitiert **Ernest Jones**, nach dessen Worten nur mediokre Geister von der Wissenschaft Sicherheit erwarten; im Gegenteil, die Aufgabe von Wissenschaft ist Verunsicherung, ihr methodisches Prinzip der Zweifel. Aber in der therapeutischen Situation sei ein solches Prinzip nicht durchzuhalten, hier müssen andere Rahmungen wirksam werden. Deshalb haben RCTs durchaus ihren Sinn, aber ihre Dominanz wird in Frage gestellt, „as it cannot account for individual experience and the evolving and continuing co-construction of meaning“ (S. 63). Man brauche einen Blick sowohl für klinisch-individuelle wie für allgemeine psychotherapiewissenschaftliche Fragen; das medizinische Modell, Psychotherapie wie ein Medikament zu prüfen wird als untauglich abgelehnt. Forschung könne aber Fragen nach „false causal

connections“ klären, etwa ob eine gegebene Intervention tatsächlich etwas bewirkt habe, ob eine Therapie bei Paaren, Gruppen oder Erwachsenen erfolgreich sei oder auch bei bestimmten Störungen – dafür wird die relevante Literatur kenntnisreich zitiert. Die klinische Neugier solle sich auch auf Forschungsfragen erstrecken.

“Grounded theory research of clients’ understanding of specific difficulties, discourse analytic research as to the presence of socio-political material in clients’ narratives and its relationship to distress and recovery, are just some projects that come to mind and may be feasible within most psychotherapy services with nothing more than some time available from the continuing professional development ‘allowance’” (S. 73)

Das ist gerade das Gute: dass man solche Forschungen über Narrative, über politische Meinungen von Patienten, über deren Beziehung zu Stressverarbeitung und Wiederherstellung mit einfachen Mitteln durchführen könne – wenn man sich nur die Zeit dafür nimmt.

Andrew Gumley zeigt im nächsten Abschnitt, der mit „Getting started“ überschrieben wird, wie man das machen und dabei quantitative und qualitative Methoden kombinieren kann. Man denke an einen Patienten und dessen *overt difficulties*, die man durch eine Hypothese mit *underlying psychological mechanisms* verbindet. In einem ersten Schritt braucht man ein Maß, um eine Ausgangsebene festzulegen, in einem zweiten Schritt muß man die Häufigkeit der Messung bestimmen. Dies wird an einem Beispiel genau beschrieben, wobei es darum geht, dass eine junge, an einer Psychose erkrankte Frau Hilfe bei der Bearbeitung ihrer depressiven Stimmungsumschwünge bekommt. Die Hypothese lautete, sie seien von äußerst negativen Gedanken über sich selbst verursacht, die von der Familie unterhalten und sogar unterstützt würden. Sie wurde darin sensibilisiert, diese negativen Gedanken überhaupt

wahrzunehmen, sie in ein Tagebuch zu schreiben und sogar eine Häufigkeitstabelle darüber zu führen und es wurde weiter angenommen, dass solche selbsttherabsetzenden Gedanken ihre Vermutung, sie sei „unheilbar psychisch krank“ nur bekräftigen würden. Sobald sie dann im therapeutischen Gespräch den Zusammenhang erkannte, reduzierte sich ihre depressive Symptomatik, aber auch ihre Überzeugung von der Unheilbarkeit – mit einem Wort, eine schöne Illustration dessen, wie eine Einzelfallstudie klinisch sinnvoll durchgeführt wurde und deren Erfolg sogar in Graphiken und Daten dargestellt werden kann.

Eine nachfolgende Studie von **Maureen Taylor, Jayne Redmond und Del Loewenthal** zeigt die Anwendung der Diskursanalyse am Beispiel von jungen Leuten, die Meinungen über Therapie äußern. Die eine Gruppe hatte keine eigene Erfahrung mit Therapie, die andere hatte durchaus. Am Beispiel von transkribierten Gruppeninterviews mit den beiden Gruppen wird der Unterschied im Verständnis gezeigt und, für psychoanalytische Ohren besonders aufschlussreich, sogar verdeutlicht, wie die Rolle des Untersuchers im Gruppeninterview in Gestalt von anspielenden Kommentaren verarbeitet wird. D.h., hier findet die Idee der Übertragung innerhalb des Forschungsprozesses Berücksichtigung.

Eine weitere Studie von **Julia Cayne und Del Loewenthal** untersucht, wie bei angehenden Therapeuten das Gefühl entsteht, „to call themselves psychotherapists“ (S. 117). Hier wurden Interviews zum Zweck der Erinnerung transkribiert, Bedeutungseinheiten isoliert, deren konstitutive Elemente definiert, individuelle deskriptive Portraits der Beteiligten formuliert und schließlich verallgemeinerungsfähige deskriptive Statements vorgeschlagen – man sieht in der Abfolge dieser methodische Schritte genau, wie eine solche phänomenologische Analyse zu ihren Schlussfolgerungen gelangt. Und es werden Überlegungen angestellt, inwieweit auch die eigene Entwicklung bestimmte Dinge zu würdigen oder nicht zu würdigen überhaupt erst ermöglicht – das ist ein Stück Realisierung der Forderung nach methodischer Reflexivität, wie sie qualitative Forscher immer wieder erhoben haben. In anderer, mehr klinischer Terminologie würde man sagen: die Forscher berücksichtigen ihre Gegenübertragung.

“If as Merleau-Ponty (1962) suggests, phenomenology has a concern with what lies outside (on the hither side) that is the context as well as the text, it will be interesting to consider how a phenomenological approach to research could allow the unspoken to be spoken and to enable the researchers to engage with their own experience as well as the participants’. Can we see the other while being subject too?” (S. 130 f.)

Um nun auch den Outcome zu messen, wende ich mich dem Schlusskapitel mit der Überschrift „Researching therapeutic Outcomes“ zu. **Richard Evans** stellt nach einigen Überlegungen zur Differenz

Blutleere Wissenschaft?

Kein Geringer als **Max Weber** hat in seinem berühmten Aufsatz „Wissenschaft als Beruf“ die folgenden markanten Sätze geschrieben:

„Die Gedankengebilde der Wissenschaft sind ein hinterweltliches Reich von künstlichen Abstraktionen, die mit ihren dürren Händen Blut und Saft des wirklichen Lebens einzufangen trachten, ohne es doch je zu erhaschen“.

Dass das keineswegs immer so sein muß belegt das neue Heft 1/2006 von „Psychotherapie und Sozialwissenschaft“, in welchem **Jörg Frommer** als Gastherausgeber schöne Texte zur „Biographieforschung in der psychosomatischen Medizin“ versammelt hat. Qualitative Forschung erbringt klinisch gehaltvolle Befunde bei Patienten mit akuter Leukämie zu deren subjektiven Krankheitsvorstellungen und Behandlungserwartungen; sie arbeitet die differentielle Rolle der Biographie heraus, die für das Erleben und Verarbeiten einer Herztransplantation Unterschiede macht; sie kann prototypische Konstellationen der Krankheitsverarbeitung bei einer Leberlebenspende herausfinden und zeigt, wie Biographieanalyse methodisch anspruchsvoll bei der Auswertung narrativer Interviews gemacht werden kann.

Qualitative Arbeiten haben freilich einen nicht vermeidbaren Nachteil: Man muß sie selbst lesen, wird freilich von den „blutleeren“ Zusammenfassungen statistischer Art, die einen anderen eigenen Wert haben, befreit und kommt durch den Nachvollzug der gebotenen Transkripte dazu, das Allgemeine des individuellen Falles, die Fallstruktur, prägnant zu erkennen. Das hat Kraft und Saft – und beides, qualitative und quantitative Forschung, so fängt man an zu grübeln, könnten gemeinsam dazu kommen, dem therapeutischen Leben nicht nur Blut zu entnehmen, sondern vielleicht belebendes Wissen zurückzugeben.

zwischen Forschung und professionellem Behandeln, das auch auch „reflexives Forschen“ bezeichnet wird, dar, wie man seine Arbeit mit einfachen Mitteln quantitativ bereichern kann. Dazu wird CORE vorgestellt, ein computergestütztes Programm, dessen Kürzel „Clinical Outcomes in Routine Evaluation“ bedeutet. Hier bei handelt es sich um einen einfachen Fragebogen mit 34 Fragen, wie Patienten sich vor der ersten Therapiesitzung fühlen oder wie ihr Leben während der letzten Woche verlief usw. Jede Frage wird wie üblich auf einer Skala von 0-4 beantwortet. Als computergestütztes Programm kann man dann gleich sehen, ob sich ein Patient unterhalb oder oberhalb bestimmter level befindet, die für andere große Patientengruppen bereits ermittelt sind und schließlich wird der Fragebogen nach der letzten Sitzung erneut appliziert – ein einfaches Vorher- / Nachher-Design, das für jeden psychosozialen Dienst von größtem Nutzen sein dürfte, denn so lassen sich einerseits Schwachstellen aufdecken, Nachweise für die Behandlung besonderer Patientengruppen führen und außerdem dokumentierte Rechtfertigungen für das gewählte Behandlungsprocedere ebenso wie für die eigene Dienstleistung erstellen. Der britische National Health Service bringt dieses Programm in weite Verbreitung und unterstützt seine Verwendung. Man kann es kostenlos downloaden unter www.coreims.co.uk Wahrscheinlich wird man eigene Erfahrungen damit machen müssen; immerhin wird CORE intensiv beforscht und dabei hat herausgestellt, dass es für die allermeisten Patientengruppen zuverlässige Indizes liefert und veränderungssensitiv ist. Aber nicht alle Sorten von Problemen werden natürlich erfasst. Evans stellt dies ausdrücklich als eines der „limits“ von CORE vor. Aber es gibt auch Studien, die CORE nach jeder Sitzung applizieren und so gerade bei jenen Patienten brauchbare Ergebnisse liefern, deren Veränderung

sonst nicht oder kaum bemerkt würde. Nun, das mag man klinisch anders sehen, aber ich stelle mir natürlich vor, alle Psychoanalytische Praxisinhaber installieren es auf ihren Rechnern, lassen es ein Jahr laufen und dann rechnet irgendjemand die Befunde mal zusammen – vielleicht wären wir Probleme des Nachweises unserer Effektivität und Effizienz dann rascher los als gedacht? Schöne Phantasie, ich weiß, ich weiß... Ein Einwand, der einem natürlich sogleich kommt, stammt vom nachfolgenden Beitrag im Buch von Loewenthal und Winter, nämlich von **Tirril Harris**. Diese Autorin ist in der britischen life-event-Forschung wohlbekannt; ihre Arbeiten mit **George Brown**, Sozialwissenschaftler am Birkbeck College mit besonderen Interessen an der Medizinsoziologie und damit an der Psychiatrie, sind in England unvergessen. Brown hatte (1958) als einer der ersten in einer großen Untersuchungen

gezeigt, dass die Rückfallquote von 229 Schizophrenen bei Entlassungen davon bestimmt war, ob sie in ihre Familien oder in Heime gehen. Diejenigen, die aus der Psychiatrie in Heime kamen, hatten die weit geringere Rückfallquote. Daran hatten sich Untersuchungen über Depressive angeschlossen, die zeigten, dass jene Depressiven, die zu ihren Ehepartnern kamen, weit höhere Rückfallquoten hatten – eine ganze Lawine von Forschungen zu „expressed emotions“ und „life events“ schloß sich hier an, wobei Brown und Harris vielfach zusammenarbeiteten und publizierten.

Im Buch von Loewenthal und Winter nun wird von Tirril Harris sogleich darauf aufmerksam gemacht, dass wir psychometrische Daten natürlich nur im Rahmen einer ätiologischen Theorie sinnvoll betrachten können; dazu wird (S. 246) am Beispiel der Depression der folgende Vorschlag gemacht:

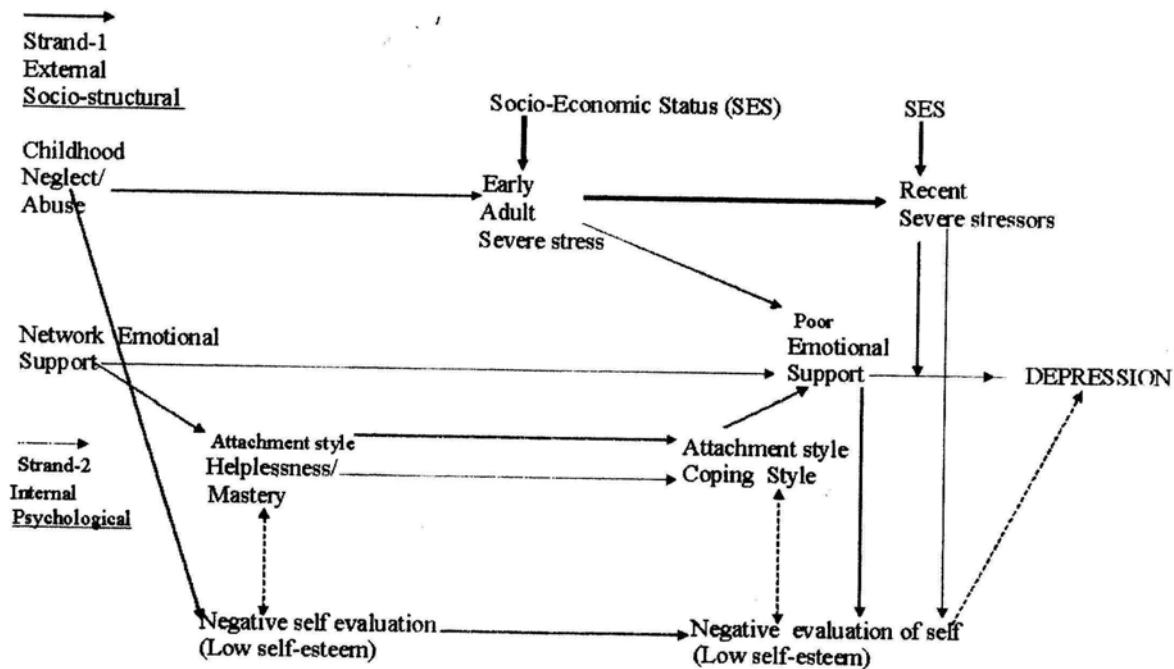


Figure 1. Key psychosocial factors in the aetiology of depression: a life-span model. Dotted lines represent impact of inner world (strand 2) on outer world (strand 1) and inner world. Continuous lines represent impact of outer world (strand 1) on inner world (strand 2).

Das sieht immer komplizierter aus, als es ist. Man muß nur sehen, dass für die Depression sowohl der soziale Status (früher und heute) auf einer waagrechten Linie dargestellt wird, dann das Netzwerk

sozialer Hilfen und schließlich die früh und später erworbenen Bindungserfahrungen eingetragen sind – dies alles hat dann Auswirkungen auf die bei Depressiven so charakteristische geringe Selbstachtung.

Ein solches Modell kann nun verschiedene Untersuchungsmöglichkeiten initiieren; man kann die Rolle der Verstärkung des sozialen Netzwerkes evaluieren, (wenn depressive Menschen in einer therapeutischen Gruppe soziale Kontakte leichter finden, verbessert sich dann ihr Zustand?); wenn das Bindungsmuster therapeutisch bearbeitet wurde, lassen sich dann ebenfalls Veränderungen feststellen usw. In einem Beitrag von **Julia Buckroyd** werden analoge Vorschläge für die psychoanalytisch informierte Erforschung von Essstörungen gemacht.

Man sieht, es lässt sich Forschung mit durchaus interessanten Fragestellungen machen – aber man muß nicht „die Methode“, man kann die eigene Tätigkeit und damit auch die eigene „Person“ selbstevaluieren. Weil wir damit wieder bei der Person angelangt sind, wird es von Interesse sein zu erfahren, wie sich **Andrew R. Arthur** die Frage: „Who do you think you are? A study of how psychotherapists' thinking styles affect orientation choice and practice“ beantwortet.

Die „persönlichen Epistemologien“ sind sein Thema; gemeint ist, was einen Therapeuten wohl dazu bringt, eher Verhaltenstherapeut als Psychoanalytiker zu werden (oder umgekehrt). Vermutet wird, dass hier Persönlichkeitsstile und Präferenzen für eine bestimmte schulische Orientierung kovariieren könnten. Einige grundlegende Dichotomien werden dazu ausgebreitet:

- Subjektivismus – Objektivismus

- Theoretische Orientierung – Faktenorientierung
- Persönlicher Wille – unpersönliche Kausalitätsannahmen
- Elementarismus – Holistisches Denken
- Quantitative Orientierung – qualitative Orientierung
- Physikalismus – Ablehnung des Physikalismus

Dazu gibt es nun schon einige Orientierungen und Fragebögen, deren Ergebnisse sich zu Stil-Profilen zusammenbündeln lassen. Ein mich natürlich erfreuender Befund lautet:

“The predicted predominant style for psychoanalysts was metaphorism as the highest scale, followed by rationalism and empiricism. For behaviorists, the prediction was empiricism as the highest scale, followed by rationalism and metaphorism” (S. 299)

Die Ergebnisse entsprechen dem im Großen und Ganzen und so erschließt sich hier eine Dimension des Persönlichen, die sogar empirisch erforscht werden kann; tatsächlich haben wohl alle Kliniker mehr oder weniger diffus eine Vermutung darüber, warum jemand dieser oder jener schulischen Orientierung sich wohl angeschlossen hat. Hier finden wir Befunde dazu und auch eine Vermutung über die großen schulischen Differenzen.

“In short, individuals' personalities mirror their overall philosophical worldviews”,

so zitiert der Autor andere Autoren zustimmend nach Durchführung einer eigenen Untersuchung mit den entsprechenden Fragebögen.

SCHULE ODER STILE ?

Natürlich kommen einem während der Lektüre solcher Studien eine Menge Gedanken. Z.B. könnte es ja sein, dass die schulischen Differenzen hinsichtlich solcher persönlicher Epistemologien von Therapeuten viel besser abgebildet und verstanden werden könnten als durch den Vergleich von „angewandten“ und „erprobten“ „Methoden und Interventionen“. Vielleicht denken Therapeuten einfach verschieden, nicht nur verschiedene Inhalte, sondern auch in verschiedenen Stilen und vielleicht haben wir alle die Neigung, den

eigenen Stil für den einzig denkbaren zu halten – was auch wiederum kein Wunder wäre, nicht?

Hier aber kann man sich über manches an anderen Denk-Möglichkeiten belehren lassen, v.a. auch darüber, dass empirische Forschung zu betreiben oder aber lieber empirische Praxis vielleicht selbst eine Präferenz für solche Dimensionen darstellen könnte? Und die Belehrung könnte lauten, dass man sich nicht vom eigenen Stil umgrenzen und einengen lassen muß, sondern über den jeweiligen Tellerrand hinausschauen könnte. Mir will scheinen, das zuletzt hier besprochene Buch mit den Zusammenfassungen der UKCP-Tagung gibt eine Fülle von praktischen Anregungen, wie beinahe Jede und Jeder sich zu einem nützlichen psychotherapiewissenschaftlichen Forscher mit *unaufwendigen* Mitteln weiter entwickeln könnte.

Daß man mit der Alternative „Methode oder Person“ beginnt und bei der Frage „Schule oder Stil?“ ein neues Feld öffnet, ist kein schlechter Hinweis dafür, dass da etwas Richtiges liegt – wenn die Forschung den Blick verändert, erweitert und auf Neues fokussiert, hat sie den Grad an Verunsicherung ermöglicht, der unsere Neugier und damit uns lebendig erhält. Mehr kann man nicht erwarten. Ich schließe also mit den Worten aus dem Dialog mit **Onkel Toby** aus dem wundervollen Klassiker „**Tristram Shandy**“:

“So, --- gerade so, meine Herren Kollegen und Mitstreiter auf diesem großen Acker der Gelehrsamkeit, dessen Ernte jetzt vor unseren Augen reift; so sind, mit langsamen Schritten gelegentlichen Wachstums, unsere physikalischen, metaphysischen, physiologischen, polemischen, nautischen, mathematischen, aenigmatischen, technischen, biographischen, literarischen, chemischen und obstretischen Kenntnisse nebst fünfzig weiteren Zweigen (die meist auch auf *ischen* endigen) während der letzten zwei Jahrhunderte oder länger schon allmählich in Richtung...ihrer Vollkommenheit emporgekrochen, von der wir, wenn man nach den Fortschritten der letzten sieben Jahre urteilen darf, unmöglich noch weit entfernt sein können“.