

PSYCHO-NEWS-LETTER NR.7

EIN KLEINER ZEITSCHRIFTENÜBERBLICK

Michael B. Buchholz

Ende April 2003

WIE VERBREITET SIND PSYCHISCHE STÖRUNGEN ?

Daß es psychische Störungen gibt, wird niemand mehr sinnvoll bezweifeln; dass Psychotherapie hilft, ebenso wenig. Das hat die hervorragende Studie von Wampold (2001), die ich an anderer Stelle rezensiert habe, nachgewiesen. Die Rezension kann man auf der Homepage der DGPT lesen, deshalb brauche ich hier Wampolds Befunde nicht zusammenzufassen. Wampolds Befund, quer über alle internationalen meta-analytischen Studien hinweg, zeigt mit größtem Nachdruck, dass das medizinische Modell „out“ ist; in Kürze also: die Idee, spezifische Methoden für spezifische Störungen zu finden, erweist sich empirisch als kontraproduktiv. Dieses Modell als Leitvorstellung für die Erforschung des *Zusammenhangs von Störung und Behandlungseffekten*, kann und muß in erhebliche empirische Zweifel gezogen werden. Meine letzten News-Letters haben deutlich gemacht, dass es höchst interessante Forschungsfragen gibt, die über die statistische Evaluation von Behandlungserfolgsmessung hinaus gehen und gleichzeitig die Beschränkung dieses thematischen Zugriffs deutlich machen. Und das bedeutet berufspolitisch erfreuliches: Ein Ende des Schulenstreits kommt in Sicht! Man muß nicht mehr die eine gegen die andere Schule ausspielen, um kleine empirische Vorteile zu gewinnen.

Empirische Methoden kommen an anderer Stelle sehr viel besser zum Einsatz, eben bei der Untersuchung der epidemiologischen Frage nach der Verbreitung psychischer Störungen.

Die neu entstehenden Psychotherapeutenkammern haben sich einen neuen und eleganten Ausweis gegeben, der berufspolitische Anliegen mit wissenschaftlichen Themen verbindet. Im Null-Heft des „Psychotherapeutenjournal“, wie sich die neue Gazette berufspolitisch einkleidet, ist eine hier relevante und interessante Studie von Wittchen und Jacobi erschienen: „Die Versorgungssituation psychischer Störungen in

Deutschland. Eine klinisch-epidemiologische Abschätzung anhand des Bundes-Gesundheitssurveys 1998“. Die Jahreszahl am Ende gibt also sogleich an, dass es hier um aktuelle Daten geht. Und die Eingangsbemerkungen belehren auch darüber, dass in *diesem* Kontext das medizinische Modell nicht „out“ sein kann. Auch wenn die Behandlung von Störungen nicht in der Analogie zur Medizin gedacht werden kann, muß man dennoch das Leiden als krankheitswertig definieren, wenn man seine Verbreitung statistisch analysieren will.

Das geschieht hier unter Berufung auf die Standards von ICD-10 bzw. DSM-IV. Diese Instrumente „sind für den Einsatz in epidemiologischen Bevölkerungsuntersuchungen geeignet“, schreiben die Autoren (S. 7) und machen damit deutlich, dass sie nicht unbedingt für die professionelle Praxis Bedeutung haben und auch nicht, wenn man Wirkungen von Interventionen erfassen will. Aber für die Erfassung von Verbreitung sind sie das sehr wohl. Hier geht es um Häufigkeiten, in den Behandlungskontexten geht es um Bedeutungen. Für die Störungen der ICD-10, die mit F1xxx, F2xxx, F3xxx bezeichnet werden (Störungen mit abhängig machenden Suchtmitteln, psychotische und affektive Störungen) sowie für Essstörungen (F50), somatoforme und Angststörungen kann man sagen, dass sie 75% der Gesamtmorbidität ausmachen. Es gibt für alle diese Störungen Nachweise der Behandlungsbedürftigkeit und der –möglichkeiten.

Definiert werden muß auch, wann ein Fall als behandlungsbedürftig gilt und wann er als versorgt angesehen werden kann. In die sog. Diagnostischen Bedarfskriterien fließen Größen wie Schweregrad, Spontanverlauf, Behinderungen und Einschränkungen ein. Schwieriger aber als diese Größen wird die Erfassung durch den Umstand, dass höchst unterschiedliche institutionelle Kontexte (Psychiatrie, andere stationäre Einrichtungen,

Hausärzte, psychotherapeutische Ambulanzen, Beratungsdienste etc.) hier eine Rolle spielen. Und wie um auch hier auf die problematische Rolle des medizinischen Modells aufmerksam zumachen, schreiben die Autoren (S. 8): „Hinzu treten, wie auch bei anderen Indikationsgebieten, zusätzliche Aspekte, die eine Gleichsetzung von Diagnose und Behandlungsbedarf trotz der expliziten diagnostischen Kriterien problematisch erscheinen lassen. Ein Beispiel hierfür ist, dass bei psychotherapeutischen Verfahren eine verfahrensspezifisch unterschiedliche ‚ausreichende Behandlungsmotivation‘ vorliegen muß“. Darauf hatte ich in den letzten Rundbriefen schon aufmerksam gemacht: die **Zielvorstellung bzw. die Verhandlungen über Zielvorstellungen** sind für Psychotherapeutischen Erfolg unverzichtbar, also ein höchst individuelles und subjektives Moment, das einen der Unterschiede zur Medizin markiert.

Sie haben ihre Effekte dann auch auf die statistische Bedarfsschätzung. Um diese Schwierigkeiten zu umschiffen, werden in dieser Studie nun keine versorgungs- oder verfahrensspezifischen Differenzierungen vorgenommen. Als „behandelt“ gilt ein Fall, wenn er wegen einer Diagnose zumindest mit einer Institution des Gesundheitssystems in Kontakt stand.

Die Daten stammen aus dem ersten Bundesgesundheitsurvey des Robert-Koch-Instituts, mit dem vom Sommer 1998 bis Frühjahr 1999 die psychische Gesundheit der erwachsenen Bevölkerung untersucht wurde. Über 4100 screening-positive Probanden wurden einer von 87 möglichen Diagnosen mit einem Computerprogramm zugeordnet.

Hauptstörungsgruppen waren Missbrauch/Abhängigkeit von psychotropen Substanzen, affektive Störungen, Angststörungen, somatoforme Störung und andere Diagnosen). Es wurden 12-Monats-Diagnosen verwendet – ob also jemand in diesem Zeitraum die vollen Kriterien einer Diagnose erfüllt oder nicht. Erfasst sind also nur voll ausgeprägte klinische Zustandsbilder.

Screening

Bei epidemiologischen Untersuchungen wird häufig in zwei Schritten vorgegangen: Man untersucht mit geeigneten Fragebogeninstrumenten einen Personenkreis – das ist das Screening. Damit erhält man einen ersten, vergleichsweise groben Überblick. Dem folgt bei den so ermittelten auffälligen Personen („screening-positiv“) ein klinisches Interview zur Sicherung bzw. Verwerfung einer Diagnose.

Ergebnisse: 32,1 % der Bevölkerung zwischen 18 und 65 Jahren erfüllen die Kriterien für eine psychische Störung; am häufigsten somatoforme (11%), phobische (7,6%), Major-Depression (8,3%) und alkoholbedingte Störungen (6,2%). Von diesen 32,1% psychischen Störungen in der Gesamtbevölkerung sind 36,4% im Sinne der obigen Definition „versorgt“, d.h. hatten mindestens einen Kontakt mit einer Institution des Gesundheitsdienstes. Substanzstörungen sind am schlechtesten, die „anderen psychiatrischen Störungen“ am besten versorgt. 13% gaben an, eine Behandlung erhalten zu haben, ohne dass an sie eine 12-Monats-Diagnose vergeben worden sei. Die Behandlungsquoten sind regional sehr unterschiedlich, in den neuen Bundesländern insbesondere bei den Angststörungen sehr viel schlechter.

In 42,3% der Fälle wurde zunächst der Hausarzt kontaktiert, aber er half selbst nur in 14,5% weiter. Dann wurden Nervenärzte, psychologische Psychotherapeuten und ärztliche Psychotherapeuten (in dieser Häufigkeitsreihenfolge) in Anspruch genommen. Man muß bei dieser Schätzung des Versorgungsgrades ausdrücklich berücksichtigen, dass hier ein sehr weiter Versorgungsbegriff angenommen wurde. Die Inanspruchnahme des Hausarztes, ein nur einmaliger Kontakt mit einer Institution – dies alles zählt statistisch bereits einen Fall als „behandelt“. Faktisch also dürften die Dinge weniger günstig liegen. Auch wird nichts über die Angemessenheit der Behandlungen untersucht. In einer anderen Untersuchung waren es nur 9%, die eine „adäquate“ Behandlung erhielten.

In ihrer **Diskussion** stellen die Autoren fest, dass diese Ergebnisse mit der internationalen Forschungslage kovariieren: psychische Störungen sind weitaus häufiger als gedacht. Die überwiegende Mehrzahl freilich wird nicht behandelt. Hier freilich muß diagnosenspezifisch differenziert werden: Angst- und Panikstörungen werden zu 67%, die somatoformen und Substanzstörungen hingegen weitaus geringer versorgt. Dafür dürften ein Bündel von Faktoren verantwortlich sein, die von öffentlicher Wahrnehmung der Störung bis zu unterschiedlichen Graden der Behandlungsbereitschaft reichen. Aber es lässt sich auch feststellen, dass der Versorgungsgrad in Regionen mit universitärer Anbindung ungleich höher ist. Das könnte Überlegungen zur Ansiedlung von Instituten der Ausbildung

und Versorgung in solchen Regionen nach sich ziehen.

DIE GROßSTADT UND DIE BULIMIE

Damit sind wir bei einem anderen, interessanten Aspekt – der Beziehung zwischen psychischen Störungen und ihrer Umwelt. Schon bei den älteren Soziologen, etwa bei Georg Simmel in seiner „Soziologie“ von 1908 finden sich bemerkenswerte Überlegungen über die von einer Großstadt produzierte „Nervosität“. Der Bielefelder Historiker Joachim Radkau (1998) hatte diesem Thema eine fabelhafte Studie „Das Zeitalter der Nervosität“ gewidmet, in welcher er vorzüglich zusammenfasst, wie die Menschen damals über solche Wirkungen diskutierten, aber auch, welche Beiträge die „Nervosität“ und ihre Diskurse zur Entstehung des ersten Weltkrieges lieferten, der ja bekanntlich von vielen Menschen als „Befreiung“ erlebt wurde – als Befreiung von eben dieser „Nervosität“. Daß damals schon die Essstörungen als Teil solcher „nervösen Störungen“ breit diskutiert wurden, verdient eigentlich mehr als nur diese Bemerkung am Rande.

Seitdem hat man gründliche Untersuchungen insbesondere zur Anorexie vorgelegt, die ziemlich einheitlich Prävalenzraten um 0,60% ermittelten. Eine Zunahme konnte – entgegen der fama – in Prävalenzstudien nicht ermittelt werden. Das verhält sich freilich anders bei der Bulimie.

Eben dieser Essstörung widmet sich eine weitere epidemiologische Untersuchung von Bahrke u.a. „Die Verbreitung pathologischen Essverhaltens unter Schülerinnen und Schülern einer Großstadt“ (PPmP 1/2003). Bulimie tritt nicht nur bei Studentinnen, sondern bereits bei Schülerinnen auf. Die Prävalenz der Bulimie *unter Schülerinnen* ist weitaus größer als gedacht, sie liegt je nach den verschiedenen Studien zwischen 5% und 8%.

Das wird hier an 2844 Schülerinnen der 10. Klassen der Stadt Halle untersucht. Unter Zusicherung von Anonymität erhielten alle Schüler und Schülerinnen das Eating Disorder Inventory (EDI) mit 64 items – aus ökonomischen Gründen blieb man bei diesem screening, ohne eine Diagnose durch Interviews zu sichern. Das Alter schwankte zwischen 14 und 18 Jahren. EDI erfasst Faktoren wie den „Drang zum Dünnsein“, „bulimisches Verhalten“, „die Wahrnehmung eigener Gefühle“, „Gefühle der Unzulänglichkeit“ etc.

Übergewichtige sind besonders häufig in den Sekundarschulen anzutreffen, aber andere Unterschiede hinsichtlich Geschlecht oder Schultyp konnten nicht gefunden werden. In der Gesamtstichprobe fand sich eine Anorexiegefährdung bei 1,4%; bulimisch gefährdet waren sehr viel mehr, nämlich 10% der Schülerinnen und 2% der Schüler. Weiter fiel den Autoren auf, dass bestimmte items („Ich fühle mich unzulänglich“ und „ich mag enge Beziehungen zu anderen Menschen“) am häufigsten unbeantwortet blieben. Das spielt methodisch eine nicht unerhebliche Rolle. Es gibt nämlich andere Studien, in denen 25% essstörungsanfällige Mädchen gefunden wurden. Auch in diesen anderen Studien waren einzelne Fragen unbeantwortet geblieben und das hatte dazu veranlasst, zu viele Probanden nach dem screening in den zweiten Schritt der Interviews mit aufzunehmen. Man riskiert dann einen Fehler Typ II, nämlich die Inkaufnahme falsch positiver Antworten. Dennoch könnte es sein, dass der „wahre Wert“, wenn denn der zweite Schritt noch getan worden wäre, höher als die genannten 10% für die Bulimie läge. Immerhin haben in

Epidemiologie – Einige wenige Grundbegriffe

Prävalenz

Damit ist die Häufigkeit eines Merkmals in einer Population zu einem bestimmten Messzeitpunkt gemeint. Die Prävalenz wird bestimmt vom **Standardmessfehler**. Hierbei wird angenommen, dass ein Messfehler umso kleiner ist, je größer die Stichprobe. Findet man in einer Stichprobe ein Merkmal zu 20% vorliegend (Prävalenz), dann ist der „wahre Wert“ bei einer Stichprobe von 1000 Personen zwischen 17,5% und 22,5%. Ist die Stichprobe hingegen nur 150 Personen groß, liegt der wahre Wert zwischen 13,5% und 26,5% (Dafür gibt es statistische Formeln...)

Die **Inzidenz** hingegen definiert die Zahl der Neuerkrankungen bezogen auf die gesamte Bevölkerung, in der diese Erkrankungen auftraten.

Neben **deskriptiven** Studien (wie die von Wittchen und Jacobi) werden **analytische** Studien durchgeführt, die soziale Einflüsse auf Entstehung und Verlauf von Erkrankungen erfassen.

Ein **relatives Risiko** (risk ratio) ergibt sich dann aus dem Verhältnis der Inzidenz bei exponierter zu dem bei nicht-exponierter Bevölkerung. Beispiel: Raucher und Krebserkrankung. Angenommen in einer Bevölkerung von 5000 Personen gibt es 100 erkrankte Raucher und 50 erkrankte Nicht-Raucher. Dann ergibt sich das relative Risiko, indem man beide ins Verhältnis zueinander setzt. Es ist der Faktor 2. Wer raucht, hat eine doppelte relative Chance zu erkranken.

Daraus lassen sich dann Möglichkeiten für die Messung der Effekte präventiver Maßnahmen errechnen. Wenn Rauchen als Risikofaktor wegfiel, könnten x% der Erkrankungen verhütet werden. Diese Angabe wird als **attributables Risiko** bezeichnet, bezieht sich also darauf, was man dem Rauchen zuschreiben kann.

der untersuchten Altersgruppe unabhängig von Schultyp oder Geschlecht mehr als 7% ein

anorektisches Körpergewicht.

ADIPOSITAS UND EINE INTERVENTION MIT DEN KASSEN

Auch die Prävalenz einer anderen Essstörung, der Adipositas steigt, insbesondere im Jugendalter und insbesondere in den westlichen Industrieländern; Prozesse gegen McDonald haben weite öffentliche Aufmerksamkeit gefunden.

Im gleichen Heft von PPM (1/2003) wird von Barnow, Bernheim, Schröder u.a. über „Adipositas im Kindes- und Jugendalter“ berichtet. Auch hier geht es um den Deutschen Osten, nämlich um Mecklenburg-Vorpommern. Diese Studie ist auf die Evaluation von Interventionen ausgerichtet.

Die Stichprobe ist deshalb sehr klein, untersucht wurden 19 Kinder zwischen 9 und 16 Jahren und diese mit einer nicht-adipösen Gruppe von $n = 38$ verglichen. Zugrunde lag der Body Mass Index (BMI), d.h. der Quotient aus Gewicht und Körperlänge.

Diese Studie wurde nun in **Zusammenarbeit mit den Kassen** durchgeführt. Es handelte sich um Kinder aus der Umgebung von Greifswald, die Eltern mussten sich schriftlich zu einer aktiven Unterstützung der „Schulung“ (S.9) sowie zur Teilnahme an den medizinisch-psychologischen Einganguntersuchungen verpflichten.

Mit verschiedenen Fragebögen wurden Parameter wie Selbstwertgefühl, Verhaltensprobleme und das Ausmaß der körperlichen Beschwerden, teils kam es zu Beurteilungen auch durch die Eltern (Child Behavior Checklist, CBCL). Die **Intervention** bestand in 2 wöchentlichen Doppelstunden, wobei mit den Kindern Sport getrieben, gesunde Ernährung und deren Erprobung beigebracht wurden sowie eine verhaltensorientierte Gruppentherapie stattfanden. Schwimmen und weitere bewegungstherapeutische Elemente wurden hinzugenommen. Dies alles dauerte ein Jahr.

Die **Ergebnisse** zeigen nun, dass die Eltern der adipösen Kinder diese anfänglich als depressiver, mit mehr Körperbeschwerden, sozialen Aufmerksamkeitsproblemen und mehr Aggressionen behaftet einschätzten als die Eltern der Kontrollgruppe. Aber durch die „Intervention“ nahm der BMI bei den adipösen Kindern **nicht** ab. Der Anstieg positiver Selbstaussagen in den Fragebögen wird als

„tendentiell“ bezeichnet, die Abnahme negativer Selbstaussagen als signifikant. Auch haben die Eltern weniger soziale Probleme und Aggressionen berichtet. Aber die Selbsteinschätzungen der Kinder haben sich nicht wesentlich verändert; das muß wohl als Tendenz der Eltern, im Sinne positiver sozialer Erwünschtheit zu antworten, verstanden werden, wie die Autoren auch andeuten (S. 12)

Die Effektstärken ihrer Maßnahmen geben die Autoren mit 0,42 bis 0,58 an; das liegt nicht, wie sie meinen, im mittleren Bereich, sondern eher im unteren. Eine Gewichtsreduktion konnte nicht erreicht werden. Bemerkenswert für die grundsätzliche methodologische Fragestellung ist die Feststellung der Autoren: „Allerdings ist zu bedenken, dass die einzelnen Studienteilnehmer unterschiedlich von der Intervention profitierten. Während einige an Gewicht verloren, gab es andere, deren Gewicht sich nicht veränderte bzw. die trotz Intervention weiter zunahmten. Weitere Analysen an größeren Stichproben sind notwendig, um herauszufinden, welche Charakteristika es sind, die eine bessere bzw. schlechtere Prognose erlauben...“ (S. 12).

Nun, das ist im Grunde ein recht erstaunliche Bemerkung. Denn sie gesteht ja zu, dass die gleiche Dosis des Psychotherapiemedikaments ganz unterschiedliche Wirkungen hatte, selbst nachdem man sich doch so sehr bemüht hatte, die diagnostischen Kriterien für die untersuchte Störung genau festzulegen. Danach hat es sich um dieselbe Störung bei allen Kindern gehandelt – aber dasselbe ist eben einfach nicht immer dasselbe. Das ist in der amerikanischen Literatur selbst wiederum gut untersucht worden.

Ich nenne ein Beispiel renommierter Autoren: Luborsky, McLellan et al (1997)¹ führten eine Untersuchung mit drogenabhängigen und depressiven Patienten durch. Die Therapeuten kamen bei allen Untersuchungsbedingungen mit der von ihnen favorisierten Methode zum Zuge. Obwohl die Patienten bezüglich der Diagnose homogenisiert und randomisiert den

¹ Luborsky, L., A.T. McLellan, L. Diguier, G. Woody und D.A. Seligman (1997): The psychotherapist matters: Comparison of outcome across twenty-two therapists and seven patient samples. *Clinical Psychology: Science and Practice*, 4, 53-65.

Therapeuten zugewiesen waren, die Therapeuten supervidiert wurden und nach Therapiemanualen arbeiteten, ergab sich nicht die erwartete Maximierung der therapeutischen Kompetenz und Minimierung der Outcome-Differenzen, die bei allen Patienten (die mit gleicher Methode behandelt) hätten annähernd gleich sein müssen. Die 22 Therapeuten hatten vielmehr eine breite Variabilität in ihren Ergebnissen, die von leichten Verschlechterungen bis zu außerordentlichen Verbesserungen bei den Patienten führten. Das Ergebnis musste vielmehr anders verstanden werden: Therapeuten, die bei einem Patienten erfolgreich waren, waren es mit großer Wahrscheinlichkeit auch bei anderen – und umgekehrt. **Was Therapie ausmacht, liegt mehr beim Therapeuten als in der Methode** – so wäre, überblickt man die Ergebnisse des Projekts MATCH, die Quintessenz zu formulieren. Diese – empirischen !!! - Ergebnisse sprechen klar gegen das medizinische Modell in der Praxis und gegen das Dosis-Wirkungs-Paradigma in der Forschung, klar gegen die Spezifitätsprogrammatik, klar gegen die EST-Kriterien, klar gegen die Kriterienliste des Wissenschaftlichen Beirates beim Psychotherapeutengesetz.

Richtiger und mit den empirischen Befunden in besserer Übereinstimmung scheint es vielmehr zu sein, professionelles Können als mit der Person des Therapeuten verbunden anzunehmen. Eine Theorie kann dann nicht „angewendet“ werden wie vielleicht in den Paradigmen der klassischen Physik; vielmehr hat eine Theorie in der Praxis ganz andere Aufgaben. Sie

- a) schafft Freiräume für die persönliche Ausgestaltung durch den professionellen Behandler und seinen Patienten/seine Patientin
- b) fördert individuelle Sensibilitäten.

Man muß dann also eher sehen, wie man diese Freiräume öffnet und offen halten und wie man Sensibilitäten verfeinern und entwickeln kann. Und es scheint klar zu sein, dass die Tendenz, Behandlungen manualisieren zu wollen, solche notwendigen Freiräume eher beschränkt. Das wäre aber dann für die psychotherapeutische Praxis sogar eher schädlich.

Was in der somatischen Medizin sinnvoll sein mag, wäre in der psychotherapeutischen Praxis dann umgekehrt eher schädlich.

AGAIN AND AGAIN – DAS DOSIS-WIRKUNGS-MODELL...

Ich kehre nun zu der Adipositas-Studie zurück. Man sieht auch hier wieder, dass therapeutische Interventionen höchst unterschiedlich aufgefasst werden können. In der pädagogischen Psychologie spricht man von der „Vorkenntnisebene“, d.h. manche Kinder profitieren von bestimmten Maßnahmen im Mathe-Unterricht, andere eben nicht. Das ist abhängig eben von der Vorkenntnisebene. Man muß daraus nicht die Schlussfolgerung ziehen, dass nur höchst individualisierte Studien hier weiter helfen können, aber ebenso wenig ist die Schlussfolgerung zwingend, dass Studien mit größeren Stichproben „notwendig“ für die Klärung der Frage nach der individuellen Wirksamkeit seien. Zwingend erscheint mir allein das Zugeständnis, dass „Dasselbe“ eben nicht immer Dasselbe ist. Bei einigen hilft die „Dosis“, also die hier konzipierte Intervention, bei anderen eben nicht. Und das, obwohl doch soviel Sorgfalt darauf verwendet wurde, die Adipositas-Diagnose so einheitlich wie nur möglich zu definieren. Manchmal, so wage ich hier etwas ironisch einzufügen, warte ich auf einen Roman, in welchem verschiedene Helden

namens Addi „Beau“ Sitas, Boo Limmy und Arno Remy gemeinsam „Durch Dick und Dünn“ gehen – so sollte der Roman dann auch heißen. Am Ende entdecken sie, so träume ich dann, ihre große Verbundenheit bei gleichzeitig großer Verschiedenheit und Einsamkeit ...

Offenbar spielen für die Untersuchung der *Zusammenhänge* zwischen irgendeiner Störung und der Wirkung einer „Intervention“ ganz andere Momente eine Rolle; ich glaube, wir müssen uns bei der psychotherapeutischen Prozessforschung in der Tat vom Dosis-Wirkungs-Modell (einer einheitlich meßbaren Dosis bei einheitlicher Störung) verabschieden, nicht aber bei epidemiologischen Untersuchungen. Der therapeutische Prozeß einerseits und das statistische Vorliegen von Störungen andererseits brauchen offenbar ganz unterschiedliche Grundkonzeptionen und Methodologien, um verstanden und analysiert zu werden. Das wird mit dem interaktiven Charakter der Psychotherapie zusammenhängen, mit der höchst vielfältigen Variation von Bedeutungen, mit dem komplexen Zusammenspiel von Maßnahmen und beteiligten

Persönlichkeitsstilen, Behandlungsmotivationen und Sinngebungen. Es erweist sich auch in diesen Studien immer mehr als Fiktion, Psychotherapie *in Analogie* zur medizinischen Gabe eines Medikamentes aufzufassen und ich finde es immer wieder bemerkenswert, dass diese Fiktion sich sogar mit empirischen Mitteln immer wieder zeigen lässt.

Mindestens ebenso bemerkenswert freilich ist, dass sich solche Äußerungen wie die über die ganz unterschiedliche Wirkung bei verschiedenen Personen (ich habe sie oben zitiert) in fast allen Studien der empirischen Psychotherapieforschung finden – und dennoch nicht mit Aufmerksamkeit besetzt werden! Es ist, als könne man sich in der empirischen Forschung nicht leicht auf die eigenen Einsichten einstellen. Die empirische Kritik am Dosis-Wirkungs-Modell richtet sich also *nicht* gegen die Annahme, dass es Störungen gibt, die als krankheitswertig angesehen werden müssen. Sie richtet sich viel enger gegen die Vorstellung, man könne Psychotherapeutische Prozesse in Analogie zur medizinischen Medikamentengabe untersuchen – das ist etwas ganz anderes! Fraglos gibt es krankheitswertige psychische Störungen und ebenso fraglos gibt es psychotherapeutische Behandlungen, die weiterhelfen – weil das so ist, bleiben die Krankenkassen zuständig. Aber ob man das, was in den Behandlungen geschieht, mit einer medizinischen Metapher untersuchen und verstehen kann, muß mittlerweile als empirisch erheblich in Zweifel gezogen angesehen werden. Das Problem ist die Einengung der Untersuchung von „Hilfe“ (bei psychischen Störungen) auf die lineare Untersuchung von Interventionseffekten. Das Problem ist die Begrenzung der Forschung auf allein biostatistische Modelle und deren basale Annahmen.

Übrigens bin ich nicht der erste und nicht der einzige, der diese Dinge beklagt. Der ehemalige Vorsitzende der internationalen „Society of Psychotherapy research“ (SPR), William Stiles, schrieb zusammen mit David Shapiro (1989) einen bemerkenswerten Aufsatz über „Abuse of the Drug-Metaphor in Psychotherapy Research“, worin er beklagte, dass die therapeutische Maßnahme gemessen werden solle und dann verglichen wie die Wirkung eines Medikamentes. Und die beiden renommierten Autoren folgerten, dass diese Metapher auch Ursache für den fatalen Wettstreit zwischen den verschiedenen Therapie-„Medikamenten“ ist.

Mittlerweile hat diese, die Forschung leitende, aber unbewusste Metapher, deutlich sichtbare Folgen. Ein Jeder, der ein neues „Medikament“ (er-)findet, prüft es und erfährt, dass er von den Vertretern herkömmlicher Medikamente aus methodischen Gründen für unglaubwürdig erklärt, mindesten aber angezweifelt wird. Die Verhaltenstherapeuten verfahren so mit den psychodynamischen Therapien, diese revanchieren sich und gemeinsam verwehrt man den systemisch Orientierten ihren Platz. Das alles hat politische, v.a. handfeste wirtschaftliche Hintergründe – aber so entsteht eine Situation wie in den Auseinandersetzungen um die Atomkraftwerke: Man musste nur das Institut kennen, aus dem ein Gutachter kam und dann wusste man, ob er pro oder contra argumentieren würde. Das ist übrigens im Bereich der Psychotherapieforschung selbst wiederum gut belegt: Es gibt Studien, die zeigen können, dass man die Ergebnisse einer vergleichenden Psychotherapiestudie vorhersagen kann, wenn man die schulische Orientierung des Autors aus dessen anderen Veröffentlichungen kennt. Der kognitive Verhaltenstherapeut Mahoney (1976)² hat das wiederum empirisch getestet. Er versandte fingierte Manuskripte an Redaktionen verschiedener wissenschaftlicher Zeitschriften. Die Einführungsteile und die Abschnitte über die angewandten Methoden waren identisch, nicht aber die Resultate. Es zeigte sich, dass die Chance zur Annahme des Manuskripts wuchs, wenn die Ergebnisse mit der theoretischen Orientierung der Zeitschrift übereinstimmten und daß sie abgelehnt wurden, wenn die Ergebnisse der theoretischen Grundorientierung widersprachen.

Auch Wissenschaftler selektieren konforme Informationen, Informationen also, die konform sind mit ihren Auffassungen. Auch Wissenschaftler sind - Subjekte. Freilich, aber die Gesamtsituation hat nun dazu geführt, dass jede Studie von vorneherein unter dem Verdacht steht, nur noch pro domo zu argumentieren. Dem Italiener ist Italien, dem Deutschen Deutschland das liebste Vaterland. Diese Treue zum Heimatland der Schule, aus der man stammt, dieser geradezu rührende Patriotismus führt aber in der öffentlichen Wahrnehmung zum **fatalen Effekt einer wachsenden Vertrauenserosion**. Zerstört wird dabei nicht nur Vertrauen in die eine oder andere Studie, oder in die eine oder andere Schule. Solange man

² Mahoney, M.J. (1976): Scientist as subject: The psychological imperative. Cambridge, MA (Ballinger).

lediglich so denkt, hofft man immer noch, eines Tages doch noch als Sieger aus dem Wettstreit hervorgehen zu können. Nein, der Prozeß der Vertrauenserosion ist viel weiter gegangen: er zerstört den Glauben an die wissenschaftliche Untersuchbarkeit der Fragen, die uns beschäftigen. Wenn aber die Wissenschaft ihre

Reputation verlöre, dann wäre das Feld frei für den direkten Zugriff der Ökonomie auf die Psychotherapie. Noch wirkt die Wissenschaft als eine Art Puffer zwischen den Forderungen der versicherungswirtschaftlichen Seite und den Verteidigungen der professionellen Psychotherapie.

REENTRY: DIE SELBSTBEOBACHTUNG DER WISSENSCHAFT

Wir bekommen übrigens aus der empirischen Forschung selbst wiederum Hinweise auf die Gründe für diese Situation. Ich erläutere dazu noch einmal in aller Kürze das Konzept des Reentry, das ich im Zusammenhang der Diskussion um das Buch von Edelman und Tononi beschrieben hatte. Reentry – das ist der Wiedereintritt des Unterschiedenen in das Unterschiedene. Man kann das leicht verstehen, wenn man etwa an den Roman von Flaubert „Madame Bovary“ denkt. Er handelt von einer jungen Frau, die durch die Lektüre von Romanen auf Seitenwege geführt und zu Seitensprüngen verführt wird – und weitere Romane antworten darauf. Der Unterschied zwischen wirklichem und erlesenem Leben – erscheint im Roman. Ein anderes Beispiel ist die Oper „Ariadne auf Naxos“ von Richard Strauss. Hier findet man eine Abendgesellschaft vor, die sich um eine Oper namens „Ariadne auf Naxos“ streitet, die Regisseure brauchen für das Reentry dann eine Bühne auf der Bühne. Oder in „Capriccio“ des gleichnamigen Komponisten streiten verschiedene Herrschaften singend darum, ob gelte „prima la parole“ oder „prima la musica“ – kommt beim Menschen zuerst die Sprache oder erst die Musik? Und dieser Unterschied zwischen Musik (Oper) und Sprache erscheint – in der Oper. Das also bezeichnet man als Reentry.

Wenn nun die Wissenschaft sich selbst beobachtet, kommt es zu dem, was man Wissenschaftswissenschaft nennt. Dieses doppelt gemoppelte Wort zeigt das Potential des Reentry an; es ist mächtig steigerungsfähig. Im letzten Heft von PPmP (Psychotherapie, Psychosomatik, medizinische Psychologie 12/2002) des Jahres 2002 erschien eine **Studie von Ada Borkenhagen, Hendrik Berth und Elmar Brähler**. Diese Autoren studieren die Veröffentlichungen der PPmP in den letzten beiden Jahren. Sie benutzen bibliometrische Methoden, um den Impact-Faktor zu messen (er ist leicht gesunken): Das ist ein Faktor, der sich daran orientiert, wie oft die untersuchte Zeitschrift von anderen Zeitschriften zitiert wird. Im „Social-Sciences-Citation-Index“ werden alle hier relevanten Zeitschriften weltweit erfasst. PPmP hat einen Impact-Faktor von 0,346 (1996 lag er bei 0,426). Das bedeutet, das ein in PPmP veröffentlichter Beitrag durchschnittlich 0,346 mal in anderen Zeitschriften zitiert wurde – immerhin!

Aber an der Studie von Borkenhagen et al. ist etwas anderes interessant: Untersucht wurde nämlich auch die Häufigkeit der Selbstzitier-Rate im Verhältnis zu den Fremdzitierreten. Und dabei finden diese Autoren, „dass die Fachvertreter der Psychologie bei einer mittleren Anzahl von Artikeln mit 32 im Mittel die geringste Anzahl von Fremdzitierungen bei höchster Eigenzitierrate (durchschnittlich 4 Eigenzitate) aufweisen“. Das bedeutet, man wird nicht so oft zitiert, zitiert bei eigenen Beiträgen 32 Artikel und im Durchschnitt davon 4 selbst verfasste – ist das eine empirische Definition unseres akademischen Narzißmus? Wir nehmen nicht so sehr viel wahr, was andere meinen, denken, schreiben, sagen; lieber wollen wir selbst wahrgenommen werden.

Aber es kommt noch schonungsloser: In der gleichen Studie wurden die PPmP-Beiträge der letzten beiden Jahre mit dem Dogmatismus-Textauswertungsverfahren (DOTA) analysiert. Dabei werden sogenannte A-Ausdrücke von B-Ausdrücken unterschieden. Erstere haben ein hohes „kognitives Prägnanzniveau“, sie lauten z.B. „endgültig, immer, alle, ausnahmslos, ausschließlich, entweder oder, müssen, nicht dürfen“ usw. Beispiele für die weicheren B-Ausdrücke sind „ab und zu, gelegentlich, unter anderem, auch, dürfen, in der Lage sein“ usw. Wer häufig A-Ausdrücke verwendet, reduziert also viel Komplexität und erreicht eine hohe kognitive Prägnanz, aber um den Preis eines kognitiven Dogmatismus. Und das Ergebnis lautet wenig schmeichelhaft: „Die Vertreter der Psychosomatik/Psychotherapie verwenden die meisten Wörter und ihre Beiträge weisen den höchsten Dogmatismusquotienten auf. Die Fachvertreter der Psychologie/Medizinpsychologie verwenden die geringste Wortzahl. Im Dogmatismusquotienten unterscheiden sich die Vertreter der Fachgebiete Psychiatrie und Psychologie/Medizinpsychologie kaum“ (S. 514).

Das Ergebnis macht nachdenklich. Manchmal scheinen wir prägnanter, also durchaus auch härter zu formulieren als nötig; damit werden dann Klarheiten fingiert, die vielleicht nicht immer gegeben sind.

Aber mit solchen A-Ausdrücken kann man viel zum Aufbau unnötiger Frontlinien tun. Umgekehrt bedeutet dies, man könnte es auch lassen. Kämpfen aber ist das Gegenstück von Lernfähigbleiben. Reentry im Fall der Wissenschaftswissenschaft, also der wissenschaftlich-empirischen Beobachtung von wissenschaftlicher Produktion könnte im günstigen Fall unsere Lernfähigkeit wieder herstellen und das könnte auch ein Beitrag zum Stop der Vertrauenserosion werden.

Gewiß ist also, dass wir in Sachen Psychotherapie bedeutend weniger wissen, als wir andere gelegentlich glauben machen können.