

# Elfenbeintürme im Treibsand oder: Was macht es so schwierig, Erkenntnisse aus der Forschung in der therapeutischen Praxis umzusetzen?

Martin Bohus

Lehrstuhl für Psychosomatische Medizin und Psychotherapie, Universität Heidelberg, Heidelberg, Deutschland;  
Klinik für Psychosomatik und Psychotherapeutische Medizin, Zentralinstitut für Seelische Gesundheit, Mannheim, Deutschland

### Schlüsselwörter

Psychotherapieforschung · Lücke zwischen Wissenschaft und Praxis · Implementierung · Wissenschaftstheorie

### Zusammenfassung

Auch 20 Jahre nach der Drucklegung von Klaus Grawes Pionierwerk ist es noch ein weiter Weg von der psychotherapeutischen Konfession zur Profession. Wie in den meisten Fachgebieten der Medizin besteht auch in der Psychotherapie ein nicht unerhebliches «science-practice gap», also eine deutliche zeitliche Verzögerung in der Umsetzung von Forschungsergebnissen in die klinische Praxis. Dabei weist die Psychotherapie einige Besonderheiten auf: Seitens der Kliniker besteht erhebliche Skepsis bezüglich der Relevanz von Top-down-Forschungsergebnissen für ihren klinischen Alltag. Dies betrifft die Repräsentativität der Stichproben, die Besonderheit der Forschungsbedingungen und die Komplexität der Programme. Im Gegenzug wird die Forderung nach einer praxisbasierten Evidenz laut, also nach Forschungsansätzen, die die alltägliche therapeutische Arbeit als Forschungsgegenstand in den Blick nehmen sollten. Der Artikel beleuchtet die jeweiligen Argumentationsstränge sowie deren immanente Logik und Relevanz. Er plädiert für moderne strukturierte, zirkulär organisierte Kooperationsmodelle zwischen Forschern und Klinikern.

### Keywords

Psychotherapy research · Science-practice gap · Implementation · Theory of Science

### Summary

*Ivory Towers Built on Quicksand, Or: What Makes It so Difficult Transferring Evidence-based Psychotherapy to Routine Care?*

Even 20 years after publication of Klaus Grawe's pioneering book, the long way from psychotherapeutic confession to psychotherapeutic profession is still fraught with obstacles. Like in most fields of medicine, psychotherapy reveals a substantial gap between science and practice. This means that most practicing psychotherapists hardly consider research evidence when making clinical decisions. The aim of this paper is to identify and to discuss the major barriers to the dissemination of evidence-based practice. There are major concerns of the clinicians regarding the limited relevance of research trials to clinical practice. This is often based on beliefs about the nature of patients selected for clinical trials, the discrepancy of the funded research settings and clinical settings, and the complexity of newly developed therapy programs. Accordingly, alternative research approaches focusing on daily activities of clinical practitioners are demanded. This paper aims to discuss the relevance of these arguments, their implicit logic background, and potential solutions. Finally, new research approaches, organized as structured collaborative projects between clinicians and researchers, are recommended.

## Problemstellung

2015 jährt sich nun 21. Mal die Drucklegung des epochalen Werks «Psychotherapie im Wandel: Von der Konfession zur Profession» [1994] von Klaus Grawe und Kollegen. Die jüngeren Kollegen mögen den Aufruhr kaum mehr verstehen, den diese erste große Meta-Analyse zur differenziellen Wirksamkeit der Psychotherapie damals nicht nur in Fachkreisen auslöste. Ich persönlich war zu diesem Zeitpunkt noch Ausbildungskandidat an einem psychoanalytischen Institut [Bohus, 2012] und wurde Zeuge tief wurzelnder Empörung im Kreise der Lehranalytiker: «So etwas machte man einfach nicht (...) man könnte ja intern diskutieren, aber die zarten Identifikationsprozesse von Psychotherapeuten und ihrer Klienten zu gefährden, indem man Wirksamkeitsstudien an die Öffentlichkeit zerrt, überschreitet die Grenzen des guten Geschmacks.» Nun, es dauerte denn auch einige Jahre, bis die «Evidence-based medicine»-Bewegung über die Umwege Großbritanniens [Sackett et al., 2005] und die USA auch die deutschsprachige Psychotherapie infiltrierte – zumindest die akademischen Kreise. Dabei ging die American Psychological Association (APA) schon 2005 so weit, ihre Mitglieder in einer Grundsatzerklärung dazu aufzufordern, sich in der therapeutischen Arbeit an den Prinzipien der «evidence-based practice» zu orientieren [APA, 2005]. Diese Prinzipien verpflichten den Therapeuten in der psychotherapeutischen Behandlung, jeweils den neuesten Stand der wissenschaftlichen Forschung, seine eigene klinische Erfahrung sowie die besonderen Charakteristika und Wertvorstellungen des jeweiligen Patienten zu berücksichtigen [APA, 2005]. Während die meisten praktisch tätigen Therapeuten wenig Schwierigkeiten haben dürften, ihre eigene klinische Erfahrung sowie die Persönlichkeitscharakteristika und Wertvorstellungen ihrer Patienten als maßgeblich zu gewichten, sieht es mit der Berücksichtigung der jeweiligen Forschungsergebnisse wohl deutlich schlechter aus. Wissenschaftliche Untersuchungen zu dieser Frage sind nicht besonders breit angelegt, jedoch weisen alle vorliegenden Studien in dieselbe Richtung: Die Umsetzung von wissenschaftlichen Ergebnissen in der klinischen Praxis geht nur sehr schleppend voran, oder um dies im Fachterminus auszudrücken, es besteht ein erhebliches «science-practice gap» [z.B. Task Force on Promotion and Dissemination of Psychological Procedures, 1995]. Ohne den Anspruch auf Vollständigkeit seien hier nur einige Studien erwähnt, die sich mit dieser Thematik befassen: So fanden etwa Beutler et al. [1995], dass nur 40% aller niedergelassenen akademischen Psychotherapeuten regelmäßig wissenschaftliche Zeitschriften lesen und der Einfluss auf ihre klinische Tätigkeit von ihnen selbst als gering eingestuft wird. In einer ähnlichen deutschen Studie jüngeren Datums [Weber et al. 2013] antworteten 62% einer angeschriebenen Stichprobe niedergelassener Tiefenpsychologen, sich entweder überhaupt nicht, minimal oder nur etwas mit der Forschungsliteratur auszukennen. Shafran et al [2009] fanden in einer groß angelegten Studie, dass nur etwa 40% der niedergelassenen Therapeuten in England wissen, dass kognitiv-behaviorale Therapie die beste Evidenz für die Behandlung von Essstörungen aufweist und nur 7% diese State-of-the-Art-Methodik in der Behandlung von Essstö-

rungen anwenden. Ehlers et al. [2010] zeigten, dass posttraumatische Belastungsstörungen (PTBS) von britischen Psychotherapeuten in aller Regel mittels unterstützender Beratung (supportive counselling) behandelt werden, obgleich dafür keinerlei Wirksamkeitsnachweis vorliegt. Expositionsbasierte Verfahren hingegen werden, trotz ausgezeichneter Evidenzlage und dezidierter Empfehlung in den nationalen Leitlinien, in der klinischen Praxis weitgehend vernachlässigt.

Auch wenn angenommen wird, dass die Umsetzung von evidenzbasierten Therapien an akademischen Ambulanzen etwas besser sein sollte, so weisen die Studien auf ähnliche Probleme hin, wie sie auch im niedergelassenen Bereich beobachtet werden: Berry und Haddock [2008] untersuchten etwa die Häufigkeit von kognitiv-behavioralen Therapieangeboten für schizophrene Patienten an akademischen Ambulanzen in England und fanden diese bei weniger als der Hälfte – obgleich die NICE-Leitlinien [National Collaborating Centre for Mental Health, 2009] dies eindeutig empfehlen. In Deutschland dürfte eine derartige Erhebung wohl noch schlechter ausfallen. Man könnte diese Liste beliebig fortsetzen, ohne auf große Widersprüche zu stoßen. Bevor wir jedoch in psychotherapeutisches Flagellantentum verfallen, sollte tröstend angemerkt werden, dass die Psychotherapie diese Kluft zwischen Wissenschaft und Praxis mit vielen Fächern der Medizin teilt: Gerd Antes, Direktor des deutschen Chochrane-Zentrums und einer der Experten auf diesem Gebiet, schätzt die durchschnittliche Zeit, bis eine neue wissenschaftliche medizinische Methode in der medizinischen Versorgungspraxis ankommt, auf 10–15 Jahre [Antes und Clarke, 2012]. Der Versuch, diesen Prozess durch die Erstellung von Leitlinien zu beschleunigen, scheint nicht besonders erfolgreich. So fanden etwa Currin et al. [2007], dass nur 4% aller britischen Ärzte sich in ihrem Handeln an Leitlinien orientieren.

Diese Daten mögen tröstlich erscheinen und insbesondere darauf hinweisen, dass es neben spezifischen Problemen in der Umsetzung des Forschungswissens in die psychotherapeutische Praxis auch allgemeine Probleme gibt, die diesen Prozess erschweren.

Bevor wir uns der Frage nach den Hindernissen nähern, die einer rascheren Umsetzung von Forschungsergebnissen in die klinische Praxis im Wege stehen, sollten wir uns vergegenwärtigen, dass diese Fragestellung aus dem Blickwinkel eines Forschers gestellt wird. Aus diesem Blickwinkel erscheint es immanent, dass die vermehrte Umsetzung von wissenschaftlich nachweisbar wirksamen Therapiemethoden in der therapeutischen Praxis tatsächlich zu einer Verbesserung von Behandlungsergebnissen führt, dass Patienten also bessere Therapieergebnisse erzielen würden, wenn ihre behandelnden Therapeuten sich an empirisch gestützten Behandlungen (empirically supported treatments; EST) [Chambless und Hollon, 1998] orientieren würden. Die meisten klinisch tätigen Therapeuten sind sich dessen nicht so sicher.

Die Hauptargumente der Kliniker gegen die Übernahme von neuen empirisch geprüften Methoden lassen sich wie folgt zusammenfassen:

- 1) Die Forschungsergebnisse insbesondere von randomisierten kontrollierten Studien (RCTs) sind für die klinische Praxis nicht relevant.

- 2) Die Persönlichkeit des Therapeuten bzw. die therapeutische Beziehung ist der zentrale Wirkfaktor in der Psychotherapie, und hierzu trägt die Forschung nichts Wesentliches bei.
- 3) Es ist nicht notwendig, komplette neue Behandlungsprotokolle zu implementieren, da es genügt, sich einige Bausteine herauszusuchen, die zum Behandlungsstil des Therapeuten passen [z.B. Tschuschke, 2005].

Neben diesen 3 Therapeuten-Argumentationssträngen besteht sicherlich noch eine Reihe von weiteren Hindernissen, wie beispielsweise mangelhafter Informationstransfer und Probleme der Aus- und Weiterbildungsstruktur, die weiter unten kurz aufgegriffen werden sollen. Doch zunächst zu den Problemen in der Akzeptanz der Validität von Forschungsergebnissen für die klinische Praxis. Der Klarheit halber seien zunächst die beiden Ansätze skizziert, die die Psychotherapieforschung derzeit prägen, weil sie unterschiedliche Probleme in der Akzeptanz und Umsetzung mit sich bringen.

#### *Top-down-Forschungsansatz*

Ganz grob gesagt, stehen sich derzeit 2 Ansätze gegenüber: Der Top-down-Forschungsansatz präferiert die Entwicklung von psychotherapeutischen Programmen und deren stufenweise Wirksamkeitsüberprüfung unter «Laborbedingungen». Die Ausgangspunkte sind meist störungsspezifisches Wissen einer Experten-Gruppe, die Kreativität eines Psychotherapieentwicklers, oder – mittlerweile häufiger – experimentell gewonnenes Wissen zu neuropsychologischen Mechanismen, das dann zu mehr oder minder komplexen Behandlungsprogrammen gebündelt, in ein Manual gegossen und an einer spezifischen Population auf Wirksamkeit überprüft wird.

Die Verortung und Umsetzung dieser im besten Fall durch randomisierte kontrollierte Studien als wirksam belegten Therapieprogramme in der klinischen Praxis erfolgt als letzter Schritt einer Entwicklungskette, die folgende Schritte abhandeln sollte: 1. Einzelfallanalysen; 2. Unkontrollierte Prä-Post-Studien; 3. Kontrollierte Studien durch den Entwickler; 4. Replikation durch unabhängige Forschergruppen; 5. Dismantling-Studien bei komplexen, multimedialen Programmen; 6. Untersuchung der Wirksamkeit in der freien Praxis (Feldstudien). In der internationalen Literatur firmiert dieses Paradigma als «evidence-based practice» [z.B. McHugh und Barlow, 2012].

Meistens entfallen die Entwicklungsschritte ab der 4. Stufe aus methodischen, finanziellen oder marktstrategischen Gründen, was zur Folge hat, dass bisweilen komplexe Therapieprogramme, deren wissenschaftliche Evidenz maximal auf einer nicht replizierten Studie basiert, dank charismatischer Entwickler mit großem Aufwand auf den Weiterbildungsmarkt drängen und damit beträchtliche finanzielle und zeitliche Ressourcen der niedergelassenen Praktiker belegen.

#### *Bottom-up-Praxisansatz*

Dem Top-down-Ansatz gegenüber steht als Bottom-up-Praxisansatz (international als «practice-based evidence» bezeichnet; z.B. Barkham et al. [2010]) das Paradigma, die alltäglich klinische Pra-

xis «so wie sie ist» als Forschungsgegenstand und Datengrundlage zu nutzen. Man geht davon aus, dass in der großen Unterschiedlichkeit von Therapeutenvariablen und der Vielzahl von angewandter Methodik eben auch optimale Behandlungsstrategien «verborgen» sind, die es zu extrahieren gilt. Methodisch wird eine möglichst hohe Varianz an Symptomatik, Patienteneigenschaften, therapeutischer Kompetenz und Behandlungssettings mittels großer Datensätze erfasst, um mithilfe neuer statistischer Verfahren, sogenannter Mehrebenen-Analysen, Zusammenhänge zu extrahieren, die anschließend in systematischen Behandlungsstudien überprüft werden können.

Dass diese beiden Forschungsansätze unterschiedliche emotionale Ladung tragen, ist immanent. Geht der Top-down-Ansatz doch davon aus, dass die therapeutische Praxis gegenwärtig unzureichend ist, die therapeutische Expertise also einer Verbesserung durch die Drainage von experimentell erworbenem, d.h. weitgehend universitärem Wissen bedarf. Der Bottom-up-Ansatz hingegen unterstellt dem Feld der klinischen Therapeuten im Gesamten die ausreichende und maximale Kompetenz und versucht, die Verbesserung der therapeutischen Expertise durch Umschichtung von Wissen zu erreichen. Es ist nicht besonders verwunderlich, dass letzterer Ansatz etwas mehr Sympathie aufseiten der Kliniker findet.

Die meisten Psychotherapieforscher haben sich heute einem der beiden Paradigmen verschrieben, sehen diese beiden Ansätze als kompetitiv und weisen den jeweiligen Vertretern des anderen Lagers wahlweise «elitäres, akademisches, weltfremdes, karriereorientiertes» Denken oder «unwissenschaftliche, populistische, fortschrittsfeindliche» Attitude vor, um dies höflich auszudrücken.

Im Folgenden sollen die – sicherlich teilweise berechtigten – Vorurteile sowie die Vor- und Nachteile diese beiden Ansätze hinsichtlich Verbesserung der therapeutischen Praxis diskutiert werden, um abschließend auf neuere bidirektionale Forschungs- und Entwicklungsansätze einzugehen.

## **Top-Down-Forschung und Entwicklung**

Die unter Klinikern weitverbreitete Annahme, dass die Ergebnisse von Psychotherapie-Studien lediglich geringe Aussagekraft für die tägliche Routine aufweisen [z.B. Tschuschke, 2005], basiert unter anderem auf Kritiken an der selektiven Patientenauswahl, der Besonderheit von üppig ausgestatteten Forschungssettings mit selektierten, spezialisierten Forschungstherapeuten und hochfrequenter Supervision bis hin zur Kritik an selektierten Messmethoden, Datenanalyse und Publikationsbias. Nehmen wir uns diese Kritiken im Einzelnen vor:

#### *Probleme der Patientenselektion*

Die gängigen Vorwürfe lauten, dass die meisten Psychotherapie-Studien an hochselektierten Patientenstichproben durchgeführt werden, die weitgehend stark motiviert, leicht führbar und nicht besonders schwer erkrankt sind. Zudem – so die Kritik –

werden die Studien primär an Patienten mit monosymptomatischen Störungsbildern durchgeführt. Diese Patienten, so der Vorwurf, sind für die typische Klientel in der Praxis nicht oder zumindest nicht ausreichend repräsentativ. Man könnte nun etwa mit Shafran et al. [2009] argumentieren, dass der Unterschied zwischen Forschungspopulationen und Patienten in der Routinepraxis deutlich geringer ist als allgemein angenommen, ja dass in aller Regel Patienten unter allgemeinen Versorgungsbedingungen häufig unter einer «Subthreshold»-Symptomatik leiden, d.h. eigentlich die strengen Kriterien einer operationalisiert nachweisbaren Achse-I-Störung oft gar nicht erfüllen [Stirman et al., 2005]. Man könnte auch argumentieren, dass in neueren Studien mittlerweile Komorbidität dezidiert eingeschlossen wird und sich oftmals vermutete negative Einflüsse von etwa schweren Persönlichkeitsstörungen auf Behandlungsverläufe unter Studienbedingungen nicht verifizieren lassen [z.B. Bohus et al., 2013].

Durch die Brille eines erfahrenen Psychotherapieforschers wäre jedoch anzumerken, dass gerade die Auswahl von schwer erkrankten Patienten dem Forscher einen Vorteil verschafft. Und zwar dahingehend, dass ein starker Behandlungseffekt zu erwarten ist, was wiederum die notwendige Fallzahl an zu untersuchenden Patienten reduziert und damit Kosten spart. Was die Komorbidität betrifft, so drängt der Methodenzwang den Forscher natürlich dazu, möglichst homogene Stichproben zu behandeln, da sich dadurch die zu erwartende Varianz reduziert, was wiederum Fallzahl und Kosten senkt. Der klassische Rat an einen Novizen in der Psychotherapieforschung lautet daher: Gestalte zu Beginn der Forschungsentwicklung das Design deiner Studie so, dass potenzielle Effekte auch deutlich sichtbar werden – und das bei beschränkten Fallzahlen und Finanzen. Sorge also dafür, dass möglichst schwer kranke Patienten mit einem möglichst einförmigen Störungsbild eingeschlossen werden. «Ideal» aus Sicht des ergebnisorientierten Forschers wäre etwa eine Kohorte weiblicher Patienten mit schwerwiegender Zwangserkrankung im Alter zwischen 25 und 40 Jahren, ohne weitere Komorbidität, und einem Mindestkrankheitsverlauf von 10 Jahren. Unter diesen Bedingungen ist eine hohe interne Validität, also Aussagekraft, zu erwarten. Die Wahrscheinlichkeit, dass sogenannte Fehler der ersten Art ( $\alpha$ -Fehler) auftreten, also eigentlich wirkungsvolle Therapie aufgrund nicht signifikanter Ergebnisse verworfen wird, ist dadurch geringer. Ungünstig wäre in diesem frühen Entwicklungsschritt etwa die Wahl einer gemischtgeschlechtlichen Stichprobe mit mittlerer Zwangssymptomatik in einer Altersspanne von 18 bis 65 Jahren sowie komorbider Depression, Alkoholabhängigkeit und möglicher psychotischer Symptomatik. Die Wahrscheinlichkeit ist hoch, dass bei knappen Ressourcen eine zu geringe Fallzahl gewählt wird, um bei der zu erwartenden hohen Varianz, also unterschiedlichem Ansprechen auf die Therapie, signifikante Ergebnisse zu verfehlen, obgleich die Therapie etwa für die reine Zwangssymptomatik durchaus wirksam wäre.

Der Vorwurf der Praktiker, dass Patienten in kontrollierten Studien häufig nicht repräsentativ seien, ist also durchaus nachvollziehbar, wenn man die immanente Logik der Forschungsentwicklung berücksichtigt. Wobei vonseiten der Forscher offengelegt werden sollte, dass es sich bei den ersten (auch kontrollierten) Studien

eben um eine erste Prüfung des Konzepts handelt. *Lege artis* sollten nach dieser ersten Studie weitere Studien folgen, die eben auf höhere externe Validität zielen und gemischte, repräsentative Stichproben einschließen. Der Fehler besteht also nicht in der Logik der Forschungsentwicklung, sondern in der unzureichenden Kommunikation. In aller Regel wird die erste «erfolgreiche» kontrollierte Studie bereits als Erfolg gefeiert und auf dem Weiterbildungsmarkt platziert, ohne weitere Studien abzuwarten.

Die Gründe dafür liegen auf der Hand: Die Entwicklung eines psychotherapeutischen Programms ist äußerst aufwendig und in aller Regel nicht oder nur sehr unzureichend finanziert. Die Entwickler investieren also Freizeit, Energie und Herzblut, getrieben von dem oft missionarischen Wunsch, ihre Entwicklung möglichst rasch einer großen Öffentlichkeit zugänglich zu machen. Selbstverständlich steigern kontrollierte Studien den Marktwert einer neuen therapeutischen Methode. Allzu oft dient daher die Wirksamkeitsforschung als notwendiges Übel, das benutzt wird, um das Produkt besser auf dem Markt zu platzieren. Und nichts ist so demotivierend wie ein mangelnder Wirksamkeitsnachweis eines über Jahre entwickelten Programms. Ist es vor diesem Hintergrund verwunderlich, dass die meisten Forscher darauf aus sind, möglichst große Effektstärken zu erzielen, indem sie die Stichproben möglichst homogen halten? Der Entschluss, nach einer ersten kontrollierten Studie mit hoher internen Validität weitere Studien aufzulegen, mit breiteren, komorbiden und damit Praxis-äquivalenten Patienten, birgt für den Forscher durchaus Gefahren und Probleme: Zum einen erfordern derartige Designs höhere Fallzahlen. Diese werden sehr selten finanziert und erfordern zudem lange Laufzeiten, multizentrische Organisation, etc. – alles sehr aufwendig. Zum anderen ist das Risiko hoch, dass die Therapieeffekte gering ausfallen und eventuell das geforderte Signifikanzniveau nicht erreichen. Da das 5%-Signifikanzniveau aus unerfindlichen Gründen nach wie vor das Maß aller Dinge darstellt, wirkt eine Irrtumswahrscheinlichkeit von 6% oft wie ein Genickschlag. Es ist ein mühsamer Kampf, die Effektstärken als Wirksamkeitsmaß zu etablieren, wobei sich die meisten Forscher einig sind, dass dieses Maß als Wirksamkeitsnachweis erhebliche Vorteile aufweist. Nicht-signifikante Studienergebnisse sind zum einen nur schwer zu publizieren, zum anderen werden sie in der Öffentlichkeit oft undifferenziert als «nicht wirksam» wahrgenommen und schaden daher der weiteren Vermarktung des Therapieprodukts. Der interne Antrieb, derartige komplexere Designs anzulegen, ist also nicht besonders hoch. Der Druck müsste von außen kommen – von den Praktikern, von den Leitlinienentwicklern, von den Gutachtern der wissenschaftlichen Journale, von den politischen Gremien und nicht zuletzt von den Kostenträgern. Es liegen durchaus Studien mit unselektierten Stichproben vor, die vielversprechende Ergebnisse aufweisen. So fanden etwa Halweg et al. [2001] in einer Feldstudie an 416 unselektierten Patienten mit der Primärdiagnose «Panikstörung und Agoraphobie», die von 52 ambulant tätigen Therapeuten behandelt wurden, signifikante Symptomverbesserungen und Effektstärken, die sich von den in Meta-Analysen errechneten Benchmark-Ergebnissen nicht unterschieden. Fairburn et al. [2013] fanden ähnliche Ergebnisse bei Essstörungen, und auch die Behandlung der PTBS

bei nicht-selektierten Zeugen des Omagh-Bombenattentats zeigte Prä-Post-Effektstärken von  $d = 2.4$ , was in etwa den Werten bei kontrollierten Studien entspricht [Gillespie et al., 2002]. Dies ist eine willkürliche Auswahl an Studien, und man sollte einen möglichen Publikationsbias nicht unterschätzen. Dennoch liegen damit erste Hinweise vor, dass sich Wirksamkeitsnachweise auch in breiter angelegten Stichproben mit ähnlicher Effektstärke zeigen lassen wie bei Studien mit hochselektierter Klientel.

#### *Fazit*

Die Glaubwürdigkeit der klinischen Forschung hängt in hohem Maße von der Repräsentativität der untersuchten Populationen ab. Nach einem ersten Wirksamkeitsnachweis bei einer hochselektierten Stichprobe sollte immanente Aufgabe des Entwicklungs- und Forschungsprozesses sein, Replikationen mit unselektierten Stichproben durchzuführen und diese auch als solche zu deklarieren. Die (Vor-)Urteile der Kliniker gegenüber der Repräsentativität der untersuchten Patienten speisen sich bisweilen auch aus der Unkenntnis des stufenweisen Vorgehens im psychotherapeutischen Forschungs- und Entwicklungsprozess. Es wäre die Aufgabe der Akademiker, dies transparenter zu machen.

#### *Besonderheiten des Forschungssettings*

Die Bedingungen, unter denen kontrollierte Studien ablaufen, sind derart artifiziell, dass sie erheblichen Einfluss auf die Therapieergebnisse nehmen und die Übertragung dieser Ergebnisse auf die Praxis erschweren bis verunmöglichen. Die Therapeuten haben mehr zeitliche Ressourcen, sind in den jeweils untersuchten Methoden ausgezeichnet ausgebildet, dabei durch die emotionale Nähe zum Therapieentwickler hoch motiviert und durch niederefrequente Verlaufs- und Erfolgsmessungen in der Lage, ihre Therapie an die jeweiligen Prozesse anzupassen und zu steuern – so die gängigen Vorwürfe der Kliniker. Schließlich wird immer wieder der Aspekt der engmaschigen Supervision bzw. der Adhärenz-Messung als einflussnehmende Variable ins Feld geführt [z.B. Tschuschke, 2005].

Gegen diese Argumente ist nicht viel einzuwenden. In den allermeisten Fällen treffen sie schlicht zu. Forschungstheoretisch handelt es sich um weitgehend unvermeidliche Experimentalvariablen, die sicherlich einen nicht unerheblichen Einfluss auf die Therapieergebnisse nehmen: Hochmotivierte Therapeuten erzielen bessere Therapieergebnisse, Rückkoppelungen von Therapieverläufen haben einen positiven Einfluss auf die jeweiligen Ergebnisse [z.B. Lambert und Shimokawa, 2011] und die Bedeutung von Supervision auf dieselben wird wohl niemand bestreiten. Diese Argumente und Kritikpunkte sollten jedoch vornehmlich die Skepsis der Kliniker gegenüber reinen Prä-Post-Studien erheblich nähren. Diese Argumente gegen die Bedeutung von RCTs ins Feld zu führen [Tschuschke, 2005], ist etwas irritierend. Gerade in RCTs sollte gewährleistet sein, dass diese Einflussvariablen für beide erforschten Therapiebedingungen in gleichem Maße zutreffen. *Lege artis* durchgeführte RCTs sorgen dafür, dass die Therapeuten der Be-

handlungsgruppen die gleiche Anzahl an Supervisionseinheiten erhalten und dass auch die Rückmeldungen und die Ausbildung parallelisiert sind. Zwischengruppen-Effektstärken sollten den Einfluss der Experimentalbedingungen also weitgehend eliminieren – vorausgesetzt, die Personen, die an den Messprozessen beteiligt sind, sind bezüglich der Zugehörigkeit zu den jeweiligen Behandlungsarmen tatsächlich «blind». Ein wichtiger und kritischer Aspekt der Psychotherapieforschung sollte jedoch an dieser Stelle kurz beleuchtet werden: Wie oben ausgeführt, haben Psychotherapieentwickler und ihre Verbündeten oft starke emotionale und nicht zu vernachlässigende finanzielle Interessen am Erfolg der jeweiligen Studie – hängen davon doch Karriere, Reputation und nicht unerhebliche Einnahmen durch Workshops und ganze Weiterbildungscurricula ab. Ein erhebliches Erwartungsbias wird also maßgeblichen Einfluss nehmen auf die Therapieergebnisse – gerade bei Studien, die vom Therapieentwickler oder engen Mitarbeitern durchgeführt werden. Die Verblindung der Rater ist nur eine Minimallösung. Um hohe Evidenzgrade zu erreichen, sollten Replikationen durch unabhängige Studiengruppen gefordert werden.

Man sollte diese Argumente jedoch noch in einer weiteren Hinsicht ernst nehmen: Da wir wissen, dass kurze Feedback-Schleifen zu Behandlungsprozessen und Therapieverläufen die Behandlungsergebnisse verbessern, sollte etwa darauf hingewirkt werden, diese Möglichkeiten auch in der Praxis zu etablieren. Zudem sollte gegenüber Kostenträgern immer wieder darauf hingewiesen werden, dass im strengen Sinne keinerlei Wirksamkeitsnachweis für Psychotherapie ohne Supervision vorliegt. Nahezu alle RCTs werden entweder supervidiert oder auf Adhärenz überprüft. Und streng genommen sollte die Finanzierung der Supervision – als integraler Bestandteil der Therapie – von den Kassen übernommen werden.

#### *Fazit*

Zusammenfassend trifft das Argument, dass die Besonderheiten des Forschungssettings auf die Behandlungsergebnisse Einfluss nehmen sicherlich zu – solange es sich um unkontrollierte Prä-Post-Studien handelt. Deshalb ist es durchaus angemessen, die oft hohen Effektstärken von unkontrollierten Studien kritisch zu hinterfragen. Studiendesigns von RCTs sollten diese Variablen jedoch parallelisieren. Um den Motivationsbias von Therapieentwicklern zu neutralisieren, sollten vom Therapieentwickler durchgeführte RCTs unbedingt von unabhängigen Forschergruppen repliziert werden, bevor ein entsprechend hoher Evidenzgrad angesetzt wird.

#### *Probleme der Messmethodik, Datenanalyse und Publikation*

Die gängigen Vorwürfe der Kliniker lauten, dass die Auswahl der zumeist störungsspezifischen Messinstrumente die Therapieergebnisse nur unzureichend und selektiert abbildet und dass die Analyse der Daten und insbesondere die Art und Weise der Publikation für die meisten Kliniker schwer verständlich, undurchsichtig und daher kaum rezipierbar erscheint.

Auch diese Argumente sind meiner Ansicht nach nicht von der Hand zu weisen: Im Bestreben, die Psychotherapieforschung an den

State-of-the-Art-Vorgaben der Psychopharmaka-Wirksamkeitsforschung auszurichten, einigte man sich mehr oder weniger implizit, zur Überprüfung der jeweiligen Hypothesen 1 bis maximal 2 primäre Outcome-Maße festzulegen. Anhand dieser wird bei geschätzten Effektstärken die notwendige Power, d.h. die Fallzahlen kalkuliert. Zudem ist man mittlerweile gehalten, diese primären Outcome-Maße in öffentlich zugänglichen Datenbanken (etwa [www.clinicaltrials.gov](http://www.clinicaltrials.gov)) vor Studienbeginn zu hinterlegen. Dieses Vorgehen birgt große Vorteile, da verhindert wird, dass nachträglich aus einer Vielzahl von erhobenen Messungen einige wenige positive Ergebnisse ausgewählt und publiziert werden. Es ergeben sich aber auch Nachteile: Die Entwicklung von störungsspezifischen Messinstrumenten, oder gar von experimentellen Methoden zur transdiagnostischen Erfassung spezifischer Pathomechanismen, führt zu einer hohen Übereinstimmung zwischen Diagnose, Intervention und Messinstrument. Das heißt, das jeweilige Instrument ist genau auf die applizierte Intervention zugeschnitten und diese wiederum ist unter optimalen Bedingungen auf den Pathomechanismus abgestimmt. Die Wahrscheinlichkeit eines starken Wirksamkeitsnachweises wird dadurch bedeutend erhöht. Definiert man etwa die Reduktion von Selbstverletzungen bei Borderline-Patienten als primäres Outcome-Maß, so wird eine Intervention, die ihr primäres Augenmerk und ihre Interventionen auf diese Symptomatik legt, sicherlich bessere Ergebnisse aufweisen als eine Intervention, die sich primär mit dem Selbstwertproblem dieser Patientenpopulation befasst und der Selbstverletzung keine Bedeutung zumisst. Die entscheidende Frage ist, ob die Selbstverletzungen tatsächlich eine zentrale Problematik dieser Population darstellen und ob deren Behandlung weitreichende Auswirkungen auf andere Symptome hat. Dies kann nur beantwortet werden, wenn man eine hohe Korrelation zwischen Reduktion von Selbstverletzung und globaleren Therapiemaßen wie etwa Steigerung der Lebensqualität, Verbesserung der sozialen Integration oder sehr essenziellen Maßen wie etwa Reduktion von Suizidraten nachweisen kann. Unterbleibt dieser Nachweis, so sollte man zumindest die Möglichkeit nicht ausschließen, dass man entweder Theorie-immanente Konstrukte oder Surrogat-Marker behandelt, die keine Auswirkung auf die Lebensdauer oder Lebensqualität der Betroffenen haben. Die Psychotherapieforschung verfiel derselben Problematik wie etwa die Pharmaka-Forschung bei der Behandlung von Schizophrenie, die jahrzehntelang die Reduktion von akustischen Halluzinationen als primäres Outcome-Maß verfolgte und dabei erhebliche Einschränkungen der Motorik und Lebensqualität ihrer Patienten in Kauf nahm. Erst im letzten Jahrzehnt hat sich Lebensqualität (quality of life) als weitgehend verbindliches Outcome-Maß durchgesetzt. Die Psychotherapieforschung täte gut daran, diesen Prozess abzukürzen und einige verbindliche, störungsübergreifende allgemeine Maße zu etablieren, die auch zu Schulen übergreifenden vergleichenden Analysen herangezogen werden könnten.

Zum Vorwurf der Verständlichkeit der Datenauswertung und Publikation sei angemerkt, dass jedwede Datenanalyse naturgemäß natürlich immer eine erhebliche Reduktion von Komplexität bedeutet. Mit Reduktion von Komplexität erfolgt Informationsverlust. Hinzu kommt, dass es sich bei den statistischen Verfahren in aller

Regel um Mittelwertvergleiche mit oft enormen Streuungen handelt, deren Aussagekraft für den einzelnen Fall notgedrungen sehr gering ist. Der klinische Praktiker kann etwa mit den Ergebnissen einer groß angelegten Multicenter-Depressionsstudie zur differenziellen Wirksamkeit von kognitiver Therapie versus kognitiver Therapie plus Achtsamkeit, die schließlich in einer einzigen Zahl mündet, («Der Unterschied zwischen den beiden Ansätzen liegt bei einer Effektstärke von 0,3») nicht viele essenzielle Informationen für seine Praxis gewinnen. Rentiert es sich nun tatsächlich, Achtsamkeit zu integrieren? Und wenn ja, für wen? Und wie viele seiner Patienten würden davon profitieren? Und ließe sich diese Kombination an allen seinen Patienten anwenden? Und was müsste er selbst investieren? Auf keine dieser Fragen bekommt der Praktiker in dieser Publikation eine Antwort – ganz abgesehen davon, dass etwa 85% der niedergelassenen Kollegen nicht wissen, was man unter einer Effektstärke versteht (eigene Schätzung aufgrund sporadisch erhobener, etwas bösartiger Umfragen in zahlreichen Workshops). Es liegt im Aufgabenbereich der Forschung, Publikationsansätze zu generieren, die auf die anstehenden Fragen der Kliniker eingehen. Die digitalen Medien ermöglichen ja mittlerweile, neben der Hauptpublikation eine Vielzahl von zusätzlichen Informationen zu hinterlegen. Dies reicht von einer verständlichen Erläuterung der Daten im Hinblick auf die klinische Praxis bis hin zu Videos von Therapiemaßnahmen, sodass für den Kliniker ersichtlich wird, um welche Interventionen es sich eigentlich handelt.

Schließlich sollte auch kurz auf die wachsende Skepsis der Kliniker bezüglich Publikationsbias eingegangen werden, also der Selektion von publizierten Studien. Eine jüngste Freiburger Studie untersuchte die Publikationsrate von klinischen Studien, die entweder von Ethikkommissionen genehmigt oder als Clinical Trials registriert waren [Schmucker et al., 2014]: Weniger als die Hälfte aller angemeldeten und registrierten Studien wurde veröffentlicht. Die Publikationsrate liegt also unter 50%. Dabei korreliert die Wahrscheinlichkeit einer Publikation mit hypothesenkonformen Ergebnissen. Die Autoren dieser Studie kommen zu dem Schluss, dass die Nicht-Publikation einem eindeutigen Bias unterliegt, oder auf gut Deutsch gesagt, in der gesamten Medizin werden erwartete und erwünschte Ergebnisse deutlich häufiger publiziert als negative Ergebnisse. Die Frage, ob dieses Publikationsbias auch für psychotherapeutische Studien zutrifft, ist derzeit noch umstritten [z.B. Cuijpers et al., 2010; Niemeyer et al., 2013]. Es wäre jedoch äußerst verwunderlich, wenn sich unser Fachgebiet hier vom Rest des medizinischen Feldes abheben würde. Eine Klärung dieser Frage wäre der Reputation und Glaubwürdigkeit unseres Faches sicherlich dienlich.

#### Fazit

Sicherlich täte die Psychotherapieforschung gut daran, neben störungsspezifischen Instrumenten breiter angelegte Messmethoden von Parametern wie Lebensqualität, Sinnerfülltheit, soziale Integration, etc. als allgemein verbindlich zu etablieren. Der (notwendige) hohe Abstraktionsgrad von Forschungsergebnissen dient zwar der Vergleichbarkeit von Daten, ist aber für die Information von Praktikern schlicht untauglich. Hier eröffnen digitale Publikationsorgane neue und didaktisch deutlich bessere Möglichkeiten.

Und schließlich sollte sich auch die Psychotherapieforschung der Problematik der selektierten Publikationen stellen, um die eigene Glaubwürdigkeit zu erhöhen.

Nachdem die kritischen Annahmen niedergelassener Psychotherapeuten bezüglich Patientenselektion, Experimentalvariablen und Publikationspraxis skizziert und diskutiert wurden, wenden wir uns kurz dem zweiten großen Themenblock der Kritiker an evidenzbasierter, top-down-entwickelter Psychotherapie zu: dem Mythos von der heilenden Kraft der therapeutischen Beziehung.

### Zum Mythos der therapeutischen Beziehung

Auch wenn wir über keine validen Umfrageergebnisse verfügen, so dürften doch die meisten Psychotherapeuten (mit Schwerpunkt in den tiefenpsychologischen Schulen) die Meinung vertreten, dass die therapeutische Beziehung die wichtigste Wirkvariable der Psychotherapie darstellt und in vielen Fällen eine ausreichende therapeutisch heilende Kraft darstellt [z.B. Tschuschke, 2005]. Angesichts dieser Bedeutung der therapeutischen Beziehung muss zwangsweise und folgerichtig die Bedeutung der therapeutischen Methodik in den Hintergrund treten. Konkret gesagt, schätzen die meisten Therapeuten intuitiv den Wert der Beziehung als hochrangiger ein als etwa den Wert von exakter Diagnostik, von Problem- und Fokuswahl, der gezielten interventionellen Veränderung von emotionalen und kognitiven Prozessen, von Kontingenz-Management, von Strukturierung, von externen Einflüssen durch Problemlösen, etc. Diese optimistische Grundannahme von der zentralen Wirkvariable therapeutischer Beziehung, getragen von weiten Teilen unseres Fachgebietes, hat erhebliche Auswirkungen, entbindet sie doch von der moralischen Verpflichtung, sich einer mühsamen Perzeption und oft zeitraubenden Aneignung differenzierter neuer methodischer Forschungsergebnisse zu unterziehen: Solange ich mich als Therapeut ausreichend um mein persönliches therapeutisches Wohlergehen kümmere und damit den wichtigsten Faktor der therapeutischen Beziehung, nämlich mich selbst, arbeite ich auch an meiner Professionalität. Wenn es diesen Beruf nicht gäbe, müsste man ihn erfinden.

Dies ist nicht der Ort für eine fundierte Auseinandersetzung mit dem Forschungsstand zur therapeutischen Beziehung, daher nur kurz und pragmatisch zusammengefasst: Die populäre Annahme, die therapeutische Beziehung sei der Hauptwirkfaktor der Psychotherapie, ist empirisch nicht nur nicht belegt, sondern greift sicherlich zu kurz.

Die moderne Forschung [z.B. Horvath et al., 2002] unterscheidet zwischen 3 Domänen der therapeutischen Beziehung: Zum einen das sogenannte Arbeitsbündnis (working alliance), also die grundlegende Expertenbeziehung für das therapeutische Arbeiten im Allgemeinen. Die zweite Domäne beschreibt die therapeutische Beziehung als heilendes Agens, etwa im Sinne der Erfahrungskorrektur bei interpersoneller Traumatisierung oder im tiefenpsychologischen Jargon im Sinne einer «Nachreifung» von Entwicklungsdefiziten. Zum Dritten wird die therapeutische Beziehung auch als interventionelles Werkzeug eingesetzt, um neue

emotionale und zwischenmenschliche Erfahrungsprozesse zu generieren. Die Literatur bietet einzig zur Bedeutung des Arbeitsbündnisses für das Gelingen psychotherapeutisch initiiert Veränderungsprozesse gute operationalisierte Messmethoden und profunde Untersuchungen (z.B. Flückinger et al. [2012]). Die mittlerweile 5 publizierten Meta-Analysen finden übereinstimmend einen schwachen Zusammenhang zwischen Arbeitsbündnis und Therapie-Outcome. Die Qualität des Arbeitsbündnisses klärt dabei etwa 7% der Varianz zum Therapie-Outcome auf [Flückinger et al., 2012]. Wobei noch einmal klargestellt werden sollte, dass der Begriff «Arbeitsbündnis» Parameter umfasst wie gemeinsame explizite oder implizite Zieldefinition, Verlässlichkeit und Expertise des Therapeuten, damit also die Grundlagen einer klassischen Expertenbeziehung. Somit unterscheidet sich die Bedeutung des Arbeitsbündnisses in der psychotherapeutischen Beziehung in keiner Weise von der Arzt-Patienten-Beziehung in der somatischen Medizin und deren Bedeutung für den Heilungsprozess [z.B. Kiessling und Langewitz, 2013]. Auch in der somatischen Medizin trägt die Qualität des Arbeitsbündnisses einen nicht unerheblichen Anteil zum Gelingen des Heilungsprozesses bei. Nur wenige überlebende somatische Mediziner würden jedoch unter Berufung auf die Bedeutung dieses Arbeitsbündnisses auf die Anwendung weiterer medizinischer Maßnahmen oder die fachliche Fortbildung verzichten. Profunde Untersuchungen zur Wirksamkeit der therapeutischen Beziehung als heilendes Agens bzw. zur Wirksamkeit der Beziehungsirritation als interventionelle Methodik liegen meines Wissens nicht vor. Daraus sollte man natürlich nicht schließen, dass diese Konzepte nicht tragfähig sind – sie entbehren lediglich der Empirie. Zudem sollten sich die Kritiker der evidenzbasierten störungsspezifischen Methoden die ersten Studienergebnisse vergegenwärtigen, die nachweisen, dass die therapeutische Beziehung bei störungsspezifischen Methoden, wie etwa der Dialektisch-Behavioralen Therapie (DBT), die ja primär auf die Vermittlung von Fertigkeiten setzen, von den Patienten deutlich besser und nachhaltiger eingeschätzt wird als bei parallel untersuchten psychodynamischen Verfahren, die primär auf die Beziehung als Wirkfaktor setzen (z.B. Bedics et al. [2012]). Aus Sicht der Beziehungsforschung machen diese Befunde durchaus Sinn, da gerade die Methodenkompetenz des Therapeuten einen wesentlichen Beitrag zum Vertrauen in den Therapeuten und damit zur Verbesserung des Arbeitsbündnisses leistet.

Eine weitere deutliche Irritation erfährt die apodiktische Bedeutungszuweisung der therapeutischen Beziehung durch den schier unaufhaltsamen Siegeszug computerbasierter Psychotherapieprogramme. Derzeit liegen über 130 randomisierte Studien zur Wirksamkeit von computerbasierten psychologischen Interventionen vor. Die minimale Überlegenheit von individuell applizierter Therapie scheint primär in der besseren Compliance begründet zu liegen, was wiederum auf die unbestrittene Bedeutung des Arbeitsbündnisses als starken motivationalen Faktor hinweist. Ansonsten scheinen die Patienten bei methodisch hochwertigen computerisierten Programmen auf die Person des Psychotherapeuten ohne erhebliche Qualitätseinbußen verzichten zu können. Nun, auch diese neuen Entwicklungen und Forschungsergebnisse werden an

der dysfunktionalen Überschätzung und Bedeutungszuweisung der therapeutischen Beziehung als zentraler Wirkfaktor wenig ändern. Zu stark ist dieses Missverständnis im Selbstverständnis der Psychotherapeuten verankert. Es liegt nicht in meiner Kompetenz, wäre aber sicherlich interessant, die individuellen psychodynamischen Wurzeln dieser kollektiven Fehlannahme zu untersuchen. Handelt es sich um Derivate narzisstischer Größenideen? Um die Manifestation einer eigenen unbestimmten Sehnsucht nach einem allmächtigen Heiler? Um die Folgen unreflektierter projektiver Identifizierungsprozesse unserer Patienten, die sich in den Selbstkonzepten der Therapeuten niederschlagen? Vielleicht handelt es sich aber auch nur um Bequemlichkeit.

#### *Fazit*

Die Bedeutung der therapeutischen Beziehung für Compliance und Therapieergebnisse ist lediglich für die Dimension des Arbeitsbündnisses im Sinne eines Expertenverhältnisses gut belegt und klärt etwa 7% der Varianz am Therapieerfolg auf. Die therapeutische Beziehung ist (damit im besten Fall) eine notwendige, aber keinesfalls eine hinreichende Bedingung für Therapieerfolg. Die Beziehung sollte als Grundlage für die Anwendung therapeutischer interventioneller Methodik gesehen werden. Die Bedeutung der therapeutischen Beziehung als Begründung für die Irrelevanz empirischer Forschung heranzuziehen, erscheint vor diesem Hintergrund wenig sinnvoll.

Kommen wir zum dritten Argumentationsstrang der Kritiker: Ist es wirklich notwendig, komplette neue Behandlungsprotokolle zu implementieren oder genügt es, sich einige Bausteine herauszusuchen, die zum Behandlungsstil des jeweiligen Therapeuten passen?

### **Modulare Ansätze**

Beschäftigt man sich mit Mechanismen von Implementierungsprozessen jedweder Art, so stößt man unweigerlich auf eine zentrale Aussage: Die Implementierung von neuen Verfahren ist in hohem Maße abhängig von den Schnittstellen zur bisherigen Erfahrung der Betroffenen. Um es klarer auszudrücken: Neue Verfahren oder Methoden werden dann, und nur dann in Systeme integriert, wenn sie einen gewissen Bekanntheitsgrad aufweisen und sich in gewissem Maße mit den bisherigen Erfahrungen und Prozessen decken. Andernfalls ist die erwartete Irritation größer als der zu erwartende Vorteil, und die neuen Verfahren werden abgelehnt [z.B. Fixsen et al., 2005]. Es ist daher nicht verwunderlich, dass Psychotherapeuten dazu tendieren, sich aus komplexen, oft geschlossen präsentierten Therapieprogrammen diejenigen Komponenten herauszupicken, die ihnen als passend erscheinen und diese in ihre eklektizistische Methodensammlung zu integrieren.

Von den Therapieentwicklern wird dieses Vorgehen häufig kritisiert (z.B. Chambless und Hollon [1998]; Shafran et al. [2009]). Führt man Methodenreinheit als Argument ins Feld, so erscheint diese Kritik zunächst auch berechtigt. Die meisten störungsspezifischen Programme sind ja modular konzeptualisiert, d.h. sie setzen

sich aus einer Vielzahl von therapeutischen Interventionen zusammen [z.B. Bohus et al., 2012]. Der Wirksamkeitsnachweis erfolgt jedoch zunächst für das gesamte Paket, und solange dieses Paket nicht dekonstruiert und damit die Wirksamkeit der einzelnen Module im Rahmen von aufwendigen Dismantling-Studien nicht nachgewiesen ist, entbehrt das «Methoden-Picking» jedweder empirischen Evidenz.

Der Praktiker steht demgemäß vor der Entscheidung, entweder seinen Bedürfnissen entsprechend einzelne, ihm sinnvoll erscheinende Komponenten eines komplexen Programms zu integrieren und sich damit außerhalb des Feldes der empirischen Evidenz zu bewegen oder sich einer nicht unerheblichen Irritation und Verunsicherung auszusetzen und diese komplexen Programme mit viel zeitlichem und finanziellem Aufwand in seine Praxis zu integrieren – nimmt es Wunder, dass er sich in aller Regel für die erste Möglichkeit entscheidet?

Es liegt also wiederum im Aufgabenbereich der Therapieentwickler, und sollte sich mit deren genuinen wissenschaftlichen Interessen decken, ihre störungsspezifischen Kompositionen zu fragmentieren, modular zu beforschen und entsprechend dem Weiterbildungsmarkt zur Verfügung zu stellen. Auch hier zeigt sich das Echo von Klaus Grawes Weitblick, der ja schon Schulen übergreifende Wirksamkeitsforschung initiiert und damit die Grundlagen einer modular konzipierten, «allgemeinen Psychotherapie» geschaffen hatte [Grawe, 2000].

Ein gutes Beispiel wäre heute etwa die Überprüfung der solitären Wirksamkeit von Skills-Training im Rahmen des multimodularen Therapiekonzeptes der DBT. Entgegen der Hypothesen der Entwickler zeigen Studienergebnisse, dass Modul-Skills-Training auch ohne Einzeltherapie hohe Wirksamkeit auf die Verbesserung der Emotionsregulation und die soziale Integration von Borderline-Patienten zeigt.

#### *Fazit*

Die aus Sicht der Forschung berechnete Forderung an die Kliniker, weitgehend komplette Therapieprogramme in ihre Praxis zu integrieren, erscheint wenig realistisch. Zielführend wäre eine modulare Ausweisung von komplexen Therapieprogrammen und eine Differenzierung zwischen bereits etablierten und neu entwickelten Therapiekomponenten sowie die Dekonstruktion der komplexen Programme und die gezielte Durchführung von Dismantling-Studien. Dies würde dem genuinen Bedürfnis der Kliniker gerecht werden, ihr therapeutisches Repertoire um entsprechende Bausteine zu erweitern und damit die Wahrscheinlichkeit einer raschen Verbreitung dieser Forschungsergebnisse deutlich erhöhen.

### **Zur Zukunft der Top-down-Forschung**

Angesichts der beschriebenen, vielfältigen Problematik sollte keinesfalls unerwähnt bleiben, dass der Top-down-Ansatz, also die Entwicklung von Therapieansätzen unter Laborbedingungen, auch unwiderlegbare Vorteile hat und in Zukunft noch stärkere Bedeutung erlangen wird. Dies betrifft insbesondere die Umsetzung von



neueren Erkenntnissen aus der experimentellen neurobiologischen und sozialpsychologischen Forschung. Es braucht nicht erwähnt zu werden, dass durch die Entwicklung der bildgebenden Verfahren ein Quantensprung in der Aufklärung von psychotherapeutisch hochrelevanten Prozessen gelungen ist, deren Ausmaß sich gerade erst erahnen lässt. Vage Konzepte wie etwa Emotionsregulation, Impulskontrolle, Suchtverhalten, Bedürfnisbefriedigung, sexuelle Appetenz, Dissoziation, emotionales Lernen, etc. werden durch die neuen Methoden allmählich greifbar, messbar und unter experimentellen Bedingungen manipulierbar [z.B. Spanagel und Bohus, 2014]. Die Psychotherapieentwicklung ist im Begriff, sich in ihrer Methodenentwicklung an diesen neuen Erkenntnissen zu orientieren und neue Mechanismen-basierte Interventionen zu generieren [z.B. Grawe, 2004; Bohus, 2013]. Auch wenn die Anfänge noch etwas holprig anmuten, so besteht meines Erachtens kein Zweifel am immensen Potenzial, den dieser Ansatz für unser Fachgebiet birgt – emanzipiert sich die Therapieentwicklung doch erstmals in der Geschichte der Psychotherapie von den wenigen, mehr oder minder genialen Charismatikern, von denen unsere Profession bislang abhing. Und – um es unmissverständlich klarzustellen – dieser Anschluss an die experimentelle psychologische Forschung kann und wird nur an Forschungszentren möglich sein. Der Top-down-Ansatz, so problematisch er sein mag, wird seine Berechtigung in den nächsten Jahren erst entfalten – umso bedeutsamer wird es sein, dass die Forschung die Überbrückung der Kluft zwischen Wissenschaft und Praxis zu ihrem Problem und Aufgabenfeld erklärt und es nicht weiterhin den Praktikern zuweist.

### Bottom-up-Forschung

Wie oben skizziert, haben sich über viele Jahre, quasi als Gegengewicht zur Top-down-Forschung, Modelle etabliert, in welchen die therapeutische Praxis «so wie sie ist» zum Forschungsgegenstand wird. Dem liegt die implizite Annahme zugrunde, dass im gesamten Feld praktischer Therapeuten eigentlich genügend Expertise vorliegt – nur ist die Verteilung etwas ungleichmäßig. Es kommt also darauf an, die «nuggets», also die hoch wirksamen Therapeuten- und interventionellen Variablen zu entdecken, zu extrahieren, zu kommunizieren und im zweiten Schritt eventuell experimentell zu überprüfen.

So charmant dieser Ansatz klingt, er birgt doch einige erhebliche Probleme: Einerseits wird vorausgesetzt, dass sich innerhalb des Feldes der Praktiker tatsächlich die optimale Expertise verbirgt, wenn auch ungleich verteilt. Dem mag so sein. Empirische Untersuchungen hingegen deuten in eine etwas andere Richtung: In einer Übersichtsarbeit zur Entwicklung von Expertise bei unterschiedlichen Berufsgruppen kommt Shanteau [1992] zu dem Schluss, dass sich für eine Vielzahl von Berufen positive Zusammenhänge zwischen Leistung und Dauer der beruflichen Erfahrung nachweisen lassen. Dies trifft insbesondere zu für Piloten, Schachmeister, Mathematiker, Astronomen und Versicherungsanalysten. Andererseits gibt es Berufsgruppen, bei denen dieser Zusammenhang nicht gezeigt werden kann, und hierzu gehören

neben Richtern und Personalberatern namentlich Psychiater und klinische Psychologen. Als zentrale Unterscheidungsmerkmale zwischen diesen beiden Berufstypen charakterisiert Shanteau neben der Vorhersagekraft von Entscheidungen insbesondere die Möglichkeit von qualitativ hochwertigen Rückmeldungen. Tracey et al. [2014] argumentieren in einer lesenswerten Arbeit zur Frage der Expertise in der Psychotherapie, dass es unserer Profession an gerade dieser Expertise mangelt, und dies in erheblichem Maße. Die psychotherapeutische Praxis, so wie sie heute organisiert ist, verfügt über so gut wie kein systematisches Feedback-System bezüglich der Wirksamkeit der jeweiligen Interventionen. Auf gut Deutsch gesagt: In aller Regel erfahren die Therapeuten nicht, ob ihre Interventionen wirksam sind oder nicht. Daher ist individuelles Lernen enorm erschwert. Vielleicht hilft es, sich diesen trockenen Sachverhalt etwas zu verbildlichen: Kein anderes Gebiet der Medizin hat in den letzten 100 Jahren so eklatante Fortschritte erbracht wie die Chirurgie. In aller Regel sieht der Chirurg unmittelbar, ob seine Schnittführung zielführend, ob die Blutstillung erfolgreich und die Gefäßnaht dicht ist. Das Feedback-System ist kurz, die Lernkurve steil. Auch nach Beendigung der Operation kann der Chirurg davon ausgehen, dass mangelhafte operative Technik zu Schmerzen, Nachblutungen oder anderweitigen Beschwerden führt, die ihm oft rascher zur Kenntnis gebracht werden, als ihm lieb ist. Er wird daher eine sorgfältige Fehleranalyse durchführen, die Technik korrigieren und beim nächsten Mal vermeiden. Dadurch wird er Renommee erwerben und durch hohen Zulauf an Patienten seine Einnahmen verbessern. Kollegen werden auf diesen Vorsprung aufmerksam und werden dessen Methodik kopieren – das gesamte Fach profitiert. Vergleicht man damit die Entwicklung der Psychotherapie im gleichen Zeitraum, so möchte man erschauern angesichts der Tatsache, dass wir uns noch immer mit Methoden herumschlagen, die vor etwa 100 Jahren entwickelt wurden. Und selbst die «modernen störungsspezifischen Verfahren» basieren in aller Regel auf Interventionen, die 30/40 Jahre alt sind, auch wenn sie in manchem Manual als neue Erfindung verkauft werden. Doch nicht nur das Fachgebiet entwickelt sich verlangsamt, auch der einzelne Therapeut pflegt Beharrlichkeit. Bislang wurde noch in keiner mir bekannten Studie der Nachweis erbracht, dass erfahrener Therapeuten bessere Therapieergebnisse erzielen als Novizen [z.B. Wampold und Brown, 2005]. In aller Regel erzielen auch unter naturalistischen Bedingungen Ausbildungskandidaten gleichwertige oder bessere Therapieergebnisse als erfahrene Therapeuten [z.B. Beutler, 1997; Beutler et al., 2004; Budge et al., 2013]. Während sich also in den meisten Berufen langjährige praktische Erfahrung in der Entwicklung einer umfangreichen und strukturierten Wissensbasis niederschlägt, gekoppelt mit einer hohen prozeduralen Kompetenz in der Entscheidungsfindung, scheinen Psychotherapeuten diesen Prozess nicht nur langsamer, sondern sogar bisweilen invers zu durchschreiten.

Auch wenn diese Ansicht vielleicht etwas akzentuiert erscheinen mag – es wäre aus wissenschaftlicher Sicht sicherlich gewagt, davon auszugehen, dass im psychotherapeutischen Feld ausreichende, sich stetig verbessernde Expertise vorhanden ist, um diese als primäre Quelle der Forschungsentwicklung zu nutzen.

Und selbst wenn dem so wäre – die Umsetzung des Bottom-up-Ansatzes erfordert enorme Methodenkompetenz und Kosten: zum einen die Entwicklung, Validierung und Auswertung differenzierter und operationalisierter Verhaltensbeobachtung, zum anderen die Auswertung enorm großer Datensätze, und schließlich können die postulierten Ergebnisse lediglich zur Hypothesengenerierung herangezogen werden und erfordern im zweiten Schritt immer noch experimentelle Studien, die dann mit denselben Problemen aufwarten wie die in Top-down-Ansätzen entwickelten Fragestellungen.

Beide beschriebenen Forschungsansätze bergen also ihre spezifischen Probleme, deren Lösungen sicherlich nicht in weiterer Polarisierung zu finden sind. Bieten sich Alternativen?

### **Bidirektionale Forschungs- und Entwicklungsansätze: Strukturierte Kollaboration zwischen Praktikern und Forschern**

Die Fachzeitschrift *Psychotherapy Research* eröffnete das Jahr 2015 mit einem Themenheft zu «Fostering collaboration between researchers and clinicians through practice-oriented research» [Gastonguay und Muran, 2015]. In einer Reihe von ausgewählten Publikationen werden Prinzipien und Beispiele von erfolgreichen strukturierten Kollaborationsprojekten zwischen niedergelassenen Praktikern und Psychotherapieforschern beschrieben. Das Grundprinzip besteht darin, dass die Kooperationspartner nicht nur an der Implementierung von top-down-entwickelten Therapieprogrammen arbeiten, sondern zirkuläre Prozesse durchschreiten. Diese beginnen mit der Entwicklung der Fragestellungen und der Auswahl der Messmethoden sowie Patientenstichproben. Nach gemeinsamer Rekrutierung und Publikation sind Implementierungsprozesse und Evaluation von Akzeptanz und Umsetzbarkeit in der Praxis integrale Bestandteile dieses Prozesses [z.B. Garland und Brookman-Fraze, 2015].

### **Literatur**

American Psychological Association (APA): Policy statement on evidence-based practice in psychology (2005). [www.apa.org/practice/guidelines/evidence-based-statement.aspx](http://www.apa.org/practice/guidelines/evidence-based-statement.aspx) (Zugriff 24.02.2015).

Antes G, Clarke M: Knowledge as a key resource for health challenges. *Lancet* 2012;379:195–196.

Barkham M, Hardy G, Mellor-Clark J: *Developing and Delivering Practice-based Evidence: A Guide for the Psychological Therapies*. Oxford, Wiley-Blackwell, 2010.

Bedics J, Atcins D, Comtoise K, Linehan M: Treatment differences in the therapeutic relationship and introject during a 2-year randomized controlled trial of dialectical behavior therapy versus nonbehavioral psychotherapy experts for borderline personality disorder. *J Consult Clin Psychol* 2012;80:66–77.

Berry K, Haddock G: The implementation of the NICE guidelines for schizophrenia: barriers to the implementation of psychological interventions and recommendations for the future. *Psychol Psychother* 2008; 81(pt 4):419–436.

Beutler L: The psychotherapist as a neglected variable in psychotherapy: an illustration by reference to the role of therapist experience and training. *Clin Psychol – Sci Pr* 1997;4:44–52.

Beutler L, Malik M, Alimohamed S, Harwood T, Talebi H, Noble S, Wong E. (2004). The neglected common factor in psychotherapy; in Lampert MH (ed): *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change*, ed 5. New York, Wiley, 2004, pp 227–306.

Beutler L, Williams R, Wakefield P, Entwistle S: Bridging scientist and practitioner perspectives in clinical psychology. *Am Psychol* 1995;50:984–994.

Bohus M: Warum ich doch kein Psychoanalytiker geworden bin; in Schneider, F. (ed): *Irgendwie kommt es anders – Psychiater erzählen*. Berlin, Springer, 2012, pp 51–61.

Bohus M: Modulare Psychotherapie – ein Werkzeugkoffer für Psychotherapeuten?; in Znoj H, Berger T (ed): *Die Kunst und Wissenschaft der Psychotherapie*. Bern, Huber, 2013, pp 116–183.

Bohus M, Dyer A, Priebe K, Krüger A, Kleindienst N, Schmahl C, Niedtfeld I, Steil R: Dialectical behaviour therapy for post-traumatic stress disorder after childhood sexual abuse in patients with and without borderline personality disorder: a randomized controlled trial. *Psychother Psychosom* 2013;82:221–233.

Bohus M, Falkai P, Herpertz S: *Modulare Psychotherapie – Rationale und Grundprinzipien*. *Die Psychiatrie* 2012;2:89–99.

Budge S, Owen J, Koptka S, Minami T, Hansen M, Hirsch G: Differences among trainees in client outcomes associated with the phase model of change. *Psychotherapy* 2013;50:150–157.

Chambless DL, Hollon D: Defining empirically supported therapies. *J. Consult Clin Psychol* 1998;66:7–18.

Cuijpers P, Smit F, Bohlmeijer E, Hollon D, Andersson G: Efficacy of cognitive-behavioural therapy and other psychological treatments for adult depression: meta-analytic study of publication bias. *Br J Psychiatry* 2010; 196:173–178.

Aus meiner Sicht bergen diese strukturierten Kooperationsmodelle erhebliche Vorteile für beide Seiten: Die Forscher sind gehalten, (oder vielleicht sollte man besser sagen, es ist ihnen ermöglicht) ihre elitäre, durch Methodenzwang und Peer-Group-Kohäsion verstärkte Sicht auf die Praxis zu revidieren, und tatsächlich «applied science», also angewandte Forschung im besten Sinne, zu praktizieren. Der Lohn wäre die rasche Umsetzung der oft mühsam erhobenen Befunde in der klinischen Praxis. Die Klinker sind gehalten, Defizite in der Behandlung zunächst zu fassen und in wissenschaftlichen Fragestellungen zu formulieren. Sie werden ihr Verständnis für die Zwänge, die Logik und damit auch für die Interpretation der Forschung erweitern und in ihrer Praxis davon profitieren. Und schließlich könnten derartige strukturierte Kooperationen zwischen verschiedenen Kulturen ja auch einigen Spaß bereiten.

### **(Disclosure) Statement**

Dieser Text ist ein Manuskript zur Klaus-Grawe-Mittagsvorlesung 2015. Er sollte nicht als wissenschaftliche Arbeit verstanden werden. Dazu fehlen sorgfältige Literaturanalysen, wissenschaftliche Bewertungskriterien und eine abwägende Grundhaltung. Vielmehr handelt es sich um eines dieser vielen pseudo-wissenschaftlichen Essays, die sich mit willkürlich aus dem Gedächtnis gefischten Publikationen schmücken, um den Anschein akademischer Relevanz zu suggerieren und der Meinung des Autors etwas mehr Gewicht zu verleihen. Und – auch dies sei vorangestellt – diese meine Meinung ist von einem erheblichen Bias geprägt, also Interessenkonflikten unterworfen, die hiermit deklariert werden sollten: Der Autor wird finanziert von einem psychiatrisch-psychotherapeutischen Forschungsinstitut, er ist gehalten, im- und explizite Zielvereinbarungen durch Drittmittelleinnahmen und Publikationen zu erfüllen und ist daher in die Wissenschaftsfabrikation sowohl ideell als auch finanziell eingebunden. Zudem erwirkt er ein wenig Zusatzeinkommen und interpersonell induzierte Freude durch Weiterbildungsveranstaltungen bei niedergelassenen Psychotherapeuten. Bewahren Sie also Ihre Skepsis.

- Currin L, Waller G, Treasure J, Nodder J, Stone C, Yeomans M, Schmidt U: The use of guidelines for dissemination of 'best practice' in primary care of eating disorders. *Int J Eat Disord* 2007;40:476–479.
- Ehlers A, Bisson J, Clark D, Creamer M, Pilling S, Richards D, Schnurr P, Turner S, Yule W: Do all psychological treatments really work the same in posttraumatic stress disorder? *Clin Psychol Rev* 2010;30:269–276.
- Fairburn CG, Cooper Z, Doll HA, O'Connor ME, Palmer RL, Dalle Grave R: Enhanced cognitive behaviour therapy for adults with anorexia nervosa: a UK-Italy study. *Behav Res Ther* 2013;51:R2–R8.
- Fixsen D, Naoom S, Blase K, Friedman R, Wallace F: Implementation Research: A Synthesis of the Literature. Tampa, University of South Florida, 2005, pp. 125. [www.ctndisseminationslibrary.org/PDF/nirmonograph.pdf](http://www.ctndisseminationslibrary.org/PDF/nirmonograph.pdf) (Zugriff 25.02.2015).
- Flückiger C, Del Re A, Wampold B, Symonds D, Horvath A: How central is the alliance in psychotherapy? A multilevel longitudinal meta-analysis. *J Couns Psychol* 2012;59:10–17.
- Garland A, Brookman-Frazee L: Therapists and researchers: advancing collaboration. *Psychother Res* 2015;25:95–107.
- Gastonguay L, Muran C: Fostering collaboration between researchers and clinicians through practice-oriented research. *Psychother Res* 2015;25:1–5.
- Gillespie K, Duffy M, Hackmann A, Clark DM: Community based cognitive therapy in the treatment of post-traumatic stress disorder following the Omagh bomb. *Behav Res Ther* 2002;40:345–357.
- Grawe K: *Psychologische Therapie*. Göttingen/New York, Hogrefe, 2000.
- Grawe K: *Neuropsychotherapie*. Göttingen/New York, Hogrefe, 2004.
- Grawe K, Donati R, Bernauer F: *Psychotherapie im Wandel: Von der Konfession zur Profession*. Göttingen/New York, Hogrefe, 1994.
- Halweg K, Fiegenbaum W, Frank M, Schroeder B, Witzleben I: Short- and long-term effectiveness of an empirically supported treatment for agoraphobia. *J Consult Clin Psychol* 2001;69:375–382.
- Horvath AO, Bedi RP: The alliance; in Norcross JC (ed): *Psychotherapy Relationships That Work: Evidence-based Responsiveness*. New York, Oxford University Press, 2002, pp 37–69.
- Kiessling C, Langewitz W: The longitudinal curriculum 'social and communicative competencies' within Bologna-reformed undergraduate medical education in Basel. *GMS Z Med Ausbild* 2013;15:30.
- Lambert M, Shimokawa K: Collecting client feedback; in Norcross J (ed): *Psychotherapy Relationships That Work: Evidence-based Responsiveness*, ed 2. New York, Oxford University Press, 2011, pp 203–223.
- McHugh R, Barlow D: *Dissemination and Implementation of Evidence-based Psychological Interventions*. New York, Oxford University Press, 2012.
- National Collaborating Centre for Mental Health: *Schizophrenia: core interventions in the treatment and management of schizophrenia in primary and secondary care*. NICE Clinical Guidelines, No. 82. Leicester, British Psychological Society, 2009.
- Niemeyer H, Musch J, Pietrowsky R: Publication bias in meta-analyses of the efficacy of psychotherapeutic interventions for depression. *J Cons Clin Psychol* 2013;81:58–74.
- Sackett DL, Straus SE, Richardson W, et al.: *Evidence-based Medicine: How to Practice and Teach EBM*, ed 3. London, Churchill Livingstone, 2005.
- Schmucker C, Schell LK, Portalupi S, Oeller P, Cabrera L, Bassler D, Schwarzer G, Scherer R, Antes G, von Elm E, Meerpohl J: Extent of non-publication in cohorts of studies approved by research ethics committees or included in trial registries. *PLoS One* 2014;9:e114023.
- Shafran R, Clark D, Fairburg C, Antz A, Barlow D, Ehlers A, Freestone M, Garety P, Hollon S, Ost L, Salkovskis P, Williams J, Wilson G: Mind the gap: improving the dissemination of CBT. *Behav Res Ther* 2009;47:902–909.
- Shanteau J: Competence in experts: the role of task characteristics. *Organ Behav Hum Dec* 1992;53:252–266.
- Spanagel R, Bohus M: Disruption of reconsolidation processes is a balancing act – can it really account for change in psychotherapy? Commentary on Richard D. Lane et al. – Memory Reconsolidation, Emotional Arousal and the Process of Change in Psychotherapy: New Insights from Brain Science. *Behav Brain Sci* 2014;15:1–80.
- Stirman S, Derubeis R, Crits-Christoph P, Rothmann A: Can the randomized controlled trial literature generalize to nonrandomized patients? *J Consult Clin Psychol* 2005;73:127–135.
- Task Force on Promotion and Dissemination of Psychological Procedures: Training in and dissemination of empirically validated psychological treatments: reports and recommendations. *The Clinical Psychologist* 1995;48:3–23.
- Tracey T, Wampold B, Lichtenberg J, Goodyear R: Expertise in psychotherapy: An elusive goal? *Am Psychol* 2014;69:218–229.
- Tschuschke V: Die Psychotherapie in Zeiten evidenzbasierter Medizin. *Fehlentwicklungen und Korrekturvorschläge*. *Psychotherapeutenjournal* 2005;2:105–115.
- Wampold B, Brown G: Estimating therapist variability: a naturalistic study of outcomes in managed care. *J Consult Clin Psychol* 2005;73:914–923.
- Weber R, Ogronizcuk J, Schultz-Venrath U, Strauß B: Zum Verhältnis von Forschung und klinischer Praxis – Ergebnisse der Mitgliederbefragung der Deutschen Gesellschaft für Gruppenanalyse und Gruppenpsychotherapie (D3G) zur Wahrnehmung von Psychotherapieforschung 1. *Gruppenpsychotherapie und Gruppendynamik* 2013;49:37–52.