

## Psychotherapieforschung – und ihre Beschränkung durch einen schulenspezifischen Bias

Jürgen Kriz

### **Zusammenfassung: Psychotherapieforschung – und ihre Beschränkung durch einen schulenspezifischen Bias**

Das Ausüben von Psychotherapie einerseits und Psychotherapieforschung andererseits gehören unterschiedlichen Handlungsbereichen an – mit jeweils unterschiedlichen Regeln und Kompetenzen. Eine Vermengung beider Bereiche wäre ebenso problematisch wie der Versuch, die Regeln eines Bereiches zur Strukturierung des anderen heranzuziehen – also therapeutisches Handeln nach bestimmten wissenschaftsmethodischen Modellen zu strukturieren oder, anders herum, Forschungsmethodik nach therapeutischen Modellvorstellungen beziehungsweise „Schulen“ auszurichten.

Ein hervorragendes Modell zur Untersuchung experimenteller Forschungsfragen ist das RCT-Design mit randomisierten kontrollierten Studien. Dies gilt aber nur dann, wenn klar definierbare Ursachen und ebenso klar definierbare Wirkungen hinreichend isolierbar sind und die Realität brauchbar abbilden. Je größer der Handlungsspielraum zur Gestaltung der Ursachen (Interventionen) ist und je komplexer die relevanten Wirkungen sind, desto weniger tauglich ist dieser Ansatz.

RCT-Designs eignen sich in der Psychotherapieforschung daher besonders dann, wenn insbesondere manualisierte *Programme angewendet* werden. Dies gilt primär für die Verhaltenstherapie. Für psychotherapeutisches Vorgehen, das wesentlich in der prozess-spezifischen *Entfaltung von Prinzipien* begründet ist, sprengt die große Variabilität der Detail-Ursachen (Interventionen) die erforderliche massenstatistische Homogenität für ein experimentelles Design. Die Forderung, Wirksamkeit vor allem mit RCT-Designs zu belegen, läuft dort auf methodische Artefakte hinaus, beziehungsweise stülpt eine bestimmte psychotherapeutische Präferenz, nämlich für die Verhaltenstherapie, auf dem Wege der Wissenschaftsmethodik anderen Psychotherapieansätzen über.

**Schlüsselwörter:** Psychotherapieforschung, RCT-Design, Efficacy, Effectiveness, Experiment

### **Abstract: Psychotherapy Research – and its limitation by a school-specific bias**

Practicing psychotherapy belongs to another realm of acting than making research of psychotherapy. Both realms have different rules and need different competences. Confusing these realms is as problematic as trying to structure one realm by using the rules and patterns of the other one. I.e. to structure psychotherapeutic activities according to scientific methodological rules or, vice versa, to adjust the methodology of research to psychotherapeutic models or “schools”.

RCT-designs – randomized controlled trials – are first-rate models in order to examine experimental research questions. However, this is only true when clear causes and clear effects can be isolated so that this design is a sufficient representation of the reality. Increasing the scope of possibilities in shaping the causes (interventions) or having more complex effects makes this approach less appropriate.

In psychotherapy research, RCT-designs are, therefore, particularly adequate when *psychotherapeutic programs are applied*, which is true for behavior therapy. In contrast, psychotherapeutic approaches which are essentially grounded on the *unfolding of therapeutic principles* show a much too large variance in detailed causes (interventions) to fulfill the mass-statistical homogeneity of an experimental design. Here the demand to prove

effectiveness by using RCT-designs ends up in research-artifacts and imposes one particular psychotherapeutic preference, i.e. behavior therapy, on other approaches by way of research methodology.

**Keywords:** Psychotherapy Research, RCT-Design, Efficacy, Effectiveness, experimental approach

**Ein Wort zum Thema Gender:** Ich bin mir der Unzulänglichkeit unserer Sprache bewusst, zwischen geschlechtsunspezifischen und sogenannten „männlichen“ Bezeichnungen zu differenzieren. Trotz Bemühens um eine neutrale Ausdrucksweise (z. B. „Angehörige therapeutischer Berufe“), ist dies nicht überall möglich, weshalb auch hier öfter von „Therapeuten“, „Patienten“ etc. gesprochen bzw. geschrieben wird. Alle mir bekannten Alternativen – von denen auch ich bisweilen Gebrauch mache – erscheinen mir gestelzt und gekünstelt. „Therapeuten“, „Patienten“ etc. sind aber als geschlechtsneutrale Bezeichnungen gemeint – d. h. „Frauen“ sind **nicht** „mit-gemeint“, weil eben auch keine „Männer“ gemeint sind, und es geht auch um keine anderen, nicht genannten, biologischen, sozialen, ökonomischen oder sonstigen Eigenschaften, Rassen oder Ethnien: Es geht um Menschen, die Bestimmtes tun, denken, meinen und dabei gegebenenfalls bestimmte Rollen einnehmen – oder eben auch nicht.

### **Einleitung: Über die Vermengung von Forschung und Psychotherapie**

Gibt man in Google die Formulierung „Forschung in der Psychotherapie“ ein, so werden rund 14.100 Hinweise aufgeführt (Google, 2015) - von Titeln für wissenschaftliche Artikel, über die Bezeichnung eines Forschungsförderungspreises bis hin zu einem Symposium an einer Universität. In der Tat ist dies eine gängige Phrase, von der man selbstverständlich zu wissen glaubt, was damit gemeint sein soll. Doch diese scheinbare Selbstverständlichkeit erweist sich bei genauerem Hinsehen als trügerisch:

Es wird nämlich kaum zu bestreiten sein, dass wissenschaftliche Forschung und professionelle Psychotherapie recht unterschiedliche Tätigkeiten sind, die in der Regel von unterschiedlichen Berufstätigen – Wissenschaftlern bzw. Forschern einerseits sowie Psychotherapeuten andererseits – ausgeübt werden. Es handelt sich somit um unterschiedliche soziale Gruppen, die ihre Fragen, Ergebnisse und Probleme in unterschiedlichen Diskurswelten kommunizieren und entsprechend nach je eigenen Regeln, Gewohnheiten und Zielen handeln. Insofern kann es also keine Forschung *in* der Psychotherapie sondern nur eine Forschung *über* Psychotherapie geben.

Bei dieser Klarstellung geht es nun aber keineswegs um eine Kritik sprachlich-begrifflicher Ungenauigkeiten. Sondern es geht um die viel tiefergreifende und für die Entwicklung der Psychotherapie bedeutsame Nichtbeachtung der Unterschiede zwischen forschendem und psychotherapeutischem Tun bzw. um die Vermengung beider Anliegen und Regeln. Eine Forschung, die sich bewusst wäre, dass sie nach ihren – den wissenschaftlichen – Regeln bestimmte Fragen

an einen Gegenstand – psychotherapeutisches Handeln – stellt, käme schwerlich auf die Idee, diese Regeln gleichzeitig dem Gegenstand überzustülpen zu wollen. Große Teile der Mainstreamforschung über Psychotherapie aber bevorzugen nicht nur einseitig nomothetische Designs nach einem reduzierten Modell der Pharmakologie-Forschung. Sondern sie erheben auch lautstark den Anspruch, dass sich das Feld und der Möglichkeitsraum von Psychotherapie aufgrund der so erbrachten Ergebnisse, die bestenfalls einen kleinen Teil psychotherapeutischen Handelns und deren zugrundeliegenden Kompetenzen erfassen, zu strukturieren habe. Und obwohl die verlässlichen Daten nur einen winzigen Ausschnitt dessen abdecken, was Psychotherapie ausmacht, werden Behauptungen von atemberaubender Verallgemeinerung über die Wirksamkeit und die Nützlichkeit bestimmter Therapieansätze aufgestellt und daraus ebenso atemberaubende Verdikte darüber gefällt, welche Richtungen bzw. Verfahren „besser“ bzw. „schlechter“ sind. In der BRD geht dies sogar so weit, dass aufgrund einer einseitigen, fragwürdigen Datenlage seit 1999 faktisch Berufsverbote für ganze Richtungen vollstreckt wurden, was wiederum zur Verdrängung gut entwickelter Lehr- und Forschungstätigkeiten führte.

Dabei ist wohl kaum zu bestreiten, dass vieles Wesentliche der psychotherapeutischen Prozesse und Beziehungen noch gar nicht erforscht ist. Das gilt ebenso für mannigfaltige Fragen, auf welche Weise und wieweit etliches dessen, was den Alltag psychotherapeutischen Handelns ausmacht, zum Erfolg oder Misserfolg beiträgt. Ein beachtlicher Teil der Fragen, die für die psychotherapeutische Praxis relevant wären, ist zudem mit den bevorzugten Designs gar nicht brauchbar erforschbar. Und anderes, was Psychotherapeuten bewegt und erforscht werden könnte, wird schlicht ignoriert, weil es nicht in jene Forschungsstrukturen passt, die im akademischen System karrierefördernd sind – etwa, weil sich Ergebnisse nicht so schnell erbringen lassen oder die zu erwartenden Befunde nicht in jene Zeitschriften passen, die einen bestimmten „impact factor“ aufweisen. Nicht klinische Relevanz sondern akademische Verwertbarkeit bestimmt daher wesentlich die Forschung.

### **RCT-Forschung als Goldstandard**

Innerhalb der Rahmenbedingungen und den damit vorfindlichen Strukturen akademischer Forschung macht es sehr viel Sinn, recht spezifische Fragen nach dem experimentellen Design zu untersuchen. In der Tat beziehen sich zahlreiche Fragen, welche im Rahmen moderner Psychologie gestellt werden, auf einen sehr spezifischen, klar umrissenen Forschungsgegenstand, bei dem einzelne Wirkvariable im Labor gut isolierbar sind. Dies ist optimal durch eine experimentelle Herangehensweise untersuchbar. Und hierauf beruht ein nicht unbeträchtlicher Teil der Reputation der Psychologie als Wissenschaft: Blickt man nämlich in die Lehrbücher moderner Psychologie, so findet man darin eine überwältigende Fülle an Befunden zu Fragen über menschliche Wahrnehmung, Lernen, Gedächtnis, Denken, Sprache usw., die ganz überwiegend

auf Experimenten beruhen. Im Kern besteht dieser Untersuchungsansatz darin, dass man „Versuchspersonen“ (Vpn) einer bestimmten Reizbedingung aussetzt („unabhängige Variable“, UV) und registriert, wie sich daraufhin bestimmte Reaktionen („abhängige Variable“, AV) verändern, was als Funktion  $f$  gefasst werden kann:

$$\text{UV} \rightarrow \text{VPN} \rightarrow \text{AV} \quad \text{bzw.} \quad \text{AV} = f(\text{UV})$$

Die große Bedeutung dieses Ansatzes für die Psychologie hängt nicht nur mit den beachtlichen Erfolgen zusammen, die bei vielen experimentell zuschneidbaren Fragen erbracht wurden. Vielmehr entspricht diese Vorgehensweise in der Forschung immer noch dem, was zumindest Laien für den Kern naturwissenschaftlichen Arbeitens halten (auch wenn dies dem Weltbild des 19. Jahrhunderts entspricht und innerhalb der modernen Naturwissenschaften seit vielen Jahrzehnten als überholt gilt). Auch die Pharmakologie bedient sich bei der Erprobung einzelner neuer Substanzen dieses Ansatzes – und erzeugt damit auch für die Erforschung von Psychotherapie einen erheblichen Konkurrenzdruck, die Wirksamkeit möglichst mit derselben Methodik unter Beweis zu stellen.

Das dabei verwendete Standardmodell wird bekanntlich als „Randomisiertes Kontrolliertes Design“ bzw. RCT-Studie (engl.: randomized controlled trial) bezeichnet und wurde lange Zeit als sogenannter „Goldstandard“ dieser Forschung angesehen. Der international renommierte Psychotherapieforscher Martin E. P. Seligman führte dazu in einer Arbeit unmittelbar vor seiner Wahl zum Präsidenten der „American Psychological Association“ (APA) aus:

Grundsätzlich wird dabei eine Behandlungsgruppe unter gut kontrollierten Bedingungen mit einer Kontrollgruppe verglichen. Aber neben der Kontrollgruppe weisen Effizienzstudien, die mittlerweile zu einem paradigmatischen Unterfangen geworden sind, noch eine hochkomplexe Methodologie auf. Im Idealfall gehören die folgenden Raffinessen dazu:

1. Zuordnung der Patienten zu der Behandlungs- und Kontrollgruppe nach dem Zufallsprinzip.
2. Strenge Kontrollen: Neben Kontrollgruppen ohne Behandlung werden Kontrollgruppen gebildet, bei denen eine Placebo-Behandlung mit potentiell therapeutischen Elementen eingesetzt wird, die für Patienten wie Therapeuten glaubwürdig sind. Damit wird der Einfluss unspezifischer Faktoren überprüft, zum Beispiel Rapport, Erfolgserwartung und teilnahmevolle Aufmerksamkeit (sog. Nonspezifika).
3. Die Behandlung ist standardisiert, mit hochgradig detaillierten expliziten Therapievorgaben. Die Einhaltung der Standards wird anhand von Videoaufnahmen der Sitzungen überprüft, Abweichungen werden korrigiert.
4. Die Behandlung erfolgt in einer festgelegten Zahl von Sitzungen.

5. Die Zielergebnisse werden sorgfältig operationalisiert (z.B. klinisch diagnostizierte DSM-IV-Störungen, Zahl der berichteten Orgasmen, Selbstauskünfte über Panikattacken, prozentueller Anteil flüssiger Äußerungen).
6. Auswerter und Diagnostiker wissen nicht, aus welcher Gruppe die Patienten kommen. (Anders als bei Medikamentenstudien sind „doppel-blinde“ Effizienzstudien in der Psychotherapie unmöglich, da Patienten und Therapeuten die Behandlungsform kennen. Wann immer Sie jemanden eine „Doppelblindstudie“ in der Psychotherapie verlangen hören, halten Sie Ihren Geldbeutel zurück!)
7. Die Patienten erfüllen die Kriterien für eine einzige diagnostizierte Störung; Patienten mit mehreren Störungen werden in der Regel nicht einbezogen.
8. Nach Behandlungsende werden die Patienten über einen festgelegten Zeitraum mit einer umfassenden Testbatterie eingeschätzt. (Seligman 1997, S. 267)

Die Notwendigkeit und Angemessenheit dieser „Raffinessen“ wird an der Prüfung der Wirkung einer antibakteriellen Medizin „M“ leicht einsichtig: Wenn die Patienten selbst aussuchen dürfen, in welche Gruppe sie kommen, könnte M als wirksamer erscheinen, weil vielleicht nur die leichten Fälle in diese Gruppe gehen, oder die mit mehr Motivation, oder es sind mehr zahlungskräftige dabei, die sich aber gleichzeitig gesünder ernähren etc. und deswegen schneller geheilt werden (1). Man muss sicherstellen, dass tatsächlich M wirkt und nicht der „Glaube an M“ oder die Flasche als Verpackung oder die Bakterien sowieso von allein verschwinden (2); ebenso, dass tatsächlich M in der Flasche ist und nicht irgendetwas anderes (3); und im Vergleich sollen natürlich nicht unterschiedliche Mengen verwendet werden (4). Bekämpft werden sollten die Bakterien (5), während z. B. Glücksgefühle bei M irrelevant sind. Die Ergebnisse sollen auch nicht durch die Erwartungen der Auswerter verfälscht werden (6). Da M gegen einen bestimmten Bakterienstamm wirken soll, sind eben auch nur Patienten mit diesem Befall relevant – bei Fußpilz würde die Wirkung von M eher unterschätzt werden, bei einer allgemeinen Immunschwäche weiß man nicht wie diese (statt M) den Verlauf beeinflusst (7). Man muss auch sicherstellen, dass die Bakterien zumindest eine gewisse Zeit beseitigt sind – und nicht nach wenigen Wochen z. B. die Krankheit wieder aufflammt (8).

### **Wirkfaktoren in der komplexen Realität**

Diese Bedingungen lassen sich, wie gesagt, für die jeweils recht speziellen Fragen in Experimenten psychologischer Grundlagenforschung hinreichend gut erfüllen – z. B. in Wahrnehmungs-, Lern-, oder Gedächtnisexperimenten. Psychotherapie aber ist nicht nur etwas anderes als Forschung. Sondern sie findet auch nicht im Labor unter den experimentellen Bedingungen einzelner isolierbarer Wirkungen statt. Vielmehr haben wir es in der Psychotherapie immer mit einem komplexen Gefüge von Einflüssen und Wirkungen zu tun, deren nichtlineare Wechsel-

wirkungen eigentlich auch ohne große Methodenkenntnis die Unangemessenheit des obigen Ansatzes offensichtlich machen sollten. Recht anschaulich hat Plaum (1999) in einer kurzen – fiktiven aber nicht untypischen – Fallgeschichte die zugrundeliegende Problematik dargestellt. Daran wird deutlich, wie fragwürdig es ist, in komplexen Prozessräumen, wie dies für Psychotherapie typisch ist, nach isolierbaren Wirkfaktoren zu suchen:

... Aufgrund der Distanz von zuhause war Frau X. in der Lage, ein anregendes Klima, eine reizvolle Landschaft, positiv auf sich wirken zu lassen, was durch die entspannende Wirkung der Bäder verstärkt worden ist; dass es Mitpatienten gibt, denen es offenbar schlechter ging als ihr, hat Frau X. neuen Lebensmut gegeben und sie offener für andere Menschen gemacht; die daraufhin möglichen Gespräche konnten ihr wiederum veränderte Lebensperspektiven nahebringen; schließlich hat der Masseur bei Frau X. erotische Gefühle ausgelöst, die wiederum in der geschilderten aktuellen sozialen Situation zur Realisierung eines ‚Kurschattens‘ führten; dieser hat aufgrund seiner anziehenden Wirkung Frau X. zu der Überzeugung bringen können, dass es doch besser sei, die Diätvorschriften einzuhalten und sich nicht etwa am Nachmittag ins Café zu setzen, um dort die Sahnetorte zu verspeisen ... (Plaum, 1999, S. 201).

Zu Recht glossiert Plaum die Suche nach Wirkfaktoren in einer komplexen Welt und fragt: „Was hat zur Verbesserung des gesundheitlichen Zustandes von Frau X. während einer Kur geführt? Bäder? Massagen? Das Klima? Die Distanz von zuhause? Der Kontakt zu Mitpatienten? Der ‚Kurschatten‘?“ (1999, S. 200). Im gewählten Beispiel muss man sagen: wohl alles „irgendwie“. Doch nichts wirkt als isolierbarer „Faktor“ als solcher – und auch nicht im Sinne varianzanalytischer Wechselwirkung „zusammen“. Denn man kann sich gut vorstellen, dass in Gruppen mit Personen unter „fast“ genau denselben Bedingungen die Gesamtwirkung völlig unterschiedlich ausfällt. Ein Forschungsansatz, bei welchem der Einfluss des Kurschattens, oder des Masseurs etc. als unabhängige Variable in einem experimentellen Design untersucht werden soll, wäre daher entsprechend lächerlich. Und bei einem anderen Patienten wirken vielleicht in derselben Kurklinik andere Aspekte zusammen. Man kann sogar davon ausgehen, dass die Patienten selbst auch etwas Kompetenz besitzen, aus den reichhaltigen Programmen und komplexen Möglichkeiten einer solchen Klinik das auszuwählen, was für sie gut ist. Und dass es vielleicht reicht, sie bei dieser Wahl und bei der Reflexion darüber zu begleiten und zu beraten sowie in der Nutzung der Möglichkeiten zu fördern.

Eine solche Sicht passt allerdings nicht in jene Forschungsdesigns, nach denen sich die Psychotherapie in ihrem Nachweis von Effektivität und empirischer Evidenz ausrichten muss. Dennoch wäre es wohl falsch, Kurkliniken als „unwirksam“ oder „wissenschaftlich nicht anerkannt“ zu brandmarken und aus der Versorgung zu verbannen (so wie es in der BRD mit vielen psychotherapeutischen Angeboten in ambulanten Praxen nach dem Psychotherapeutengesetz von 1999 geschehen ist).

## Die Fragwürdigkeit des experimentellen Schemas für die Psychotherapieforschung

Der Unterschied zwischen Forschung *über Psychotherapie* zur experimentellen Forschung *in der Psychologie* wird deutlich, wenn man die komplexen Prozesse einer Psychotherapie in dieses klassische experimentelle Schema einzuordnen versucht, um damit – im üblichen Sinne – psychotherapeutische Wirksamkeitsnachweise zu erbringen. Das obige Schema (UV → VPN → AV) sieht dann wie folgt aus:

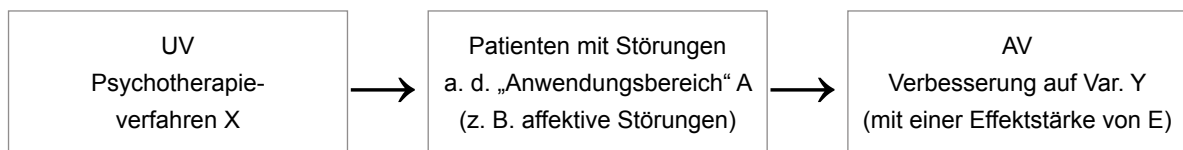


Abbildung 1: (Kriz, 2010, S. 132)

Auf den ersten Blick wirkt auch diese Vorgehensweise recht plausibel. Immerhin will man doch prüfen, ob ein bestimmter Ansatz der Psychotherapie (X) bei Patienten mit Störung A auf einer Zielvariablen Y eine Verbesserung erzielt, und zwar mit Effektstärke E. Allerdings enthält dieses Schema einige *implizite* Annahmen, von denen fraglich ist, ob sie von Psychotherapieforschern auch *explizit* so unterstellt werden oder aber einfach nicht bewusst sind. Für Psychotherapeuten jedenfalls ist klar, dass diese Annahmen in hohem Maße unzutreffend sind. Diese zentralen Annahmen sind:

1. **Es „gibt“ X:** Dafür ist notwendig, dass die Therapie nicht situationsspezifisch aus allgemeinen Prinzipien entfaltet wird, weil dann X je nach Situation gegebenenfalls unterschiedlich sein würde. Vielmehr muss X quasi manualisiert „verabreicht“ werden. In der Tat führt dies genau zum o. a. „Goldstandard“ der Psychotherapieforschung (Kriz, 2004; Seligman, 1997) mit randomisierten, kontrollierten Studien (RCT-Design), wo eine klare Vorgehensweise X mit einer festen Stundenzahl (oft 10-15 Sitzungen) sicherstellen soll, dass X eben nicht zu beliebig wird und somit gar nicht klar ist, was als UV in dem Schema wirkt.

Das klingt plausibel. Doch wenn selbst für die Verhaltenstherapie (VT) ein prominenter VT-Institutsleiter betont: „VT stellt eben nicht die Behandlung von Diagnosen mit medikationsähnlichen Standardinterventionen dar“ (Zarbock, 2010, S. 126), fragt man sich, für welche Therapie dann Ergebnisse aus dem oben angeführten Schema Gültigkeit beanspruchen können. Denn es ist klar, dass selbst bei solchen Vorgehensweisen in der VT, die recht stark manualisiert und quasi als „Programm“ vorgegeben sind, in jedem konkreten Einzelfall vieles im therapeutischen Raum geschieht, das gar nicht untersucht wurde. Schon so einfache praktische Fragen, ob man Patienten mit Handschlag begrüßen oder einer weinenden Patientin ein Taschentuch reichen sollte oder nicht, sind nicht RCT-erforscht (siehe auch Punkt 3 unten).

2. **Es „gibt“ A:** Da X als UV natürlich nicht einfach auf alles wirkt, muss ein genaues A definiert werden – dies sind im ICD oder DSM beschriebene Diagnose-Kategorien. Allerdings ist das Anliegen klinischer Diagnostik, Störungen kategoriell zu beschreiben, wobei für einen bestimmten Patienten meist mehrere Kategorien herangezogen werden müssen. Man spricht dann – unsinnigerweise – von Komorbidität bzw. Multimorbidität, so als hätte dieser Mensch zwei oder mehrere Störungen (wie Fußpilz und Rückenleiden). In Wirklichkeit fällt aber hier nur *eine* Störung in mehrere Kategorien – was etwas völlig anderes ist. Ein fataleres Missverständnis liegt allerdings darin, dass zwar – statistisch gesehen – das „Gewicht“ oder die „Körpergröße“ eines Menschen als „Realisation“ einer Verteilung der Gewichte oder Größen einer sinnvoll definierten Population mit „wahren Mittelwerten“ gesehen werden kann. Eine als „Depression“ diagnostizierte Störung eines Menschen ist aber nicht sinnvoll als „Realisation“ einer Depressions-Verteilung mit „wahrem Depressionswert“ anzusehen. Es sei auch nochmals an das oben angeführte Zitat von Zarbock erinnert, dass eben nicht „Diagnosen“ behandelt werden.

Beachtenswert ist auch, dass ohnedies psychopathologische Diagnosen wie „Depression“ überaus grobe Cluster von Symptomen darstellen. So geht beispielsweise Florian Holsboer, langjähriger Direktor des Max-Planck-Instituts für Psychiatrie in München, davon aus, dass man wahrscheinlich zwischen zehn oder mehr Unterformen mit ganz unterschiedlichen, krankmachenden Mechanismen differenzieren müsse (Lakotta, 2009). Doch macht dann das Schema in der Psychotherapieforschung noch Sinn? Über welches A reden wir bzw. worüber wird der RCT-Nachweis erbracht?

3. **X ist hinreichend isolierbar:** Eine weitere Voraussetzung ist, dass außer X keine weiteren bedeutsamen Wirkvariablen im Experiment wirken. Wie ist es aber beispielsweise mit der „therapeutischen Beziehung“ in der VT? Inzwischen wird deren Bedeutsamkeit auch für die VT nicht bestritten. Doch wo findet sich in der VT-Theorie eine genaue Herleitung dessen, was mit „therapeutischer Beziehung“ in diesem Ansatz gemeint ist? Wo sind die Dismantling-Studien, welche den Einfluss dieser Wirkvariablen von der „eigentlichen“ VT bestimmen können? Das heißt, wir dürfen davon ausgehen, dass letztlich keineswegs die X allein oder auch nur die wesentlichen Anteile der Wirkung hervorbringen - wie ohnedies unter dem Thema „unspezifische Wirkfaktoren“ von den meisten Forschern eingeräumt wird und oben schon unter Punkt 1 angesprochen wurde.

4. **Y ist aggregierbar:** Die letzte der hier (keineswegs erschöpfend - siehe Kriz 1996, 2003, 2012) aufzuführenden impliziten Voraussetzungen ist die Aggregierbarkeit der Daten. Denn im Rahmen der üblichen statistischen Auswertungsmethoden muss sichergestellt werden, dass dieselbe „Dosis“ von X nicht einmal fast nichts, ein andermal extrem viel verändert – weil dies z. B. bei Mittelwerten dann weitgehend Artefakte der zufälligen Mischung unterschiedlicher Phasen abbilden würde und nicht die Wirkung von X im Vergleich z. B.



zur Kontrollgruppe oder zu einem anderen Ansatz Z. Dies wiederum setzt voraus, dass die Symptom- und auch die Heilungsdynamik hinreichend linear verläuft – zumindest ohne große Instabilitäten und Nichtlinearitäten.

Wie aber die umfassende Literatur über therapeutische und andere Entwicklungsprozesse übereinstimmend zeigt, sind solche Verläufe typischerweise (und nicht etwa ausnahmsweise) nichtlinear, mit bedeutsamen Entwicklungs“schüben“ und –„stagnationen“ in der Zeit.

Die in den Punkten 1 - 4 zusammengefassten Argumente werfen erhebliche Zweifel auf, wie weit der sehr erfolgreiche linear-experimentelle Ansatz entsprechend dem oben angeführten Schema verallgemeinerbar ist. Insbesondere bei komplexeren Prozessen, die nicht hinreichend im Sinne der X, Y und A idealisierbar sind und bei denen sich vor allem die durch Rückkopplungen verursachten Nichtlinearitäten nicht gut ausblenden lassen, sind die Voraussetzungen so inadäquat, dass Forschungsartefakte statt Ergebnisse befürchtet werden müssen. Nur dort, wo diese Bedingungen hinreichend erfüllt sind – was im Rahmen streng kontrollierter, isolierbarer und unter Ausschluss von Rückkopplungen (hinreichend) im Rahmen von Grundlagenforschung möglich ist – gibt das oben angeführte Schema tatsächlich auf klare Weise brauchbare Antworten. Und dies deshalb, weil dort UV, AV und f spezifiziert werden können.

Damit aber lässt sich die Frage auch anders herum stellen: Welche Bedingungen müssen gegeben sein, damit das Modell der experimentellen Forschung und das RCT-Design zumindest hinreichend brauchbar zur Realität passen?

### **Der schulenspezifische Bias des RCT-Ansatzes**

Geht man dieser eben aufgeworfenen Frageperspektive nach, so ist jenseits des Labors ein solches Modell für die (Psychotherapie)Praxis dann hinreichend brauchbar, wenn diese Praxis eben entsprechend den Bedingungen gestaltet werden kann bzw. typischerweise so gestaltet wird. Dies ist besonders bei der Entwicklung und Evaluation von *Programmen* der Fall, in denen die Aspekte in der Seligman Liste eben auch zu standardisierten Vorgehensweisen in der Praxis umgesetzt sind – also auch bei psychotherapeutischen *Programmen*: Sind z. B. in einer Klinik für den 6-wöchigen Aufenthalt 10 Sitzungen für eine bestimmte Störung vorgesehen, in denen jeweils ein sehr klar vorgegebenes Programm durchgeführt wird, so sind die oben angeführten Bedingungen 3, 4, 5 und 7 per se erfüllt. Und 1, 2, 6 und 8 ließen sich leicht für eine Untersuchung realisieren – bzw. stehen für die reine Praxis in keinem Widerspruch zum ehemaligen Forschungsdesign. *Um die Wirksamkeit der Anwendung von Programmen zu prüfen, sind RCT-Studien somit gut brauchbar.*

Wie brauchbar ist dieser Ansatz aber zur Beurteilung der Wirksamkeit von Psychotherapie, bei welcher die Behandlung typischerweise nicht nach einem vorgegebenen Programm erfolgt, sondern auf einer dynamischen Passung von eher allgemeinen Prinzipien an die gegenwärtige Problem-Situation beruht (wobei in das, was mit „Problem-Situation“ gemeint ist, recht viele Aspekte einfließen können: von der genetischen Disposition über biographische Strukturierungen, Lerngeschichten, aktuelle Belastungen, Symptomstabilisierungen im sozialen Umfeld etc. etc.). Wird also entsprechend der spezifischen aktuellen Situation das konkrete therapeutische Vorgehen aus bestimmten Prinzipien erst entfaltet und in passende Handlungen umgesetzt, so sind die zentralen Bedingungen 3, 4 und 7 typischerweise nicht anwendbar. Bedingung 7 wird sogar als Artefakt gesehen, wenn man „komorbide“ Patienten so versteht, dass diese (meist) nicht „mehrere Störungen haben“, sondern *eine* „Störung“, deren Symptome gegebenenfalls mehreren diagnostischen Kategorien zugeordnet werden können. Über Bedingung 5 kann man sich einigen – sie sollte nicht zu sehr an den essentiellen Aspekten der Störung vorbeigehen. Ob also z. B. der Fokus auf der „Zahl der berichteten Orgasmen“ eher als ein Heilungskriterium angesehen wird oder selbst als Ausdruck einer Störung (wie schon Wilhelm Reich 1972 betonte), bedarf einer *inhaltlichen* Diskussion.

Fasst man diese Argumente zusammen, so lässt sich sagen: *Um die Wirksamkeit von Therapieformen zu prüfen, die in spezifischer Passung Kriterien entfalten, sind RCT-Studien somit unangemessen.*

Natürlich kann man auch diese Therapieformen auf programmartige Vorgehensweisen, unter den oben genannten Standardbedingungen reduzieren und damit die Forderung nach RCT-Studien für diese „Varianten“ erfüllen. Gezwungenermaßen wurden solche Studien auch für Vorgehensweisen durchgeführt, die nicht zu dieser Forschungslogik passen – denn auch in den USA ist eine Gruppe sehr einflussreich, die unter dem Begriff „Empirically Supported Treatment (EST)“ (empirisch gestützte Behandlungen) nur RCT-Studien als einzige Form der empirischen Forschung ansieht (Chambless & Hollon, 1998) – obwohl die Lehrbücher über empirische Forschungsmethoden auch in den USA zahlreiche andere Ansätze enthalten. Es sollte aber klar sein, dass ein solcher Forschungsansatz eben unangemessen ist und das Wesentliche verfehlt. Es ist ein wenig so, als wenn die Autolobby lediglich den Nachweis gelten lässt, wie schnell man auf Autobahnen vorankommt – und dabei das Essentielle ignoriert, dass dies für die Beurteilung von PKW und LKW sinnvoll sein mag, aber nicht für Pferde – auch wenn man diese zwingen kann, sich am Wettrennen auf Autobahnen zu beteiligen. Doch dies ist nicht nur unangemessen, sondern entwertet auch die Leistung der Pferde, z. B. über Hecken, Zäune und in unwegsamem Gelände voranzukommen und somit auch solche Menschen zu erreichen, die nicht an den Hauptstraßen angesiedelt sind.

## Das weite Feld an Forschungsdesigns jenseits des RCT-Fokus

Nicht alle Wissenschaftler sind einseitig auf RCT-Forschung fixiert. Die obige Zusammenstellung der „Raffinessen“ der RCT-Forschung (sogenannte „Efficacy“-Studien) durch Seligman (1997) stammt nämlich bereits aus einem eher kritischen Beitrag, obwohl er zunächst viele Jahre selbst vehement den „Goldstandard“ in der Psychotherapie vertreten hatte. Seligman führt im Beitrag aber aus, wie ihn die Beteiligung an der „consumer report study“ von 1995 – einer großen amerikanischen Verbraucher-Umfrage über die Wirkungen von Behandlungen im realen psychotherapeutischen Feld – eines Besseren belehrt habe. Selbst bei Einräumung aller methodischer Probleme (von denen allerdings auch die Laborforschung, wie alle Forschung, nicht frei ist) wichen die Ergebnisse dieser Studie sehr deutlich von den aus Laboruntersuchungen abgeleiteten Behauptungen über die Wirksamkeitsunterschiede zwischen den psychotherapeutischen Ansätzen ab. Seligman schwor daraufhin vehement seinem bisherigen Irrtum öffentlich ab: „Durch diese Studie wurde mir klar, dass die Entscheidung, ob eine bestimmte Methode unter streng kontrollierten Bedingungen einer anderen Methode oder der Kontrollgruppe überlegen ist, etwas anderes ist als die Frage nach dem, was in der Praxis wirkt (...). Ich halte „Efficacy“-Studien seitdem nicht mehr für die einzige, ja nicht einmal für die beste Möglichkeit, um festzustellen, welche Methoden in der Praxis tatsächlich wirksam sind. Ich bin zu der Überzeugung gekommen, dass „Effectiveness“-Studien, mit denen die Erfahrungen der Patienten unter den realen Bedingungen in der Praxis erhoben werden, eine brauchbare und glaubwürdige „empirische Validierung“ von Psychotherapie und Medikation ermöglichen“ (Seligman 1997, S. 271). Wie bereits erwähnt, wurde Seligman ein Jahr nach dieser (in etlichen Publikationen betonten) Kehrtwendung seiner bisherigen Position zum Präsidenten der mächtigen APA gewählt. Dies zeigt, dass in den USA heterogenere Positionen einflussreich vertreten werden, als oft in jenen Diskursen behauptet wird, welche von „Wissenschaft“ sprechen und dabei allen anderen die Ansicht einer bestimmten Therapieschule aufdrücken wollen.

Doch zwischen RCT-Studien einerseits und dem „Effectiveness“-Ansatz von Seligman gibt es ein breites Spektrum an weiteren Möglichkeiten, stichhaltige Aussagen über die Wirksamkeit von Psychotherapien zu gewinnen. So wurde in der BRD – um nur ein Beispiel zu nennen – nur dann jemand als „Gesprächspsychotherapeut“ zertifiziert, der nach umfassender Ausbildung mindestens fünf erfolgreiche Psychotherapien ausführlich dokumentiert hatte. Dabei wurde die therapeutische Vorgehensweise mit Bandausschnitten gegenüber der Supervisionsgruppe und deren Leiter belegt und der Fortschritt anhand diagnostischer Tests und weiterer Unterlagen so nachgewiesen, dass diese Gruppe in ausführlicher Evaluation den Erfolg der Behandlung bewerten konnte. Allein diese erfolgreichen Zertifizierungsfälle in diesem einen Ansatz der Humanistischen Psychotherapie liegen bei über zehntausend Fällen.

Wenn solche umfassenden Nachweise der Wirksamkeit von manchen Gruppen und Gremien (wie dem deutschen Wissenschaftlichen Beirat Psychotherapie [WBP] und dem Gemeinsamen Bundesausschuss [G-BA]) bei ihren Werturteilen über Psychotherapie als „irrelevant“ ignoriert werden, während Belegen aus oft nur kleinen Gruppen von 50 oder weniger Patienten unter genormten Bedingungen allein „wissenschaftliche“ Beweisfähigkeit zugesprochen wird, zeigt dies eine Mystifikation von bestimmten methodischen Zugängen und einen fast magischen Glauben an die Bedeutung statistischer Zahlen, die leider nicht untypisch für manche Diskurse in der Psychotherapieforschung ist, und die z.B. Kromrey (1999) zu Recht scharf kritisierte.

Nicht einmal die Erforschung der Wirksamkeit von Psychotherapie kann sich daher so einseitig auf den RCT-Ansatz stützen, wie dies gegenwärtig geschieht – denn dabei wird zu sehr die (berechtigte) Sichtweise *einer* Therapieschule (Verhaltenstherapie) allen anderen inadäquat übergestülpt. Mindestens ebenso problematisch ist aber, dass die anfangs skizzierten Forschungsfragen, die für psychotherapeutisches Handeln wichtig sind, – und die daher Praktiker weit mehr interessieren als die „Rechtfertigungsforschung“ gegenüber der Konkurrenz aus Pharmaindustrie und Medizin – aufgrund dieser Einseitigkeit stark vernachlässigt werden. Es ist sowohl aus wissenschaftlicher Sicht wie auch im Hinblick auf die Interessen der Therapeuten und ihrer Patienten überaus bedenklich, welche ungeheuren Ressourcen in den letzten Jahrzehnten für den Nachweis aufgewendet wurden, dass ein bestimmter Ansatz mindestens so wirksam ist wie die Konkurrenz (d. h. andere Ansätze und/oder Psychopharmaka). Und wie wenig vergleichsweise den vielfältigen Fragen zur Wirkungsweise von zahlreichen Aspekten im psychotherapeutischen Prozess nachgegangen wurde.

## Literatur

Chambless, D.L. & Hollon, S.D. (1998). Defining empirical supported therapies. *Journal of Consulting and Clinical Psychology*, 66(1), 7-18.

Google. (2015). *Google Deutschland*. Zugriff am 16.08.2015. Verfügbar unter <https://www.google.de>

Kriz, J. (1996). Grundfragen der Forschungs- und Wissenschaftsmethodik. In R. Hutterer-Krisch, J. Kriz, E. Parfy, U. Margreiter, W. Schmetterer & G. Schwentner, *Psychotherapie als Wissenschaft – Fragen der Ethik* (S. 15-160). (Band 5 der „Serie Psychotherapie“, Hrsg. G. Sonneck). Wien: Facultas.

Kriz, J. (2003). Gutachten über den Begriff der Wissenschaftlichkeit in der Psychotherapie. *Punktum. Zeitschrift des SBAP*, Sonderdruck, Mai 2003. Zürich: Schweizerischer Berufsverband für Angewandte Psychologie.

Kriz, J. (2004). Methodologische Aspekte von „Wissenschaftlichkeit“ in der Psychotherapieforschung. *Psychotherapie und Sozialwissenschaft*, 6(1), 2004, 6-31.

- Kriz, J. (2010). Was leistet das Psychologiestudium und was fehlt ihm im Hinblick auf eine psychotherapeutische Ausbildung und Tätigkeit? *Psychotherapeutenjournal*, 9(2), 130-140.
- Kriz, J. (2012). Evidenzbasiertes Entscheiden – zwischen guter Theorie und missbräuchlicher Praxis. *Kerbe - Forum für Sozialpsychiatrie*, 30(1), 22-24.
- Kromrey, H. (1999). Von den Problemen anwendungsorientierter Sozialforschung und den Gefahren methodischer Halbbildung. *Sozialwissenschaften und Berufspraxis*, 22(1), 58-77.
- Lakotta, B. (2009). SPIEGEL-GESPRÄCH - "Weltformel der Seele": Der Psychiater Florian Holsboer über seine Suche nach den Genen der Depression, die molekularbiologischen Ursachen psychischer Krankheiten und die Therapie noch nicht ausgebrochener Hirnleiden. *DER SPIEGEL* 18/2009, 136-140. Auch in *SPIEGEL ONLINE*, 27.04.2009. Zugriff am 26.8.2015. Verfügbar unter <http://www.spiegel.de/spiegel/a-621733.html>
- Plaum, E. (1999). Weshalb fährt der IC 781 am 26. Geburtstag von Sabine M. um 13.49 Uhr mit einer Geschwindigkeit von 82,5 km/h durch den Bahnhof Eichstätt? Oder: Das Elend mit der Suche nach reinen Wirkfaktoren in einer hochkomplexen Realität. *Gestalt Theory*, 21(3), 191-207.
- Reich, W. (1972). *Die Entdeckung des Orgons I*. Frankfurt: Fischer.
- Seligman, M.E.P. (1997): Die Effektivität von Psychotherapie. Die Consumer Reports-Studie. *Integrative Therapie*, 22(4), 264-288. (Originalarbeit erschienen 1995 in *American Psychologist*, 50(12), 965-974).
- Zarbock, G. (2010). Der Master in Psychologie als optimale Basisqualifikation für Psychotherapeuten. *Psychotherapeutenjournal*, 9(2), 125-129.

### Angaben zum Autor

**Jürgen Kriz**, Emeritus für Psychotherapie und Klinische Psychologie an der Universität Osnabrück; Gastprofessuren in Wien, Zürich, Riga, Moskau und den USA. Psychologischer Psychotherapeut, Ehrenmitglied mehrerer therapeutischer Fachverbände; 20 Bücher und 250 Fachpublikationen, 2005 Viktor-Frankl-Preis der Stadt Wien, 2014 Award der AGHP (Arbeitsgemeinschaft Humanistische Psychotherapie).

**Kontakt:** [kriz@uos.de](mailto:kriz@uos.de)

### Zitationsempfehlung

Kriz, J. (2015). Psychotherapieforschung – und ihre Beschränkung durch einen schulenspezifischen Bias. *Resonanzen. E-Journal für biopsychosoziale Dialoge in Psychotherapie, Supervision und Beratung*, 3(2), 101-113. Zugriff am 15.11.2015. Verfügbar unter <http://www.resonanzen-journal.org>