



Wirkfaktoren und metaanalytische Fundierung der Wirksamkeit psychotherapeutischer Interventionen

Mag. Christian Frauscher, BSc

Publikation Mai 2020



Abstract

Die vorliegende Arbeit beschäftigt sich mit der Frage nach Wirksamkeitsfaktoren der Psychotherapie und der metaanalytischen Fundierung der Wirksamkeit von Psychotherapie. Es scheint unumstritten, dass Psychotherapie wirkt, die Frage wie sie wirkt, ob unspezifisch oder spezifisch, kann jedoch noch nicht eindeutig geklärt werden. Das wissenschaftliche Interesse, das die Psychotherapie als angewandte Disziplin auf sich gezogen hat, resultierte in zahllosen Studien zu deren Wirksamkeit, die spätestens seit den 1970er-Jahren systematisch zu Metaanalysen und Reviews zusammengefasst werden. Die Befundlage muss aus heutiger Sicht als uneinheitlich beurteilt werden: bis in die 2000er-Jahre hinein zeigte die Psychotherapie bei verschiedenen Störungsbildern hohe bis sehr hohe Effektstärken, neuere Analysen kommen durch verfeinerte statistische Methoden und sukzessives Korrigieren von Verzerrungen nur mehr zu geringen bis mittleren Effektstärken. Verschiedene Wirkfaktoren werden dargestellt, metaanalytische Befunde zur Wirksamkeit der Psychotherapie anhand des Störungsbildes der Depression referiert und Schlussfolgerungen für Wissenschaft und Praxis gezogen.

Abstract

The present work deals with the question of effectiveness factors of psychotherapy and the meta-analytical foundation of the effectiveness of psychotherapy. It seems undisputed that psychotherapy works, but the question of how it works, whether unspecific or specific, cannot yet be clearly clarified. The scientific interest that psychotherapy has attracted as an applied discipline has resulted in numerous studies on its effectiveness, which have been systematically summarized into meta-analyses and reviews since the 1970s at the latest. From today's perspective, the findings have to be judged as inconsistent: until the 2000s, psychotherapy showed high to very high effect sizes for various disorders, more recent analyses show only low to medium effect sizes due to refined statistical methods and the gradual correction of distortions. Various impact factors are presented, meta-analytical findings on the effectiveness of psychotherapy are presented based on the mental disorder of depression, and conclusions are drawn for science and practice.



Inhaltsverzeichnis

1	Einleitung	4
2	Wirksamkeit und Wirkfaktoren der Psychotherapie	5
2.1	Wirkt Psychotherapie?	5
2.1.1	Was wir sicher wissen	6
2.1.2	Was wir mehr oder weniger wissen.....	7
2.1.3	Wo noch Unklarheit besteht.....	8
2.2	Wirkfaktoren der Psychotherapie	8
2.3	Die zwei Hemisphären der Psychotherapieforschung.....	9
2.3.1	Allgemeine und spezifische Wirkfaktoren der Psychotherapie	10
2.3.2	Die 5 Wirkfaktoren der Psychotherapie nach Klaus Grawe	13
3	Das Verfahren der Metaanalyse	15
3.1	Wozu Metaanalysen?.....	15
3.2	Die Wirksamkeit der Psychotherapie als metaanalytisches Forschungsfeld.....	16
3.3	Statistische Maßzahlen und die Frage nach der Validität	16
3.3.1	„Efficacy“ und „Effectiveness“ – Interne und externe Validität.....	17
4	Metaanalysen in der Psychotherapie	18
4.1	Metaanalytische Fundierung der Psychotherapie bei Depressionen.....	21
4.1.1	Efficacy Studies – Studien im Labor	21
4.1.2	Effectiveness Studies – Studien im Feld.....	24
4.2	Hatte Eysenck doch recht?	25
5	Diskussion	29
6	Literaturverzeichnis	31
7	Autoreninformation	35



Einleitung¹

Die vorliegende Arbeit beschäftigt sich mit der Frage, warum Psychotherapie wirkt und welche Wirkfaktoren angenommen werden sowie mit der metaanalytischen Fundierung der postulierten Wirksamkeit. Die Gliederung erfolgt inhaltlich in zwei große Blöcke.

Zum einen beschäftigt sich das Kapitel 2 mit der Wirksamkeit und den Wirkfaktoren der Psychotherapie. Dabei wird der Frage nachgegangen, ob Psychotherapie tatsächlich „wirkt“, was wir über die Wirksamkeit wissen oder zu wissen glauben und wo noch deutliche Wissenslücken vorhanden sind. Es folgt eine Beschäftigung mit konkreten Wirkfaktoren, die an das therapeutische Vorgehen angelehnt sind, und somit hohe Praxisrelevanz besitzen.

Zum anderen, in Kapitel 4, wird die metaanalytische Befundlage der Psychotherapieforschung zur Wirksamkeit der Psychotherapie anhand ausgewählter metaanalytischer Studien dargestellt. Besonders wichtig ist hier die Unterscheidung von Labor- und Feldstudien, die eng mit der Frage nach der Validität verknüpft sind.

Zwischen diesen zwei thematischen Blöcken ist das Kapitel 3 eingeschoben, das sich mit theoretischen Grundlagen der Metaanalyse auseinandersetzt. Dies war notwendig, um das theoretische Rüstzeug zum Verständnis der dargestellten Metaanalysen zu legen. Dabei geht es um statistische Kennzahlen und deren Interpretation, Fragen nach der Validität von Studien und den Unterschied von Labor- und Feldstudien im Rahmen psychotherapeutischer Forschung.

Die Diskussion in Kapitel 5 rundet die Darstellung ab. Der Autor referiert hier noch einmal in gebotener Kürze die wesentlichen methodischen Schwächen vor allem älterer Metaanalysen, stellt Forderungen an die wissenschaftliche Durchführung von Studien und appelliert, den Schulenstreit innerhalb der Psychotherapie zu überwinden und stattdessen die Ergebnisse wissenschaftlich fundierter, aussagekräftiger Forschung als Grundlage des therapeutischen Argumentierens und Handelns heranzuziehen.

¹ Werden Personenbezeichnungen aus Gründen der besseren Lesbarkeit lediglich in der männlichen oder weiblichen Form verwendet, so schließt dies das jeweils andere Geschlecht mit ein.



1 Wirksamkeit und Wirkfaktoren der Psychotherapie

1.1 Wirkt Psychotherapie?

1952 erschien im *Journal of Consulting Psychology* ein einflussreicher Artikel von Hans Jürgen Eysenck mit dem Titel „*The effects of psychotherapy*“, der die Gemeinschaft der Psychotherapeuten erschütterte (Eysenck, 1992). Der Autor untersuchte in seiner Arbeit die Wirksamkeit von psychotherapeutischen Behandlungen. Die Vorgehensweise war denkbar einfach und unterscheidet sich deutlich von der Methodik heutiger Metaanalysen, dennoch wird die Eysenck-Arbeit als Quasi-Vorläufer der ersten „richtigen“ Metaanalysen der Psychotherapieforschung gewertet.

Eysenck zog insgesamt 19 psychotherapeutische Outcome-Studien heran, die er in zwei Behandlungskategorien einteilte: Psychoanalytisch und eklektisch, also die Kombination von verschiedenen therapeutischen Methoden. Dabei wurden 5 Studien der psychoanalytischen Tradition zugeordnet, die restlichen Studien dem eklektischen Typ. Den Erfolg jeder Studie bewertete er erneut nach kategorialer Klassifizierung: alle Teilnehmer jeder Studie wurden einer von 4 Kategorien zugeordnet, entweder (a) *Cured, or much improved*, (b) *Improved*, (c) *Slightly improved* oder (d) *Not improved, died, discontinued treatment, etc.* Alle inkludierten Studien beschäftigten sich mit der Pathologie der „Neurose“, ohne weiter zu klären, was darunter genau zu verstehen ist; Schizophrenie, manisch-depressive Störung oder paranoide Erkrankungen wurden auf jeden Fall ausgeschlossen. Die eigentliche „statistische Analyse“ beschränkte sich darauf, Prozentzahlen zu berechnen, anhand derer die Erfolgsquote der Psychotherapie bestimmt wurde. Das Ergebnis der Auszählung sieht wie folgt aus: Von den Patienten, die mittels Psychoanalyse behandelt wurden, besserten sich die Beschwerden bei 44%, bei diejenigen, die mit eklektischen psychotherapeutischen Methoden behandelt worden waren, war die Erfolgsquote bei 64%; bei denjenigen Patienten, die nur pflegerisch oder vom Allgemeinmediziner behandelt wurden, erreichten 72% eine Verbesserung. Es zeigt sich daher ein negativer Zusammenhang zwischen der Besserung der Beschwerden und der Psychotherapie: je mehr Psychotherapie, desto geringer der Grad der Genesung. Eysenck (1992, S. 662) zieht folgende Schlussfolgerung:

A survey was made of reports on the improvement of neurotic patients after psychotherapy, and the results compared with the best available estimates of recovery without benefit of such therapy. The figures fail to support the hypothesis that psychotherapy facilitates recovery from neurotic disorder.

Nach Pfammatter und Tschacher (2012) ist die Befundlage zur Wirksamkeit der Psychotherapie jedoch eindeutig: *Psychotherapie wirkt*. Je nach herangezogener Metaanalyse erreicht die Psychotherapie gegenüber unbehandelten Kontrollgruppen eine allgemeine Effektstärke von $d = 0.75$ bis 1.11 (Grissom, 1996; Wampold et al., 1997). Im



Vergleich zu Placebobedingungen finden sich Effektstärken von 0.42 (Prioleau, Murdock & Brody, 1983) bis 0.58 (Grissom, 1996). Die Effektstärken spontaner Remission liegen zwischen 0.10 und 0.20 (Grawe, 1995). Psychotherapie zeigt demnach eine mindestens vierfach höhere Besserungsrate als Spontanheilung und eine eineinhalbmal bis doppelt so starke Wirkung wie Placebo.

Die gute Wirkung von Psychotherapie bestätigt sich auch, wenn statt der Höhe von Effektstärken oder der zugrundeliegenden statistischen Signifikanz die klinische Relevanz der Verbesserungen zum Gradmesser genommen wird. Nach Lambert und Ogles (2004, zit. nach Pfammatter & Tschacher, 2012) zeigen eine Reihe von klinischen Studien, dass je nach Behandlungsdauer 50 bis 75% der Patienten nach der Therapie in einem psychopathologisch unauffälligen Bereich liegen. Darüber hinaus zeigen auch Metaanalysen, dass die Wirkung der Psychotherapie in vielen Fällen weit über die Dauer der Behandlung hinaus anhält.

Zu ähnlichen Einschätzungen kommt auch Wampold (2015) in seinem Werk „*The Great Psychotherapy Debate*“. Darin beschreibt der Autor systematisch, was hinsichtlich der Wirksamkeit von Psychotherapie als gesichert gelten kann, was wahrscheinlich sicher ist, und wo noch Unklarheit besteht. Die in diesem Zusammenhang wichtigsten Punkte sollen in den folgenden drei Unterkapiteln wiedergegeben werden.

1.1.1 Was wir sicher wissen

- Psychotherapie ist bemerkenswert effektiv. Die Effekte der Psychotherapie sind größer als diejenigen vieler medizinischer Interventionen. Dazu gehören Impfstoffe, die meisten Interventionen in der Kardiologie und Behandlungen gegen Asthma, wobei neben hohen Kosten auch signifikante Nebenwirkungen zum Tragen kommen. Psychotherapie ist bei psychischen Störungen genauso effektiv wie Medikamente, allerdings ohne Nebenwirkungen. Darüber hinaus halten die Effekte der Psychotherapie länger an als die der Medikation, das heißt, es gibt weniger Rückfälle nach Beendigung der psychotherapeutischen Maßnahme.
- In der Praxis erzielen Psychotherapeuten Ergebnisse, die denen von randomisierten klinischen Tests vergleichbar sind, oftmals sogar in weniger Sitzungen.
- Wenn Klienten/Patienten die Möglichkeit erhalten, soviel Therapie in Anspruch zu nehmen, wie sie sich wünschen, dann beenden sie automatisch die Therapie, wenn sie sie nicht mehr benötigen. Manche Patienten machen rasche Fortschritte, andere brauchen mehr Zeit, aber Letztere benötigen nicht außergewöhnlich mehr Therapiesitzungen, in der Regel etwa 7 bis 9 Sitzungen (!).
- Therapeuten unterscheiden sich in ihrer Effektivität: manche erreichen konstant bessere Ergebnisse mit ihren Klienten als andere, sowohl im klinischen Versuchssetting als auch in der Praxis.
- Die Therapietreue, Empathie, Erwartungen, Psychoedukation und andere „unspezifische“ Faktoren sind fest mit dem Ergebnis der Psychotherapie verknüpft.



- Darüber hinaus erzielen diejenigen Therapeuten bessere Ergebnisse, die ein „Arbeitsbündnis“ mit den Patienten eingehen können, die hochentwickelte interpersonale Fertigkeiten besitzen, die Effektivität ihrer Behandlungen kritisch reflektieren und Anstrengungen unternehmen, um ihre Behandlungen stets zu verbessern.
- Die Verfügbarkeit von Informationen über den Fortschritt des Klienten verbessert die Qualität der Psychotherapie. Informationen über den Fortschritt des Klienten beugen Behandlungsfehlern vor und heben die Wahrscheinlichkeit, dass der Klient Fortschritte macht. Ein Klienten-Monitoring verbessert die Qualität der Behandlung („*practice-based evidence*“).

1.1.2 Was wir mehr oder weniger wissen

- Therapeuten werden im Laufe der Zeit durch mehr Erfahrung *nicht* besser. Es scheint, dass sich Therapeuten über den Verlauf einer professionellen Karriere hinweg nicht verbessern, wenn Verbesserung als „*to achieve better outcomes*“ definiert wird.
- Die Bewertung der Kompetenz eines Therapeuten hinsichtlich einer umschriebenen Behandlungsform korreliert nicht mit besseren Behandlungsergebnissen. Anders gesagt, Bewertungen von Therapeuten durch Videoanalysen mit Fokus auf theoriespezifischen Verhaltensbeobachtungen differenzieren nicht zuverlässig zwischen effektiveren und weniger effektiven Therapeuten. Diese Beobachtung kommt deshalb zustande, weil sich erstens die benutzten Rating-Skalen auf die Kompetenz der Umsetzung therapiespezifischer Vorgehensweisen beschränken (zum Beispiel Fertigkeiten der kognitiven Verhaltenstherapie) und nicht auf unspezifische Wirkfaktoren beziehen. Zweitens haben die Klienten einen Einfluss darauf, wie kompetent Therapeuten erscheinen; interpersonal aggressive Klienten beispielsweise lassen Therapeuten relativ inkompetent wirken.
- Die Adhärenz (Therapietreue) einem Behandlungsprotokoll gegenüber steht in keinerlei Verbindung zum Ergebnis der Behandlung. Das heißt, dass Therapeuten, die sich streng an das Behandlungsregiment halten beziehungsweise sich strikt daran orientieren nicht mehr oder weniger erfolgreich sind als jene, die davon abweichen. Es scheint eher so zu sein, dass Therapeuten, die sich strikt an die Methodik halten, egal, wie der Klient darauf reagiert, eher schlechtere Behandlungsergebnisse erzielen als diejenigen, die weniger strikt vorgehen. Die besten Behandlungserfolge erzielen demnach Therapeuten, die eine flexible Behandlung verfolgen.
- Unstrukturierte und unüberlegte Behandlungen sind weniger effektiv als Behandlungen, die überlegte Handlungen und Schritte setzen, die sich an den Bedürfnissen des Klienten orientieren. Zwei Aspekte, die besonders wichtig sind, um gute Behandlungserfolge zu erzielen, sind zum einen, dem Klienten psychoedukativ Erklärungen für seinen Disstress zu liefern und ein Störungsmodell zu vermitteln, zum



anderen diejenigen Behandlungsschritte so erklären zu können, dass der Klient versteht, wie ihm die Vorgehensweise helfen soll.

1.1.3 Wo noch Unklarheit besteht

- Sind bestimmte Therapieformen effektiver als andere? Meistens wird postuliert, dass die kognitive Verhaltenstherapie anderen Therapieformen überlegen ist. In der Regel sind die Effekte für die Überlegenheit der kognitiven Verhaltenstherapie klein und begrenzt auf umschriebene Symptome, die von der Therapie angezielt werden. Dies ist noch keine Überlegenheitsaussage bezüglich der gesamten psychotherapeutischen Behandlung. Es scheint so, dass es keine Unterschiede zwischen den verschiedenen Behandlungsformen beziehungsweise psychotherapeutischen Schulen hinsichtlich der Lebensqualität, dem Wohlbefinden oder dem generellen psychologischen Funktionsniveau gibt.
- Befürworter von spezifischen Behandlungseffekten behaupten, dass die Behandlungen effektiv sind, weil sie spezifische Methoden enthalten und zur Anwendung bringen. Wenn in Studien die spezifischen Methodeneffekte jedoch herauspartialisiert werden, reduziert sich die Behandlungseffektivität kaum. Daher wird angenommen, dass spezifische Therapiebausteine nicht die entscheidenden Erfolgsfaktoren sind.
- Es ist davon auszugehen, dass evidenzbasierte Behandlungsformen die Qualität der psychotherapeutischen Versorgung verbessern. Basierend auf der Annahme, dass Behandlungen ungefähr gleich effektiv sind, kann die Schlussfolgerung gezogen werden, dass die Behandlungen auch in klinischen Settings (und nicht nur in klinischen Trials) erfolgreich sind und dass die Art und Weise, wie die Versorgung stattfindet wichtiger ist als die Fokussierung auf eine bestimmte Methode.

1.2 Wirkfaktoren der Psychotherapie

Über verschiedene Störungsbilder und Ergebnisvariablen hinweg finden Metaanalysen nur geringfügige Wirksamkeitsunterschiede zwischen verschiedenen Therapieformen. Grissom (1996) ermittelte bei der Zusammenführung von 46 Metaanalysen einen durchschnittlichen Effektstärkeunterschied von 0.23 und Luborsky et al. (2003) berichten auf der Grundlage von 18 Metaanalysen eine Abweichung von 0.20. Auch die allein auf die Ergebnisse vergleichender Psychotherapiestudien gerichtete Metaanalyse von Wampold et al. (1997) deutet mit einer durchschnittlichen Differenz der Effektstärken von 0.18 auf nur geringe Wirksamkeitsunterschiede hin. Die Befundlage spricht also dafür, dass mehr von allgemeiner als spezifischer Wirksamkeit der Psychotherapie gesprochen werden kann. Im Folgenden sollen die beiden wissenschaftlichen Positionen zu „allgemeinen“ versus „spezifischen“ Wirkfaktoren dargestellt werden.



1.3 Die zwei Hemisphären der Psychotherapieforschung

Ein Teil der Psychotherapieforscher hat das Scheitern, bedeutende Unterschiede in der allgemeinen Wirkung verschiedener Verfahren aufzuzeigen mit der Annahme der Nullhypothese reagiert und sind von der Gültigkeit des sogenannten „Dodo-Verdikts“ überzeugt. Der Schiedsspruch des Dodovogels aus dem Märchen Alice im Wunderland, demnach alle gewonnen haben und deshalb einen Preis verdienen, ist in der Psychotherapieforschung zur vielzitierten Metapher für die vergleichbare Wirksamkeit verschiedener Psychotherapieformen geworden. Das Dodo-Verdikt wird auch als Äquivalenzparadoxon der Psychotherapie bezeichnet, da angesichts der Unterschiedlichkeit der Krankheits- und Veränderungstheorien sowie unterschiedlicher psychotherapeutischer Schulen und Methoden vergleichbare Wirksamkeiten erzielt werden (Pfammatter & Tschacher, 2012).

Den Anhängern des Dodo-Verdikts steht eine Fraktion von Forschern gegenüber, die dessen Gültigkeit anzweifeln oder zumindest dessen Geltungsbereich einschränken. Werden (Meta-)Analysen nämlich störungsspezifisch betrachtet, zeigen sich regelmäßig klare Wirksamkeitsunterschiede zwischen verschiedenen Psychotherapiemethoden. Diese Befunde (DeRubeis, Brotman & Gibbons, 2005) widersprechen nicht nur der Annahme einer äquivalenten Wirksamkeit, sondern bedeuten gleichzeitig, dass bei bestimmten Störungsbildern spezifische Wirkfaktoren hervortreten.

Sowohl das Fehlen von Wirksamkeitsunterschieden als auch das Auftreten solcher werden von den jeweiligen Vertretern der entsprechenden Sichtweise als methodische Artefakte beiseitegeschoben. Die immer wieder beobachtete Überlegenheit kognitiver Verhaltenstherapie wird mit der Präferenz der Untersucher für diese Therapierichtung, dem sogenannten *Allegiance*-Effekt (Luborsky et al., 1999), mit der stärkeren Reaktivität von Instrumenten zur Verhaltensmessung oder mit der Berücksichtigung von Therapieanalogiestudien an subklinischen Untersuchungsteilnehmern bei der Evaluation verhaltenstherapeutischer Ansätze begründet (Lambert & Ogles, 2004, zit. nach Pfammatter & Tschacher, 2012). Es gibt aber auch eine Reihe methodischer Gründe, das Dodo-Verdikt in Frage zu stellen: viele Psychotherapievergleichsstudien weisen eine zu geringe Teststärke auf, die Aggregation der Effekte über verschiedene Störungen und Ergebnismaße in umfassenden Metaanalysen oder die fehlende Beobachtung von Interaktionseffekten (Aptitude by Treatment-Interaktionen), das heißt, der Interaktionseffekte zwischen individuellen Personmerkmalen und spezifischen therapeutischen Vorgehensweisen, können tatsächlich bestehende Wirkungsunterschiede verdecken und so den Mythos der homogenen Therapieergebnisse begründen (Beutler et al., 2004, zit. nach Pfammatter & Tschacher, 2012). Ein weiteres Gegenargument zum Dodo-Verdikt ist das so genannte Äquifinalitätsprinzip, also die Frage nach den Wirkmechanismen. Zwar erzielen verschiedene



Psychotherapieverfahren äquivalente Ergebnisse, erreichen dies aber über unterschiedliche spezifische Wirkmechanismen (Mediatorvariablen).

1.3.1 Allgemeine und spezifische Wirkfaktoren der Psychotherapie

Als spezifische Wirkfaktoren werden die in einem Psychotherapiekonzept explizit verankerten therapeutischen Methoden oder Therapietechniken bezeichnet. Spezifische Wirkfaktoren sind einerseits Standardtechniken der verschiedenen Psychotherapieschulen wie beispielsweise Übertragungsdeutung in der psychoanalytischen Psychotherapie, die Verbalisierung emotionaler Erlebnisinhalte in der Gesprächspsychotherapie, die Technik des leeren Stuhls in der Gestalttherapie, die Reizkonfrontation in der kognitiven Verhaltenstherapie oder paradoxe Interventionen in der systemischen Therapie. Zum anderen fallen darunter auch störungsspezifische psychotherapeutische Vorgehensweisen wie zum Beispiel die Planung angenehmer Aktivitäten bei depressiven Störungen. Spezifische Wirkfaktoren werden also allein unter einem bestimmten Psychotherapiemodell oder zur Behandlung bestimmter psychischer Störungen eingesetzt. Sie werden vor dem Hintergrund der jeweiligen Veränderungstheorie der unterschiedlichen psychotherapeutischen Schulen oder basierend auf einem bestimmten Störungskonzept als kausal therapeutische Komponenten betrachtet. Diesen Annahmen liegt das medizinische Störungsbild zugrunde, das von umschriebenen Krankheitsentitäten ausgeht, die auf bestimmte, isolierbare Krankheitsursachen zurückzuführen sind und entsprechend nur durch bestimmte Therapiemethoden eingedämmt oder beseitigt werden können, die spezifisch an diese verursachenden krankhaften Prozesse anknüpfen (Tschuschke, 2005).

Unter allgemeinen Wirkfaktoren (auch: gemeinsame, kommunale oder universelle Wirkfaktoren) werden hingegen diejenigen Therapievariablen verstanden, die implizit in allen Therapieverfahren auftreten und zu therapeutischen Veränderungen beitragen. Sie sind immanente Bestandteile jeglicher Form von Psychotherapie. Sie kommen weder exklusiv bei nur einem bestimmten Verfahren vor, noch spielen sie bei der Behandlung spezifischer psychischer Störungen eine besondere Rolle. Mit der Verwendung des Begriffs „unspezifischer Wirkfaktor“ wurden die allgemeinen Wirkfaktoren häufig mit Placebowirkungen gleichgesetzt. Außerdem kann „unspezifisch“ sowohl verfahrens- als auch störungsspezifisch bedeuten. Die Placebowirkung beruht auf der Erwartung auf Besserung. Die Abgrenzung von Erwartungseffekten gegenüber den Effekten des aktiven Wirkstoffs (Verum) ist eine konstituierende Bedingung der medikamentösen Behandlung, ergibt aber im Zusammenhang mit Psychotherapie keinen Sinn: Erwartungen sind nicht reines Nebenprodukt einer Therapie, sondern werden durch tatsächliche Therapiefaktoren vermittelt, die konstruktive Veränderung der Erwartungen des Patienten verlangt nach einem glaubwürdigen Behandlungsangebot. Diese Glaubwürdigkeit ist bei einem



Scheinmedikament durch identische Form und Farbe mit dem echten Medikament leicht herzustellen. Im Gegensatz zur pharmakologischen Forschung gibt es in der Psychotherapie allerdings kein glaubwürdiges Placebo. Der Aufbau positiver Therapieerwartungen basiert wesentlich auf der Vermittlung einer glaubwürdigen therapeutischen Vorgehensweise und der Etablierung einer vertrauensvollen Therapiebeziehung. Es ist daher konzeptionell angemessener, die Induktion von Besserungserwartungen nicht als Placeboeffekt, sondern als allgemeinen Wirkfaktor zu betrachten.

Tabelle 1-1 enthält eine Zusammenstellung und Beschreibung der in der Literatur diskutierten allgemeine Wirkfaktoren in der Psychotherapie, übernommen von Pfammatter und Tschacher (2012).

Tabelle 1-1: Wirkfaktoren der Psychotherapie nach Pfammatter und Tschacher (2012, S. 71)

Allgemeine Wirkfaktoren	Beschreibung
Therapiebeziehung	Therapeut und Patient haben eine vertrauensvolle, kooperative Therapiebeziehung, die gekennzeichnet ist durch gegenseitige Verbundenheit, Übereinstimmung in den Therapiezielen und therapeutischen Aufgaben sowie, auf Seiten des Therapeuten, durch Wertschätzung und emotionale Wärme für die Person des Patienten, wohlwollende Zuwendung und aufrichtiges Interesse an dessen Problemen
Abschwächung sozialer Isolation	Der Patient erlebt eine Reduktion seiner sozialen Entfremdung und Isolation
Erklärungssystem	Der Patient erhält ein für ihn plausibles Erklärungssystem für seine Probleme, das ihm ermöglicht, seine problematischen Erfahrungen einzuordnen und ihm einen nachvollziehbaren, an das therapeutische Vorgehen geknüpften Lösungsweg aufzeigt
Besserungserwartung	Beim Patienten wird Hoffnung auf eine Besserung seiner Schwierigkeiten und die Erwartung, dass die Therapie Erfolg haben wird, aufgebaut
Veränderungsbereitschaft	Der Patient entwickelt die Bereitschaft, seine Situation und sein Verhalten zu verändern
Aktive Patiententeilnahme	Der Patient nimmt aktiv an der Therapie teil, ist engagiert
Ressourcenaktivierung	In der Therapie werden die Stärken, Fähigkeiten und Ressourcen des Patienten angesprochen
Affektives Erleben	Der Patient erlebt mit seinen Problemen verbundene Emotionen und Affekte
Freisetzung unterdrückter Emotionen (Katharsis)	Der Patient erlebt Emotionen und Affekte, die bisher unterdrückt und verdrängt waren
Problemaktualisierung	Der Patient wird mit seinen Problemen konfrontiert, erlebt diese und setzt sich damit auseinander



Desensibilisierung	Der Patient erlebt durch Habituation ein zunehmendes Nachlassen, eine Abschwächung aversiver Gefühlsreaktionen in problematischen Situationen
Korrektive emotionale Erfahrung	Der Patient macht Erfahrungen, die seinen bisherigen negativen Erfahrungen in problematischen Situationen und den damit verbundenen Erwartungen, Befürchtungen widersprechen
Achtsamkeit	Der Patient entwickelt die Fähigkeit zu nicht-bewertendem Gewahrwerden seiner Gedanken, Wahrnehmungen, emotionalen Zustände; es gelingt ihm, ganz im Hier und Jetzt zu sein und sich seiner inneren Vorgänge bewusst zu sein, ohne sie zu beurteilen
Affektregulation	Der Patient lernt, seine Emotionen und Affekte in bestimmten Situationen besser wahrzunehmen, auszudrücken und zu kontrollieren
Klärung und Einsicht	Der Patient entwickelt ein Problemverständnis, erkennt Zusammenhänge und wiederkehrende Muster in seinen Problemen
Assimilation problematischer Erfahrungen	Der Patient gleicht die problematischen Erfahrungen an seine sonstigen Wahrnehmungs- und Verarbeitungsmuster an und kann sie sich dadurch besser zu eigen machen
Kognitive Umstrukturierung und Bewältigung	Der Patient passt schrittweise (Akkommodation) seine Wahrnehmungs- und Verarbeitungsmuster an eine vom Therapeuten vorgeschlagene Problemsicht an (Internalisierung), was ihm eine Neueinschätzung und stimmigere Einordnung problematischer Erfahrungen in ein System von Wahrnehmungs- und Verarbeitungsschemata (Integration) erlaubt
Mentalisierung und Selbstreflexivität	Der Patient entwickelt die Fähigkeit, die mentalen Zustände (Stimmungen, Intentionen, Wünsche etc.) seiner Interaktionspartner angemessener wahrzunehmen und einzuschätzen (Theory of Mind)
Verhaltensregulation	Der Patient eignet sich neue instrumentelle und soziale Verhaltenskompetenzen an und erfährt dadurch eine erhöhte Handlungskontrolle
Vermittlung von Bewältigungserfahrungen	Der Patient macht die Erfahrung erfolgreicher Problembewältigung
Erhöhung der Selbstwirksamkeitserwartung	Der Patient erhöht seine Erwartung in die eigene Wirkungskraft, überzeugt sich von Möglichkeiten eigener Einflussnahme
Neue Narration des Selbst	Der Patient entwickelt eine neue, kohärentere Wahrnehmung der eigenen Person in Bezug auf die Vergangenheit, Gegenwart und Zukunft sowie in Beziehung zu seiner Umwelt

1.3.2 Die 5 Wirkfaktoren der Psychotherapie nach Klaus Grawe

Nach Grawe, Donati und Bernauer (2001) war die wissenschaftliche Befundlage zur Wirksamkeit der Psychotherapie noch im Jahr 1994, als der Erstautor, der Schweizer Psychotherapieforscher Klaus Grawe, damit begann, die Wirksamkeit der Psychotherapie systematisch zu erforschen, zunächst sehr ernüchternd. Für die überwiegende Mehrzahl der Psychotherapieformen lagen zu diesem Zeitpunkt wenige bis keine Studien vor, die die Mindestanforderungen an wissenschaftliche, methodische und statistische Qualitätsansprüche erfüllten. Gerade die psychoanalytisch orientierten Psychotherapien, die für sich in Anspruch nahmen, in der Wirksamkeit allen anderen Therapien überlegen zu sein, wiesen eine äußerst spärliche Studienlage auf, wobei auch die wissenschaftlich methodische Qualität oft unzureichend war. Am besten untersucht und validiert waren die verhaltenstherapeutischen Psychotherapieformen, die auch gute Ergebnisse vorweisen konnten, allerdings oft nur bei spezifischen Störungsbildern.

In der Zwischenzeit sind jedoch sehr viele neue Studien hinzugekommen, sodass metaanalytische Betrachtungen der Studienlage möglich wurden. Grawe hatte das Ziel, den Schulenstreit zu überwinden und Grundlagen einer allgemeinen Psychotherapie, unabhängig von der spezifischen Ausrichtung, zu schaffen. Deshalb beschäftigte er sich in seinen letzten 10 Lebensjahren bis zum Jahr 2005 mit der Entwicklung und empirischen Überprüfung der Grundlagen einer psychologischen Psychotherapie, für die er die Besonderheiten der tiefenpsychologischen und verhaltenstherapeutischen Ansätze herausarbeitete, ihre gemeinsamen Wirkfaktoren extrahierte und als gemeinsame theoretische Grundlage seine Konsistenztheorie entwickelte. Bis kurz vor seinem Tod arbeitete Grawe an der Validierung von fünf Wirkfaktoren, die therapieschulenübergreifend notwendige Voraussetzungen für das Gelingen einer Psychotherapie sind (Dreher, 2012). Die fünf Wirkfaktoren werden in Tabelle 1-2 dargelegt.

Tabelle 1-2

Die 5 Wirkfaktoren der Psychotherapie nach Grawe (2000)

Wirkfaktor	Beschreibung
Therapeutische Beziehung	Qualität der Beziehung zwischen Patient und Therapeut
Ressourcenaktivierung	Bereits vorhandene Ressourcen explorieren und nutzbar machen
Problemaktualisierung	Probleme im Hier und Jetzt erfahrbar machen
Problembewältigung	Konkrete Maßnahmen zur Problembewältigung erarbeiten
Motivationale Klärung	Das eigene Denken, Handeln und Fühlen wahrnehmen und reflektieren



Therapeutische Beziehung. Die Qualität der Beziehung zwischen dem Psychotherapeuten und dem Patienten, also ob die Chemie zwischen beiden stimmt, trägt wesentlich zum Therapieergebnis bei. Je besser die Qualität der Beziehung ist, desto besser ist das Therapieergebnis.

Ressourcenaktivierung. Nicht der Therapeut verändert den Patienten, sondern er ist selbst für seine Veränderung zuständig. Im Rahmen der Ressourcenaktivierung werden die beim Patienten bereits vorhandenen Ressourcen nutzbar gemacht. Dabei wird die Aufmerksamkeit auf das gerichtet, was der Patient bereits erreicht hat, was gut funktioniert, welche Stärken und Fähigkeiten er besitzt, auf welche Erfahrungen er zurückgreifen kann und was er bereits alles überwunden hat. Die Eigenarten, die die Patienten in die Therapie mitbringen, werden als positive Ressource für das therapeutische Vorgehen genutzt. Das betrifft vorhandene motivationale Bereitschaften, Fähigkeiten und Interessen des Patienten. In der Verhaltenstherapie ist ein mögliches Vorgehen die sogenannten „lösungsorientierte Gesprächsführung“. Die Ressourcenperspektive bestimmt das *Wie* der Veränderung, wohingegen die Problemperspektive das *Was* der Veränderung beschreibt.

Problemaktualisierung. „Aktualisierung“ beschreibt den Prozess der unmittelbaren Erfahrung, das Erleben eines Problems im Hier und Jetzt. Das bloße Gespräch über Erleben und Verhalten, ohne dass es im Moment prozessual abläuft, ist nicht in der Lage, Veränderungen herbeizuführen. Es kann aber benutzt werden, um gewünschte Veränderungen vorzubereiten. Damit diese Veränderungen jedoch auch tatsächlich eintreten, braucht es einen Moment des unmittelbaren Erlebens. In der Verhaltenstherapie (VT) erreicht man dies mittels Konfrontation beziehungsweise Exposition, in psychoanalytisch orientierten Therapieformen mittels Übertragung und Gegenübertragung. Auch andere Methoden wie Rollenspiele oder Visualisierungsübungen sind möglich.

Problembewältigung. Hierbei geht es um die Erarbeitung von problemspezifischen Maßnahmen beziehungsweise die konkrete Handlungsorientierung. Der Patient soll darin unterstützt werden, positive Bewältigungserfahrungen zu machen. In der VT kann dies durch die therapeutische Begleitung während der Exposition realisiert werden.

Motivationale Klärung. Der Therapeut unterstützt seinen Patienten dabei, sein Erleben und Verhalten hinsichtlich seiner unbewussten und bewussten Ziele und Bedürfnisse zu verstehen. Wie bei der Problemaktualisierung ist die unmittelbare Erfahrung wichtig. Der Patient soll sich bewusst werden, was er wahrnimmt, denkt, fühlt, wie er handeln und was er vermeiden möchte. Methoden der Wahl sind beispielsweise der Sokratische Dialog oder die 2-Stühle-Technik (Grawe, 2000; Grawe et al., 2001).



2 Das Verfahren der Metaanalyse

2.1 Wozu Metaanalysen?

Hilft Psychotherapie? Wie wirkt sich die Attraktivität von Personen auf ihre Beurteilung durch andere Personen aus? Gibt es Geschlechtsunterschiede bei Persönlichkeitseigenschaften? Zu dieser und anderen Fragestellungen der akademischen Psychologie liegen zahllose von empirischen Studien vor. Nun sind die Ergebnisse von Einzelstudien oftmals alles andere als einheitlich: in machen Studien findet man, dass Psychotherapie große Besserungen bei Depressivität erzielt, dass attraktive Personen als intelligenter eingeschätzt werden und das Männer weniger ängstlich sind als Frauen. In anderen Studien werden wiederum nur kleine Effekte berichtet und manchmal kommt man gar zu gegenteiligen Schlüssen: Psychotherapie schadet, weniger attraktive Personen werden als intelligenter eingeschätzt und Männer sind ängstlicher als Frauen. Wie kann diese Uneinheitlichkeit und Gegensätzlichkeit in der Studienlage erklärt werden? Eine Möglichkeit ist, nach der passenden „Geschichte“ hinter den Daten zu suchen, also nach bedingenden Variablen, sogenannten „Moderatoren“: Psychotherapie wirkt nur, *wenn* die Patienten eine bestimmte Vorgeschichte haben, bei den Auswirkungen der Attraktivität spielt das Alter der wahrgenommenen Personen eine Rolle und die Ängstlichkeit hängt auch vom Bildungsniveau ab. Moderatorvariablen können tatsächlich eine große Rolle spielen, aber eine große Variation bei empirischen Ergebnissen kann man rein durch den Zufall (also den Stichprobenfehler) erwarten, wenn tatsächlich immer derselbe Populationseffekt untersucht wird aber unterschiedliche Stichproben gezogen werden. Man sollte daher versuchen, die Ergebnisse zu quantifizieren und zunächst bestimmen, ob ein Ergebnis signifikant war oder nicht. Allerdings spiegeln Signifikanztestergebnisse nicht unbedingt die Größe beziehungsweise Stärke von Effekten wider. Hätte man beispielsweise viele Studien mit niedriger Teststärke² vorliegen, sind die Chancen für signifikante Ergebnisse auch bei substanziellen Effektgrößen sehr niedrig: man würde nur einen kleinen Anteil an signifikanten Ergebnissen erhalten und sich mit hoher Wahrscheinlichkeit in den Schlussfolgerungen irren. Ein Ausweg aus dem Dilemma der mangelnden Aussagekraft vieler Einzelstudien ist die Synthese von Forschungsergebnissen, die die in den jeweiligen Studien gefundenen Effektgrößen als Grundlage für die Analyse zu benutzen und eine so genannte „Metaanalyse“ durchzuführen (Sedlmeier & Renkewitz, 2011).

Metaanalysen fassen Daten von Primärstudien zusammen, die für ähnliche Fragestellungen erhoben wurden. Oftmals wird eine Theorie oder ein aktuell diskutiertes Problem von mehreren Forschergruppen empirisch untersucht, wobei sich die Methodik, die untersuchten Populationen, die Erhebungsinstrumente und die Datenanalysen von Studie zu

² Die Teststärke (Power) bezeichnet die Wahrscheinlichkeit, mit der ein Test einen Effekt (ein signifikantes Ergebnis) aufdeckt, falls es in der untersuchten Population tatsächlich einen Effekt gibt. Sie ist abhängig von der Anzahl der Versuchsteilnehmer (pro Gruppe), der postulierten Effektstärke sowie dem angenommenen Signifikanzniveau.

Studie mehr oder weniger stark unterscheiden. Um die Datenbasis zu vergrößern und dadurch den Grad der Zuverlässigkeit (Reliabilität) der Ergebnisse zu steigern, kann man die Daten verschiedener Untersuchungen zum gleichen Untersuchungsgegenstand zusammentragen und mittels geeigneter statistischer Verfahren erneut auswerten. Die Ergebnisse solcher Metaanalysen sind nicht nur zuverlässiger als die Ergebnisse jeder einzelnen Studie; unter Umständen können Metaanalysen auch erklären, warum die Ergebnisse einzelner Studien voneinander abweichen (Eid, Gollwitzer & Schmitt, 2017).

2.2 Die Wirksamkeit der Psychotherapie als metaanalytisches Forschungsfeld

Der Ausdruck „Metaanalyse“ geht nach allgemeiner Auffassung auf Glass (1976) zurück, der zusammen mit einem Kollegen auch die erste Metaanalyse zur Wirksamkeit der Psychotherapie publizierte (Smith & Glass, 1977). Das Hauptergebnis einer Metaanalyse ist die präzise Schätzung des jeweiligen Populationseffekts einer Intervention. Die Schätzung erhält man durch die Berechnung eines gewichteten Mittelwerts aller vorliegenden Effektstärken. In der ersten metaanalytischen Studie von Smith und Glass (1977) war das zentrale Ergebnis der mittlere, mit den Stichprobengrößen gewichtete Unterschied zwischen den Gruppen, die Psychotherapie erhalten hatten und den jeweiligen Kontrollgruppen, ausgedrückt in Standardabweichungen, mit einem Effekt von $d = .68$. Die seinerzeit allgemein akzeptierte Interpretation war, dass der Effekt der Psychotherapie zwar nicht überwältigend hoch ist, dass aber ihre Wirkung zweifelsfrei nachgewiesen werden konnte. Statistisch wurde bei dieser Analyse vereinfachend angenommen, dass das Ausmaß der psychischen Gesundheit in den untersuchten Populationen normalverteilt ist: bei den meisten Menschen ist das Ausmaß der psychischen Gesundheit „mittelmäßig“, einigen geht es aber sehr schlecht und einigen sehr gut. Die Psychotherapie verändert nicht die Verteilung an sich, das heißt, die Glockenkurve bleibt erhalten, verschiebt sich jedoch auf der Abszisse nach rechts. Das Ausmaß der Verschiebung wurde als Effektstärke d interpretiert³ (Vgl. Abbildung von Smith & Glass, 1977, S. 754): die gefundene Effektgröße, die nichts anderes ist, als $.68$ Standardabweichungseinheiten der Normalverteilung, bedeutet, dass die Gruppe der Therapierten durchschnittlich einen Gesundheitswert erreicht, den nur 25% der Nichttherapierten erreichen (Sedlmeier & Renkewitz, 2011).

2.3 Statistische Maßzahlen und die Frage nach der Validität

Die quantitative Metaanalyse („quantitativ“ in Abgrenzung zum qualitativen, mehr subjektiv gefärbten „narrativen Review“) wird im Allgemeinen zur Entwicklung klinischer Richtlinien herangezogen und bietet die beste Grundlage für die Evaluation von Behandlungseffekten. Metaanalysen geben den aktuellsten und gründlichsten Überblick über ein Forschungsgebiet, benutzen state-of-the-art statistische Techniken (zum Beispiel gewichtete kleinste Quadrate,

³ Die Effektgröße, die Smith und Glass (1977) benutzten ist nicht gleichzusetzen mit dem Effektstärkemaß Cohen's d . Sie unterscheidet sich dadurch, dass im Nenner nicht die gepoolte Standardabweichung von Therapie- und Kontrollgruppe benutzt wurde, sondern die Standardabweichung der Kontrollgruppe (Sedlmeier und Renkewitz (2011))

Modellierung mit Zufallseffekten, Korrekturen für mangelnde Reliabilität) und sichern die akkurate Synthese von Ergebnissen vieler Einzelstudien. Metaanalysen kombinieren die Ergebnisse von Einzelstudien, indem sie eine gemeinsame Metrik verwenden, die sogenannte „Effektstärke“. Diese beschreibt, wie stark ein beobachteter Unterschied zwischen Gruppen tatsächlich ist. Dabei werden wiederum zwei Typen von Effektstärken unterschieden: Zum einen diejenigen, die Unterschiede zwischen Gruppenmittelwerten quantifizieren, wobei die Effektstärke den Unterschied zwischen den Gruppenmittelwerten der Posttestwerte, dividiert durch die gepoolte Stichprobe der beiden Gruppen, beschreibt. Die statistischen Kennwerte werden als Cohen's *d*, Hedges' *g* oder *SMD* (*standard mean difference*) bezeichnet. Die drei Kennwerte können als gleichwertig betrachtet werden und spiegeln lediglich unterschiedliche statistische Traditionen wider. Zum anderen gibt es einen Typ von Effektstärke, der den Vergleich von Gruppen in Form eines Verhältnisses oder einer Wahrscheinlichkeit beschreibt und üblicherweise in epidemiologischen Studien Verwendung findet, weniger in der psychologischen oder psychotherapeutischen Forschung. Das Quotenverhältnis (*odds ratio, OR*) wird berechnet, um die Beziehung zwischen Bedingung (treatment versus no-treatment) und einer kategorialen, binären abhängigen Variablen zu quantifizieren (zum Beispiel das Auftreten oder Nicht-Auftreten eines Rückfalls). Das Risiko-Verhältnis (*risk ratio, RR*) wird berechnet, um die Wahrscheinlichkeit des Auftretens eines Ereignisses in jeder Gruppe zu berechnen (Hunsley, J., Elliott K. & Therrien, 2013).

2.3.1 „Efficacy“ und „Effectiveness“ – Interne und externe Validität

Eine weitere wichtige Unterscheidung in der metaanalytischen psychotherapeutischen Wirksamkeitsforschung (outcome research) ist, ob als Qualitätskriterium die interne Validität (*treatment efficacy*) oder die externe Validität (*treatment effectiveness*) im Mittelpunkt des Interesses steht. Die Unterscheidung dieser beiden Validitätsformen ist zentral für die erfolgreiche Übertragung von Forschungsergebnissen auf den klinischen Anwendungskontext. Die interne Validität sagt aus, ob ein gemessener Effekt tatsächlich auf die Intervention zurückgeführt werden kann und bezieht sich auf die Merkmale des Designs, insbesondere auf die randomisierte Zuteilung von Untersuchungsteilnehmern auf Experimental- und Kontrollgruppe. Mit der Frage nach der Übertragung beziehungsweise Übertragbarkeit von experimentellen Settings auf den klinischen Kontext der praktischen Anwendung der Psychotherapie beschäftigt sich die externe Validität. Beide Validitätsformen sind wichtig für die Beurteilung der Wirksamkeit einer Therapie: ohne interne Validität kann keine Aussage getroffen werden, ob eine Intervention erfolgreich war, ohne externe Validität ist eine Übertragung auf den praktischen Kontext nur schwer möglich. Es sollten in der metaanalytischen Beurteilung der Wirksamkeit von Psychotherapie also vor allem diejenigen Studien herangezogen werden, die beide Validitäten angemessen berücksichtigen, also Studien, die sowohl strengen methodischen Standards genügen und gleichermaßen in einen klinischen Kontext eingebettet sind. Behandlungen, die unter hoch kontrollierten Bedingungen durchgeführt werden haben einen vergleichbar hohen



praktischen klinischen Effekt, wenn sie unter regulären klinischen Bedingungen stattfinden (Hunsley, J. et al., 2013).

In der Psychotherapieforschung bezieht sich „efficacy“ und „effectiveness“ aber nicht nur auf die Frage nach interner und externer Validität, sondern schlicht auf die Durchführungsbedingungen beziehungsweise das Setting der Studien. „Efficacy“ bezieht sich auf die meist an Universitäten durchgeführten, organisatorisch aufwendigen und methodisch rigorosen Therapiestudien, mit randomisierter Zuweisung von Patienten zu Behandlungs- und Kontrollgruppen, weitgehend manualisierten Behandlungen, einer festgelegten Anzahl von Behandlungen, klar operationalisierten Behandlungszielen und klar abgegrenzten Störungsbildern. Solche Studien herrschen üblicherweise an Universitäten vor, eben weil für Forschung die Universitäten und vergleichbare Institutionen zuständig sind und nicht Berufsverbände und Vereine und eine methodisch saubere Forschung in die Karrieremuster dieser Institutionen passt. Im Gegensatz zu diesen methodisch stark kontrollierten Bedingungen der universitären Studien findet reale Therapie meist mit Patienten statt, die gezielt einen bestimmten Therapeuten wählen, die ein oft nicht klar umschriebenes Störungsbild mitbringen beziehungsweise keine eindimensionale Störung aufweisen, deren Therapieziele sich im Laufe der Behandlung verändern können und deren Therapeuten sich durch große Freiheitsgrade in der detaillierten Vorgehensweise auszeichnen und nicht durch strikte Befolgung eines Manuals. Die Frage nach der Wirksamkeit ist aber auch für diese Therapien bedeutsam. Seligman, M. E. (1995) zieht, basierend auf den „Consumer Reports“, einer groß angelegten Studie in den USA, in denen die „effectiveness“ realer therapeutischer Praxis untersucht wurde, folgende Schlussfolgerungen, die sich von denen der „efficacy“-Studien unterscheiden:

- Für längere Behandlungen lässt sich ein höherer Besserungsgrad als für kurze Behandlungen belegen,
- Psychotherapie wird durch Medikamente (in der Regel) *nicht* verbessert,
- Keine Therapieform konnte allgemein oder mit Hinblick auf bestimmte Störungsbilder eine höhere Wirkung als andere nachweisen,
- Psychologen, Psychiater und Sozialarbeiter waren gleich effektiv, aber deutlich besser als Eheberater und Behandlungen durch nicht fachtherapeutische Ärzte.

Die zentrale Schlussfolgerung von Seligman, M. E. (1995) ist jedoch, dass die in efficacy-Studien benutzten Kriterien zur Abschätzung der Wirksamkeit einer Psychotherapie zu einseitig sind.

3 Metaanalysen in der Psychotherapie

Seit der ersten summarischen Analyse von Effekten der Psychotherapie von Eysenck im Jahre 1952 über die erste „echte“ Metaanalyse von Smith und Glass im Jahre 1977 hat sich die Wissenschaft sowohl in methodischer, als auch statistischer Hinsicht dramatisch

weiterentwickelt und zahllose Einzelstudien wie auch Metaanalysen zur Effektivität der Psychotherapie hervorgebracht. Mit Stand 19. April 2019 ergab eine Suche in der Wissenschaftsdatenbank PubMed⁴ zu dem Suchbegriff „psychotherapy & meta-analysis“ 3049 Treffer. Die Befundlage ist also unüberschaubar, zeigt aber auch, wieviel wissenschaftliches Interesse die Psychotherapie auf sich gezogen hat. Bei Durchsicht der Suchergebnisse wird auch deutlich, in welche Forschungsfelder die Psychotherapie als Behandlungsform hineinwirkt, von genuin psychotherapeutischen Problemstellungen bis hin zu genetischen Einflüssen und medizinischen Behandlungen, beispielsweise bei Übergewicht. Es fällt auch auf, dass viele Metaanalysen neueren Datums sind und erst in den letzten Jahren veröffentlicht wurden. Der Trend, die Effektivität von Behandlungsformen beziehungsweise Interventionen über zusammenfassende Aussagen der (quantitativen) Studienlage zu evaluieren (evidenzbasierte Therapie), ist also ungebrochen.

In den folgenden Kapiteln werden einige ausgewählte Metaanalysen zur Wirksamkeit der Psychotherapie bei Depressionen rudimentär vorgestellt und ihre zentralen Ergebnisse benannt. Einen Überblick über ausgewählte Metaanalysen für verschiedene psychische Störungen gibt

Tabelle 3-1 (nach Hunsley, J. et al., 2013, S. 29). Auffällig sind die große Variabilität in den Effektstärken und die große Inkonsistenz der Maßzahlen.

Tabelle 3-1

Metaanalytische Ergebnisse zur psychotherapeutischen Behandlung unterschiedlicher Störungsbilder, basierend auf kontrollierten Studien mit randomisierter Gruppenzuteilung (RCT)

Metaanalytische Studien	ES	NNT
Depression		
Psychotherapy vs. no treatment, effects post-treatment (Cuijpers, P., Andersson, Donker & van Straten, 2011); (Driessen, Cuijpers, Hollon & Dekker, 2010)	$d = .40$ to $.88$	2.1-4.5
Psychotherapy vs. no treatment, effects post-treatment for older adults (> 55) (Peng, Huang, Chen & Lu, 2009)	$d = .92$	2.1
Psychotherapy vs. no treatment, effects post-treatment for youth (Weisz, McCarty & Valeri, 2006)	$d = .34$	5.2
Brief CBT vs. TAU, effects post-treatment (Cape, Whittington, Buszewicz, Wallace & Underwood, 2010)	$d = .33$	5.9

⁴ <https://www.ncbi.nlm.nih.gov/pubmed/> [19-4-2019]



Bipolar Disorder

Adjunctive psychotherapy vs. medication only, relapse rate (Szentagotai & David, 2010) *OR* = .53 ---

Generalized Anxiety Disorder

CBT vs. TAU or no-treatment, effects post-treatment (Hunot, V., Churchill, Teixeira & Lima, 2007) *SMD* = .82 2.3

Social Anxiety Disorder

Psychotherapy vs. no treatment, effects post-treatment (Acarturk, Cuijpers, van Straten & Graaf, 2009) *d* = .80 2.3

CBT vs. no treatment, effects post-treatment for youth (Segool & Carlson, 2008) *d* = .86 2.2

Obsessive-Compulsive Disorder

Psychotherapy vs. no treatment, effects post-treatment (Rosa-Alcázar, Sánchez-Meca, Gómez-Conesa & Marín-Martínez, 2008) *d* = 1.08 1.9

Psychotherapy vs. no treatment, effects post-treatment for youth (Watson & Rees, 2008) *d* = 1.45 1.8

PTSD

Trauma-focused CBT vs. no treatment, effects post-treatment (Bisson et al., 2007) *d* = 1.40 to 1.70 1.7-1.8

Prolonged Exposure (CBT) vs. no treatment, effects post-treatment (Powers, Halpern, Ferenschak, Gillihan & Foa, 2010) *g* = 1.08 1.9

Trauma-focused CBT vs. no treatment, effects post-treatment for youth (Kowalik, Weller, Venter & Drachman, 2011) *d* = .33 6.0

Specific Phobias

Exposure-based CBT vs. no treatment, effects post-treatment (Wolitzky-Taylor, Horowitz, Powers & Telch, 2008) *d* = 1.05 1.9

Panic Disorder

Anmerkungen. RCT = randomized controlled trial; ES = effect size, NNT = number needed to treat (Anzahl der notwendigen Behandlungen), d = Cohen's d , g = Hedges' g , OR = odds ratio (Quotenverhältnis), SMD = standard mean difference, „---“ = keine Angabe verfügbar.

3.1 Metaanalytische Fundierung der Psychotherapie bei Depressionen

Im Folgenden wird die metaanalytische Befundlage zum Störungsbild der Depression beispielhaft vorgestellt. Für dieses spezielle Störungsbild liegen die meisten Metaanalysen hinsichtlich der Wirksamkeit einer Psychotherapie vor. Die Auswahl der Studien orientiert sich größtenteils an den Darstellungen von Hunsley, J. et al. (2013), ergänzt um wichtige aktuellere Befunde.

3.1.1 Efficacy Studies⁵ – Studien im Labor

Cuijpers, P. et al. (2011) überprüften in ihrer Metaanalyse die Gesamtwirkung der Psychotherapie und ob es Verbindungen zwischen Charakteristika der Patienten oder der Behandlung gibt, die einen Einfluss auf das Ergebnis der Depressionsbehandlung bei Erwachsenen haben. Basierend auf 147 Studien und mehreren tausend Teilnehmern wurde eine Effektstärke für den Gesamteffekt der Psychotherapie von $d = .66$ errechnet (und später auf $d = .53$ revidiert, nachdem Ausreißerwerte identifiziert und vom Datenset entfernt wurden). Verschiedene methodische Ausrichtungen von Psychotherapie (CBT⁶, IPT⁷, problem-solving therapy, non-directive supportive therapy, behavioral activation therapy) wurden in die Analyse inkludiert. Es zeigte sich, dass alle Psychotherapieformen in ihren Behandlungsergebnissen denen der Kontrollgruppen überlegen waren mit Effektstärken⁸ von $d = .57$ bis $d = .87$. Die Autoren berichteten, dass es einen kleinen, aber signifikanten Effekt zugunsten der Einzel- gegenüber der Gruppentherapie gab mit dem zusätzlichen Ergebnis, dass der Dropout im Einzelsetting geringer ist ($OR^9 = .56$: die Wahrscheinlichkeit, dass der Klient die Therapie im Einzelsetting beendet, ist etwa halb so groß wie im Gruppensetting).

Diessen et al. (2010) prüften in ihrer Metaanalyse die Fragestellung, ob die häufige Annahme, dass psychologische Interventionen bei schweren Depressionen nur einen geringen Effekt haben, belegbar ist. Die Fragestellung lautete daher, ob der Schweregrad

⁵ Bei Durchsicht verschiedener Studien fiel auf, dass es oftmals begriffliche Unschärfen gibt zwischen „psychological“ und „psychotherapeutical“, „counselling“, „intervention“ und anderen. Es ist hier davon auszugehen, dass es sich stets um psychotherapeutische Interventionen handelt (meistens CBT), wenn dies an anderer Stelle nicht explizit anders erwähnt wird.

⁶ CBT = Cognitive Behavioral Therapy

⁷ IPT = Interpersonal Psychotherapy

⁸ Es werden für alle Metaanalysen Effektstärkemaße berichtet (Cohen's d , Hedges' g und SMD), die zugehörigen Konfidenzintervalle werden mangels Relevanz nicht gesondert genannt.

⁹ OR = Odds Ratio



depressiver Symptome vor Behandlungsbeginn einen moderierenden Effekt auf die Wirksamkeit einer psychotherapeutischen Intervention hat. Die Autoren wählten 132 Studien mit insgesamt $N = 10.134$ Patienten aus, bei denen ambulante Patienten mit

schweren depressiven Störungen mit Kontrollgruppen verglichen wurden. Als statistisches Verfahren wurde die Meta-Regression eingesetzt. Die psychologische Behandlung erwies sich den Kontrollbedingungen überlegen ($d = .40$ bis $d = .88$). Die Forscher konnten kein Indiz dafür finden, dass die Vortestwerte (*depression scores*) einen Einfluss auf die gefundenen Effektstärken haben, auch nicht nach Herauspartialisieren von relevanten Variablen, die in Zusammenhang mit Studien-charakteristiken standen. Teilweise zeigte sich jedoch, dass bei *within subject*-Designs höhere Effektstärken zustande kommen, wenn Patienten unter schwereren Depressionen litten ($d = .63$) als dies bei weniger schweren Depressionen der Fall war ($d = .22$). Die Autoren ziehen aus den Ergebnissen die Schlussfolgerung, dass, entgegen der gängigen wissenschaftlichen Meinung, die psychologische Behandlung bei schweren Depressionen besser hilft als bei leichten. Eine Schwere x Treatment-Interaktion wurde nicht geprüft, für zukünftige Studien aber gefordert.

In ihrem systematischen Review untersuchen Peng et al. (2009) metaanalytisch die Wirksamkeit psychotherapeutischer Interventionen (CBT, Reminiszenz-Therapie und GPT¹⁰), deren Wirksamkeit gegen Depressionen belegt ist, für ältere Menschen, das heißt älter als 55 Jahre. Sie stützten ihre Analyse auf 14 randomisierte kontrollierte Studien. Die Ergebnisse zeigten, dass, verglichen mit Placebo, alle Formen von Psychotherapie erfolgreich in der Reduktion von Depressionswerten sind ($SMD = -.92$). Subgruppenanalysen konnten zeigen, dass CBT, Reminiszenz-Therapie und GPT alle effektiver sind als Placebo-Behandlungen, allerdings erhöhte Psychotherapie als zusätzliche Therapie zur medikamentösen Behandlung mit Antidepressiva die Wirksamkeit nicht. Zwischen CBT und Reminiszenz-Therapie zeigte sich kein Unterschied in der Effektivität der Depressionsbehandlung. Die Autoren schlussfolgern, dass verschiedene Formate von Psychotherapie wirksam in der Behandlung von Depressionen bei älteren Menschen sind, obwohl Psychotherapie die Wirksamkeit von antidepressiver Medikation nicht signifikant erhöhen kann.

Die ersten Folgewirkungen von jugendlichen Depressionen und Bedenken über die Sicherheit von Medikamenten bei Heranwachsenden veranlasste Weisz et al. (2006), sich mit der Wirksamkeit der Psychotherapie bei depressiven Jugendlichen metaanalytisch zu beschäftigen. In vorhergehenden Metaanalysen zeigten sich Effektstärken von durchschnittlich 0.99, die damit einen bedeutsamen Effekt hervorbrachten, der über den Effekten anderer Metaanalysen für ähnliche Fragestellungen steht. Die Autoren wandten strenge analytische Methoden für die größte bis dato verfügbare Stichprobe an und errechneten einen durchschnittlichen Effekt von 0.34, was signifikant unterhalb vorheriger Analysen lag. CBT schnitt nicht besser ab als alle anderen psychotherapeutischen nicht-

¹⁰ GPT = General psychotherapy



kognitiven Interventionen. Die Analysen zeigten, dass die Interventionen generell (Reduktion von Angst) und spezifisch wirken (Ausnahme: Externalisierung von Problemen), wobei die Behandlungserfolge allerdings nur kurzzeitig, jedoch nicht längerfristig anhalten. Die Schlussfolgerung der Autoren lautet, dass die psychotherapeutische Depressionsbehandlung

signifikante Effekte produziert, diese aber bescheiden in ihrer Stärke sind. Außerdem ist die Breite der Wirksamkeit der Verfahren gering, ebenso wie die zeitliche Stabilität der Effekte.

Feng et al. (2012) führten eine Metaanalyse zur kognitiven Verhaltenstherapie im Gruppensetting durch, die sich auf Daten von 32 Studien stützte (mit einer Altersspannweite von $R^{11} = 19.5-75.2$ Jahren). Die typischen Gruppengrößen waren 6-10 Personen, mit Interventionen von 8-12 Wochen Dauer mit einer Stunde Therapie wöchentlich. Zum post-treatment-Messzeitpunkt waren die Depressionslevels signifikant vermindert ($g = .40$), verglichen mit unbehandelten Kontrollgruppen, und dieser Effekt hielt auch an, gemessen nach dem 6-Monats-Follow-up ($g = .38$).

Huntley et al. (2012) führten eine systematische Prüfung der Wirksamkeit gruppenbasierter Psychotherapie für Depressionen durch. Sie führten 23 Studien zusammen, wobei die Mehrheit der Studien beträchtliche Verzerrungen aufwies. Die Analyse der gruppenbasierten kognitiv-verhaltenstherapeutischen Therapie (CBT) versus bloße Betreuung zeigte einen signifikanten Effekt zugunsten der Gruppentherapie, gemessen unmittelbar nach der Behandlung ($SMD^{12} = -.55$). Evidenz gibt es auch für den Erhalt der Behandlungseffekte, sowohl kurzfristig ($SMD = -.47$), als auch mittel- bis langfristig ($SMD = -.47$). Beim Vergleich von Gruppen-CBT mit Therapie im Einzelsetting (ebenso CBT) zeigt sich ein moderat höherer Behandlungseffekt zugunsten des Einzelsettings unmittelbar nach der Behandlung ($SMD = .38$), mittel- und langfristig gibt es allerdings keine Unterschiede mehr. Die Schlussfolgerungen der Autoren lauten daher, dass gruppenbasierte CBT bei Depressionen eine erfolgreiche Intervention ist, und den Effekt einer bloßen Betreuung übertrifft. Verhaltenstherapie im Einzelsetting ist erfolgreicher als Gruppentherapie, drei Monate nach der Therapie gibt es hier allerdings keine Unterschiede mehr. Die Qualität der Evidenz wird von den Forschern als schlecht beurteilt, außerdem bemängeln sie die begrenzte Studienlage für Psychotherapien, die nicht auf der CBT basieren.

Cape et al. (2010) beschäftigten sich in ihrer Metaanalyse mit dem Vergleich der Wirksamkeit von unterschiedlichen Arten von psychologischen Kurzzeitinterventionen, die in der Erstversorgung bei der Behandlung von Angst und Depressionen (sowie der Kombination dieser beiden, wenn eine trennscharfe Abgrenzung nicht gelingt, sogenannte „mixed disorders“) eingesetzt werden. Vergleiche wurden angestellt zu Patientengruppen, die keine psychotherapeutische Behandlung, sondern ein „*treatment as usual* (TAU)“ erhielten. Ihre Analyse beruht auf 34 Studien mit insgesamt 3962 Patienten. Meistens wurde eine

¹¹ R = Range

¹² SMD = standardized mean difference



kurzzeitige CBT eingesetzt, gefolgt von Beratung¹³ oder Verfahren zur Problemlösung. Für Angststörungen ist die CBT am effektivsten ($d = -1.06$), für Depressionen immerhin noch moderat effektiv mit einer Effektstärke von $d = -0.33$ und für mixed disorders weniger effektiv mit $d = -0.26$. Beratung erzielte bei Depressionen und mixed disorder einen Effekt

von $d = -0.32$ und Problemlösungsintervention bei Depressionen und mixed disorder $d = -0.21$. Die letztgenannten Interventionen sind also ebenso effektiv, wenn auch geringer als CBT. In der Regression, wenn die Diagnose als zusätzlicher Prädiktor in das Modell aufgenommen wird, zeigen sich keine Unterschiede in der Effektivität zwischen den Verfahren. Die Studienautoren schlussfolgern, dass Kurzzeitinterventionen in der Erstversorgung sinnvoll sind, deren Effektstärken im Vergleich zu längeren Interventionen aber niedrig sind. Einzige Ausnahme hiervon ist die kurze CBT bei Angststörungen, die auch bei kurzen Dauern bereits vergleichbare Effektstärken zeigt wie längere Interventionen.

3.1.2 Effectiveness Studies – Studien im Feld

Hans und Hiller (2013) führte eine Metaanalyse mit dem Ziel durch, die Wirksamkeit und den Dropout von individueller und Gruppentherapie (beides CBT) für Erwachsene mit der Diagnose „unipolare Depression“ in der klinischen Praxis zu erforschen. Sie führten eine Metaanalyse von 34 nicht-randomisierten Feldstudien zu ambulanter individueller und Gruppentherapie (CBT) durch. Ambulante CBT erweist sich als effektiv in der Reduktion von depressiven Symptomen: diejenigen Patienten, die die Behandlung komplett abschlossen („completers“), zeigten Verbesserungen der Symptomatik mit einer Effektstärke von $d = 1.13$ und intention-to-treat-Stichproben (RTCs: *randomized control trials*) einen Effekt von $d = 1.06$. Die Completer-Analysen zeigten, dass es nach der Behandlung zu signifikanten Reduktionen in dysfunktionalem Denken, genereller Angst, psychologischem Distress und der Funktionsbeeinträchtigung kommt ($d = 0.67$ bis $d = 0.88$). In einer nachfolgenden weiteren Diagnostik (*follow-up assessment*), 6 bis 12 Monate nach der Behandlung, konnten die Behandlungseffekte gehalten werden oder verbesserten sich sogar. Unterschiede zwischen individueller und gruppenbasierter CBT verfehlten die Signifikanz, allerdings zeigten post-hoc Teststärkeanalysen, dass die Power zu gering war, um kleine oder moderate Effekte aufzudecken. Daher sollten die nicht signifikanten Ergebnisse mit Vorsicht interpretiert werden. Insgesamt zeigt die Metastudie von Hans und Hiller (2013) sehr eindrucksvoll, dass Depressionen bei Erwachsenen effektiv mit CBT in typischen klinischen Settings erfolgreich behandelt werden kann, und zwar sowohl im Einzel- als auch im Gruppensetting. Allerdings zeigen Patienten im klinischen Setting weniger Besserung ihrer Symptome als in kontrollierten randomisierten Studien; außerdem ist die Dropout-Quote in der klinischen Versorgung beträchtlich. Die Studienautoren beklagen darüber hinaus die geringe methodische Qualität der verfügbaren Studien.

¹³Der Begriff „Beratung“, engl. „counselling“ wird in der Literatur uneinheitlich und widersprüchlich verwendet. Es ist aber davon auszugehen, dass es sich um eine therapeutische Intervention handelt.



Hunsley, J. und Lee (2007) prüften in ihrem Review, ob die Ergebnisse von Labor- (*efficacy*) und Feldstudien (*effectiveness*) bei gleichen Behandlungsmethoden zu denselben Ergebnissen führen. Die Literatursuche der Studienautoren ergab 35 Feldstudien, wobei sich 21 mit Störungen des Erwachsenenalters und weitere 14 mit Störungen des Kindesalters beschäftigten. Als Richtwert des Vergleiches wurden kontrollierte randomisierte Studien herangezogen. Die Abbruchquoten der Settings waren vergleichbar, auch bei der Besserung

der Beschwerden lassen sich in klinischen Settings vergleichbare Ergebnisse erzielen. Problematisch bei den Analysen war die methodischen Einschränkungen und – abermals – die mangelnde Qualität der *effectiveness*-Studien; dennoch liefert dieser Review ermutigende Ergebnisse bezüglich der Übertragbarkeit wissenschaftlicher Forschungsergebnisse von universitären Einrichtungen auf die klinische Praxis.

Lee, Horvath und Hunsley (2013) präsentieren einen Review von Wirksamkeitsstudien im Feld zu psychologischen Problemen bei Kindern und Heranwachsenden. Die Motivation der Autoren bestand darin, dass es trotz der Verfügbarkeit von zahllosen Wirksamkeitsstudien in der wissenschaftlichen Literatur kaum Wissen darüber gibt, inwiefern sich dieses Wissen auch in der Praxis umsetzen lässt. Die Forscher trugen 20 Forschungsarbeiten zur Wirksamkeit psychologischer Interventionen im Feld zusammen, die sich mit der Behandlung von Angststörungen, Depressionen und Verhaltensauffälligkeiten beschäftigten. Sie verglichen die Daten von diesen Studien mit den als Richtwert benutzten *efficacy*-Studien auf den Dimensionen „Anzahl der Klienten, die die Behandlung abgeschlossen haben“ und „Verbesserungen, die sich bei diesen Klienten eingestellt haben“. Alle Studien zur Behandlung von internalisierten Störungen berichteten Abschlussquoten von über 80%, die Interventionen für Eltern verhaltensauffälliger Kinder wurden von mehr als 75% der Eltern abgeschlossen. Die Verbesserungsraten bei Feldstudien für internalisierte Probleme waren vergleichbar mit jenen der klinischen Studien. Größere Variabilität in den Ergebnissen zeigten sich für Verhaltensauffälligkeiten, wobei mehrere Studien die klinischen Studien sogar übertrafen, eine kleinere Anzahl hingegen die Richtwerte nicht erreichte. Diese vielversprechenden Ergebnisse des Reviews sollten die Praktiker davon überzeugen, dass evidenz-basierte Behandlungen, die in Forschungssettings unter kontrollierten Bedingungen durchgeführt werden, effektiv in die Praxis des klinischen Settings übertragen werden können.

3.2 Hatte Eysenck doch recht?

In einer erst kürzlich vorgelegten Metaanalyse zu Depressionen bei Erwachsenen stellen Cuijpers, P., Karyotaki, Reijnders und Ebert (2019) die Ergebnisse bisheriger Metaanalysen zur Wirksamkeit der Psychotherapie generell in Frage. Der Artikel mit dem provokanten Titel „*Was Eysenck right after all? A reassessment of effects of psychotherapy for adult depression*“ kritisieren die Studienautoren die bisherige Vorgehensweise bei der Erstellung von Metaanalysen. Kern der Kritik ist die oft nicht durchgeführte statistische Korrektur



(*adjustment*) von Verzerrungen (*Bias*), die zur systematischen Überschätzung der Effekte von Psychotherapie führen. Zu diesen Verzerrungen gehören das „*garbage in, garbage out*“ out-Prinzip, das „*apple and oranges*“-Problem, das „*file drawer*“-Problem und der nicht aufrichtige Umgang mit Kontrollgruppen.

Garbage in, garbage out. Dieses Prinzip besagt, dass, wenn die Qualität der Einzelstudien, die in die Metaanalyse Eingang finden, schlecht ist, die Metaanalyse selbst dies nicht mehr korrigieren könne und ihre Ergebnisse nicht besser sein können als die Ergebnisse der Einzelstudien.

Apple and oranges. Äpfel mit Birnen-Vergleiche bedeutet, dass sich die Studien zur Psychotherapiewirkung untereinander erheblich unterscheiden, in Form von Patientenrekrutierung, individuellen Charakteristiken, Setting, angewandten therapeutische Methoden und der Person des Therapeuten selbst. Ist es wirklich möglich, die Ergebnisse von so unterschiedlichen Studien zusammenzufassen und in eine gemeinsame Effektgröße zu gießen?

File drawer problem. Dieses Problem weist auf die Veröffentlichungspraxis der Forscher und Verlage hin. Nicht alle Studien, die gemacht wurden, wurden auch tatsächlich veröffentlicht und sind den Forschern zugänglich. Besonders Studien mit negativem Ausgang werden oft nicht veröffentlicht, daher führt die Integration vieler Studien zu einer systematischen Überschätzung der wahren Effekte.

Kontrollgruppen. Die Kontrollgruppenbedingungen in der Psychotherapieforschung sind schwer zu realisieren und wurden auch oft nicht makellos durchgeführt. In pharmakologischen Studien ist es möglich, die Patienten darüber im Unklaren zu lassen, ob sie den Wirkstoff oder ein Placebo bekommen. Das ist in psychotherapeutischen Forschungssettings nicht möglich, wo Studienteilnehmer randomisiert entweder einer Wartekontrollgruppe oder einer üblichen Behandlung (TAU)¹⁴ zugewiesen werden. Die Patienten wissen, ob sie die tatsächliche Behandlung bekommen oder nicht. Das erzeugt Erwartungen und Hoffnung auf Besserung, was aber mit der eigentlichen Therapie nichts zu tun hat. Wartekontrollgruppen stimulieren Patienten, nichts gegen ihre Probleme zu unternehmen, weil sie wissen, die Behandlung nach der Warteperiode zu bekommen. Jüngste Metaanalysen (Furukawa et al., 2014) legen nahe, dass Wartekontrollgruppen wie ein Nocebo-Effekt wirken und damit die Effektstärken von Metaanalysen künstlich überhöhen. TAU oder „*care as usual*“ als „Behandlung“ der Kontrollgruppen ist ebenso

¹⁴ TAU = treatment as usual



problematisch, da sich diese Interventionsformen zwischen den Studien erheblich unterscheiden und somit Vergleiche schwierig werden.

In den Anfangstagen der Metaanalyse (Glass, 1976) waren diese Probleme bereits bekannt und eingestanden, aber der Enthusiasmus über die neue Methode war zu groß, um hier noch gegenzusteuern; außerdem waren zur damaligen Zeit die statistischen Methoden zur Korrektur dieser Probleme noch nicht verfügbar. So blieb die damalige Schlussfolgerung, dass Psychotherapie wirksam ist, als Kernbotschaft bis in die heutige Zeit hinein im

Gedächtnis vieler Forscher und Verantwortlicher „hängen“. Darüber hinaus wurden gerade in den Anfangstagen der Metaanalyse alle verfügbaren Studien in einen Topf geworfen, nur um die Teilnehmerzahl so groß wie möglich zu bekommen. So wurden gänzlich unterschiedliche Störungsbilder miteinander „gepoolt“, darunter Depressionen, Angststörungen, psychotische Störungen, Delinquenz und zahlreiche andere psychische Probleme.

Cuijpers, P. et al. (2019) korrigierten in ihren Analysen statistisch für die Heterogenität (Apfel und Birnen-Problem), den Typ von Kontrollgruppe, das Risiko von Verzerrungen (garbage in, garbage out) und für den Publikations-Bias. Nach Ausschluss von Studien mit Wartekontrollgruppen und solchen, die ein Risiko für Bias beinhalteten, nach Korrektur für den Publikations-Bias, fielen die Effektstärken beträchtlich. CBT, die am besten wissenschaftlich überprüfte Therapiemethode, verlor ebenso an Effektstärke wie Beratung und Problemlösungstherapie, die beide nur mehr auf kleine Effektstärke kommen. Psychodynamische Therapien waren nicht mehr statistisch signifikant. Die Autoren machen klar, dass Metaanalysen, die alle veröffentlichten Studien einfach aufnehmen, ohne auf Verzerrungen zu korrigieren, zu dramatischen Überschätzungen der Effekte kommen. Wie auch immer, es zeigte sich dennoch, dass die Gesamtwirkung der Psychotherapie statistisch bedeutsam bleibt, der Effekt allerdings gering ist. Die Frage ist nun, welche Schwelle an

Wirksamkeit überhaupt überschritten werden muss, um einem klinischen relevanten Effekt sprechen zu können. Cuijpers, P., Turner, Koole, van Dijke und Smit (2014) geben darauf eine Antwort und benennen eine Schwelle von $g = .24$ als Untergrenze, ab der von einem klinisch relevanten Effekt gesprochen werden kann. Die vier in dieser Metaanalyse untersuchten Therapieformen erreichten Effektstärken von $g = .20$ bis $g = .35$. Problemlösungstherapie und psychodynamische Therapie erreichten die Schwelle nicht, und die Effektstärken für CBT und counselling (*supportive psychotherapy*) sind nahe dieser Schwelle.

Nach Ansicht der Autoren ist klar bewiesen, dass aufgrund mangelnder Reliabilitäten der Einzelstudien, die durch methodische Mängel beziehungsweise Verzerrungen zustande kommen, die bisher errechneten Effekte nicht die wahre Wirksamkeit der Psychotherapie abbilden.

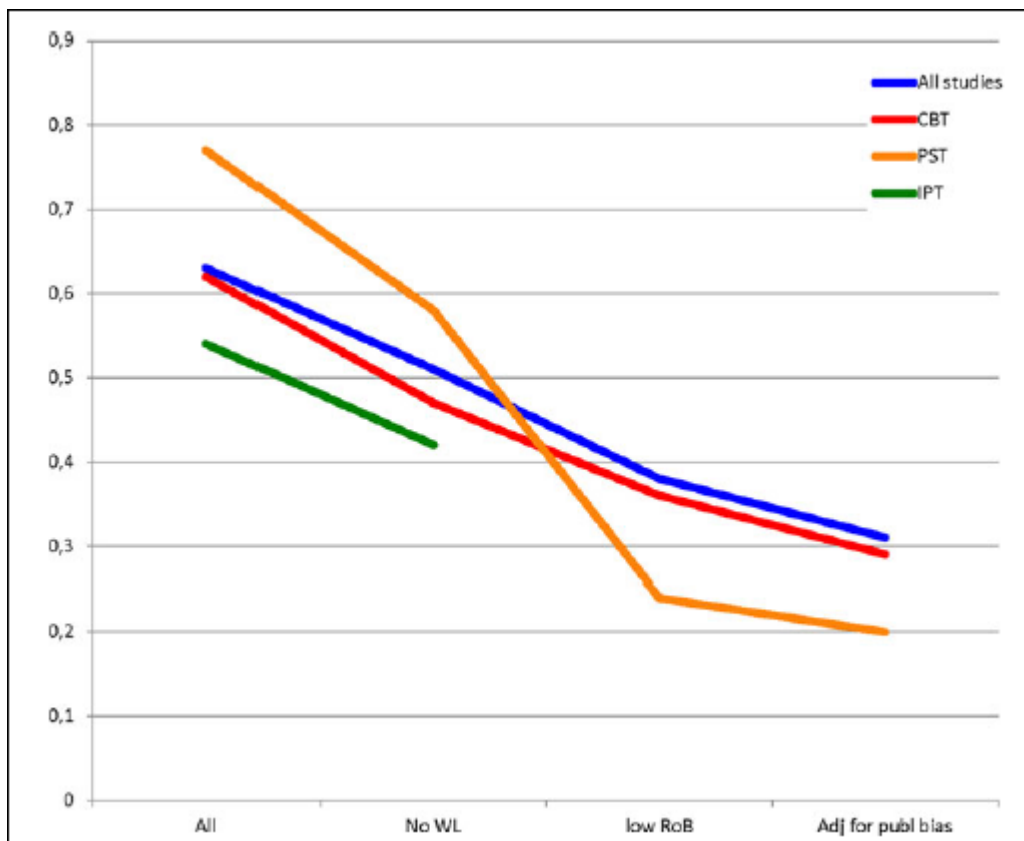


Abbildung 3-1 zeigt den Abfall der Effektstärken nach Korrektur von Verzerrungen.

Abbildung 3-1
Effektstärken für die drei am besten überprüften Typen von Psychotherapie für Depression bei



Die finale Schlussfolgerung der Autoren lautet (Cuijpers, P. et al., 2019, S. 29):

[...] it seems safe to conclude that the effects of psychotherapies for depression are considerably overestimated when all studies are included and no further analyses of risk of bias, publication bias and type of control group are done. After adjustment for these methodological issues, the effect sizes found for psychotherapies drop considerably from moderate to small.

4 Diskussion

Im Mittelpunkt der vorliegenden Arbeit stand zum einen die Fragestellung nach den Wirkfaktoren der Psychotherapie, zum anderen nach der metaanalytischen Fundierung der Wirksamkeit von Psychotherapie. Nach anfänglicher Zuversicht über die umfassende Studienlage zur Wirksamkeit, die die Psychotherapie und Psychotherapiewirkungsforschung im Laufe der Jahrzehnte hervorgebracht haben, machte sich zuletzt etwas Ernüchterung breit. Was die nachgewiesene Wirksamkeit betrifft, stößt man sowohl in Einzelstudien als auch in Metaanalysen und Reviews vor allem auf methodische Mängel, die deren (überwiegend positive) Ergebnisse oft in ein schlechtes Licht rücken und einen fahlen Beigeschmack hinterlassen. Der Autor schließt sich der Sichtweise an, die zuletzt von Cuijpers, P. et al. (2019) vorgetragen wurde: die Effektstärken von Psychotherapie sind entweder unterhalb des (von den Autoren definierten) signifikanten Bereichs oder knapp oberhalb, jedenfalls sind sie gering und unterscheiden sich deutlich von denjenigen älterer Analysen. Je genauer und methodisch strenger die Vorgehensweise bei der Studienauswahl und der Analyse, desto geringer die Effekte. Es fällt, beispielsweise im Vergleich zu den Metaanalysen der Arbeits- und Organisationspsychologie (zum Beispiel Hunter & Schmidt, 2015), auf, dass sehr viel weniger auf methodische Strenge geachtet wird, wodurch es zu künstlich überhöhten Zusammenhängen beziehungsweise Effektstärken kommt. Es wird zukünftig nötig sein, sich auf universitärem Boden zu einer wissenschaftlichen Neubewertung der Wirkung von Psychotherapie durchzuringen. Dazu ist es vor allem nötig, sich von Vorurteilen und Erwartungen in beide Richtungen zu befreien und insbesondere strenge methodische Standards, die die akademische Psychologie und die Statistik prinzipiell zur Verfügung stellen, walten zu lassen und zur Anwendung zu bringen.

Zunächst ist bei der Auswahl der Studien darauf zu achten, dass möglichst homogene Konstrukte beziehungsweise eng umschriebene Krankheitsbilder und Störungen

¹⁵ Für IPT waren nicht genügend Studien verfügbar, um nach Wegnahme der Studien mit Wartekontrollgruppen noch Effektstärken berechnen zu können.

¹⁶ PST = Problem-solving therapy

¹⁷ Cuijpers, P., Karyotaki, E., Reijnders, M. & Ebert, D. D. (2019). Was Eysenck right after all? A reassessment of the effects of psychotherapy for adult depression. *Epidemiology and psychiatric sciences*, 28 (1), 28, reproduced with permission.



aufgenommen werden. So macht es wenig Sinn, Ängste und Depressionen zu vermischen, obwohl differenzialdiagnostisch eine trennscharfe Abgrenzung nicht immer möglich ist, wie die Testkonstruktion auch immer wieder gezeigt hat (zum Beispiel im State-Trait-Angstinventar, Laux, L., Glanzmann, P., Schaffner, P. & Spielberger, C.D., 1981). Des Weiteren sind Teststärkeanalysen bereits bei der Planung von Einzelstudien dringend erforderlich: wenn die Anzahl an Studienteilnehmern sehr gering ist, kann kaum ein spezifischer Effekt aufgedeckt werden, die Konfidenzintervalle sind zu groß und die Schätzungen damit zu ungenau. Diese Fehler können dann auch nicht mehr ausreichend korrigiert werden, wenn viele Studien „gepoolt“ werden, da die Qualität der Aussage einer Metaanalyse niemals über die Aussagekraft von Einzelstudien hinausgehen kann. Sind die Einzelstudien mangelhaft, trägt die Metaanalyse wenig zum Erkenntnisgewinn bei. Besonders die mangelnde Reliabilität vieler Einzelstudien ist problematisch.

Weiters sind die Bedingungen in Form von Moderatoranalysen zu prüfen, unter denen Psychotherapie wirksam ist. Zu den Bedingungen gehören Fragen wie Patienten- und Therapeutencharakteristika, aber auch biographisch-ätiologische Variablen, aufrechterhaltende Umweltbedingungen, sozioökonomische Variablen oder genetische Varianten, wahrgenommene Selbstwirksamkeit, dispositionaler Optimismus und Erwartungshaltungen sowie motivationale Aspekte. Dazu gehört aber auch die Frage, ob Studien streng kontrolliert unter Laborbedingungen oder im Feld durchgeführt wurden. Dafür müsste es streng genommen zwei verschiedene Arten von Metaanalysen geben, jene, die sich mit *efficacy* (Labor) und jene, die sich mit *effectiveness* (Feld) beschäftigen. Die Erfahrung hat allerdings gezeigt, dass die Ergebnisse von kontrollierten Laborstudien gut auf das Praxissetting übertragen werden können, sodass eine Zweigleisigkeit hier gar nicht notwendig ist. Auch die Frage nach den Wirkmechanismen (Mediatoranalysen) sollte noch stärker untersucht werden. Worin ist begründet, dass eine Psychotherapie wirkt, wenn sie wirkt? Sind die Wirkmechanismen von der Schule oder therapeutischen Methode abhängig? Es gibt hierauf, wie oben dargestellt, bereits Antworten, die dem Praktiker an die Hand gegeben werden können, von der wissenschaftlichen Seite her besteht allerdings noch deutlicher Klärungsbedarf: so ist die Frage, „spezifische Wirkmechanismen versus Dodo-Verdikt“ noch immer nicht zufriedenstellend geklärt.

Ein großes Dilemma ist die Benutzung von Kontrollgruppen. Freilich ist es das Anliegen jedes (quantitativen) Forschers, einen Interaktionseffekt im Vergleich von Experimental- mit Kontrollgruppe zu erzielen, und somit einen Effekt des Treatments eindeutig belegen zu können. Leider sind unbehandelte Kontrollgruppen, wie oben dargestellt, in der Psychotherapieforschung kaum zu realisieren und Störvariablen, wie Erwartungshaltungen oder plötzliche Verhaltens- und Einstellungsänderungen der Kontrollgruppenteilnehmer, nicht zu verhindern. Dazu kommt, dass aus ethischen Gründen stets Wartekontrollgruppen verwendet werden, und individuelle Erlebens- und



Verhaltensänderungen nicht ausreichend kontrolliert beziehungsweise auspartialisiert werden können. Im Endeffekt entsteht hier das klassische Problem der Sicherstellung der internen Validität: kann ein Behandlungseffekt tatsächlich auf die Behandlung zurückgeführt werden? Entsteht die Wirksamkeit einer Behandlung auch wirklich durch die experimentelle Manipulation, das heißt, durch die Manipulation der unabhängigen Variable „Treatment versus no-treatment“? Unmittelbar kann das Problem nur minimiert werden, indem umfangreiche psychologische Inventare zur Exploration des Erlebens und Verhaltens eingesetzt werden, die dann als Kovariaten in die statistischen Modelle aufgenommen werden. Darüber hinaus ist – zumindest im Laborsetting – eine strikte randomisierte Zuteilung zu Behandlungs- und Kontrollgruppe zwingend erforderlich. Dies erhöht die interne Validität ungemein und erleichtert darüber hinaus auch den Einsatz von standardisierten statistischen Auswerteprozeduren.

Sinnvoll wären auch Studien, die sich stärker auf die neurobiologischen Wirkmechanismen, ähnlich der pharmakologischen Forschung, konzentrieren. Dazu gehören Fragestellungen nach geänderten Neurotransmitterhaushalten, strukturell-anatomische und physiologische Wirkmechanismen, biologische Stabilität der Effekte oder Neuroplastizität als Wirkprinzip der Psychotherapie. Und natürlich die hochrelevante Frage, wie Psychotherapie bei organischen Erkrankungen wie Herz-Kreislauf-Erkrankungen, Koronarerkrankungen, essenzieller Hypertonie oder onkologischen Erkrankungen sinnvoll additiv eingesetzt werden kann. Hier geht es um die Rehabilitation ebenso wie um Prävention von psychischen Komorbiditäten.

Ein besonderes Anliegen der Wissenschaft wie auch der Vertreter verschiedener psychotherapeutischer Richtungen sollte es sein, den sogenannten „Schulenstreit“ zu überwinden und tiefgehende methodisch strenge und saubere wissenschaftliche Forschung zu fordern, auf deren Realisierung zu drängen und schließlich zu betreiben, um auf die genannten Fragestellungen Antworten zu geben und somit zu einer Weiterentwicklung der Disziplin im Sinne der Wissenschaft und letztlich auch im Sinne der Patienten, beizutragen.

5 Literaturverzeichnis

- Acarturk, C., Cuijpers, P., van Straten, A. & Graaf, R. de. (2009). Psychological treatment of social anxiety disorder. A meta-analysis. *Psychological medicine*, 39 (2), 241-254.
- Bisson, J. I., Ehlers, A., Matthews, R., Pilling, S., Richards, D. & Turner, S. (2007). Psychological treatments for chronic post-traumatic stress disorder. Systematic review and meta-analysis. *The British journal of psychiatry: the journal of mental science*, 190, 97-104.



- Cape, J., Whittington, C., Buszewicz, M., Wallace, P. & Underwood, L. (2010). Brief psychological therapies for anxiety and depression in primary care. Meta-analysis and meta-regression. *BMC medicine*, 8, 38.
- Cuijpers, P., Andersson, G., Donker, T. & van Straten, A. (2011). Psychological treatment of depression. Results of a series of meta-analyses. *Nordic journal of psychiatry*, 65 (6), 354-364.
- Cuijpers, P., Karyotaki, E., Reijnders, M. & Ebert, D. D. (2019). Was Eysenck right after all? A reassessment of the effects of psychotherapy for adult depression. *Epidemiology and psychiatric sciences*, 28 (1), 21-30.
- Cuijpers, P., Turner, E. H., Koole, S. L., van Dijke, A. & Smit, F. (2014). What is the threshold for a clinically relevant effect? The case of major depressive disorders. *Depression and anxiety*, 31 (5), 374-378.
- DeRubeis, R. J., Brotman, M. A. & Gibbons, C. J. (2005). A Conceptual and Methodological Analysis of the Nonspecifics Argument. *Clinical Psychology: Science and Practice*, 12 (2), 174-183.
- Dreher, J. (2012). *Die 5 Wirkfaktoren der Psychotherapie nach Klaus Grawe*. Zugriff am 19.04.2019. Verfügbar unter <https://psychiatrietogo.de/2012/02/04/die-5-wirkfaktoren-der-psychotherapie-nach-klaus-grawe/>
- Driessen, E., Cuijpers, P., Hollon, S. D. & Dekker, J. J. M. (2010). Does pretreatment severity moderate the efficacy of psychological treatment of adult outpatient depression? A meta-analysis. *Journal of consulting and clinical psychology*, 78 (5), 668-680.
- Eid, M., Gollwitzer, M. & Schmitt, M. (2017). *Statistik und Forschungsmethoden* (5., korrigierte Auflage). Weinheim: Beltz.
- Eysenck, H. J. (1992). The effects of psychotherapy. An evaluation. *Journal of consulting and clinical psychology*, 60 (5), 659-663.
- Feng, C.-Y., Chu, H., Chen, C.-H., Chang, Y.-S., Chen, T.-H., Chou, Y.-H. et al. (2012). The effect of cognitive behavioral group therapy for depression. A meta-analysis 2000-2010. *Worldviews on evidence-based nursing*, 9 (1), 2-17.
- Furukawa, T. A., Noma, H., Caldwell, D. M., Honyashiki, M., Shinohara, K., Imai, H. et al. (2014). Waiting list may be a placebo condition in psychotherapy trials. A contribution from network meta-analysis. *Acta psychiatrica Scandinavica*, 130 (3), 181-192.
- Glass, G. V. (1976). Primary, Secondary, and Meta-Analysis of Research. *Educational Researcher*, 5 (10), 3.
- Grawe, K. (1995). Grundriß einer Allgemeinen Psychotherapie. *Psychotherapeut*, 40 (3), 130-145.
- Grawe, K. (2000). *Psychologische Therapie* (2., korrigierte Aufl.). Göttingen: Hogrefe Verl. für Psychologie.
- Grawe, K., Donati, R. & Bernauer, F. (2001). *Psychotherapie im Wandel. Von der Konfession zur Profession* (5., unveränd. Aufl.). Göttingen: Hogrefe.
- Grissom, R. J. (1996). The magical number .7 +/- .2. Meta-meta-analysis of the probability of superior outcome in comparisons involving therapy, placebo, and control. *Journal of consulting and clinical psychology*, 64 (5), 973-982.
- Hans, E. & Hiller, W. (2013). Effectiveness of and dropout from outpatient cognitive behavioral therapy for adult unipolar depression. A meta-analysis of nonrandomized effectiveness studies. *Journal of consulting and clinical psychology*, 81 (1), 75-88.



- Hunot, V., Churchill, R., Teixeira, V. & Lima, M. S. d. (2007). Psychological therapies for generalised anxiety disorder. *Cochrane Database of Systematic Reviews* (1). Verfügbar unter <https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD001848.pub4/full>
- Hunsley, J., Elliott K. & Therrien, Z. (2013). *The Efficacy and Effectiveness of Psychological Treatments*. Ottawa, Ontario: Canadian Psychological Association.
- Hunsley, J. & Lee, C. M. (2007). Research-informed benchmarks for psychological treatments: Efficacy studies, effectiveness studies, and beyond. *Professional Psychology: Research and Practice*, 38 (1), 21-33.
- Hunter, J. E. & Schmidt, F. L. (Eds.). (2015). *Methods of meta-analysis. Correcting error and bias in research findings* (3. ed.). Thousand Oaks, Calif.: Sage Publ.
- Huntley, A. L., Araya, R. & Salisbury, C. (2012). Group psychological therapies for depression in the community. Systematic review and meta-analysis. *The British journal of psychiatry : the journal of mental science*, 200 (3), 184-190.
- Kowalik, J., Weller, J., Venter, J. & Drachman, D. (2011). Cognitive behavioral therapy for the treatment of pediatric posttraumatic stress disorder. A review and meta-analysis. *Journal of behavior therapy and experimental psychiatry*, 42 (3), 405-413.
- Laux, L., Glanzmann, P., Schaffner, P. & Spielberger, C.D. (1981). *Das State-Trait-Angstinventar. (Testmappe mit Handanweisung, Fragebogen STAI-GForm X1 und Fragebogen STAI-GForm X 2)*. Weinheim: Beltz.
- Lee, C. M., Horvath, C. & Hunsley, J. (2013). Does it work in the real world? The effectiveness of treatments for psychological problems in children and adolescents. *Professional Psychology: Research and Practice*, 44 (2), 81-88.
- Luborsky, L., Diguier, L., Seligman, D. A., Rosenthal, R., Krause, E. D., Johnson, S. et al. (1999). The Researcher's Own Therapy Allegiances. A "Wild Card" in Comparisons of Treatment Efficacy. *Clinical Psychology: Science and Practice*, 6 (1), 95-106.
- Luborsky, L., Rosenthal, R., Diguier, L., Andrusyna, T. P., Levitt, J. T., Seligman, D. A. et al. (2003). Are Some Psychotherapies Much More Effective Than Others? *Journal of Applied Psychoanalytic Studies*, 5 (4), 455-460. Verfügbar unter <https://doi.org/10.1023/A:1026023715740>
- Peng, X.-D., Huang, C.-Q., Chen, L.-J. & Lu, Z.-C. (2009). Cognitive behavioural therapy and reminiscence techniques for the treatment of depression in the elderly. A systematic review. *The Journal of international medical research*, 37 (4), 975-982.
- Pfammatter, M. & Tschacher, W. (2012). Wirkfaktoren der Psychotherapie – eine Übersicht und Standortbestimmung. *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie*, 60 (1), 67-76.
- Powers, M. B., Halpern, J. M., Ferenschak, M. P., Gillihan, S. J. & Foa, E. B. (2010). A meta-analytic review of prolonged exposure for posttraumatic stress disorder. *Clinical psychology review*, 30 (6), 635-641.
- Prioleau, L., Murdock, M. & Brody, N. (1983). An analysis of psychotherapy versus placebo studies. *Behavioral and Brain Sciences*, 6 (02), 275.
- Rosa-Alcázar, A. I., Sánchez-Meca, J., Gómez-Conesa, A. & Marín-Martínez, F. (2008). Psychological treatment of obsessive-compulsive disorder. A meta-analysis. *Clinical psychology review*, 28 (8), 1310-1325.



- Sánchez-Meca, J., Rosa-Alcázar, A. I., Marín-Martínez, F. & Gómez-Conesa, A. (2010). Psychological treatment of panic disorder with or without agoraphobia. A meta-analysis. *Clinical psychology review*, 30 (1), 37-50.
- Sedlmeier, P. & Renkewitz, F. (2011). *Forschungsmethoden und Statistik in der Psychologie* (Ps, Psychologie, [Nachdr.]. München: Pearson Studium.
- Segool, N. K. & Carlson, J. S. (2008). Efficacy of cognitive-behavioral and pharmacological treatments for children with social anxiety. *Depression and anxiety*, 25 (7), 620-631.
- Seligman, M. E. (1995). The effectiveness of psychotherapy. The Consumer Reports study. *The American psychologist*, 50 (12), 965-974.
- Smith, M. L. & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 32 (9), 752-760.
- Szentogotai, A. & David, D. (2010). The efficacy of cognitive-behavioral therapy in bipolar disorder. A quantitative meta-analysis. *The Journal of clinical psychiatry*, 71 (1), 66-72.
- Tschuschke, V. (2005). Die Psychotherapie in Zeiten evidenzbasierter Medizin. Fehlentwicklungen und Korrekturvorschläge. *Psychotherapeutenjournal*, 2, 106-115.
- Wampold, B. E. (2015). *The Great Psychotherapy Debate. The Evidence for What Makes Psychotherapy Work* (Second edition). New York, NY: Routledge.
- Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K. & Ahn, H.-n. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies. Empirically, "all must have prizes.". *Psychological bulletin*, 122 (3), 203-215.
- Watson, H. J. & Rees, C. S. (2008). Meta-analysis of randomized, controlled treatment trials for pediatric obsessive-compulsive disorder. *Journal of child psychology and psychiatry, and allied disciplines*, 49 (5), 489-498.
- Weisz, J. R., McCarty, C. A. & Valeri, S. M. (2006). Effects of psychotherapy for depression in children and adolescents. A meta-analysis. *Psychological bulletin*, 132 (1), 132-149.
- Wolitzky-Taylor, K. B., Horowitz, J. D., Powers, M. B. & Telch, M. J. (2008). Psychological approaches in the treatment of specific phobias. A meta-analysis. *Clinical psychology review*, 28 (6), 1021-1037.



6 Autoreninformation

Mag. Christian Frauscher, BSc

Anton Bruckner Privatuniversität Linz

Arbeitsbereich Studio Angewandte Psychologie

Raum 2.315 und 2.315a (Labor Bio- und Neurofeedback)

Hagenstraße 57, 4040 Linz | +43 650 74 33 199 | christian.frauscher@bruckneruni.at