

***Eine Meta-Analyse zur Effektivität verschiedener  
Behandlungsstrategien bei Angststörungen:  
Ein Vergleich von behavioraler, kognitiv-behavioraler  
und pharmakologischer Therapie***

**Dissertation  
zur Erlangung des akademischen Grades**

doctor philosophiae (Dr. phil.)

**vorgelegt dem Rat der Fakultät für Sozial- und Verhaltenswissenschaften  
der Friedrich-Schiller-Universität Jena**

von Dipl. Psych. Kristin Mitte  
geboren am 18.03.1977 in Annaberg-Buchholz



## Danksagung

An dieser Stelle möchte ich mich bei all denjenigen bedanken, die durch ihre vielfältige Unterstützung zum Gelingen dieser Arbeit beigetragen haben.

Herrn Prof. Dr. Peter Noack sei aufs herzlichste gedankt für die unkomplizierte Übernahme der Betreuerfunktion, die wertvollen Anregungen und sein engagiertes Interesse. Auch Herrn Prof. Dr. Eckart Straube danke ich für die geschaffenen Rahmenbedingungen. Herrn Prof. Dr. Martin Hautzinger danke ich für die Übernahme der Gutachterfunktion. Ich bedanke mich ganz herzlich bei der Vergabekommission für die Gewährung eines Landesgraduiertenstipendiums.

Frau Anna-Maria Eberhardt danke ich für ihre freundschaftliche Hilfe und die anregenden Gespräche. Ein besonderer Dank gilt Frau Dr. Regina Steil für Ihre Unterstützung durch Rat und Tat. Herrn Dr. Christof Nachtigall danke ich für die Hilfe in statistischer Not.

Frau Ellen Kremberg, Frau Anna Watzke und Frau Peggy Braun seien gedankt für ihren Zeiteinsatz und ihre Unterstützung. Frau Una-Lena Halitschke und Frau Julia Wache danke ich für die Eingabe eines Teils der Literatur.

Ich danke allen WissenschaftlerInnen, die mir Daten zur Verfügung gestellt haben, für ihre freundliche Bereitschaft, die Arbeit zu unterstützen. Den MitarbeiterInnen der Thüringer Universitäts- und Landesbibliothek danke ich für die Bearbeitung der vielen Fernleihen.

Mein wärmster und herzlichster Dank gilt Kersten Schäfer für seine liebevolle Unterstützung und Ermutigung während der Zeit meiner Promotion. Nicht nur seine emotionale sondern auch seine praktische Hilfe war von unschätzbarem Wert. Diese Arbeit wäre nicht in dem Zeitrahmen möglich gewesen ohne seine vielen Stunden Einsatz für die Entwicklung des Auswertungsprogramms und die Hilfestellung bei technischen Problemen. Wohltuende Gelassenheit und materielle Unterstützung erhielt ich von meiner Familie.

<b>1 Grundlagen</b>	<b>7</b>
<b>1.1 Störungsbilder</b> .....	<b>8</b>
1.1.1 Generalisierte Angststörung (GAS) .....	8
1.1.2 Panikstörung mit und ohne Agoraphobie, Agoraphobie ohne Panikstörung in der Vorgeschichte (PDA) .....	9
1.1.3 Posttraumatische Belastungsstörung (PTB) .....	11
1.1.4 Soziale Phobie (SoP) .....	12
1.1.5 Spezifische Phobie (SpP) .....	13
1.1.6 Zwangsstörung (OCD) .....	13
<b>1.2 Therapieansätze</b> .....	<b>14</b>
1.2.1 Behaviorale und kognitiv-behaviorale Therapien (VT, KVT) .....	14
1.2.2 Pharmakotherapie (PhT) .....	17
<b>1.3 Stand der Forschung und Kritik</b> .....	<b>18</b>
<b>1.4 Fragestellungen</b> .....	<b>20</b>
<b>2 Methode</b>	<b>21</b>
<b>2.1 Die Meta-Analyse als integrierendes Verfahren</b> .....	<b>21</b>
<b>2.2 Selektion der Primärstudien</b> .....	<b>22</b>
2.2.1 Literatursuche .....	22
2.2.2 Codierungskatalog .....	23
2.2.3 Untersuchungsdesigns der Primärstudien .....	25
2.2.4 Selektionskriterien .....	28
<b>2.3 Statistische Methoden</b> .....	<b>38</b>
2.3.1 Berechnung der Effektstärken für eine Studie .....	38
2.3.2 Integration der Effektstärken über alle Studien .....	43
2.3.3 Weitere Aspekte .....	54
2.3.4 Beschreibung der untersuchten methodischen Variablen .....	56

<b>3</b>	<b>Ergebnisse</b>	<b>61</b>
<b>3.1</b>	<b>Empirische Exploration des Schätzers der Effektivität verschiedener Kontrollgruppen.....</b>	<b>61</b>
3.1.1	Ergebnisse .....	61
3.1.2	Schlussfolgerung .....	62
<b>3.2</b>	<b>Überblick über die eingeschlossenen Primärstudien.....</b>	<b>63</b>
<b>3.3</b>	<b>Generalisierte Angststörung .....</b>	<b>66</b>
3.3.1	Deskriptive Analyse der Primärstudien .....	66
3.3.2	Vergleich der Prä-Werte .....	67
3.3.3	Psychotherapie .....	68
3.3.4	Pharmakotherapie.....	75
3.3.5	Vergleich Psycho- und Pharmakotherapie .....	79
3.3.6	Diskussion.....	82
<b>3.4</b>	<b>Panikstörung mit und ohne Agoraphobie, Agoraphobie ohne Panikstörung in der Vorgeschichte.....</b>	<b>85</b>
3.4.1	Deskriptive Analyse der Primärstudien .....	85
3.4.2	Vergleich der Prä-Werte .....	86
3.4.3	Psychotherapie .....	87
3.4.4	Pharmakotherapie.....	97
3.4.5	Vergleich Psycho- und Pharmakotherapie .....	101
3.4.6	Vergleich Psychotherapie mit Kombination aus Psycho- und Pharmakotherapie.....	104
3.4.7	Diskussion.....	108
<b>3.5</b>	<b>Posttraumatische Belastungsstörung .....</b>	<b>112</b>
3.5.1	Deskriptive Analyse der Primärstudien .....	112
3.5.2	Vergleich der Prä-Werte .....	113
3.5.3	Psychotherapie .....	113
3.5.4	Pharmakotherapie.....	118
3.5.5	Vergleich Psycho- und Pharmakotherapie .....	120

3.5.6 Diskussion.....	122
<b>3.6 Soziale Phobie.....</b>	<b>125</b>
3.6.1 Deskriptive Analyse der Primärstudien .....	125
3.6.2 Vergleich der Prä-Werte .....	126
3.6.3 Psychotherapie .....	127
3.6.4 Pharmakotherapie.....	134
3.6.5 Vergleich Psycho- und Pharmakotherapie .....	137
3.6.6 Diskussion.....	140
<b>3.7 Spezifische Phobie .....</b>	<b>143</b>
3.7.1 Deskriptive Analyse der Primärstudien .....	143
3.7.2 Vergleich der Prä-Werte .....	144
3.7.3 Psychotherapie .....	144
3.7.4 Pharmakotherapie.....	147
3.7.5 Diskussion.....	148
<b>3.8 Zwangsstörung .....</b>	<b>150</b>
3.8.1 Deskriptive Analyse der Primärstudien .....	150
3.8.2 Vergleich der Prä-Werte .....	151
3.8.3 Psychotherapie .....	151
3.8.4 Pharmakotherapie.....	154
3.8.5 Vergleich Psycho- und Pharmakotherapie .....	157
3.8.6 Vergleich Psychotherapie mit einer Kombinationsbehandlung.....	158
3.8.7 Diskussion.....	161
<b>4 Diskussion</b>	<b>164</b>
<b>4.1 Zusammenfassende Betrachtung der Ergebnisse .....</b>	<b>164</b>
4.1.1 Effektivität der behavioralen und kognitiv-behavioralen Verfahren .....	164
4.1.2 Akzeptanz behavioraler und kognitiv-behavioraler Verfahren.....	166
4.1.3 Effektivität der medikamentösen Therapien .....	166
4.1.4 Akzeptanz medikamentöser Therapien.....	167

4.1.5	Vergleichende Effektivität von Psycho- und Pharmakotherapie .....	167
4.1.6	Vergleichende Effektivität von Psychotherapie mit einem integrativen Ansatz aus Psycho- und Pharmakotherapie .....	168
4.1.7	Zusammenfassung der Ergebnisse zur Effektivität der Therapien .....	168
4.1.8	Einschätzung der Verteilung gemeinsamer und spezifischer Faktoren für die Wirksamkeit der untersuchten Therapien .....	169
4.1.9	Einfluss methodischer Variablen .....	170
<b>4.2</b>	<b>Einschränkungen der Arbeit.....</b>	<b>173</b>
4.2.1	Besonderheiten der eingeschlossenen Primärstudien.....	173
4.2.2	Meta-Analytische Besonderheiten .....	180
<b>5</b>	<b>Zusammenfassung</b>	<b>188</b>
<b>6</b>	<b>Literatur</b>	<b>190</b>
6.1	Eingeschlossene Studien Generalisierte Angststörung .....	205
6.2	Eingeschlossene Studien Panikstörung .....	212
6.3	Eingeschlossene Studien Posttraumatische Belastungsstörung.....	224
6.4	Eingeschlossene Studien Soziale Phobie.....	227
6.5	Eingeschlossene Studien Spezifische Phobie.....	232
6.6	Eingeschlossene Studien Zwangsstörung.....	235
<b>7</b>	<b>Anhang</b>	<b>239</b>
7.1	Anhang A: Bisherige Meta-Analysen .....	240
7.2	Anhang B: Codierungskatalog.....	250
7.3	Anhang C: Beschreibung des Programms.....	264
7.4	Anhang D: Berechnung der Effektstärken.....	268
7.5	Anhang E: Retest-Reliabilitäten .....	271
7.6	Anhang F: Ausreißeranalyse.....	272
7.7	Anhang G: RE-Schätzer .....	273

Schon wieder eine Meta-Analyse. Seit der wegleitenden Arbeit von Smith, Glass und Miller (1980) steigt die Anzahl veröffentlichter Meta-Analysen stetig. Auch innerhalb des Themengebietes der Behandlung von Angststörungen wurden schon einige Arbeiten publiziert. Dies spiegelt natürlich die potentiellen Vorteile des Verfahrens wider: neben einem zusammenfassenden Überblick über ein Forschungsfeld sind im Rahmen von Meta-Analysen Aussagen möglich, die einer Einzeluntersuchung verwehrt bleiben. Insbesondere im Bereich der Psychotherapieforschung ist die Generalisierbarkeit von Ergebnissen bedeutsam; diesbezüglich ist das Potenzial von Meta-Analysen erheblich höher, als es einer Primärstudie je möglich ist. So fließen in die folgende Arbeit Daten von über 30 000 Patienten ein, dies sind Patienten aus verschiedenen Kohorten, mit unterschiedlichen Eigenschaften, die von verschiedenen Therapeuten in unterschiedlichen Settings mit verschiedenen Therapien behandelt wurden, Patienten, die an Untersuchungen teilnahmen, bei denen Wissenschaftler differierende Entscheidungen hinsichtlich des Designs getroffen haben, Patienten, deren Symptomatik mit den unterschiedlichsten Messinstrumenten erfasst wurde. Eigenschaften einer Studie können somit in Zusammenhang gebracht werden mit den beobachteten Behandlungseffekten. Bestimmte Fragestellungen, wie der Einfluss von Therapiedauer oder Abbrecherrate auf die beobachtete Effektivität sind im Rahmen von Meta-Analysen möglich, nur schwer oder gar nicht dagegen in Primärstudien.

Beelmann und Bliesener (1994) schrieben bereits vor 10 Jahren, dass der Höhepunkt der wissenschaftlichen Arbeit zu technischen Fragen der Durchführung einer Meta-Analyse vorbei ist. Gleichzeitig werden aber diese methodischen Weiterentwicklungen auch heute nur selten angewandt. Ein Beispiel dazu ist, dass seit vielen Jahren zur Integration der Effektstärken ein statistisches Modell vorliegt, welches im Besonderen für Fragen der Psychotherapieforschung relevant ist. Es erlaubt sowohl Generalisierungen der Ergebnisse über die eingeschlossenen Studie hinaus und es ignoriert nicht den Einfluss wichtiger, aber leider oft nicht beobachtbarer Variablen auf die Effektstärke (z.B. die Güte der Therapeuten-Klienten-Interaktion). Trotz dieser bedeutsamen Vorteile wurde nur in einem sehr geringen Prozentsatz der bisherigen Meta-Analysen dieses Modell eingesetzt, in der vorliegenden Arbeit wird es verwendet. Auch weitere methodische Besonderheiten der meta-analytischen Technik sollen beachtet werden, so dass eine fehlerfreiere Einschätzung der Effektivität der Behandlungsverfahren möglich ist.



Meta-analytische Techniken sind zu einem etablierten Verfahren geworden und gehören zu den bedeutendsten Instrumenten, die einem Wissenschaftler zur Verfügung stehen. Ihre Anwendung bei der Behandlung einer der häufigsten psychischen Störungen soll in dieser Arbeit erfolgen.

## 1 Grundlagen

Angststörungen zählen zu den häufigsten psychischen Störungen. So fanden Kessler und Kollegen in einer epidemiologischen Untersuchung an über 8000 Personen aus der Allgemeinbevölkerung der USA (des National Comorbidity Survey), dass fast jeder fünfte Mann und fast jede dritte Frau im Leben an einer Angststörung erkrankt (Kessler et al., 1994)<sup>1</sup>. In einer Kombination der Daten dieser mit einer weiteren wichtigen epidemiologischen Studie, der Epidemiologic Catchment Area Studie, ergab sich, dass 16 % aller Personen innerhalb eines Jahres an einer Angststörung erkrankt waren, im Vergleich mit 7 % an einer Affektiven Störung (n. Barlow, 2002). Obwohl das Ausmaß an Symptomen bei unbehandelten Angststörungen fluktuiert, sind vollständige Remissionen selten. So lag die Rate der Spontanremission über 7 Jahre in einer deutschen Untersuchung von Wittchen (1991) unter 20 %. Ähnliche Zahlen fanden sich auch bei internationalen epidemiologischen Studien (z.B. Yonkers, Warshaw, Massion & Keller, 1996). Höhere Raten ergaben sich für die Panikstörung ohne Agoraphobie mit ca. 43 % für einen Zweijahreszeitraum (Hirschfeld, 1996). Angststörungen sind hoch komorbide untereinander, aber auch mit anderen Störungen insbesondere Affektiver Störungen und Substanzmissbrauch. Brown und Kollegen fanden bei 37 % (Soziale Phobie) bis 71 % (Generalisierte Angststörung) der Patienten eine zusätzliche Diagnose einer Angststörung, bei 36 % (Spezifische Phobie) bis 85 % (Posttraumatische Belastungsstörung) die einer Affektiven Störung und bei 7 % (Spezifische Phobie) bis 31 % (Posttraumatische Belastungsstörung) die von Alkoholabhängigkeit bzw. -missbrauch für die Lebenszeit (Brown, Campbell, Lehman, Grisham & Mancill, 2001).

Angststörungen stellen sowohl eine außerordentliche Belastung für die Betroffenen selbst als auch für die Gesellschaft dar (z.B. Mendlowitz & Stein, 2001). Personen mit einer Angststörung zeigen im Vergleich mit Gesunden ein höheres Risiko für ein geringes Einkommen und Fluktuationen in der Beschäftigung. Sie leben häufiger allein und sind häufiger geschieden. Amerikanische Studien zeigten, dass Angststörungen einen Aufwand von 42 bis 46 Milliarden US-Dollar für die Jahre 1990 bzw. 1991 verursachten und somit als kostenintensivste psychische Störung anzusehen sind (Greenberg et al., 1999; Rice & Miller, 1998). Diese ergeben sich aus direkten Ausgaben, wie z.B.

---

<sup>1</sup> Zwangsstörung und Posttraumatische Belastungsstörung nicht eingeschlossen.

Kosten für Behandlungen (sowohl psychiatrisch als auch nichtpsychiatrisch medizinisch), und indirekten Ausgaben, wie eingeschränkte Produktivität.

## **1.1 Störungsbilder**

Im Bereich der Angststörungen können folgende Störungsbilder unterschieden werden: Generalisierte Angststörung (GAS), Panikstörung mit und ohne Agoraphobie sowie Agoraphobie ohne Panikstörung in der Vorgeschichte (PDA), Posttraumatische Belastungsstörung (PTB), Soziale Phobie (SoP), Spezifische Phobie (SpP) und Zwangsstörung (OCD). Im Folgenden sind diese kurz beschrieben.

Für die einzelnen Störungen nicht erwähnt, aber gemeinsam für die Diagnose ist eine bedeutsame Einschränkung des Alltagslebens bzw. der Leistungsfähigkeit (Ausnahme PDA/PA), sowie dass die Symptome nicht durch eine andere, auch medizinische, Störung bzw. direkte Effekte von Substanzen besser erklärt werden können (Diagnostic and Statistical Manual of Mental Disorders, 4. Ed. (DSM-IV), American Psychiatric Association (APA), 1994).

### **1.1.1 Generalisierte Angststörung (GAS)**

Die Generalisierte Angststörung ist nach DSM-IV (APA, 1994) definiert durch exzessive Angst oder Sorgen hinsichtlich einer Reihe an Ereignissen oder Aktivitäten verbunden mit Symptomen der Ruhelosigkeit, leichter Ermüdbarkeit, Konzentrationschwierigkeiten, Reizbarkeit, Muskelverspannungen oder Schlafstörungen. Die Sorgen können nur schwer von der Person kontrolliert werden. Die Symptome dauern mindestens über einen Zeitraum von 6 Monaten an.

Empirisch wurde gefunden, dass sich die Inhalte der Sorgen nicht von denen nicht-klinischer Stichproben unterscheiden und sich auf Finanzen, Arbeit oder das Wohlbefinden der Familie beziehen mit Ausnahme der Alltagsprobleme („daily hassles“), die häufiger bei Patienten mit GAS zu finden sind (Craske, Rapee, Jackel & Barlow, 1989). Bedeutsamer erscheint die Zeit der Beschäftigung mit den Sorgen. So fanden z.B.

Dupuy und Kollegen (2001), dass Patienten mit GAS täglich 310 Minuten mit Sorgen verbringen im Vergleich zu 55 Minuten bei einer nicht-klinischen Stichprobe.

Die Lebenszeitprävalenz beträgt bei GAS für neuere epidemiologische Studien (ab DSM-III-R) um 5 % (z.B. Kessler et al., 1994). Der Anteil an Frauen ist ca. doppelt so hoch wie der an Männern (Wittchen, Zhao, Kessler & Eaton, 1994).

### **1.1.2 Panikstörung mit und ohne Agoraphobie, Agoraphobie ohne Panikstörung in der Vorgeschichte (PDA)**

Im DSM-IV (APA, 1994) wird zwischen Panikstörung mit bzw. ohne Agoraphobie sowie Agoraphobie ohne Panikstörung in der Vorgeschichte unterschieden. Panikattacken selbst sowie agoraphobisches Verhalten sind keine codierbaren Störungen. Panikattacken sind definiert als ein abgegrenzter Zeitraum intensiver Furcht oder des Unbehagens mit verschiedenen Symptomen physiologischer (z.B. Schwindel, Schwitzen) und kognitiver Natur (z.B. Angst zu sterben, die Kontrolle zu verlieren). Die Symptome setzen dabei plötzlich ein und entwickeln ihre volle Schwere innerhalb 10 Minuten. Kennzeichen der Panikstörung ohne Agoraphobie sind wiederholt unerwartete Panikattacken („aus dem Blauen heraus“), in deren Folge mindestens über einen Monat lang über deren antizipiertes Auftreten oder ihre Konsequenzen Besorgnis besteht (z.B. einen Herzinfarkt bekommen) bzw. die zu einer Verhaltensänderung führen (z.B. häufigere Arztbesuche). Panikstörung mit Agoraphobie wird diagnostiziert bei zusätzlicher Angst vor Orten oder Situationen, in denen eine Flucht schwer möglich oder peinlich (z.B. beim Frisör) bzw. in denen keine Hilfe erreichbar wäre und diese Orte oder Situationen vermieden oder nur ertragen werden mit der Befürchtung, panik- oder panikähnliche Symptome zu haben. Das Vermeidungsverhalten kann dabei so ausgeprägt sein, dass nur noch die eigene Wohnung als sicherer Platz erlebt wird. Ein Verlassen der Wohnung ist dann oft nur noch möglich bei Einbezug verschiedener Sicherheitsaspekte, wie z.B. die Begleitung des Partners oder das Mitführen von Medikamenten.

Liegt Agoraphobie vor, werden aber die Kriterien für eine Panikstörung nicht erfüllt, z.B. bzgl. Symptomanzahl oder Zeitverlauf, entspricht dies einer Agoraphobie ohne Panikstörung in der Vorgeschichte. In einer Studie der Arbeitsgruppe um Goisman

berichteten 57 % dieser Patienten über eingeschränkte Panikattacken, die Autoren schlussfolgern, dass eher eine quantitative denn qualitative Unterscheidung angemessen wäre (Goisman et al., 1995), andere Autoren kritisieren dagegen den engen diagnostischen Zusammenhang zwischen Agoraphobie und Panik, der daraus resultiert, dass auch bei Agoraphobie ohne Panikstörung in der Vorgeschichte Befürchtungen bzgl. panikähnlicher Symptome vorliegen müssen (Salkovskis & Hackmann, 1997). Die Kategorie gilt somit als sehr umstritten, eine eindeutige Entscheidung, ob es sich um eine eigenständige Störung handelt, steht noch aus. In der vorliegenden Arbeit werden alle Therapiestudien, die eine dieser Störungen untersuchten, zusammengefasst, da a) auch bei Panikstörung ohne Agoraphobie oft leichtes agoraphobisches Verhalten vorliegt (vgl. Brown & Barlow, 2002), b) in einigen Studien verschiedene Typen zusammen untersucht werden und c) bisherige aktuelle Meta-Analysen keinen Einfluss des Vermeidungsverhaltens auf die Therapieeffektivität fanden (z.B. Gould et al., 1995).

Die Lebenszeitprävalenz beträgt für Panikstörung mit und ohne Agoraphobie um 3 %, für Agoraphobie ohne Panikstörung in der Vorgeschichte um 5 % (Kessler et al., 1994). Allerdings gilt letztere Kategorie als sehr umstritten, finden sich doch diskrepante Befunde bzgl. Häufigkeit der Störung im Vergleich zu Panikstörung mit Agoraphobie abhängig von der untersuchten Stichprobe: in epidemiologischen Studien der Allgemeinbevölkerung eher hohe Häufigkeiten vs. in Studien von Klinikern an Patienten mit eher geringen Häufigkeiten (White & Barlow, 2002). Als mögliche Gründe für diese Unterschiede nennen die Autoren, dass Patienten mit Panik eher eine Behandlung suchen, dass klinische Stichproben häufiger komorbide Störungen haben bzw. dass in epidemiologischen Studien fälschlicherweise einige spezifische Phobien der Agoraphobie zugeordnet wurden. Wie für andere Angststörungen auch ergibt sich ein Geschlechterverhältnis von ca. 2:1 mit mehr Frauen, die eine dieser Störungen entwickeln (Kessler et al., 1994).

### **1.1.3 Posttraumatische Belastungsstörung (PTB)**

Nach extrem belastenden Erlebnissen, wie z.B. Vergewaltigung, Unfall oder Krieg, kann sich eine Posttraumatische Belastungsstörung (PTB) entwickeln. Sie definiert sich im Wiedererleben des Ereignisses, in einer Vermeidung traumarelevanter Reize bzw. einer Reduktion der emotionalen Reagibilität und in einer Erhöhung der Erregung (APA, 1994). Als wichtigstes Symptom zur Differenzierung zu anderen Störungen gilt dabei das intrusive Wiedererleben des Ereignisses im Wachzustand (Creamer, 93). Die Patienten können dabei empfinden, als würden sie das Trauma im Moment erfahren. Vorherrschend sind dabei sensorische Eindrücke, hauptsächlich visueller Natur (Ehlers & Steil, 1995), aber auch affektive und physiologische Reaktionen, die denen während des Ereignisses entsprechen. Es können Aspekte des Traumas wiedererlebt werden, für die den Patienten eine bewusste Erinnerung fehlt („Emotion ohne Erinnerung“: Ehlers, 1999, S. 16). Daneben treten häufig Alpträume auf. Bei Einschätzung der Intrusionen als sehr aversiv und stressreich, versuchen die Betroffenen durch Vermeidung traumarelevanter Reize eine Aktivierung der Erinnerungen zu unterdrücken (Ehlers et al., 1995). Dies kann z.B. der Ort des Geschehens sein, Gespräche über das Erlebte oder auch bestimmte Gerüche. Es wird versucht, das Erlebnis aus dem Gedächtnis und dem eigenen Leben quasi zu verbannen. Befürchtungen bestehen nicht nur bzgl. traumarelevanter Reize, sondern – ähnlich bei Patienten mit PDA – auch für Konsequenzen der Symptomatik selbst (z.B. Intrusionen als Zeichen des Verrücktwerdens). Daneben treten Symptome der emotionalen Taubheit auf, es liegt ein vermindertes Interesse an vormals wichtigen Dingen vor. Zusätzlich zeigen die Betroffenen Symptome der autonomen Übererregung, wie übertriebene Schreckreaktionen oder Aggressivität. Bei Dauer der Symptome von mehr als drei Monaten wird eine chronische PTB diagnostiziert (DSM-IV, APA, 1994)

Die Lebenszeitprävalenz liegt bei ca. 8 % (Kessler, Sonnega, Bromet, Hughes & Nelson, 1995). Obwohl mehr Männer ein traumatisches Ereignis erleben, entwickeln Frauen doppelt so häufig eine PTB. Die Ereignisse, die am häufigsten eine PTB auslösten, waren Vergewaltigung (sowohl bei Männern als auch Frauen), Kampfeinsätze (bei Männern) und Missbrauch in der Kindheit (sowohl bei Männern als auch Frauen). So entwickelten 65 % der Männer und 46 % der Frauen eine PTB nach einer Vergewaltigung.

#### **1.1.4 Soziale Phobie (SoP)**

Unter Sozialer Phobie versteht man eine Angst vor sozialen Situationen, in denen man mit Fremden konfrontiert oder einer Bewertung durch andere ausgesetzt ist, z.B. in der Kantine essen oder vor einer Menschengruppe sprechen (DSM-IV, APA, 1994). Es wird gefürchtet, in der Situation bestimmtes Verhalten bzw. Angstsymptome zu zeigen, was peinlich oder demütigend sein könnte. Die Situationen werden vorrangig vermieden oder nur unter starker Angst ertragen. Die Person selbst schätzt die Angst als übertrieben und unbegründet ein. Können die Situationen selbst nicht vermieden werden, werden oft verschiedene Strategien eingesetzt, von denen der Patient glaubt, dass diese eine Wahrnehmung durch andere als peinlich reduzieren können. Beispiele für solches Sicherheitsverhalten als Aspekt von Vermeidungsverhalten sind das Tragen von Jacken damit der Gesprächspartner nicht das Schwitzen bemerkt oder das mentale Proben von Sätzen bevor man sie ausspricht.

Das DSM-IV (APA, 1994) ermöglicht zusätzlich die Codierung einer generalisierten SoP, wenn in vielen verschiedenen sozialen Situationen Ängste auftreten. Es wird in der Literatur kontrovers diskutiert, wie die Subtypen angemessen operationalisiert werden können. Es liegen sowohl Klassifikationen bzgl. einer Anzahl an Situationen vor, z.B. wird oft der generalisierte Typus über mindestens drei Situationen definiert (n. Stangier & Heidenreich, 1997), die Arbeitsgruppe um Heimberg schlägt aufbauend auf der Anzahl sogar drei Kategorien vor: generalisiert, nicht-generalisiert und spezifisch (Heimberg, Holt, Schneier, Spitzer & Liebowitz, 1993). Andere Autoren befürworten dagegen eine Aufteilung nach gefürchtetem Situationstyp, z.B. Diagnose des generalisierten Typus bei Angst vor generellen Situationen, wie Gesprächen, nicht aber bei Angst vor einer oder mehreren Leistungssituationen (Turner, Beidel & Townsley, 1992).

Die Lebenszeitprävalenz beträgt für neuere epidemiologische Studien (ab DSM-III-R) zwischen 7 % und 16 % (Wittchen & Fehm, 2001). SoP kommt häufiger bei Frauen vor mit einem Verhältnis von ca. 3:2 (Kessler et al., 1994).

### **1.1.5 Spezifische Phobie (SpP)**

Spezifische Phobie ist nach DSM-IV (APA, 1994) definiert als intensive und anhaltende Furchtreaktion, die bei fast jeder Konfrontation mit einem bestimmten Stimulus ausgelöst wird. Die Symptome können dabei das Ausmaß einer Panikattacke annehmen. Die Furcht ist irrational und der Betroffene weiß, dass sie unbegründet oder exzessiv ist. Es wird versucht, die Situation zu vermeiden. Die Symptome liegen mindestens 6 Monate vor. Das DSM-IV unterscheidet nach Art des Stimulus verschiedene Subtypen der SpP: den Tier-Typus (z.B. Spinnen, Hunde), den Umwelt-Typus (z.B. Höhen, Dunkelheit), den Blut-Spritzen-Verletzungs-Typus (z.B. Blut sehen, eine Injektion erhalten), den Situativen Typus (z.B. Brücken, Fahrstühle) sowie den Anderen Typus (z.B. Pflanzen, laute Geräusche). Diese Aufteilung wird zum Teil kritisiert, zum einen ist eine Zuordnung nicht immer eindeutig (z.B. ist die Angst vor Brücken tatsächlich ein Situativer Typus oder doch ein Umwelt-Typus bezogen auf den Aspekt der Höhe?), zum anderen erscheint der Nutzen dieser Klassifikation gegenüber einer einfachen Nennung des gefürchteten Stimulus gering (Antony & Barlow, 2002).

Eine häufige Annahme besteht in einer nur geringen Schwere der SpP im Vergleich zu anderen psychischen Störungen. Dies ist abhängig vom gefürchteten Stimulus, z.B. Spinnen vs. öffentliche Transportmittel. Die Störung sollte aber nicht bagatellisiert werden. So kann auch SpP zu einer bedeutsamen Beeinträchtigung des sozialen und beruflichen Lebens führen (z.B. Magee, Eaton, Wittchen, McGonagle & Kessler, 1996). Die Lebenszeitprävalenz beträgt um 11 % (Kessler et al., 1994). Die am häufigsten gefürchteten Stimuli sind Tiere und Höhen mit jeweils ca. 20 % (Curtis, Magee, Eaton, Wittchen & Kessler, 1998). Frauen leiden häufiger an dieser Störung als Männer mit einem Geschlechterverhältnis von 2:1, welches in mehreren epidemiologischen Studien gefunden werden konnte (vgl. z.B. Fredrikson, Annas, Fischer & Wik, 1996). Spezifische Phobie ist nur selten bei Vorliegen komorbider Störungen die Hauptdiagnose, häufiger zusätzliche Diagnose (Antony et al., 2002).

### **1.1.6 Zwangsstörung (OCD)**

Als Zwangsstörung bezeichnet man wiederkehrende Zwangsgedanken oder Zwangshandlungen, bei denen die Person erkennt, dass sie übertrieben oder unbegründet sind (DSM-IV, APA, 1994). Zwangsgedanken sind anhaltende Gedanken, Impulse oder



Vorstellungen, die als fremd und unangemessen wahrgenommen werden und ausgeprägte Angst oder Leiden verursachen. Der Patient versucht, diese zu ignorieren, zu unterdrücken oder mit anderen Gedanken oder Handlungen zu neutralisieren. Zwangsgedanken stellen dabei nicht übermäßige Besorgtheit gegenüber alltäglichen Problemen dar, wie bei der GAS. Ein Beispiel sind Gedanken einer Mutter, ihr Kind töten zu können. Zwangshandlungen sind rigide und stereotype Verhaltensweisen oder geistige Handlungen (z.B. Zählen) zur Reduktion verbundenen Unwohlseins oder zur Verhinderung möglicher negativer Konsequenzen.

Die Lebenszeitprävalenz liegt bei ca. 2 % bis 3 %, der Frauenanteil ist leicht höher (n. Bebbington, 1998). Die häufigsten Untergruppen sind Kontrollzwänge und Waschzwänge (z.B. Hodgson & Rachman, 1977). Kontrollzwänge beziehen sich auf Befürchtungen vor Schaden durch eingeschaltete Haushaltsgeräten (z.B. Herd, Bügeleisen), offene Türen oder Fenster nach Verlassen der Wohnung oder Unachtsamkeit im Straßenverkehr, so wird z.B. eine Strecke wiederholt abgefahren, zur Kontrolle niemanden verletzt zu haben. Bei Waschzwängen liegen Befürchtungen vor bzgl. einer Ansteckung mit Krankheiten oder Verunreinigungen, z.B. durch Händeschütteln oder das Anfassen von Türklinken. Intensives, zeitaufwendiges Waschen ist die Folge, wobei oft eine Ritualisierung des Waschvorgangs erfolgt, zur Sicherheit, dass Sauberkeit erreicht wird.

## **1.2 Therapieansätze**

### **1.2.1 Behaviorale und kognitiv-behaviorale Therapien (VT, KVT)**

Grundlage der hier untersuchten psychologischen Therapien bilden behaviorale und kognitive Modelle der Angst. Da der Schwerpunkt der vorliegenden Arbeit auf einer Anwendung der Meta-Analyse für die Therapieforschung bei Angststörungen liegt, wird auf eine Beschreibung der Modelle verzichtet. Diskutiert werden vornehmlich konfrontative und kognitive Verfahren. Die Wirksamkeit konfrontativer Therapien ergibt sich nach den Modellen aus a) einer Aktivierung und Veränderung der angstrelevanten Gedächtnisstrukturen aufgrund Habituation und/oder Löschung und/oder ein Überprüfen von Annahmen und somit eine Änderung der Outcome-Erwartungen (z.B. Beck & Clark, 1996; Foa & Kozak, 1986), b) einer angemessenen Auswahl an Coping-

verhalten (z.B. Beck et al., 1996), c) einer Steigerung der Selbstkontrolle und somit einer Änderung der Effektivitäts-Erwartungen (Bandura, 1977) und d) einer Aufmerksamkeitsfokussierung (Marks & Dar, 2000). Daneben ermöglichen kognitive Verfahren zusätzlich e) eine direkte Veränderung von kognitiven Strukturen einschließlich der Fehlannahmen und Überprüfung adäquater Annahmen (z.B. Beck et al., 1996; Brewin, 1996) bzw. in Terminologie behavioraler Modelle eine Neubewertung des UCS (Davey, 1992). Es sei allerdings darauf verwiesen, auch hinsichtlich Pharmakotherapien, dass die Ergebnisse der folgenden Arbeit nicht dazu geeignet sind, die Angemessenheit der zugrunde liegenden Modelle zu evaluieren.

Im Folgenden sollen die am häufigsten in den untersuchten Studien eingesetzten Verfahren bzw. Komponenten kurz beschrieben werden.

### **Systematische Desensibilisierung (SD)**

Im Rahmen der SD erfolgt eine Kopplung von gering angstauslösenden Reizen mit einer mit Angst unvereinbaren Reaktion, in der Regel Entspannung (Wolpe, 1958). Die Patienten führen ein Entspannungstraining durch und stellen eine Angsthierarchie auf, zur Einteilung gefürchteter Situationen abhängig von ihrer Bedrohlichkeit. Während der SD soll sich der Patient in einen Entspannungszustand versetzen und es folgt eine graduierte Konfrontation mit den angstauslösenden Reizen meist in der Vorstellung, beginnend mit einem nur gering bedrohlichen Stimulus. Dies wird wiederholt, bis der Patient keine Angst mehr empfindet, dann wird zur nächsten Situation auf der Angsthierarchie übergegangen.

### **Angewandte Entspannung (AE)**

Ziel des Verfahrens ist die Anwendung der Entspannung, wenn der Patient frühe Symptome der Angstsymptomatik erkennt (Öst, 1987). Der Patient lernt, mit der Angst umzugehen. Durch Selbstbeobachtung lernt der Patient, verschiedene Anzeichen von Angst, z.B. zunehmender Herzschlag, in unterschiedlichen Situationen zu identifizieren. Nach einem Entspannungstraining, mit Ziel einer schnellen, in verschiedensten Situationen herbeigeführten Entspannung, erfolgen Übungen mit graduierter Konfrontation in sensu und in vivo. Der Patient setzt in diesen Situationen seine Entspannungsübungen ein.

**Expositionsbehandlung**

Unter Exposition versteht man ein Aufsuchen genau der Situationen, die für den Patienten bedrohlich sind (Hand, 1996). Die Konfrontation kann dabei in der Vorstellung (in sensu) oder in der Realität (in vivo) erfolgen. Der Patient bleibt solange in der Situation, bis eine deutliche Abnahme der Angst eingetreten ist. Die Auswahl der Situation kann dabei nach einem graduierten Vorgehen (d.h. einer langsamen Steigerung der Bedrohlichkeit und damit des Schwierigkeitsgrades) oder nicht-graduierten Vorgehen (d.h. früh in der Therapie stark gefürchtete Situationen) erfolgen. Auch Systematische Desensibilisierung und Angewandte Entspannung sind konfrontative Methoden, ist im Folgenden von Exposition die Rede, werden diese Verfahren nicht einbezogen.

**Training Sozialer Kompetenz (TSK)**

Ziel ist eine Steigerung verbaler und nonverbaler sozialer Fertigkeiten (z.B. Beginn eines Gespräches, Augenkontakt halten). Komponenten sind Modellernen vom Therapeuten oder anderen Patienten bei Gruppentherapien, Rollenspiele verschiedener sozialer Situationen, Feedback über positives und negatives Verhalten auch unter Einsatz moderner Medien, wie z.B. Videoaufnahmen, sowie eine Erprobung des gelernten Verhaltens in der Realsituation. Neuere Konzepte enthalten zudem kognitive Elemente, wie Kognitive Restrukturierung (Heimberg et al., 1990).

**Kognitive Restrukturierung**

Zur Identifikation, Überprüfung der Angemessenheit und Veränderung inadäquater Kognitionen liegt eine Vielzahl an Techniken vor (Beck, Emery & Greenberg, 1985), z.B. Entkatastrophisieren oder das Überprüfen von Evidenzen. Zusätzlich werden oft Verhaltensexperimente durchgeführt, in denen die Genauigkeit von Vorhersagen überprüft wird. Dieses Verfahren enthält meist konfrontative Elemente mit den angstausslösenden Stimuli.

### 1.2.2 Pharmakotherapie (PhT)

Ziel der Pharmakotherapie ist eine Regulation des Neurotransmitterhaushaltes einschließlich der beteiligten Rezeptoren. Es werden verschiedene Substanzklassen unterschieden, eine Übersicht bietet z.B. Klerman et al. (1994), die im Folgenden kurz beschrieben ist. Die Klassifikation der Substanzen bezieht sich dabei zum einen auf deren Wirksamkeit bei verschiedenen Störungsbildern (z.B. Antidepressiva, Anxiolytika, Antipanische Medikamente sowie Medikamente gegen Zwangsgedanken). Dies basiert allerdings auf Ergebnissen empirischer Effektivitätsstudien, die zum Teil die Stichprobe der vorliegenden Arbeit bilden. Ein Erstellen von Hypothesen würde somit an der gleichen Stichprobe erfolgen wie ein Testen der Hypothesen. Zum anderen erfolgt die Klassifikation nach der unterschiedlichen Wirkungsweise der Substanzen. Die am häufigsten in den vorliegenden Studien in ihrer Wirkung untersuchten Substanzklassen sind:

a) Benzodiazepine (BZD)

Zur Gruppe der Anxiolytika und Antipanischen Medikamente gehörend. Die Hauptwirkung scheint eine Beeinflussung des hemmenden Neurotransmitters Gamma-Aminobuttersäure (GABA) zu sein im Besonderen in limbischen Bereichen des Gehirns. Beispiele sind Alprazolam, Diazepam und Lorazepam.

b) Selektive Serotoninwiederaufnahmehemmer (SSRI)

Zur Gruppe der Antidepressiva und Medikamente gegen Zwangsgedanken gehörend. Die primäre Wirkung scheint eine durch Hemmung der Serotoninwiederaufnahme erhöhte synaptische Konzentration von Serotonin im Besonderen im Raphe Nucleus zu sein. Beispiele sind Fluoxetin, Fluvoxamin und Paroxetin.

c) Tricyclische Antidepressiva (TCA)

Zur Gruppe der Antidepressiva, antipanischen und Medikamente gegen Zwangsgedanken gehörend. Die Wirkung ist komplex und nicht eindeutig, Hauptwirkung scheint eine Hemmung der Wiederaufnahme verschiedener Neurotransmitter wie Noradrenalin und Serotonin zu sein. Beispiele sind Clomipramin und Imipramin.

### 1.3 Stand der Forschung und Kritik

Es liegen schon mehrere Meta-Analysen vor, die die Wirksamkeit einer oder mehrerer Therapieformen bei einer Angststörung betrachteten. Alle Arbeiten belegen die Wirksamkeit der behavioralen, kognitiv-behavioralen oder pharmakologischen Therapie, wenn auch im unterschiedlichen Ausmaß. Bzgl. Unterschiede zwischen den Therapieansätzen, im Besonderen zwischen Psycho- und Pharmakotherapie, ergaben sich divergente Befunde.

Die methodische Durchführung der bisherigen Meta-Analysen erscheint allerdings zum Teil sehr problematisch, so dass sich bedeutsame Probleme in der Interpretation der Ergebnisse zeigen<sup>2</sup>. Circa die Hälfte der Meta-Analysen setzten Prä-Post-Effektstärken ein. Diese überschätzen jedoch die wahre Wirksamkeit einer Therapie. Die Berechnung der mittleren Effektstärke erfolgte nur bei ca. einem Drittel der bisherigen Meta-Analysen mit einer Gewichtung über die Stichprobengröße. Dies reduziert aber den Einfluss des Stichprobenfehlers. Studien mit einem geringen Stichprobenumfang haben eine geringere Genauigkeit der Ergebnisse. Zentrales Problem fast aller bisherigen Meta-Analysen mit Ausnahme zweier ist das zugrunde liegende statistische Modell. Bisher fand meist das Fixed Effects-Modell Verwendung. Dies erlaubt aber keine Generalisierbarkeit der Ergebnisse über die eingeschlossenen Primärstudien hinaus. Das heißt, dass nach diesem Modell keine Aussagen zu treffen sind hinsichtlich einer allgemeinen Wirksamkeit einer Therapieform bei einer Angststörung, sondern nur Aussagen darüber, wie wirksam die Therapie in den bisherigen Primärstudien war. Zudem ist für dieses Modell ein Homogenitätstest unabdingbar für Aussagen darüber, ob die gefundenen Effektstärken der Studien Ausprägungen eines gemeinsamen Wertes sind und somit die mittlere Effektstärke als Populationseffektstärke interpretiert werden darf. Über die Hälfte der Meta-Analysen mit Fixed Effects-Modell überprüften allerdings nicht die Homogenität, von denen, die eine Berechnung durchführten, ignorierte zudem ca. ein Drittel ein signifikantes Ergebnis. Es sind somit oft keine Aussagen möglich, inwieweit die berechneten Effektstärken dem Populationseffekt entsprechen. Annähernd die Hälfte der bisherigen Meta-Analysen untersuchte keine Prädiktorvariablen, wurde

---

<sup>2</sup> Zum Verständnis der Kritikpunkte sei auf Kapitel 2 verwiesen, in dem eine Diskussion der Methode der vorliegenden Arbeit erfolgt.

der Einfluss von Moderatoren auf die Effektstärke untersucht, geschah dies meist mittels univariater Verfahren. Auch Unterschiede zwischen verschiedenen Therapieformen wurden überwiegend ohne Kontrolle methodischer Variablen durchgeführt, so dass unklar ist, ob die bisherigen Befunde auf die Effektivität einer Therapie oder möglicherweise auf unterschiedliche Eigenschaften der Studien zurückzuführen sind. Zudem wurden in über drei Viertel der Meta-Analysen zur Bestimmung der Bedeutsamkeit des mittleren Effektes bzw. hinsichtlich des Einflusses von Moderatoren standardstatistische Verfahren eingesetzt. Diese sind wegen unterschiedlichen Stichprobengrößen zu kritisieren, welche zu heterogenen Varianzen führen (Hedges & Olkin, 1985). In einigen Meta-Analysen wurde der Aspekt berücksichtigt durch Einsatz von Verfahren, die eine Homogenität dieser nicht als Voraussetzung haben. Eine Gewichtung mit der Stichprobengröße wurde dagegen weitestgehend ignoriert. Ein weiterer Kritikpunkt sind fehlende Aussagen zur Stabilität der Ergebnisse. Sensitivitätsanalysen wurden bisher nicht durchgeführt. Diese ermöglichen jedoch eine Einschätzung des Einflusses, den der Wissenschaftler durch seine Entscheidungen auf das Resultat der durchgeführten Meta-Analyse hat. Zudem fehlten in ca. drei Viertel der Arbeiten Diskussionen über einen möglichen Publikationsbias mittels statistischer Verfahren, wie z.B. das Fail-Safe N.

Insgesamt ist festzustellen, dass die Validität der bisherigen Meta-Analysen eingeschränkt ist. Meist werden nur die grundlegenden Berechnungsmöglichkeiten durchgeführt und Besonderheiten der meta-analytischen Integration missachtet. Eine Darstellung der Methode sowie ein Überblick über die Ergebnisse bisheriger Meta-Analysen finden sich in Anhang A.

## 1.4 Fragestellungen

In der vorliegenden Arbeit sollen folgende Fragestellungen untersucht werden:

1. Überprüfung der unmittelbaren Effektivität von behavioraler, kognitiv-behavioraler und pharmakologischer Therapie bei den einzelnen Angststörungen einschließlich folgender Fragestellungen:
  - a. Sind behaviorale und kognitiv-behaviorale Verfahren ebenso effektiv wie die häufiger eingesetzte Pharmakotherapie?
  - b. Sind konfrontative Verfahren ausreichend zur Therapie der Angststörungen oder ist die kognitiv-behaviorale Behandlung effektiver?
  - c. Ist eine Kombination aus Pharmako- und Psychotherapie effektiver als Psychotherapie allein?
2. Überprüfung der zeitlichen Stabilität der Effekte nach Beendigung der Therapie
3. Überprüfung der Akzeptanz der untersuchten Therapieformen operationalisiert über die Anzahl an Patienten, die vorzeitig die Studie verlassen<sup>3</sup>
4. Überprüfung des Einflusses methodischer Variablen

---

<sup>3</sup> Es sei anzumerken, dass nicht bei allen Abbrechern die Akzeptanz für die Therapie gering ist im Sinne eines unangemessenen Kosten-Nutzen-Verhältnisses aufgrund mangelnder Wirksamkeit der Therapie oder zu hoher Nebenwirkungen, sondern auch äußere Umstände, wie z.B. Umzug, ein Ausschluss durch den Wissenschaftler, z.B. aufgrund geringer Kooperation, oder, bei Psychotherapie, eine schon frühe sehr ausgeprägte Symptomverbesserung zum Verlassen der Studie führen können. Dennoch erscheint aufgrund der vorliegenden Daten dies die angemessenste Operationalisierung.

## 2 Methode

### 2.1 Die Meta-Analyse als integrierendes Verfahren

Unter Meta-Analysen werden Verfahren zur statistischen Integration von relevanten Einzelergebnissen zusammengefasst. Weisen mehrere Studien in einem Themengebiet unterschiedliche Ergebnisse auf, sinkt durch die steigende Komplexität die Möglichkeit, eindeutige Aussagen zu treffen. Durch eine Zusammenfassung der verschiedenen Resultate wird eine Strukturierung versucht. Ein besonderer Vorteil integrativer Verfahren liegt in der Generalisierbarkeit der Ergebnisse über verschiedene Stichproben, Behandlungsbedingungen oder Erhebungsverfahren: „strong generalizations are rarely – if ever – possible based on single studies. Instead, generalizations are best justified by programmatic reviews of findings from many studies“ (Matt, in press, S.2). Im Gegensatz zu anderen integrativen Verfahren, wie narrative Literaturübersichten oder die Box-Score-Methode (das Auszählen von Signifikanzen), erfolgt im Rahmen von Meta-Analysen eine quantitative Zusammenfassung der statistischen Kennwerte der vorliegenden Studien. Dies erhöht die Objektivität der Ergebnisse: Meta-Analysen sind im Vergleich zu anderen integrativen Verfahren weniger abhängig von der Power<sup>4</sup> der eingeschlossenen Studien, da primär die Berechnung der Kennwerte einer Meta-Analyse (Effektstärken; ES; als Maß der Effektivität) über Mittelwerte und Standardabweichungen<sup>5</sup> erfolgt. Eine vollständige Unabhängigkeit von der Power der Primärstudien ist auch bei Meta-Analysen nicht gegeben, z.B. im Hinblick auf den Publikationsbias sowie fehlender Angaben in Primärstudien (Mittelwerte und Standardabweichungen), die eine Berechnung über Kennwerte statistischer Tests notwendig machen. Die Objektivität der Ergebnisse wird zudem erhöht durch explizite Regeln für die Integration und dadurch einer Reduktion des Einflusses des Wissenschaftlers auf die Ergebnisse im Vergleich zu den anderen genannten Verfahren. Auch eine Bereinigung der Resultate von Störfaktoren kann in anderen integrativen Methoden nicht durch-

---

<sup>4</sup> Power ist die Wahrscheinlichkeit des Zurückweisens der Nullhypothese (z.B. Es besteht kein Unterschied), wenn diese falsch ist. Je höher die Power ist, desto wahrscheinlicher wird ein Unterschied festgestellt, wenn ein wahrer Unterschied besteht. Die Höhe der Power ist abhängig von drei Faktoren: der Effektstärke, der Stichprobengröße und der Irrtumswahrscheinlichkeit. Die Power sinkt mit geringerem Effekt, kleinerer Stichprobe und strengerem Alphafehlerniveau (Cohen, 1988).

<sup>5</sup> Eine Berechnung über Korrelationen spielt im Bereich der Psychotherapieforschung keine Rolle.



geführt werden, für Meta-Analyse liegen dagegen statistische Möglichkeiten vor, z.B. hinsichtlich des Stichprobenfehlers.

Meta-Analysen ermöglichen Aussagen über die Größe eines Effektes anstatt nur darüber, ob ein Effekt vorhanden ist, sowie Aussagen über den Einfluss von Variablen auf diesen Effekt, die zum Teil in Primärstudien nicht oder nur schwer untersucht werden können. Eigenschaften der vorliegenden Untersuchungen werden zu Prädiktoren des Therapieerfolges, z.B. Therapiedauer oder Abbrecherquote. Lipsey und Wilson (2001) sprechen in diesem Zusammenhang von der Meta-Analyse als „... a quasi-experimental alternative to the experiment on experiments“ (S. 414). Meta-Analysen erlauben somit auch Einblicke in den Zusammenhang zwischen Entscheidungen der Wissenschaftler der Primärstudien, z.B. hinsichtlich des eingesetzten Designs und dem Ergebnis dieser Untersuchungen, als klassisches Beispiel ob die Ergebnisse randomisierter Effektivitätsstudien denen nicht-randomisierter gleichen. Meta-Analysen gehören somit zu den stärksten und bedeutsamsten Verfahren, die einem Wissenschaftler zur Verfügung stehen.

## **2.2 Selektion der Primärstudien**

### **2.2.1 Literatursuche**

Es wurden verschiedene Quellen zur Identifikation relevanter Literatur genutzt. Primär erfolgte die Suche mittels Datenbanken. Dabei wurden Psyclit, Psyn dex und Medline ab des ersten jeweils verfügbaren Jahres bis Mai 2002 nach folgenden Stichwörtern durchsucht: „treatment“, „\*therapy“, „\*phob\*“, „panic“, „obsessive compulsive“, „posttraumatic“, „generalized anxiety“, „anxiety“. Eine zweite Methode stellte die Verwendung von Verweisen aus existierenden Zusammenfassungen und empirischen Studien dar. Zusätzlich wurde eine Handsuche ab 1997 in folgenden Zeitschriften durchgeführt: American Journal of Psychiatry, Archives of General Psychiatry, Behavior Therapy, Behaviour Research and Therapy, British Journal of Psychiatry, Clinical Psychology Review, Journal of Abnormal Psychology, Journal of Consulting and Clinical Psychology, Verhaltenstherapie und der Zeitschrift für Klinische Psychologie. Eine systematische Suche in Konferenzbänden entfiel aufgrund mangelnder Verfügbarkeit. Es wurde im Internet mit Hilfe von Suchmaschinen, hauptsächlich Google, recherchiert.

Abschließend wurde Kontakt mit Wissenschaftlern, Forschungszentren und Pharmafirmen aufgenommen.

Die letzten beiden Strategien dienten der Reduktion des "file-drawer"-Problems (vgl. z.B. Beelmann et al., 1994; Sharpe, 1997). Meta-Analysen können nur Ergebnisse derjenigen Studien integrieren, die ihnen zur Verfügung stehen, in der Regel sind dies publizierte Studien. Ist eine Generalisierung der so gewonnenen Aussagen auf die Forschungswelt insgesamt, also einschließlich der Untersuchungen „in Schubladen“ möglich? Kritiker stützen sich auf die Annahme, dass unpublizierte Studien im Vergleich zu publizierten Studien geringere ESn aufweisen. Es liegen indirekte Belege dazu vor (n. Sohn, 1996): in einer Auswahl an Journalen konnte festgestellt werden, dass mehr als 90 Prozent der Artikel signifikante Resultate zeigten, teilweise dadurch erklärbar, dass Forscher achtmal häufiger bereit sind, Artikel mit signifikanten Ergebnissen bei Zeitschriften einzureichen. Lipsey und Wilson (1993) fanden in einer Meta-Meta-Analyse, d.h. einer statistischen Integration von Ergebnissen bisheriger Meta-Analysen, eine höhere mittlere ES für publizierte Studien von 0.53 im Vergleich zu einer ES von 0.39 für unpublizierte Studien an 92 Meta-Analysen. Lösungen des Problems scheinen kaum gegeben, da auch trotz intensiver Suche nicht jede durchgeführte Untersuchung zu finden ist und ebenfalls nicht eingeschätzt werden kann, wie viele Studien tatsächlich noch in Schubladen liegen. Ein Wandel in der Bedeutung des 5 %-Niveaus, wie von einigen Autoren gefordert (z.B. Cohen, 1994) könnte zu einer Verringerung dieses Problems führen.

### **2.2.2 Codierungskatalog**

Ziel war die Erstellung eines Codierungskataloges zur Erfassung methodischer und inhaltlicher Variablen, die einen Einfluss auf die Validität der Studie hatten oder der Beschreibung des Designs dienten bzw. die zum Ausschluss der Untersuchung führten. Dabei wurde sich auf das System von Cook und Campbell (1979; vgl. Kapitel 2.2.4) bezogen. Grundlage zur Formulierung der einzelnen Items bildeten dabei die Vorschläge von Grawe, Donati und Bernauer (1994) und Wortman (1994). Auf eine Codierung

von Items von „high-inference“<sup>6</sup> wurde zur Erhöhung der Objektivität der Ergebnisse verzichtet. Die Codierung der Studien erfolgte in der folgenden Arbeit durch die Autorin selbst sowie mehrere Studentinnen im Rahmen ihrer Diplomarbeit. Für Items von „high-inference“ sind aber mehrere methodische Experten, d.h. Wissenschaftler mit mehrjähriger Erfahrung im Bereich der Therapieforschung bzw. der psychologischen Methodenlehre vorzusetzen, die das Rating blind von den Ergebnissen durchführen (Hall, Tickle-Degnen, Rosenthal & Mosteller, 1994). Der vollständige Codierungskatalog befindet sich in Anhang B.

Da die Codierung durch mehrere Rater erfolgte, wurde ein intensives Training durchgeführt bis die Interrater-Reliabilität gemittelt über alle Variablen größer .75 und bei keiner Variable unter .50 war. Nach Fleiss (1981) gilt eine Interrater-Reliabilität größer .75 als exzellente Übereinstimmung und größer .40 als Übereinstimmung, die gut bzw. gut genug ist. Die Interrater-Reliabilität wurde bestimmt über eine Codierung von 10 zufällig ausgewählten Primärstudien. Bei kontinuierlichen Variablen wurde die Reliabilität erhoben über die Korrelation nach Pearson, bei kategorialen Variablen über Kappa mittels SPSS (SPSS inc., Version 10). Die in der vorliegenden Arbeit untersuchten Variablen wurden vollständig durch die Autorin codiert.

Der Codierungskatalog wurde als computergestützte Version erstellt, so dass die Codierung über Internet Explorer (Microsoft) erfolgen konnte. Vorteil war die Minimierung potentieller Fehler, die bei einer manuellen Übertragung einer Papierversion in ein datenverarbeitendes Programm entstehen können. Nach einem ersten Test an mehreren zufällig gezogenen Studien erfolgte eine Modifikation des Kataloges. Modifikationen des Kataloges während der Codierung betrafen die Hinzunahme weiterer Behandlungsgruppen und Behandlungsverfahren. Die zur Berechnung der Effektstärken relevanten Maße wurden direkt mittels des Berechnungsprogramms erfasst (Study Input, Mitte & Schäfer, 2003, vgl. Anhang C).

---

<sup>6</sup> Dies sind Informationen, die nicht direkt in den Primärstudien zu finden sind, sondern eine globale Bewertung von Studieneigenschaften benötigen, z.B. eine Einschätzung der internen Validität.

### 2.2.3 Untersuchungsdesigns der Primärstudien

Die Berechnung der ES<sub>n</sub> kann entweder als Prä-Post-ES erfolgen oder im Rahmen eines kontrollierten Designs als ES zwischen Gruppen. In beiden Fällen ergeben sich sowohl Vor- als auch Nachteile.

Die fundamentalen Arbeiten von Campbell und Stanley (1963) bzw. in einer Weiterentwicklung von Cook et al. (1979) legen potentielle Schwierigkeiten bei einem unkontrollierten Design dar. Das Hauptproblem der Prä-Post-ES<sub>n</sub> ergibt sich aus ihrer eingeschränkten internen Validität, d.h. dass Veränderungen in der Treatment-Gruppe nicht eindeutig auf die Behandlung zurückgeführt werden können. Mögliche alternative Erklärungen, die durch den Einsatz einer Kontrollgruppe ausgeschlossen werden können, beziehen sich auf den Retest-Effekt<sup>7</sup>, d.h. Faktoren, die durch wiederholte Messungen auftreten, wie z.B. eine Beeinflussung der Ergebnisse zu Zeitpunkt 2 durch Bearbeiten des Fragebogens zu Zeitpunkt 1 oder auf zeitliche Ereignisse einschließlich Spontanremission, Erwartungen oder Regression zur Mitte: „all sorts of interpretations other than treatment effects can explain the pre- to posttreatment changes“ (Kazdin, 1999, S. 11). Hiermit soll nicht Hunter und Schmidt (1990) widersprochen werden, dass auch Studien mit Prä-Post-Designs existieren, die keine Verletzung der Validität aufweisen. Allerdings zeigen bisherige Arbeiten, dass die Prä-Post-ES<sub>n</sub> von Behandlungsgruppen nicht ausschließlich auf die durchgeführte Behandlung zurückzuführen sind, da auch Wartelistenpatienten eine systematische Verbesserung in ihrer Symptomatik haben. So untersuchte Arrindell (2001) Veränderungen bei Patienten in einer Wartelistenbedingung. Er konnte an zwei verschiedenen Stichproben (1. ambulante Patienten mit PDA und 2. stationäre Patienten mit PDA, OCD oder Depression) zeigen, dass es einen systematischen positiven Effekt über die Zeit gab, d.h. nach der Warteperiode wiesen die Patienten eine geringere Symptomatik auf als bei Beginn der Untersuchung. Dies konnte über fast alle eingesetzten Messinstrumente gezeigt werden, bedeutsame Effekte waren ausschließlich positiv<sup>8</sup>. Allerdings war die Höhe der Prä-

---

<sup>7</sup> Als Retest-Effekt werden im Allgemeinen nur Veränderungen verstanden aufgrund der Erhebung zu Zeitpunkt 1 und nicht Veränderungen z.B. durch geschichtliche Ereignisse. Zur Vereinfachung wird der Retest-Effekt im Folgenden definiert über alle Veränderungen von Zeitpunkt 1 zu Zeitpunkt 2.

<sup>8</sup> Neben den von Cook et al. (1979) dargelegten potentiellen Einflussfaktoren diskutiert Arrindell (2001) speziell Ursachen der Beeinflussung der Ergebnisse zu t<sub>2</sub> aufgrund der Bearbeitung des Fragebogens zu

Post-ESn sowohl abhängig von den jeweiligen Messinstrumenten (Streuung der ES in Stichprobe 1 von Hedges  $g = 0.15 - 0.36$ , in Stichprobe 2 von Hedges  $g = -0.01 - 0.32$ ) als auch von der Stichprobe bei Betrachtung eines Instrumentes (z.B.: SCL90R Subskala Angst: Stichprobe 1: Hedges  $g = 0.30$  vs. Stichprobe 2: Hedges  $g = 0.15$ ). Auch bisherige Meta-Analysen zeigten mittlere positive Effekte für Wartelistenbedingungen, z.B.  $ES = 0.44$  für PTB-Symptomatik bei Patienten mit PTB (Van Etten et al., 1998), andere Arbeiten dagegen geringere ES von  $ES = 0.03$  für die Gesamtsymptomatik bei Selbstbeurteilungsinstrumenten bei SoP (Federoff et al., 2001). Auch zwischen den Meta-Analysen zeigt sich somit, je nach Methode und eingeschlossenen Studien und somit Störungen, Stichproben bzw. Messinstrumenten, ein unterschiedliches Ausmaß der Veränderung bei Wartelistenpatienten. Dies hat Implikationen für die Durchführung von Meta-Analysen. Die ausschließliche Betrachtung von Prä-Post-ESn der Behandlungsgruppe überschätzt den wahren Effekt in unbekannter Höhe. Eine Alternative bietet die von Becker (1988) vorgeschlagene Differenz der mittleren Prä-Post-ESn von Behandlungs- und Kontrollgruppen aus den verschiedenen Studien als Möglichkeit der Kombination von kontrolliertem und unkontrolliertem Design. Allerdings ist hier, aufbauend auf den Ergebnissen von Arrindell (2001), eine ausreichend hohe Anzahl an Studien mit Wartelistendesign (bzw. bei zusätzlicher Betrachtung unspezifischer Faktoren mit einer Placebobedingung) notwendig zur Berechnung eines angemessenen Schätzers des Retest-Effektes, da eine Vielzahl verschiedener Messinstrumente an verschiedenen Stichproben einbezogen werden muss<sup>9</sup>. Bei einer nur geringen Anzahl an Studien mit Wartelistendesign erlaubt auch dieses Verfahren keine angemessene Berechnung einer von Fehlern durch wiederholte Messung bereinigten Effektstärke. Aufbauend auf den bisherigen Ausführungen wurden in der vorliegenden Arbeit ausschließlich Studien betrachtet, die einem kontrollierten Design entsprechen.

---

t1. Das Ergebnis widerspricht übrigens einer häufig zu findenden Annahme (z.B. Basham, 1986), dass bei Patienten der Wartelistenbedingung aufgrund von Demoralisierung eine Verschlechterung der Symptomatik gegeben sei.

<sup>9</sup> Obwohl verschiedenste Bemühungen bestehen, einen goldenen Standard für den Einsatz von Messinstrumenten bei Therapieforschung einzuführen (z.B. Shear & Maser, 1994), ist dennoch von einer Heterogenität diesbezüglich auszugehen.

Allerdings ergeben sich auch in Bezug auf kontrollierte Studien Probleme. Neben ethischen Bedenken bei Zuteilung von Patienten auf Wartelistenbedingung (man enthält ihnen über einen bestimmten Zeitraum eine wirksame Behandlung vor), sind spezielle meta-analytische Probleme vorhanden. Der Ausschluss unkontrollierter Studien führt zu einem starken Absinken der Anzahl an Untersuchungen und möglicherweise zu einer Einschränkung der Repräsentativität der Ergebnisse. Eine weitere Kritik an ESN zwischen Gruppen sind mögliche Unterschiede zwischen den Gruppen zu Prä, so dass die ESN zu Post nicht interpretiert werden können. Bisher wurde dieser Aspekt im Rahmen von Meta-Analysen vor allem bzgl. der Frage nach einer Vergleichbarkeit von randomisierten vs. nicht-randomisierten Untersuchungen überprüft (z.B. Shadish, Matt, Navarro & Philipps, 2000). In der vorliegenden Meta-Analyse wurde unabhängig von der Art der Zuweisung der Patienten auf die verschiedenen Bedingungen zusätzlich eine ES zwischen Gruppen zu Prä berechnet, um die Baseline-Vergleichbarkeit der Gruppen zu kontrollieren. Aufbauend auf der stochastischen Theorie ist die Qualität der Randomisierung abhängig von der Anzahl an Ziehungen, d.h. je größer die Anzahl an Personen einer Studie ist, desto wahrscheinlicher ist eine Gleichverteilung von (Stör)Variablen auf die Gruppen: „Randomisierung garantiert lediglich asymptotisch (für  $N \rightarrow \infty$ ) Unabhängigkeit zwischen Treatmentbedingungen und Störvariablen... Randomisierung führt nicht zwangsläufig zur Elimination von Störfaktoren in endlichen Stichproben“ (Erdfelder, 1983, S. 85). Dies bedeutet allerdings, dass bei einer nur geringen Anzahl an Personen (wie häufig in der Therapieforschung vorzufinden) auch bei Einsatz von Randomisierung Unterschiede zwischen den Gruppen zu Prä möglich sind. Hsu (1989) und Kraemer (1981) kommen über Simulationen bzw. Erfahrungen aus klinischer Forschung zu dem Schluss, dass ca. ab 20 Personen pro Gruppe Randomisierung zu adäquaten Ergebnissen führt, eine Stichprobengröße, die oft nicht in (Psycho)Therapiestudien zu finden ist. Eine Kontrolle der Vergleichbarkeit der untersuchten Gruppen zu Prä sollte somit auch in Meta-Analysen erfolgen<sup>10</sup>. Ein weiteres Problem von Prä-Post-ESN, diskutiert von Hunter et al. (1990), ist, dass im Rahmen von Meta-Analysen eine potentielle Behandlungs-Patienten-Interaktion, d.h. dass die

---

<sup>10</sup> Es wurde für jede Studie codiert, ob eine Überprüfung der Vergleichbarkeit der Gruppen stattfand und welches Ergebnis resultierte. Dies erschien aber nicht ausreichend, da die Power bei kleinen Stichprobengrößen gering ist. Es wäre eine Anpassung der Irrtumswahrscheinlichkeit in den Primärstudien zu fordern oder eine Überprüfung über die Effektstärke.

Behandlung einen unterschiedlichen Effekt für verschiedene Personen hat, kaum in kontrollierten Studien untersucht werden kann. Dies sei zwar für Prä-Post-Designs möglich, allerdings werden hierfür Daten benötigt, die nicht vollständig vorliegen, z.B. die Reliabilität des jeweiligen Messinstrumentes.

#### **2.2.4 Selektionskriterien**

Zur Auswahl der Studien wurden folgende Selektionskriterien genutzt:

1. deutsche oder englische Sprache
2. VT, KVT oder PhT (bzw. deren Kombination) bei Angststörungen
3. Stichprobe: junges bis mittleres Erwachsenenalter
4.  $N > 3$  pro Gruppe
5. keine Analogpopulation
6. vorhandene Kontrollgruppe und/oder Vergleichstherapie (VT-KVT; PsT-PhT; PsT-Kombination)
7. pro Stichprobe nur ein Datensatz (pro Messzeitpunkt)
8. Variablen aus den Bereichen Angst, Depression, Lebensqualität, Klinische Signifikanz
9. Daten zu Berechnung einer ES

##### 1. Deutsche oder englische Sprache

Studien, die in einer anderen Sprache als Deutsch oder Englisch vorlagen, wurden ausgeschlossen.

##### 2. VT, KVT, PhT (bzw. deren Kombination) bei Angststörungen

Pharmakotherapie wurde definiert als Gabe von pharmakologischen Substanzen zur Reduktion der Symptomatik der Angststörungen. Studien, in denen nur der Firmenschlüssel angegeben war, wurden ausgeschlossen, d.h. eine Benennung der Substanz über die INN-Nomenklatur (International non-proprietary name; Internationaler Freiname für Arzneimittel) sollte erfolgen. Die Dauer der Medikation sollte mindestens zwei Wochen betragen als durchschnittlicher Wert für beginnende therapeutische Wirkungen. Nach Brewin (1996) wurden behaviorale Verfahren definiert als „direct attempts to reduce dysfunctional emotions and behavior by altering behavior“ (S. 34)

und kognitive Verfahren als „attempts to reduce dysfunctional emotions and behavior by altering individual appraisals and thinking patterns“ (S. 34). Kognitiv-behaviorale Verfahren setzten beide Techniken ein. Eine Differenzierung zwischen rein kognitiver und kognitiv-behavioraler Therapie wurde in der vorliegenden Arbeit nicht durchgeführt, da eine eindeutige Codierung oft nicht reliabel durchgeführt werden konnte<sup>11</sup>. So enthalten Verhaltensexperimente oft konfrontative Elemente oder Autoren deklarierten das Verfahren als kognitiv, setzten aber in der Beschreibung Exposition ein.

In der Studie sollten eine oder mehrere dieser Behandlungen an Personen mit Angststörungen untersucht worden sein<sup>12</sup>. Im optimalen Fall erfolgte die Diagnose nach einem diagnostischen System (ICD, DSM bzw. Forschungskriterien, die denen eines dieser Systeme entsprechen). Es wurden zudem Studien eingeschlossen, die bei der Benennung der Störung die exakte Bezeichnung nach einem diagnostischen System verwandten (z.B. „outpatients with generalized anxiety“ nicht aber „anxiety outpatients“, „rape victims with posttraumatic stress disorder“ nicht aber „females after a traumatic incidence“). Zusätzlich wurden für SoP Studien an Patienten mit Vermeiender Persönlichkeitsstörung eingeschlossen. Die Unterscheidung zwischen diesen beiden Störungen scheint wenig valide, die diagnostischen Kriterien überschneiden sich stark. So fanden z.B. Turner et al. (1992), dass beide Gruppen in einer Vielzahl an Variablen (z.B. Verhalten während einer Rede oder allgemeines Stressniveau) nicht bedeutsam differieren. Unterschiede zeigten sich in der Schwere der sozialen Angst, die Befunde sprechen eher für eine quantitative denn qualitative Differenz. Auch Hofmann und Barlow (2002) sprechen in einer Zusammenfassung bisheriger Ergebnisse gegen eine Abgrenzung der Störungen.

---

<sup>11</sup> Auch eine Differenzierung zwischen behavioraler und kognitiv-behavioraler Therapie erscheint zum Teil künstlich, wird aber aufgrund der etablierten Gliederung beibehalten.

<sup>12</sup> Wurde in einer Studie als Inklusionskriterium eine systematische Nichtwirksamkeit einer speziellen anderen Behandlung eingesetzt, so wurde – konservativ – diese Studie ausgeschlossen, um eine Vergleichbarkeit der Patientenpopulation mit anderen Untersuchungen zu ermöglichen. Aufgrund der geringen Anzahl an Studien mit gleichzeitiger starken Differenzen in den Kriterien (z.B. Nichtwirksamkeit eines bestimmten Medikamentes vs. einer bestimmten Therapieform) konnte keine statistische Überprüfung der Vergleichbarkeit durchgeführt werden.



### 3. Stichprobe: junges bis mittleres Erwachsenenalter

Es wurden nur Studien eingeschlossen, die die Wirksamkeit der betrachteten Therapieformen bei Patienten im jungen bis mittleren Erwachsenenalter untersuchten. Sowohl die Behandlung von Kindern und Jugendlichen, als auch von Personen im höheren Erwachsenenalter stellen eigenständige Therapieformen dar, die nicht im Fokus der vorliegenden Arbeit liegen. Die Exklusion der Studien erfolgte bei Angabe der folgenden Begriffe bzgl. der Stichprobenbeschreibung: „children“, „adolescents“ und „older adults“. Zudem wurden Studien ausgeschlossen, bei denen ein Altersmittelwert von unter 18 Jahren oder über 64 Jahren angegeben war.

### 4. N>3 pro Gruppe

Studien mit einer Stichprobengröße von unter vier Personen pro Gruppe wurden ausgeschlossen.

### 5. Keine Analogpopulation

„The study is an analogue of clinical treatment in that sense that college students as a group may differ on several dimensions from clinical populations typically treated for anxiety disorders ...“ (Kazdin, 1978, S. 678). Besonderen Stellenwert hat das Kriterium für den Bereich der Spezifischen und Sozialen Phobien. Für spezifische Phobie wurden Studien, die ausschließlich Studenten untersuchten, ausgeschlossen. Besonders im Bereich der spezifischen Phobie wurde eine Vielzahl an Studien mit Analogpopulation durchgeführt, so dass eine Inklusion zu einer starken Verzerrung der Ergebnisse geführt hätte.

Ein gestaffeltes Vorgehen wurde bei sozialer Phobie gewählt. Stangier et al. (1997) sehen als wesentliches Unterscheidungsmerkmal zwischen subklinischen und klinischen sozialen Ängsten den Schweregrad der Störung und den Grad an Beeinträchtigung der Lebensführung. Es liegen Befunde vor, dass je größer die Anzahl an gefürchteten Situationen ist, desto höher sind die Einschränkungen (Stein, Torgrud & Walker, 2000). Als Grundlage der Auswahl der Studien diente somit zum einen die Menge der gefürchteten Situationen. Es wurden Studien eingeschlossen mit Studenten, die eine

Diagnose der sozialen Phobie generalisierter Typus<sup>13</sup> erhielten. Wurde dagegen die Angst vor einer spezifischen sozialen Situation (z.B. Redeangst) untersucht oder keine Diagnose der generalisierten sozialen Phobie nach DSM bzw. einem anderen diagnostischen System gestellt, wurden als Operationalisierung des Kriteriums der Beeinträchtigung der Lebensführung diejenigen Studien eingeschlossen, bei denen sich die Studenten selbst für die Behandlung gemeldet haben. So wurden z.B. Untersuchungen ausgeschlossen, in denen Wissenschaftler Veranstaltungen besuchten und eine Selektion der Studenten über einen kritischen Wert eines Fragebogens durchführten.

#### 6. Vorhandene Kontrollgruppe und/oder Vergleichstherapie (VT-KVT; PsT-PhT; PsT-PsT+PhT)

Aufgrund der Nachteile von Prä-Post-ESn wurden nur Studien eingeschlossen, die eine Kontrollgruppe (KG) enthielten oder aber einen direkten Vergleich der untersuchten Behandlungen ermöglichten. Als KGn wurden folgenden Gruppen zugelassen: a) unbehandelte Gruppen einschließlich Wartelisten (zusammengefasst als WL) und b) Gruppen, in denen gemeinsame Wirkfaktoren realisiert wurden einschließlich Pillen-placebo (PP) (entsprechend einer Bezeichnung in den Primärstudien als „placebo“) und Gruppen mit Therapieplacebo (TP). TP wurde dabei nicht, wie in manchen Untersuchungen üblich, anhand eines mangelnden Wirksamkeitsbeleges in bisherigen Studien definiert. Dies ist zu kritisieren, da dieser möglicherweise aus einer geringen Power der zugrunde gelegten Untersuchungen resultieren kann. TP wurde wie oben erwähnt in der vorliegenden Arbeit, basierend auf Lambert und Bergin (1994), definiert als eine Realisierung gemeinsamer Faktoren in einer Therapiesituation (z.B. positive Erwartungen oder Beziehungsvariablen, wie Empathie und Aufmerksamkeit). Beispiele für TP waren nondirektive Gesprächspsychotherapie oder das Anhören eines Musiktonbandes, welches angeblich vorbewusste Botschaften enthielt. Gruppen, in denen verhaltenstherapeutische Techniken, auch z.B. Entspannungstraining, angewandt wurden, zählten in der vorliegenden Arbeit nicht zu TP. Der Vergleich zwischen Kombinationsbehandlung und Psychotherapie allein wurde nur über Studien durchgeführt, die beide Verfahren direkt miteinander verglichen.

---

<sup>13</sup> In der vorliegenden Arbeit wird das Kriterium eingesetzt als Operationalisierung einer großen Anzahl an gefürchteten Situationen, es ist somit quantitativer denn qualitativer Natur.

### 7. Pro Stichprobe nur ein Datensatz (pro Messzeitpunkt)

Vor allem bei sehr aufwendigen Untersuchungen werden oft mehrere Publikationen veröffentlicht, die sich auf eine Stichprobe stützen. Um eine Inflation des Einflusses dieser Studien auf das Ergebnis zu vermeiden, wurde versucht, nur einen Datensatz pro Stichprobe pro Messzeitpunkt einzuschließen. So wurden Publikationen ausgeschlossen, die vorläufige Ergebnisse präsentierten, wenn zusätzlich Daten zur Berechnung einer ES für die vollständige Stichprobe vorlagen. Konnten Studien nicht eindeutig einer Hauptpublikation zugeordnet werden, obwohl sie vorläufige Daten oder spezielle Unterfragestellungen betrachteten, wurden sie eingeschlossen. Es wurde so zwar ein potentieller Fehler für die mittlere ES in Kauf genommen (einschließlich einer künstlichen Erhöhung der Studienanzahl und Inflation des Alpha-Fehlers), schwerwiegender erschien aber der Verlust einer Untersuchungseinheit.

### 8. Erhobene Variablen

Es sollten Daten vorliegen (erhoben mittels Verhaltenstests, Selbst- oder Fremdeinschätzung), die einer der folgenden Kategorien zuzuordnen waren: Angst, Depression (als wichtige komorbide Symptomatik), Lebensqualität oder Klinische Signifikanz. Klinische Signifikanz als Maß für eine klinisch wichtige Veränderung kann unterschieden werden in Responder- und Endstate-Status. Responder bezieht sich dabei auf eine relative Verbesserung zu Prä, Endstate auf einen Vergleich mit der gesunden Normpopulation. Klinische Signifikanz wurde in der vorliegenden Arbeit nur bzgl. der Angstsymptomatik ausgewertet. Physiologische Werte, wie z.B. Herzfrequenz, wurden aufgrund der geringen Anzahl an Studien, die diese Erhebungsmethode einsetzten, ausgeschlossen. Zudem bemerkt Barlow (2002), dass bei PDA, für die physiologische Erhebungen einen besonderen Stellenwert einnehmen, „research in past decades ... suggested that physiological assessments are generally unreliable measures of change“ (S. 352). In Zusammenhang damit stehen Ergebnisse verschiedener Studien, die zeigten, dass bei Patienten mit PDA trotz subjektiver Wahrnehmung physiologischer Erregung objektiv keine Veränderung in physiologischen Kennwerten gefunden werden konnte (Taylor et al., 1986). Auch bei SoP zeigten z.B. Edelman und Baker (2002) in einer experimentalpsychologischen Untersuchung, dass Patienten zwar in sozial phobischen Situationen höhere Werte für schwitzende Hände oder Herzrasen im Rahmen einer Selbstbeurteilung angaben, in den physiologischen Erhebungen sich aber kein bedeut-

samer Unterschied zu einer nicht-klinischen Kontrollgruppe zeigte. Ausgeschlossen wurden des Weiteren Studien, die rein den Einfluss einer Behandlung auf experimental-psychologische Variablen erhoben, z.B. Stroop-Interferenz.

## 9. Daten zur Berechnung einer ES

In Studien mit Crossover-Design mussten Daten vor dem Wechsel der Therapiebedingung vorliegen.

Es wurden Studien ausgeschlossen, die keinen Gruppenvergleich anhand folgender Werte ermöglichten: Mittelwerte und Standardabweichungen,  $t$ - und  $F$ -Werte, Veränderungswerte, Häufigkeiten sowie Signifikanzniveaus. In der vorliegenden Arbeit erfolgte die Berechnung der ESn somit aus einer Vielzahl an Informationen. In Kontrast dazu stehen einige bisherige Meta-Analysen (z.B. Ruhmland, 2001), die die ESn nur anhand von Mittelwerten und Standardabweichungen ermittelten. Ray und Shadish (1996) unterschieden algebraische Äquivalente zur ursprünglichen Berechnung ( $t$ - und  $F$ -Werte)<sup>14</sup> sowie Berechnungen über Schätzverfahren (z.B. Signifikanzlevel oder Veränderungswerte). Sie verglichen die Berechnungsvarianten von ESn verschiedener Quellen an Studien, in denen mehrere Informationen vorlagen. Nicht alle Schätzverfahren korrelierten hoch mit der standardisierten Mittelwertsdifferenz, am schlechtesten schnitten geschätzte Signifikanzniveaus mit einer Korrelation von  $r = .57$  ab zurückzuführen auf Rundungsfehler (z.B. Angabe „ $p < .05$ “ → Inklusion des  $t$ -Wertes des 5%-Niveaus, Angabe „nicht signifikant“ → Inklusion von  $ES = 0$ ). Die mittlere ES lag für dieses Verfahren bei  $ES = 0.31$  im Vergleich zu  $ES = 0.60$  bei einer ursprünglichen Berechnung. Auch die Varianz in den ESn war bedeutend kleiner ( $\sigma^2 = 0.12$  vs.  $\sigma^2 = 0.48$ ) Die Autoren schlussfolgern, dass mögliche Variationen durch unterschiedliche Berechnungsvarianten stärkere Aufmerksamkeit erfordern, ein Ausschluss der Berechnungsalternativen aber „might be more extreme than needed“ (S. 1322). Es ist zu beachten, dass eine Exklusion von Studien, die keine Mittelwerte und Standardabweichungen berichten, eine Verzerrung der Stichprobe darstellt. Ein Ausschluss der Untersuchungen kann potentiell zu einem stärkeren Fehler führen als der Einschluss einer nur marginal genauen Effektstärke. Pigott (2001) verdeutlicht dieses Problem: vor

---

<sup>14</sup> Diese können somit problemlos mit der standardisierten Mittelwertsdifferenz kombiniert werden.

allem hinsichtlich eines nichtsignifikanten Unterschiedes zwischen den Gruppen wäre der Ausschluss der Studie abhängig von der Effektstärke und somit nicht randomisiert. Die verbleibenden Studien entsprächen keiner repräsentativen Stichprobe. ESN aus alternativen Berechnungen stellen die besten Schätzungen dar, die dem Meta-Analysten für diese Studien zur Verfügung stehen.

Einige Autoren fordern als zusätzliches Exklusionskriterium eine geringe methodische Qualität der Studien (z.B. Eysenck, 1978). Dies soll im Folgenden diskutiert werden.

### **Das „Garbage in - Garbage out“-Argument**

Ein Vorwurf am Verfahren der Meta-Analyse stellt das „Garbage in - Garbage out“-Argument dar (z.B. n. Sharpe, 1997). Kritiker zweifeln die Validität der Ergebnisse an, da methodisch schwache Primärstudien die Effekte verzerren würden<sup>15</sup>. Sie fordern, diese Studien auszuschließen. Dieses Vorgehen erscheint problematisch. Zum einen wäre es notwendig, die methodische Qualität einer Studie zu bestimmen und mit dieser adäquat zu arbeiten. Zum anderen müsste feststehen, ob ein Ausschluss generell sinnvoll ist, d.h. ob eine methodisch schlechtere Studie tatsächlich andere Ergebnisse aufweist im Vergleich zu einer methodisch besseren Studie.

Die Ansätze zur Bestimmung der Qualität einer Studie müssen bislang als unbefriedigend angesehen werden (Sharpe, 1997; Wortman, 1994). Methodische Qualität kann operationalisiert werden über verschiedene Validitätsarten (z.B. Cook et al., 1979) und ergibt sich nach diesem System aus:

- a) der Konstruktvalidität (Angemessenheit bzgl. „generalizations about higher-order constructs from research operations“; S. 38),
- b) der internen Validität (Angemessenheit bzgl. „causal relationship from one variable to another in the form in which the variables were manipulated or measured“; S. 38),
- c) der externen Validität (Angemessenheit bzgl. einer Generalisierung „across different types of persons, settings, and times“; S. 37) und

---

<sup>15</sup> Vgl. allerdings Hunter et al. (1990): „... no study can be ‘without methodological inadequacy’“ (S. 495)

- d) der statistischen Validität (Angemessenheit bzgl. „covariation“ zwischen „presumed independent and dependent variables“; S. 37)

Es zeigt sich die Forderung nach dem Wissen, was man wie misst und angemessen auswertet, einer eindeutigen Interpretation der Ergebnisse sowie deren Generalisierbarkeit.

Für jede Validitätsart zeigen die Autoren potentielle Gefahrenquellen für Einbußen auf, z.B. Regression zur Mitte, Abbrecher bzw. Testung in Bezug zur internen Validität, Reliabilität der Messungen oder der Behandlung in Bezug zur statistischen Validität, inadäquate Operationalisierung oder Erwartungen des Versuchsleiters bzw. der Patienten in Bezug auf die Konstruktvalidität. Methodische Qualität setzt sich somit aus mehreren Elementen zusammen, die sich wiederum in verschiedene Aspekte aufteilen lassen. In den bisherigen Versuchen zur Bestimmung der Validität wurden nur einzelne Komponenten einbezogen. So wurde von Smith et al. (1980) die interne Validität als hoch geratet, wenn Randomisierung eingesetzt wurde und die Abbrecherrate unter 15 % gleich für beide Gruppen lag. War letzteres nicht gegeben, wurde die Studie trotzdem als hoch intern valide eingeschätzt, wenn es sich um eine Intent to treat-Analyse handelte oder Abbrecher und Beender sich nicht in bedeutsamen Variablen unterschieden. Weitere Kriterien bzgl. Randomisierung und Abbrechern lagen vor hinsichtlich mittlerer und geringer interner Validität. Es ergibt sich als Problem, dass sowohl die Auswahl der Elemente aus der Vielzahl an Aspekten der internen Validität als auch die Bestimmung der Grenzwerte für hohe, mittlere und niedrige Validität bzw. einer Gewichtung der Komponenten bei Bildung eines Gesamtindex für die Qualität im subjektiven Ermessen des Meta-Analysten liegen. Zudem erfordert die Bestimmung solch einer zusammengesetzten Variablen, dass für sämtliche Komponenten auch Informationen in den Primärstudien vorliegen bzw. diese angemessen geschätzt werden können. Als gebräuchliche Alternative gilt die Einzelbetrachtung von Komponenten, deren Einfluss auf die ESN untersucht wird, wie z.B. die Art der Zuweisung der Patienten auf die Gruppen (vgl. Sharpe, 1997). Ein Vorteil besteht in der Möglichkeit Aussagen zu treffen, welche speziellen methodischen Aspekte die Effekte verändern können.

Nach Glass (1978) ist die Frage, ob es einen Zusammenhang zwischen Qualität und Ergebnis einer Untersuchung gibt, nicht a priori zu klären. Beelmann et al. (1994) kommen zu dem Schluss, dass „bislang kein genereller systematischer Effekt der

methodischen Qualität auf die Studienergebnisse feststellbar ist“ (S. 218). Auch Hunter et al. (1990) sprechen dagegen, methodische schwache Studien ohne weiteres auszuschließen. Sie räumen allerdings ein, dass bei guter a priori-Evidenz Untersuchungen mit bestimmten Merkmalen verworfen werden können. Dies sei allerdings selten der Fall. Somit sollte in jeder Meta-Analyse eine eigene Prüfung dieses Zusammenhangs durchgeführt werden, um adäquate Aussagen treffen zu können. Bei heterogenen Befunden kann dann eine Betrachtung getrennt nach Qualität durchgeführt oder diese als Kovariate in ein Regressionsmodell aufgenommen werden.

### **Ein Beispiel für methodische Qualität: Randomisierung**

Unter Randomisierung versteht man die zufällige Aufteilung der Patienten auf die verschiedenen Treatmentbedingungen. Sie dient dem Versuch einer Kontrolle und Gleichverteilung von potentiellen Störfaktoren. Obwohl zur Durchführung „echter“ Experimente eine randomisierte Aufteilung der Patienten als notwendig erachtet wird und dadurch eine Steigerung der internen Validität erreicht werden kann (z.B. Cook et al., 1979), wird a priori keine Exklusion nicht-randomisierter Studien durchgeführt.

Die Frage, ob randomisierte Studien andere Ergebnisse aufweisen als nicht-randomisierte Studien, ist nicht eindeutig zu beantworten. Lipsey et al. (1993) fassten die Ergebnisse von 74 Meta-Analysen, die die Differenzen zwischen diesen Designs betrachteten, zusammen. Die Unterschiede waren im Durchschnitt Null mit sowohl positiven als auch negativen Ausprägungen. Probleme dieses Vorgehens sind aber die möglicherweise unterschiedliche Operationalisierung der Designvariablen oder verschiedene Berechnungen der Effektstärken zwischen den Meta-Analysen. Shadish und Ragsdale (1996) konnten in einer eigens zu diesem Thema angefertigten Meta-Analyse unter Einbezug von 100 Untersuchungen aus dem Bereich der Familien- und Eheberatung zeigen, dass randomisierte Studien signifikant höhere Effektstärken mit geringerer Varianz aufweisen als nicht-randomisierte Studien. Dieses Ergebnis bestätigte sich auch bei Kontrolle von Drittvariablen, wie z.B. aktive vs. passive KG. Sie schlussfolgerten, dass es nicht angebracht erscheint, diese Designs in einer Meta-Analyse ohne adäquate Analysen miteinander zu kombinieren. Demgegenüber steht eine Meta-Analyse von Heinsman und Shadish (1996), die sich auf 98 Studien hauptsächlich zu pädagogischer Intervention stützt. Es zeigten sich auch hier signifikant

höhere Effektstärken mit geringerer Varianz für randomisierte Studien. Sie finden jedoch bei Einschluss konfundierender Variable, dass die Effektstärken von randomisierten vs. nicht-randomisierten Studien beinahe gleich sind, wenn beide „were equally well designed and executed“ (Heinsman et al., 1996, S. 162). Dies bezieht sich vor allem auf die Problematik der Selbstselektion und der Gleichheit der Präwerte zwischen den Gruppen. Es liegen somit sehr unterschiedliche Ergebnisse vor, abhängig von den eingeschlossenen Studien, eingesetzten meta-analytischen Methoden oder möglicherweise Fehler in der Codierung. Es erscheint aber voreilig, nicht-randomisierte Studien a priori auszuschließen. Bei einer ausreichenden Anzahl an Untersuchungen sollte der Einfluss des Designs überprüft werden und bei bedeutsamen Ergebnissen entweder zu einem Ausschluss nicht-randomisierter Studien führen oder eine explorative Analyse der Gründe dieser Unterschiede stattfinden. Aufgrund der geringen Anzahl nicht-randomisierter Studien in der vorliegenden Arbeit wurde allerdings auf eine Evaluation potentieller Differenzen verzichtet. Zudem sei angemerkt, dass Randomisierung nicht unweigerlich gleichzusetzen ist mit besserer Qualität und präziseren Resultaten: eine

quasi- oder gar präexperimentelle Studie kann also durchaus hohe interne Validität besitzen, wenn die Beobachtungsbefunde nur auf eine Weise erklärt werden können ..., und umgekehrt bleibt ein „echtes“ Experiment schwach, solange es für das Zustandekommen der Befunde gleichermaßen plausible konkurrierende Erklärungen gibt (was in den Verhaltens- und Sozialwissenschaften selbst bei Anwendung randomisierter Versuchspläne eher die Regel als die Ausnahme ist) (Brandtstädter, J. (1990), zit.n. Hager, 1998).



## 2.3 Statistische Methoden<sup>16</sup>

### 2.3.1 Berechnung der Effektstärken für eine Studie

Bei gegebenen Mittelwerten und Standardabweichungen von Kontroll- und Behandlungsgruppe (KG; BG)<sup>17</sup> wurde als ES Hedges  $g$  berechnet. Sie wird ermittelt über die Differenz der Mittelwerte zwischen den Gruppen geteilt durch die gepoolte Standardabweichung (Gleichung 1; Hedges et al., 1985). Im Gegensatz zu Glass, McGaw und Smith (1981), die nur die Standardabweichung der KG einsetzen, werden die Standardabweichungen der beiden Gruppen als gleich für die Population angesehen. Vorteile der gepoolten Standardabweichung legt Hunter et al. (1990) dar: „The within-group [gepoolte] standard deviation has only about half the sampling error of the control-group standard-deviation.“ (Hunter et al., 1990, S. 271); zudem sei die ES nach Glass et al. (1981) nicht aus  $t$ - und  $F$ -Werten zu berechnen, da hier sowohl die Daten von KG als auch BG einfließen. Kritiker der gepoolten Standardabweichung weisen auf eine potentielle höhere Standardabweichung für die BG nach der Behandlung hin, so dass nur die Standardabweichung der KG einen angemessenen Schätzer bilde (z.B. Smith et al., 1980). Demgegenüber argumentieren Hartmann et al. (1995) basierend auf der klinischen Interpretation der ES<sub>n</sub>, dass bei höherer Varianz in der BG zu Post eine geringere ES adäquat sei, da diese Behandlung negativer zu bewerten sei, als eine mit gleicher Mittelwertsdifferenz aber geringerer Variation. Der Fokus der Autoren liegt zwar bei Prä-Post-Effekten, die Darlegung lässt sich auf ES<sub>n</sub> zwischen Gruppen übertragen. Bei Berechnung von Hedges  $g$  wurde eine Korrektur für kleine Stichprobenumfänge durchgeführt (Gleichung 2; Hedges et al., 1985), eine weitere Reduktion des Einflusses des Stichprobenfehlers.

Bei gegebenen  $t$ -,  $F$ - oder Veränderungswerten wurden diese in Hedges  $g$  überführt (Gleichungen 3 - 7; Ray et al., 1996). Zur Berechnung von  $g$  aus Veränderungswerten wird zusätzlich die jeweilige Korrelation zwischen den Messzeitpunkten benötigt. Diese war allerdings nicht direkt in den eingeschlossenen Studien gegeben. Als Schätzer

---

<sup>16</sup> Der folgende Text enthält Verweise auf Gleichungen. Diese finden sich in Anhang D.

<sup>17</sup> Aufgrund der Einfachheit wird im Folgenden immer vom Vergleich zwischen KG und BG gesprochen; die statistischen Methoden gelten auch für den Vergleich zwischen verschiedenen Behandlungen.

fungierte die mittlere Retestrelabilität von 23 Messinstrumenten, die häufig in den eingeschlossenen Primärstudien eingesetzt wurden (soweit diese anhand von Primärliteratur erhoben werden konnte; vgl. Anhang E). Nachteile dieser Methode sind allerdings eine potentielle Überschätzung der Prä-Post-Korrelation durch die Retestrelabilität bei Vorliegen einer Behandlungs-Personen-Interaktion (Hunter et al., 1990), sowie dass nur ein Schätzer für alle Studien eingesetzt wurde, unabhängig von Stichprobe und Messinstrument sowie der unterschiedlichen Länge der miteinander korrelierten Zeiträume von Reliabilitätsstudien und Therapiestudien. Da allerdings die Effektstärke aus Veränderungswerten umso höher ist, je geringer die einberechnete Korrelation, wurde – konservativ – der relativ hohe Schätzer von  $r = .81$  eingesetzt. Dieser ist höher als der von Smith et al. (1980) empfohlene für Zeiträume von zwei bis sechs Monaten mit  $r = .50$ . Ray et al. (1996) vermuten allerdings, dass dieser Schätzer möglicherweise zu niedrig sei, da in ihren Vergleichen die ES aus Veränderungswerten die ES aus Mittelwerten und Standardabweichungen überschätzten. Ein angemessener Schätzer scheint somit nicht gegeben, eine Lösung dieses Problems erfordert eine Änderung der Publikationspraxis hinsichtlich berichteter Informationen.

In einigen Studien musste das Signifikanzniveau geschätzt werden, da der Versuch, mit den Autoren Kontakt aufzunehmen und Informationen direkt zu erhalten, erfolglos blieb bzw. Studien schon älter waren, so dass auf eine Kontaktaufnahme verzichtet wurde. Es wurden die gebräuchlichen Schätzungen eingesetzt, auch wenn diese fehlerbehaftet sind (Ray et al., 1996): für Tendenzen wurde das Signifikanzniveau auf 10% festgelegt, für einen signifikanten Unterschied auf 5%, für einen hochsignifikanten auf 1%. War mitgeteilt, dass es keinen signifikanten Unterschied zwischen den Bedingungen gab, wurde die ES auf Null gesetzt. Dies birgt neben einer geringeren mittleren ES die Gefahr einer Unterschätzung der Varianz der Untersuchungen in der Meta-Analyse (Pigott, 2001).

Bei gegebenen Kategorialdaten wurde Odds Ratio berechnet, dies war in der Symptomkategorie Klinische Signifikanz notwendig (Gleichung 8; Haddock, Rindskopf & Shadish, 1998). Odds Ratio bietet gegenüber anderen Effektstärken für dichotome Daten den Vorteil einer Invarianz gegenüber verschiedenen Designs und unterschiedlichen Stichprobengrößen und gilt als geeignetstes Maß für Vierfeldertafeln (Haddock

et al., 1998). Durch Anpassung der logistischen in eine Standardnormalverteilung wird Odds Ratio in  $g$  überführt (Gleichung 9; Haddock et al., 1998; Hasselblad & Hedges, 1995). Die relative Unabhängigkeit der Effektstärken vom Cut-off wird diskutiert von Hasselblad et al. (1995).

Entsprechend einer allgemeinen Konvention schlägt Cohen (1988) eine Interpretation der ES von  $ES = 0.2$  als klein,  $ES = 0.5$  als mittel und  $ES = 0.8$  als groß vor.

Da nach Abschluss der untersuchten Behandlungsgruppe Wartelistenpatienten eine Therapie erhalten, ist die Berechnung der ES zwischen Gruppen im Follow-up nicht möglich. Es wurde für Aussagen zur Stabilität der Ergebnisse, d.h. ob sich nach Beendigung der Therapie eine Verschlechterung oder aber Verbesserung der Symptomatik ergeben hat, eine ES für die Behandlungsgruppen zwischen post und den nachfolgenden Messzeitpunkten berechnet (Gleichung 10)<sup>18</sup>. Bei Vergleich verschiedener Therapieformen in einer Studie (mit Ausnahme des Vergleichs zur Placebo-bedingung) wurde zusätzlich zwischen diesen Gruppen die ES berechnet. Bei jeder Berechnung der Effektstärken zwischen Messzeitpunkten wurde die gepoolte Standardabweichung eingesetzt entsprechend den Empfehlungen von Hartmann et al. (1995).

Die Richtung der Effekte wurde standardisiert: in den Vergleichen BG zu KG, VT zu KVT, PhT zu PsT bzw. PsT zu PsT+PhT zeigt ein negativer Wert ein besseres Ergebnis für die jeweils erstgenannte Gruppe, ein positiver Wert ein besseres Ergebnis der jeweils Zweitgenannten. Für die Betrachtung des Follow-up ist eine positive ES eine weitere Verbesserung, eine negative ES dagegen eine Verschlechterung über die Zeit.

---

<sup>18</sup> Die Erhebungen sollten dabei mindestens einen Monat nach Beendigung der Therapie erfolgen (als Operationalisierung von Stabilität). Lagen Daten für frühere Katamneszeiträume vor, wurden diese ausgeschlossen. Für Pharmakotherapiestudien wurde dabei zur Vergleichbarkeit der Daten zwischen verschiedenen Messzeitpunkten als auch zu den Daten der Psychotherapiestudien folgende Bedingungen eingeführt: getrennte Daten für Placebo- und aktive Substanzgruppe, Einschluss sämtlicher verfügbarer Patienten anstatt nur der Responder (dies führt ansonsten zu einer Verringerung der Varianz und damit einer Erhöhung der ES), sowie ein Absetzen des Medikamentes.

### **Fehlende Informationen**

Die Kernfrage im Umgang mit fehlenden Informationen ist, ob diese randomisiert entfallen oder es einen Zusammenhang zwischen der Ausprägung der Daten und der Wahrscheinlichkeit des Auslassens gibt (Pigott, 2001). Je nach Art der fehlenden Information wurde aufbauend auf dieser Frage ein anderes Verfahren eingesetzt.

Bei fehlenden Mittelwerten und Standardabweichungen zu Prä wurde keine Effektstärke berechnet. Ziel war die Untersuchung möglicher Unterschiede in den Prä-Werte mit einer möglichst hohen Power. Bei Einschluss sämtlicher in den Primärstudien mitgeteilten nichtsignifikanten Tests mit  $g = 0$  wäre dieses Ziel potentiell verletzt.

Bei fehlenden Angaben zur Anzahl an Personen, die die Teilnahme an der Studie vorzeitig abbrechen, wurde verzichtet, einen Wert zu implementieren. Zwei Alternativen sind möglich, weshalb diese Information ausgelassen wurde: a) besonders in Studien mit sehr hoher Dropout-Rate wird diese nicht genannt bzw. b) in keiner der Gruppen haben die Patienten die Untersuchung abgebrochen. Da nach Erachten der Autorin keine dieser beiden Möglichkeiten eindeutig zu präferieren ist, wurden Studien ohne Angabe der Abbrecherrate in den jeweiligen Analysen ausgeschlossen. Weitere fehlende Daten, die zum Ausschluss einer Studie in der jeweiligen Analyse führten, betrafen die Therapiedauer, wenn nicht mindestens die Anzahl an Therapiesitzungen angegeben war.

Bei fehlenden Angaben zur Stichprobengröße und Dauer einer Therapiesitzung wurde angenommen, dass es keinen Zusammenhang zwischen Ausprägung und Wahrscheinlichkeit des Auslassens der Information gab, so dass hier der Mittelwert aus den restlichen Studien der Therapierichtung eingesetzt wurde ( $n = 17$  pro Gruppe bei Psychotherapie; 84 Minuten als durchschnittliche Länge für eine Therapiesitzung). Waren die einzelnen Gruppengrößen nicht mitgeteilt, wurde von einer Gleichverteilung der Patienten auf die Bedingungen ausgegangen.

War nicht das diagnostische System angegeben, wurde das jeweilige System eingesetzt, welches dem Publikationsjahr der Studie minus zwei Jahre<sup>19</sup> entsprach.

### **Extremwerte**

Obwohl in der vorliegenden Meta-Analyse das Random Effects-Modell (vgl. Statistische Modelle) eingesetzt wird, das die ungeklärte Varianz einbezieht, wurde eine Ausreißer-Analyse durchgeführt. Im Vergleich zu Primärstudien ist der Weg von den Rohdaten zum Endergebnis bei einer Meta-Analyse bedeutend länger<sup>20</sup>. Auf jedem Teilabschnitt können sich Fehler ergeben, die einen Einfluss auf das Ergebnis zeigen, z.B. Fehler bei der Datenübertragung von den eingesetzten Instrumenten in ein Datenverarbeitungsprogramm oder Fehler bei der Übertragung dieser Kennwerte in die Publikation. Trotz sorgfältiger Durchführung der statistischen Auswertung durch den Meta-Analysten (bzw. nach Hunter et al., 1990: „it is well known that meta-analysts do not make errors“ S.262), ist eine direkte Kontrolle der vorgeschalteten Fehlerquellen nicht möglich.

Zur Bestimmung der Extremwerte wurde der jeweilige ungewichtete Mittelwert aller einzelnen Effektstärken innerhalb jeder Angststörung und Behandlung (PsT bzw. PhT) und Vergleichgruppe (WL, Placebo, andere Therapie) hinsichtlich der verschiedenen Symptomkategorien und KGs gebildet. Als Ausreißer wurde ein Wert definiert, der über oder unter zwei Standardabweichungen von diesem Mittelwert gerundet auf den nächsten .5-Wert lag. Berechnung und Exklusion erfolgten auf der untersten Ebene, d.h. bzgl. der Einzel-ESn der Messinstrumente und nicht über die gemittelten Werte in den Symptomkategorien.

Durch den Einsatz einer Ausreißer-Analyse wird die Varianz zwischen den ESn gesenkt und es vergeben sich somit mögliche Varianzaufklärungen durch die Regressions-

---

<sup>19</sup> Die Differenz wurde gebildet, um die zeitliche Verzögerung zwischen Durchführung der Studie und Publikation zu beachten.

<sup>20</sup> Daneben existieren natürlich auch potentielle Fehler im Weg zu den Rohdaten, z.B. Schwierigkeiten der Patienten oder Interviewer in der Beantwortung der Messinstrumente. Auch diese liegen nicht unter der Kontrolle des Meta-Analysten.

analyse (so könnten die oberen ES<sub>n</sub> eine gemeinsame Subpopulation innerhalb des Universums an ES<sub>n</sub> darstellen). Aufbauend auf Empfehlungen von Hunter et al. (1990) soll aber durch den Ausschluss an Extremwerten zumindest eine indirekte Kontrolle über potentielle Fehler in der Datenerhebung erfolgen.

### **2.3.2 Integration der Effektstärken über alle Studien**

Ein Problem bei der Integration von Effektstärken ist der Umgang mit stochastisch abhängigen Effekten. Diese ergeben sich, wenn mehr als eine Effektstärke pro Studie berechnet werden kann, also z.B. mehrere Messinstrumente pro Studie eingesetzt oder mehrere Behandlungsgruppen im Vergleich zu einer Kontrollgruppe untersucht wurden.

Es wurde eine Aufteilung entsprechend den Messzeitpunkten in den Primärstudien vorgenommen. Neben der Erhebung am Ende der Behandlung (Post), erfolgte eine Gliederung der Daten nach folgenden Kategorien: Kurzzeit (< 6 Monate), Mittel (6 – 11 Monate) und Langzeit (>11 Monate). Lagen in einer Studie Daten mehrerer Erhebungen innerhalb einer Kategorie vor, so wurden die Daten des jeweils spätesten Messzeitpunktes eingeschlossen, es sei denn, speziell in der Kategorie Langzeit, dieser differierte sehr stark zu den restlichen Studien (Bsp. Studien 1-10: 12-24 Monate, Studie 11: 12 Monate und 48 Monate → es wurden von Studie 11 die Daten des 12-Monats-Follow up eingeschlossen).

Es wurden mehrere Symptomkategorien gebildet. Dies beruht auf einer häufigen Kritik an Meta-Analysen, die als Unterkategorie des „Apfel und Birnen“-Problems angesehen werden kann. Dies ist die Kombination von Messinstrumenten verschiedener Symptomkategorien, die einen unterschiedlichen Zeitverlauf in der Veränderung zeigen. Die Angemessenheit dieser Annahme konnte in empirischen Studien gezeigt werden (z.B. Kopta, Howard, Lowry & Beutler, 1994). So könnten störungsspezifische Variablen (z.B. Vermeidungsverhalten) schneller höhere Effektstärken zeigen, als Daten von globaler Natur (z.B. Lebensqualität). Es erfolgte eine Aufteilung der Messinstrumente nach den Hauptkategorien Angst, Depression, Lebensqualität und Klinische Signifikanz. Zusätzlich wurde in der Symptomkategorie Angst eine Unterteilung in Verhalten, Kognitionen und Erregung vorgenommen. In allen Angststörungen war, mit

Ausnahme bei OCD, die behaviorale Komponente der Angst das Vermeidungsverhalten; bei OCD dagegen Zwangshandlungen. Bei PTB war Schwerpunkt des kognitiven Aspektes Intrusionen, bei OCD Zwangsgedanken. Es sei allerdings zu bedenken, dass die Aufteilung der Erhebungsinstrumente in die Unterkategorien teilweise umstritten ist, gehört z.B. die FNE in die Kategorie Kognition oder zu Angst im Allgemeinen (vgl. z.B. Diskussion bei Taylor, 1996)? Eine Darstellung der Symptomkategorien mit Beispielinstrumenten, einschließlich einer kurzen Beschreibung befindet sich in Tabelle 1. Konnte ein Messinstrument keiner dieser Kategorien zugeordnet werden, wurden die Daten ausgeschlossen. Bzgl. mehrerer Effektstärken pro Symptom- oder Unterkategorie aufgrund verschiedener Messinstrumente wurde als konservatives Verfahren die Bildung eines Mittelwertes genutzt. Für die Kategorie Angst wurde dabei nicht über die Unterkategorien gemittelt, sondern über die einzelnen Erhebungsverfahren. Somit wurden alle in einer Studie eingesetzten Messinstrumente innerhalb der Kategorien gleich gewichtet. Obwohl einige Verfahren adäquatere Ergebnisse bezüglich der Integration stochastisch abhängiger Effektstärken liefern (Gleser & Olkin, 1994), werden hierfür zusätzliche Informationen über Interkorrelationen der Messinstrumente benötigt. Diese stehen in diesem Fall nicht zur Verfügung bzw. könnten nur ungenau geschätzt werden. Es sei angemerkt, dass im Gegensatz zu der vorliegenden Arbeit in einigen Meta-Analysen eine Auswahl der Effektstärken in Abhängigkeit z.B. von der Güte des Messinstrumentes erfolgte. Matt (1989) zeigte aber die Subjektivität dieser Entscheidungen und Einschränkungen hinsichtlich der Reproduzierbarkeit der Ergebnisse.

Tabelle 1: Beschreibung der Symptomkategorien mit Beispielinstrumenten

Hauptkategorie	Unterkategorie	Beschreibung	Beispiel
Angst		Erfassung allg. Ängstlichkeit und störungs-spezifischer Aspekte der Angsterkrankungen	Hamilton Anxiety Scale (HAMA, Hamilton, 1959), Impact of Event Scale (IES; Horowitz, Wilner & Alvarez, 1979),
	Verhalten	Erfassung des von Aspekten des Verhaltens, im Besonderen des Vermeidungsverhaltens	Subskala Vermeidung der Liebowitz Social Anxiety Scale (LSAS; Liebowitz, 1987), agoraphobische Subskala des Fear Questionnaire (FQ-AG; Marks & Mathews, 1979)
	Kognition	Erfassung angstspezifischer Symptomatik speziell auf der kognitiven Ebene	Fear of Negative Evaluation (FNE; Watson & Friend, 1969), Agoraphobic Cognitions Questionnaire (ACQ; Chambless, 1984)
	Erregung	Erfassung der mit der Angststörung assoziierten körperlichen Erregung	Body Sensation Questionnaire (BSQ; Chambless, 1984), Subskala Erregung der IES (Horowitz et al., 1979)
Depression		Erfassung der depressiven Symptomatik	Beck Depression Inventory (BDI; Beck, 1967), Hamilton Depression Scale (HAMD; Hamilton, 1960)
Lebensqualität		Erfassung der Beeinträchtigung in Aspekten des Lebens durch die Störung	Sheehan Disability Scale (SDS; Sheehan, 1983)
Klinische Signifikanz		Erfassung des Responder- bzw. Endstate-Status bezogen auf die Angstsymptomatik	Verbesserung in einem bestimmten Messinstrument um mind. 25 %, keine Panikattacken



Bzgl. mehrerer möglicher Gruppenvergleiche wurde folgendes Vorgehen gewählt. Wurden in Pharmakotherapiestudien mehrere Substanzen überprüft, wurde jeder dieser Vergleiche in die Endanalyse einbezogen. Wurde dagegen eine Substanz in mehreren Dosierungen überprüft, wurde diejenige Gruppe einbezogen, welche die höchste Effektivität zeigte. Wurden in Psychotherapiestudien mehrere Behandlungen durchgeführt, erfolgte die Bildung eines Mittelwertes auf unterer Ebene, d.h. Exposition in sensu wurde mit Exposition in vivo gemittelt, nicht aber Exposition mit Exposition und zusätzlicher Entspannung (vgl. auch Codierungskatalog). Nachteil ist eine künstliche Erhöhung der Stichprobe und damit eine Erhöhung des Fehlers erster Art. Lagen sowohl Daten für Completer- als auch Intent to treat-Analyse vor, wurde letztere eingesetzt, da durch Einschluss der Patienten, die die Studie abbrechen eine höhere Angemessenheit der Ergebnisse gegeben ist. Ausnahme bildeten Studien, in denen für die Completer-Auswertung Mittelwerte und Standardabweichungen vorlagen, nicht dagegen für die ITT-Analyse.

Lagen für eine Studie sowohl Daten vor bzgl. eines Vergleiches zu einer Gruppe, die PsT erhielt und zu einer Gruppe, die PsT mit zusätzlichen Pillenplacebo erhielt, wurde nur erstere in die Analyse eingeschlossen.

### **Statistisches Modell**

„[meta-analytic] modelling would be improved by the increased use of random effects models in preference to the current default of fixed effects models“ (National Research Council, 1992, S. 185, zit. n. Erez, Bloom & Wells, 1995, S. 3).

Die Berechnung der mittleren Effekte und die Berechnung des Einflusses dritter Variablen kann unter Einsatz zweier statistischer Modelle erfolgen: das Fixed und das Random Effects-Modell (FEM; REM) (Hedges, 1994; Hedges et al., 1985; Hedges & Vevea, 1998).

Im FEM sind die beobachteten ESn Ausprägungen einer gemeinsamen Populations-effektstärke. Unterschiede zwischen den Studien ergeben sich nur durch die verschiedenen Stichproben (within-studies variance). Eine mögliche Varianz durch weitere,

---

unbekannte Faktoren wird ignoriert. Mittels eines Homogenitätstests kann diese Annahme (d.h. die beobachtete Varianz übersteigt nicht die aufgrund des Stichprobenfehlers zu erwartende Varianz) überprüft werden. Bei heterogenen Effekten darf die mittlere ES im FEM nicht interpretiert werden (Hedges et al., 1985; Shadish & Haddock, 1994). Eine Überprüfung der Homogenitätsannahme muss bei Einsatz des FEM somit unbedingt erfolgen<sup>21</sup>.

Bei Vorliegen von Heterogenität zwischen den ESn wird in der Regel durch das Bilden von Untergruppen versucht, die Variation zwischen den Studien aufzuklären. Dabei ist zu beachten, dass das Modell impliziert, dass nur Variablen einen Einfluss haben, die aus den Primärstudien codiert werden können. Andere Variablen haben nach den Modellannahmen keinen bedeutsamen Einfluss auf die ES, z.B. die Therapeuten-Klienten-Beziehung als kaum codierbarer Prädiktor. Zudem kann trotz der Bildung von (Unter-)Untergruppen weiterhin Heterogenität vorliegen, ein Problem mit dem im FEM nicht umgangen werden kann.

Die Modellannahmen bedingen auch den Grad der Generalisierbarkeit der Ergebnisse. Die Ergebnisse des FEM sind nur bzgl. der eingeschlossenen Studien interpretierbar: "inferences apply to this collection of studies and say nothing about other studies that may be done later, could have been done earlier, or may have already been done but are not included among the observed studies" (Hedges et al., 1998, S. 487). Die Ergebnisse des FEM sind somit anfälliger gegenüber Schwierigkeiten in der Lokalisation potentieller (publizierter) Primärstudien (vgl. allerdings Voraussetzungen des REM). Zudem darf eine Generalisierung der Ergebnisse nicht erfolgen. D.h. allerdings, wünscht man Aussagen zu (realen) Situationen, die nicht identisch sind zu denen der betrachteten Untersuchungen - man stellt also die Frage nach der allgemeinen Effektivität einer Intervention - so sind die Ergebnisse des FEM diesbezüglich nicht interpretierbar. Trotzdem erfolgt gerade dies in den meisten Meta-Analysen.

---

<sup>21</sup> Es ergibt sich hier allerdings zusätzlich das Problem, dass die Power des etabliertesten Homogenitätstests Q oft gering ist (vgl. Harwell, 1997, Hunter et al., 1990). Somit kann sich der Meta-Analyst fälschlicherweise „auf der sicheren Seite“ wähnen und das FEM einsetzen, obwohl eine Variation zwischen den ESn existiert.

Im REM wird im Gegensatz dazu ein Universum an ES<sub>n</sub> angenommen, d.h. es gibt nicht nur eine Ausprägung der ES sondern zufällige Variationen. Die Untersuchungen unterscheiden sich neben ihrer Stichprobe auch in ihren Charakteristika und in ihrer Effektstärke, d.h. die (ungeklärte) Heterogenität der Primärstudien wird in die Analyse einbezogen. Es wird eine zusätzliche Varianz eingeführt, die sich aus der Ziehung der Studien aus dem Universum ergibt (auch between-studies oder random effects Varianz genannt). Homogenität der ES<sub>n</sub> ist zur Interpretation der Ergebnisse nicht notwendig, aber möglich: ist die between-studies Varianz Null, entsprechen die Ergebnisse denen des FEM, da nur noch die within-studies Varianz in die Rechnung einfließt. Im Falle von Heterogenität ergibt sich die Varianz einer Studie aus der Summe von within- und between-studies Varianz. Durch diese zusätzliche Varianzkomponente sind auf Grund breiterer Konfidenzintervalle im Vergleich zum FEM konservativere Ergebnisse die Folge. Auch die mittlere ES unterscheidet sich zwischen den Modellen aufgrund der unterschiedlichen Varianzgewichte: je größer die between-studies Varianz ist, desto geringeren Einfluss zeigen die unterschiedlichen Stichprobengrößen der Primärstudien. Die Modelle sind in Abbildung 1 dargestellt.

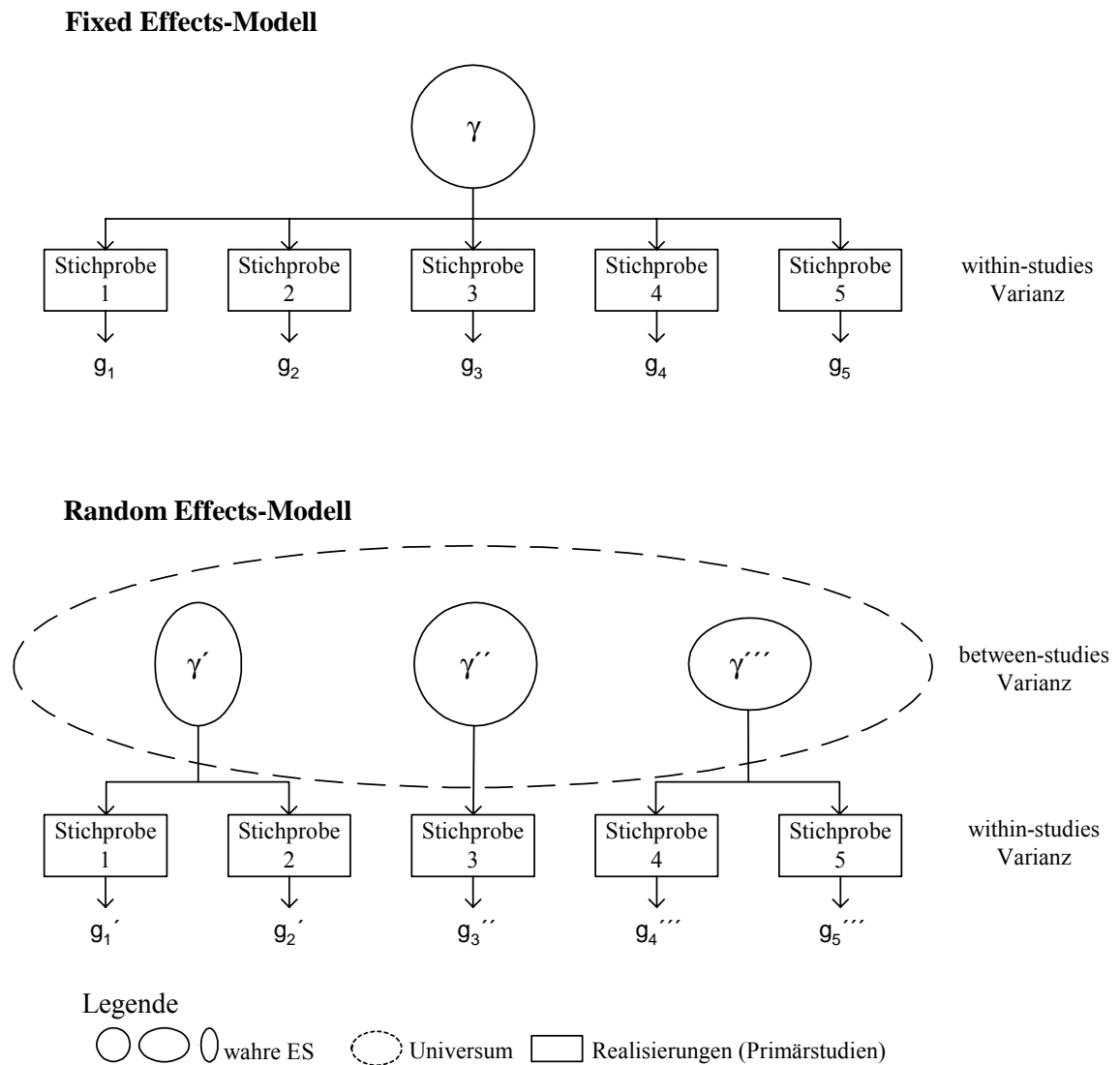


Abbildung 1: Darstellung von Fixed und Random Effects-Modell

Ist die Annahme einer between-studies Varianz gerechtfertigt? In der Psychotherapieforschung werden eine Vielzahl an Faktoren unterschieden, die in unterschiedlichem Ausmaß einen Einfluss auf die Effektivität einer Therapie haben: Variablen des Patienten, des Therapeuten, der Therapie und der Patienten-Therapeuten-Interaktion. Eine der wohl umfassendsten und anregendsten Zusammenfassungen bildet das Handbuch von Bergin und Garfield (1994). Für eine große Zahl dieser Aspekte gilt, dass sie nur in den seltensten Fällen aus den Effektivitätsstudien extrahiert werden können, z.B. die Güte der Therapeuten-Patienten-Beziehung oder das Ausmaß der Kooperation des Patienten. Auch wenn der Einfluss dieser Variablen gemittelt über alle Patienten einer Studie möglicherweise gering ist, erscheint eine vollständige Ignorierung dieser Faktoren, wie

im FEM, nicht sinnvoll. Lipsey und Wilson (2001a) untersuchten die Ergebnisse von 319 Meta-Analysen aus der Effektivitätsforschung. Es wurde versucht, die Varianz zwischen den Effektstärken zu erklären durch den Stichprobenfehler (entspricht der within-studies Varianz), Eigenschaften der Intervention (z.B. Dauer/Intensität) und Eigenschaften der Methode (z.B. Art der KG). Es zeigte sich, dass jede der drei Komponenten ca. 25 % Varianz aufklärte. Zum Vergleich: das FEM geht davon aus, dass nur die within-studies Varianz Unterschiede zwischen den Effekten erklärt, die restliche Varianz wird ignoriert! Zudem konnte ein weiteres Viertel der Varianz nicht durch die eingeschlossenen Studiencharakteristika erklärt werden, dies entspricht der between-studies Varianz nach Kontrolle potentieller Prädiktoren. Allerdings stellt die Verteilung der Varianz auf die Studieneigenschaften aufgrund einiger methodischer Probleme der Untersuchung nur approximative Ergebnisse dar. So hatten einige Primärstudien einen stärkeren Einfluss auf Ergebnis als andere, da sie in mehreren Meta-Analysen eingeschlossen wurden. Die unterschiedliche Berechnung der ES<sub>n</sub> zwischen den Meta-Analysen wurde nicht in Betracht gezogen. Dies könnte möglicherweise einen weiteren Varianzanteil aufklären. Ein Nachteil sämtlicher integrierender Verfahren ist zudem natürlich die Abhängigkeit von der Ergebnisdarstellung in den eingeschlossenen Untersuchungen (hier Meta-Analysen).

Das REM erlaubt Generalisierungen, so dass die Ergebnisse auch auf Ausprägungen des Treatments, die nicht genau den Situationen der Studien entsprechen, bezogen werden können. Die Annahmen des REM können als hilfreich angesehen werden, wenn nicht alle wichtigen Eigenschaften der Untersuchungen berichtet wurden bzw. davon ausgegangen werden kann, dass eine Anzahl unkontrollierbarer Faktoren einen Einfluss auf die Durchführung des Treatments oder der Messung haben und dies sich in den ES<sub>n</sub> niederschlägt (Overton, 1998; Raudenbush, 1994).

Aufgrund der unterschiedlichen Varianzberechnung und aufbauend auf dieser die Bestimmung der Konfidenzintervalle ist das FEM wie oben erwähnt im Vergleich zum REM weniger konservativ. Nach Overton (1998) wird dadurch allerdings die Angemessenheit der Aussagen verfälscht, es wird ein als sicher geltendes Ergebnis vorgestellt und die Unsicherheit der Resultate der Meta-Analysen unterschätzt. Dies wird unterstützt durch statistische Modelltests (Hedges et al., 1998; Overton, 1998), die zeigen konnten, dass bei Einsatz des FEM bei unbedingten Aussagen, d.h. bei der Intention zur

Generalisierung der Befunde über die eingeschlossenen Studien hinaus, und unvollkommener Homogenität zwischen den Studien, eine erhöhte Wahrscheinlichkeit des Zurückweisens der Nullhypothese zugunsten der Alternativhypothese gegeben war, d.h. eine Inflation des Fehlers erster Art. Auch Erez et al. (1995) stellten in Simulationen fest, dass zwar weder FEM noch REM exakte Ergebnisse lieferten, dass REM aber genauer war. Je größer die between-studies Varianz war, desto ungenauer wurden die FEM-Ergebnisse; die Konfidenzintervalle des FEM wurden zu eng berechnet, so dass im Extremfall nicht mehr der wahre Populationswert eingeschlossen wurde. Eine Interpretation der Ergebnisse des REM wird allerdings erst ab fünf Studien empfohlen, bei einer geringeren Anzahl ergaben die Modelltests nur approximative Ergebnisse (Hedges et al., 1998). Die optimale Studienanzahl scheint ab ca. 15 Studien gegeben (Field, 2001).

Das REM beinhaltet allerdings einige Probleme. Wie das Universum an Studien aussieht, ist nur anzunehmen. Eine Diskussion dieses Aspekts, d.h. eine Definition des Universums, auf welches die Ergebnisse zu beziehen sind, sollte aufgrund theoretischer Überlegungen bzw. zumindest der Eigenschaften der einbezogenen Untersuchungen erfolgen. Die Annahme der Zufallsauswahl der betrachteten Primärstudien aus diesem Universum ist nicht nachzuprüfen (Raudenbush, 1994) und erscheint aufgrund des File Drawer-Problems potentiell verletzt. Overton (1998) geht in diesem (wohl der Realität entsprechenden) Fall von einer möglichen Überschätzung der Variabilität zwischen den Effekten im REM aus, während das FEM diese unterschätzt.

Es sei angemerkt, dass in der Literatur auch ein Fixed Effects-Modell zu finden ist, welches Heterogenität zwar anerkennt, diese aber nicht modelliert (vgl. z.B. Overton, 1998). Hedges et al. (1998) sieht die Homogenitätsannahme somit nicht notwendigerweise als grundlegende Charakteristik des FEM. Auch in diesem Modell können aber die Ergebnisse nicht auf Situationen generalisiert werden, die nicht denen der Primärstudien entsprechen, sondern nur auf Studien, die in ihren Eigenschaften denen gleichen, die eingeschlossen wurden.

Zusammenfassend ist aus bisherigen Modelltests festzustellen, dass die optimale Berechnung der statistischen Kennwerte im Rahmen von Meta-Analysen noch nicht

---

existiert. Alle vorliegenden Schätzer sind im bestimmten Ausmaß fehlerbehaftet. Allerdings lassen sich drei bedeutsame Vorteile des REM im Vergleich zum FEM bei Meta-Analysen im Bereich der Psychotherapieforschung aufzeigen:

1. die Intention sind generalisierbare Aussagen (oder nach Matt (in press): „Will it work in Münster?“ (S.1)); diese lassen sich aber nur im REM überprüfen
2. die Annahme einer nicht-existenten between-studies Varianz ist a priori nicht aufrechtzuerhalten
3. bei einer between-studies Varianz von Null ergeben sich die gleichen Ergebnisse im FEM und REM, d.h. auch in diesem Fall kann das REM eingesetzt werden; ist sie dagegen größer Null werden im FEM fehlerhafte Kenngrößen berechnet

Die in der Literatur meist diskutierten Ansätze zur Berechnung der RE-Varianz sind zum einen der Momentmethodenschätzer sowie als iteratives Verfahren der Maximum Likelihood Schätzer. Simulationsstudien haben gezeigt, dass beide in etwa annähernde Ergebnisse liefern (G. Knapp, persönliche Mitteilung, 21.10. 2001), der Momentmethodenschätzer möglicherweise etwas genauer (Overton, 1998). Dieser wurde in der vorliegenden Arbeit eingesetzt. Zur Bestimmung erfolgte über die erste Methode nach Shadish et al. (1994). Wurde eine RE-Varianz größer Null berechnet, wurde im Rahmen einer Sensitivitätsanalyse überprüft, ob diese nur aufgrund einer einzelnen Studie vorlag<sup>22</sup>. Als Extremum wurde die größte Abweichung vom ungewichteten Mittel der ES<sub>n</sub> definiert. Zur Bestimmung der Konfidenzintervalle wurde die *t*-Verteilung statt der *z*-Verteilung eingesetzt nach den Empfehlungen von Overton (1998), wobei sich die Anzahl an Freiheitsgraden im FEM aus der Summe der einzelnen Stichprobengrößen dagegen im REM aus der Anzahl an Studien ergibt (jeweils subtrahiert mit Eins). Im Rahmen einer Sensitivitätsanalyse werden zusätzlich die Ergebnisse des FEM präsentiert (Berechnung entsprechend Shadish et al., 1994). Auch die Kennwerte der Regressionsanalysen wurden mit der Methode der Momente berechnet (Raudenbush, 1994). Die Varianzaufklärung bezieht sich hierbei auf die between-studies Varianz, so

---

<sup>22</sup> Dies wurde als pragmatisches Vorgehen eingesetzt zur Überprüfung der Bedeutsamkeit der RE-Varianz. Als Testverfahren gilt der Homogenitätstest Q, der aber, wie schon erwähnt, oft eine geringe Power aufweist. Zudem ist zu beachten, dass die Entscheidung, welches Modell in dieser Arbeit eingesetzt wird, nicht aufgrund des Ergebnisses der Homogenitätstests getroffen wurde, sondern a priori das Random Effects-Modell festgelegt wurde.

---

dass eine Berechnung der Regressionen nur durchgeführt wurde, wenn eine RE-Varianz größer Null (auch in den Sensitivitätsanalysen) gegeben war. Regressionen wurden nur für die Symptomkategorie Angst berechnet, da diese die bedeutsamste Kategorie darstellte. Mittels multipler Regressionsanalysen wurde der Einfluss der in Kapitel 2.3.4 genannten Variablen, mit Ausnahme des Ratings, auf die ES geprüft. Allerdings führt eine multiple Regression zum Ausschluss sämtlicher Studien, in denen mindestens ein Prädiktor fehlende Werte enthält, da ansonsten die Korrelationsmatrizen nicht definiert sind. Auf eine Berechnung von Interaktionen wurde entsprechend den Empfehlungen der Arbeitsgruppe um Shadish verzichtet (Shadish et al., 2000).

Für die Überprüfung der Vergleichbarkeit der Gruppen zu Prä werden nur die Ergebnisse des FEM präsentiert. Dies basiert auf folgenden Punkten: a) die Ergebnisse werden nicht generalisiert und b) Ziel ist die Überprüfung mit einer möglichst hohen Power. Da die Konfidenzintervalle enger für FEM- im Vergleich zu REM-Analysen und nicht-signifikante Ergebnisse somit stärker zu gewichten sind, wird hier das FEM eingesetzt.

Die Integration der Effektstärken über die Studien sowie die Regressionsanalysen erfolgten mittels einer Gewichtung über das Reziproke der Summe aus der Varianzkomponenten (Hedges et al., 1985; Raudenbush, 1994). Da die bedingte Varianz der Effektstärken abhängig von der Stichprobengröße ist, wird durch die Gewichtung auch die unterschiedliche Genauigkeit verschiedener Stichprobenumfänge berücksichtigt und der Stichprobenfehler gesenkt. Dabei erhalten größere Gruppengrößen höhere Gewichte. Shadish und Kollege (1994) fordern für Gewichtungsschemata folgende Eigenschaften: a) es ist theoretisch begründbar, dass bestimmte Studieneigenschaften zu mehr oder weniger fehlerbehafteten Ergebnissen führen, b) die Richtung dieses Fehlers ist a priori bekannt und c) es können angemessene Gewichte konstruiert werden. Sie schließen, dass es neben einer Gewichtung nach Varianz bzw. Reliabilitäts- und Validitätskoeffizienten anhand des heutigen Forschungsstandes keinen weiteren gesicherten und eindeutigen Aspekt gäbe, der die drei Punkte erfüllen könne (vgl. z.B. Kapitel 2.2.4 zur nicht eindeutigen Beziehung von Randomisierung und Ergebnis). Die Reliabilitäts- und Validitätskoeffizienten standen hier nicht zur Verfügung. Dementsprechend wurden keine weiteren Gewichtungen vorgenommen.



### **Prüfung der Homogenität**

Da in der vorliegenden Arbeit im Rahmen einer Sensitivitätsanalyse zusätzlich die Ergebnisse des FEM präsentiert werden, erfolgte eine Diagnose der Variabilität der Effektstärken über den Homogenitätstest Q (Shadish et al., 1994). Seine Prüfgröße Q folgt dabei einer Chi-Quadrat-Verteilung. Bei einem signifikanten Ergebnis muss die Nullhypothese der Gleichheit der Effektstärken zurückgewiesen werden. Aufgrund der geringen Power dieses Tests wurde a priori eine Irrtumswahrscheinlichkeit von 10 % festgelegt.

Da keine der vorliegenden Software zufrieden stellend die Organisation großer Datensätze sowie die notwendigen Berechnungen durchführte, wurde aufbauend auf Excel (Microsoft, Version XP) ein eigenes Programm entwickelt und an mehreren vorliegenden Datensätzen überprüft. Eine Beschreibung der Software ist in Anhang C zu finden.

### **2.3.3 Weitere Aspekte**

#### **Evaluation des Publikationsbias**

In der vorliegenden Arbeit wurde ein etabliertes Verfahren zur Einschätzung der Stabilität der Ergebnisse gegenüber einem möglichen Publikationsbias eingesetzt, das so genannte „Fail-Safe N“ (Orwin, 1983). Es besagt, wie viele Studien mit einer ES von  $g = 0$  benötigt würden, um einen beobachteten bedeutsamen Effekt in einen kleinen Effekt zu verändern<sup>23</sup>. Je höher Fail-Safe N ist, desto stabiler gelten die Ergebnisse der Meta-Analyse; Rosenthal (1979; n. Carson & Schriesheim, 1990) schlägt als Grenze das Fünffache der eingeschlossenen Studien plus 10 vor. Auf eine Exploration des Publikationsbias mittels Funnel-Plots (Begg, 1994), d.h. der graphischen Abbildung von ES<sub>n</sub> zu Stichprobenumfang, wurde verzichtet, da nach Ansicht der Autorin eine Interpretation der Darstellung möglicherweise nur gering reliabel ist und dies potentiell die Objektivität der Ergebnisse einschränkt.

---

<sup>23</sup> Dies ist die klassische Art des Einsatzes des Fail-Safe N. Die Berechnung erlaubt auch andere Fragestellungen, z.B. wie viele Studien mit einem großen Effekt benötigt würden, um einen beobachteten kleinen Effekt in einen mittleren zu verändern.

**Artefakte**

Die Arbeitsgruppe um Hunter (Hunter et al., 1990) zeigte mehrere systematische und unsystematische Artefakte bei ESn. So können sich Unterschiede zwischen ESn ergeben nur aufgrund der unterschiedlichen Reliabilität der Messinstrumente. Eine Korrektur wird empfohlen, wenn diese für die Mehrzahl an Studien vorliegt. Aufgrund der fehlenden Informationen wurde – der meta-analytischen Praxis entsprechend – auf eine Korrektur, auch bzgl. der anderen systematischen Artefakte, verzichtet. Der Einfluss unsystematischer Artefakte wurde dagegen wie oben beschrieben reduziert: a) Stichprobenfehler durch Bilden der gepoolten Standardabweichung, Korrektur für kleine Stichprobenumfänge und Gewichtung mit dem Reziproken der Varianz sowie b) Fehler in den Daten durch eine Ausreißeranalyse.

**Besonderheiten des Vergleichs von Pharmako- mit Psychotherapie**

Für einen Vergleich verschiedener Therapieformen, insbesondere von PhT und PsT, ergibt sich aus ESn zwischen Gruppen das Problem der Abhängigkeit der ES von der „Wirksamkeit“ der KG. Das heißt je stärker die KG ist, desto geringer ist die ES zwischen Gruppen zu post (unter der Annahme gleicher Prä-Werten). Problematisch ist dies für einen Therapievergleich bei systematisch unterschiedlichen KGn zwischen den untersuchten Verfahren. Leider lagen für fast alle Angststörungen keinerlei Studien vor, die, um die relative Wirksamkeit der verschiedenen KGn einschätzen zu können, einen direkten Vergleich insbesondere von WL und PP ermöglichten. Es wurden für jede Störung die Prä-Post-ESn von WL und PP in der Symptomkategorie Angst berechnet. Als Schätzer wurde die mittlere Differenz der Prä-Post-ESn berechnet. Die ES zwischen Gruppen der Psychotherapiebedingung wurde um diesen Wert angeglichen. Die so ermittelten neuen ES bildeten die Grundlage für den Therapievergleich. Es ergeben sich zwei Probleme bei diesem Verfahren: a) inwieweit stimmen die Berechnungsvarianten überein und b) es gelten die gleichen Probleme wie allgemein für Effekte aus Prä-Post-Berechnungen, d.h. vor allem bei einer nur geringen Anzahl an Untersuchungen sind Unterschiede in den Prä-Post-ESn verschiedener Gruppen möglicherweise auf verschiedenen ausgeprägte Fehler durch wiederholte Messung in den Messinstrumenten zurückzuführen (vgl. auch Kapitel 2.2.3), so dass zur Interpretation der Ergebnisse zusätzlich die Annahme eingeführt werden muss, dass zwischen Psycho- und Pharmako-

therapiestudien kein systematisch unterschiedlicher Retest-Effekt vorliegt. Zur Klärung der Übereinstimmung der Berechnungsvarianten wurde eine explorative Analyse durchgeführt.

Das Vorgehen stellt somit eine Adaption des Vorschlages von Becker (1988) dar. Im Gegensatz zu ihrer Methode muss die Annahme eines gleichen Retest-Effektes zwischen Gruppen erst für den Vergleich von Pharmako- und Psychotherapie eingeführt werden und nicht für die anderen Vergleiche. Welchen Vorteil bietet dieses Vorgehen im Gegensatz zu einem direkten Vergleich der Prä-Post-ESn der Behandlungsgruppen? Zum einen können auch die vorliegenden Studien ohne vorhandene Prä-Werte bzw. ohne Angaben zur Veränderung der Behandlungsgruppe einbezogen werden. Zum anderen erlaubt dieses Vorgehen eine Schätzung der Wirkung gemeinsamer Faktoren der Pillenplacebogruppe. Unter der Annahme einer additiven Zusammensetzung der einzelner Faktoren (vgl. z.B. Horvath, 1988) ergibt sich die Änderungen in der Symptomatik für die Warteliste aus dem Retest-Effekt, für die behandelten Kontrollgruppe aus Retest-Effekt und gemeinsamen Faktoren und für die Behandlungsgruppe aus Retest-Effekt, gemeinsamen Faktoren und spezifischen Faktoren.

#### **2.3.4 Beschreibung der untersuchten methodischen Variablen**

Die Auswahl der berücksichtigten methodischen Variablen erfolgte nach methodischen Gesichtspunkten, Ergebnissen bisheriger Meta-Analyse, sowie ausreichender Datenlage und Variabilität in den Primärstudien. Die Variablen dienten dabei auch im Vergleich verschiedener Therapieformen als Kovariate, die mit den untersuchten Therapien möglicherweise konfundiert waren. Es wurden folgende Prädiktoren eingeschlossen:

- a) Diagnostisches System
- b) Analyseart
- c) Dropouts
- d) Rating
- e) Stichprobengröße
- f) Placebowashout bei Pharmakotherapiestudien
- g) Art der Dosierung der Pharmakotherapiestudien
- h) Therapiedauer bei Psychotherapiestudien.

#### a) Diagnostisches System

Über die verschiedenen diagnostischen Systeme (vom DSM-III bis zum DSM-IV<sup>24</sup>; APA, 1980, APA, 1994) änderten sich die Kriterien zur Diagnose der Störungen. Markantes Beispiel ist die Generalisierte Angststörung. Im DSM-III (APA, 1980) waren zur Diagnose verschiedene psychophysiologische Symptome im Zusammenhang mit Angst oder Sorgen über verschiedene Lebensumstände für die Dauer von einem Monat ausreichend; GAS wurde als Restkategorie aufgefasst, die nicht codiert werden konnte, wenn eine andere Angststörung vorlag. Ab DSM-III-R (APA, 1987) wurde GAS als unabhängige Störung angesehen und es erfolgte bis zum DSM-IV eine Betonung der kognitiven Komponente im Sinne eines übermäßigen Sich Sorgen bei einer immer kleiner werdenden Anzahl somatischer Symptome sowie eine Verlängerung des Zeitraums auf 6 Monate. Es ergeben sich somit potentielle systematische Unterschiede in der Stichprobe zwischen Studien, die verschiedene Diagnosesysteme einsetzten. Bzgl. einer Generalisierbarkeit der gefundenen Ergebnisse soll der Einfluss der Diagnosestellung auf die Effektivität der untersuchten Therapien überprüft werden.

#### b) Analyseart

Wenn Patienten vorzeitig eine Studie verlassen, ergibt sich als Problem der Umgang mit ihren Daten. Bei der so genannten Completer-Analyse werden nur diejenigen Patienten eingeschlossen, die die Untersuchung beenden, das heißt Abbrecher werden in der Ergebnisanalyse ignoriert. Eine andere Möglichkeit ist die so genannte Endpoint- oder auch Intent-to-treat-Analyse (ITT). Hier werden die letzten verfügbaren Daten einer Person in die Endanalyse einbezogen. Letztere berücksichtigt bei den Ergebnissen somit auch die Fälle, in denen die Therapie gescheitert ist, sei es aufgrund geringer Effektivität, einer zusätzlichen benötigten Behandlung oder Nebenwirkungen. Unter der Annahme, dass die untersuchte Behandlung wirksam ist, sollten Patienten, die früher die Studie verlassen, im Mittel eine höhere Symptomatik aufweisen als Patienten am Ende einer Studie. So sollte sich bei ITT-Analysen aufgrund einer höheren mittleren Symptomatik sowie größerer Varianz in der Symptomatik zwischen den Patienten einer Gruppe eine geringere ES ergeben im Vergleich zur Completer-Analyse.

---

<sup>24</sup> Studien mit einer Klassifikation der Angststörungen nach DSM-II (APA, 1968) wurden weitgehend ausgeschlossen, da die neurotischen Störungen des DSM-II nicht eindeutig einer der aktuellen Kategorien der Angststörungen zuzuordnen sind.

### c) Dropouts

Es soll untersucht werden, welcher Zusammenhang zwischen der Anzahl an Abbrechern der Behandlungsgruppe zur ES besteht. Die bisherigen Befunde sind sehr divergent. Matt et al. (1997) fanden von neun Meta-Analysen, die diesen Zusammenhang untersuchten, bei vier Meta-Analysen keinen bedeutsamen Zusammenhang, bei drei einen negativen und bei zwei einen positiven. Generelle Vorhersagen scheinen somit sehr schwierig. Dies steht natürlich im Zusammenhang mit den sehr unterschiedlichen Gründen für ein Ausscheiden aus der Studie. Des Weiteren gilt der Umfang an Dropouts als Aspekt der Validität einer Studie. Da die restlichen Patienten potentiell nicht einer repräsentativen Stichprobe entsprechen, ist die Validität umso geringer, je mehr Personen abbrechen.

### d) Rating

Einige der bisherigen Meta-Analysen für Angststörungen zeigten höhere Effektstärken für eine Erhebung der Symptomatik durch den Therapeuten bzw. andere Fremdeinschätzer im Vergleich zu Selbstbeurteilungsinstrumenten (z.B. Abramowitz, 1996). Beide Verfahren unterliegen mehreren Bias, mögliche Einflussfaktoren betreffen z.B. einen Versuchsleitereffekt bei unvollständiger Blindheit gegenüber der Gruppenzugehörigkeit eines Patienten oder auch eine unterschiedliche Änderungssensitivität, zum Teil hervorgerufen durch eine potentiell höhere Varianz in der Einschätzung gleichen Verhaltens bei Selbstbeurteilung im Vergleich zur Fremdbeurteilung. Zur Einschätzung der Abhängigkeit des Therapieerfolges von der Art der Erhebung erfolgte eine Analyse des Einflusses von Fremd- vs. Selbstbeurteilung auf die Höhe der ES.

#### e) Stichprobengröße

Obwohl eine Reduktion des Stichprobenfehlers, im Besonderen relevant für kleine Stichproben, durch Einsatz verschiedener statistischer Verfahren erfolgte, wurde zusätzlich der Einfluss der Stichprobengröße betrachtet. Mehrere Meta-Analysen zeigten einen negativen Zusammenhang zwischen diesen Variablen (Matt et al., 1997). In der folgenden Meta-Analyse wurden trotz sorgfältiger Literatursuche überwiegend publizierte Studien eingeschlossen. Der Publikationsstatus scheint allerdings abhängig von den Ergebnissen einer Studie zu sein, im Besonderen von der Gegebenheit von Signifikanz. Diese wiederum ist abhängig von der Power einer Studie. Je größer die Stichprobe, desto höher ist die Power, geringe Effekte zu erkennen. Somit ist ein negativer Zusammenhang zwischen Stichprobengröße und Effektstärke zu erwarten.

#### f) Placebo Run-in bei Pharmakotherapiestudien

Bei einem Teil der vorliegenden Untersuchungen wurde allen Patienten bei Beginn der Studie für einen kurzen Zeitraum (in der Regel ein bis zwei Wochen) ein Pillenplacebo gegeben. Erfüllten die Patienten weiterhin die Selektionskriterien, wurden sie in die Studie eingeschlossen. Ziel ist eine Reduktion der Reaktion auf Placebo. Allerdings ist zu beachten, dass eine geringere Reaktion auf ein Placebo sich nicht nur in der Pillenplacebogruppe niederschlagen kann, sondern äquivalent in der Gruppe mit aktiver Substanz, in der ein Teil der Wirksamkeit auch unspezifische Faktoren sind. So fanden zwei bisherige Meta-Analysen keinen bedeutsamen Einfluss des Einsatzes eines Placebo Run-ins (Oosterbaan, 2001 bei SoP, Trivedi & Rush, 1994 bei Depression). Eine systematische Untersuchung für alle Angststörungen im Rahmen einer multiplen Regression wurde bislang nicht durchgeführt. Entgegen methodischer Intention ist allerdings kein bedeutsamer Einfluss des Placebo Run-in zu erwarten.

#### g) Art der Dosierung bei Pharmakotherapiestudien

Es liegen zwei Typen der Dosierung in Pharmakostudien vor, zum einen feste Dosierung, d.h. jeder Patient erhält die gleiche Menge an Substanz, zum anderen flexible Dosierung, d.h. die Patienten erhalten abhängig von Effektivität und Nebenwirkungen eine unterschiedliche Menge an Substanz. Dies hat Implikationen zum einen für eine Generalisierbarkeit der Ergebnisse auf die Praxis (in der flexibel dosiert wird), zum anderen aber auch für den Vergleich mit Psychotherapie, da diese verglichen

werden soll mit einer „best available“ Pharmakotherapie. Der Einfluss der Dosierungsart sollte in der vorliegenden Arbeit minimiert werden, so dass bei Studien mit mehreren Gruppen fester Dosierung nur diejenige mit der höchsten Effektivität eingeschlossen wurde. Es stellt sich nun die Frage, inwieweit sich die Wirksamkeit von Studien mit fester Dosierung von denen mit flexibler unterscheidet.

#### h) Therapiedauer bei Psychotherapiestudien

Die Befunde zum Einfluss der Therapiedauer auf die Effektivität einer Therapie sind divergent. In einer Zusammenfassung der Ergebnisse verschiedener Meta-Analysen zur Effektivität von Psychotherapie fanden Matt et al. (1997) in der Hälfte von 28 Meta-Analysen, die diese Variable untersuchten, einen positiven Zusammenhang. In den restlichen zeigte sich gleich verteilt ein negativer oder nicht bedeutsamer Zusammenhang. In der vorliegenden Arbeit soll deshalb erneut eine Überprüfung erfolgen auch im Hinblick auf eine Kontrolle einer potentiell unterschiedlichen Therapiedauer behavioraler und kognitiv-behavioraler Verfahren.

## 3 Ergebnisse

### 3.1 Empirische Exploration des Schätzers der Effektivität verschiedener Kontrollgruppen

Für den Vergleich von Psycho- und Pharmakotherapie war es notwendig, einen Schätzer für die relative Wirksamkeit der Kontrollbedingungen aus der Differenz der Prä-Post-ESn zu ermitteln und diesen in Beziehung zu setzen zu den berechneten ESn zwischen Gruppen. Es ist zu verifizieren, inwieweit diese Berechnung mit einer ES zwischen Gruppen zu vergleichen ist, wenn es ausreichend Studien gäbe die beide Kontrollgruppen betrachteten.

Basierend auf den vorliegenden Daten der Primärstudien wurden randomisiert 100 Fälle ausgewählt, in denen bzgl. eines Messinstrumentes der Kategorie Angst sowohl Mittelwerte und Standardabweichungen zu Prä und Post vorlagen. Bei mehr als zwei Gruppen erfolgte eine randomisierte Auswahl eines Gruppenvergleiches. Es wurde dann berechnet a) die ES zwischen Gruppen nach Hedges (Hedges et al., 1985) und b) die Differenz der Prä-Post-ESn der beiden Gruppen.

#### 3.1.1 Ergebnisse

Die Korrelation zwischen beiden Berechnungsvarianten war  $r = .79$ , die mittlere quadrierte Abweichung betrug 0.37. Die Verfahren korrelieren hoch miteinander, es bestehen aber Unterschiede. In Abbildung 2 ist ein Scatterdiagramm der beiden Varianten dargestellt. Die mittlere ES für die Berechnung aus den Post-Werten betrug  $g = 0.28$  ( $\sigma^2 = 0.007$ , 95%-CI: 0.11 – 0.43), die mittlere ES der Differenzen betrug  $g = 0.31$  ( $\sigma^2 = 0.011$ , 95%-CI: 0.10 – 0.52). Die beiden Verfahren unterscheiden sich über alle Studien nicht bedeutsam voneinander ( $\beta = 0.03$ , n.s.,  $R^2 = .00$ ). Bei einer genügend großen Anzahl an Studien scheinen sich somit Unterschiede zwischen den Berechnungsvarianten auszumitteln.



### 3.1.2 Schlussfolgerung

Bei einer ausreichend großen Anzahl an Studien scheint die Differenz der Prä-Post-ESn eine angemessene Lösung zur Schätzung der differentiellen Wirksamkeit der verschiedenen Kontrollgruppen zu sein. Dies basiert allerdings auf der Überprüfung der beiden Berechnungsvarianten an gleichen Stichproben und Messinstrumenten. Die Kombination der beiden Verfahren erscheint somit angemessen unter der Annahme eines gleichen Retest-Effekts.

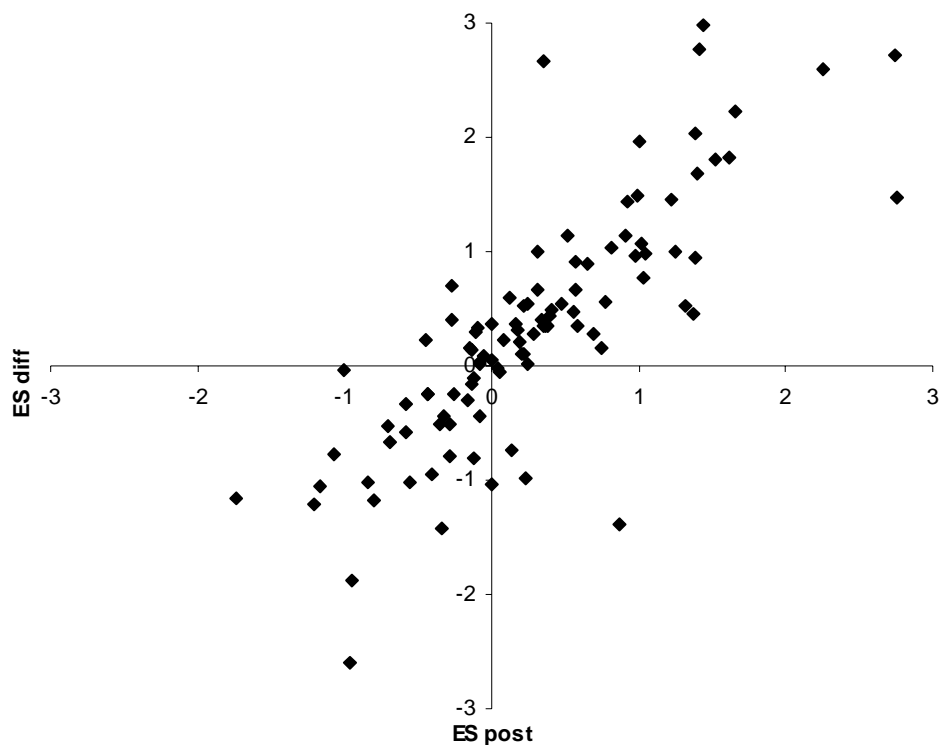


Abbildung 2: Scatterplot zum Zusammenhang verschiedener Berechnungsalternativen der Effektstärke (ES) direkt aus Postwerten bzw. ermittelt über die Differenz der Prä-Post-ES

### 3.2 Überblick über die eingeschlossenen Primärstudien

Insgesamt konnten die Daten von 340 Studien eingeschlossen werden. Als bestuntersuchte Angststörung kann die PDA mit ca. einem Drittel der Studien zählen. Ein Fünftel der Untersuchungen betrachteten Patienten mit GAS. PTB und SpP waren die am wenigsten untersuchten Störungen.

Betrachtete man die untersuchten Therapien, ergab sich eine deutliche Dominanz der medikamentösen Verfahren mit ca. zwei Drittel der eingeschlossenen Studien. Etwa jede fünfte Untersuchung betrachtete Psychotherapie allein. Der Unterschied in der Häufigkeit der Behandlung zeigte sich noch drastischer auf Patientenebene. Insgesamt wurden in den Studien über 30 000 Patienten untersucht. Nicht einmal jeder Zehnte nahm dabei an reinen Psychotherapiestudien teil. Neun von zehn Patienten wurden dagegen mit einer pharmakologischen Substanz einschließlich Pillenplacebo behandelt. Der hohe Prozentsatz an Pharmakotherapien fand sich übrigens auch in einer repräsentativen Bevölkerungserhebung zu Angstsyndromen in Deutschland von Margraf und Poldrack (2000). Von 2948 befragten Personen wurden 394 Personen<sup>25</sup> wegen Angstsyndromen behandelt, in 89.1 % der Fälle mit einer medikamentösen Therapie.

Innerhalb der einzelnen Angststörungen zeigten sich zum Teil starke Differenzen in der Häufigkeit der Therapien. Ergab sich in etwa eine Gleichverteilung an Studien bei PDA, PTB und SoP, wurden bei SpP häufiger psychologische Verfahren untersucht, bei GAS und OCD dagegen häufiger medikamentöse Behandlungen.

In Abbildung 3 findet sich die prozentuale Verteilung der Studien hinsichtlich der einzelnen Angststörungen, in Abbildung 4 ist dargestellt, wie viel Prozent der Studien welche Therapie untersuchten, sowie eine Aufteilung der Patienten auf die verschiedenen Therapien.

---

<sup>25</sup> Es sei zu beachten, dass fast zwei Drittel der Betroffenen subklinische Fälle waren definiert über einen BAI-Wert kleiner 11.

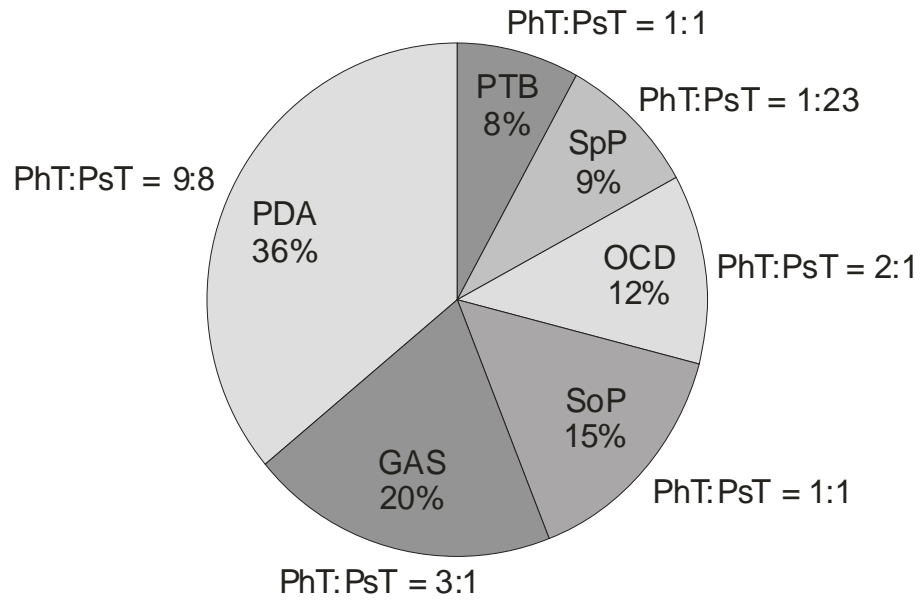


Abbildung 3: Prozentuale Aufteilung der Studien auf die einzelnen Angststörungen

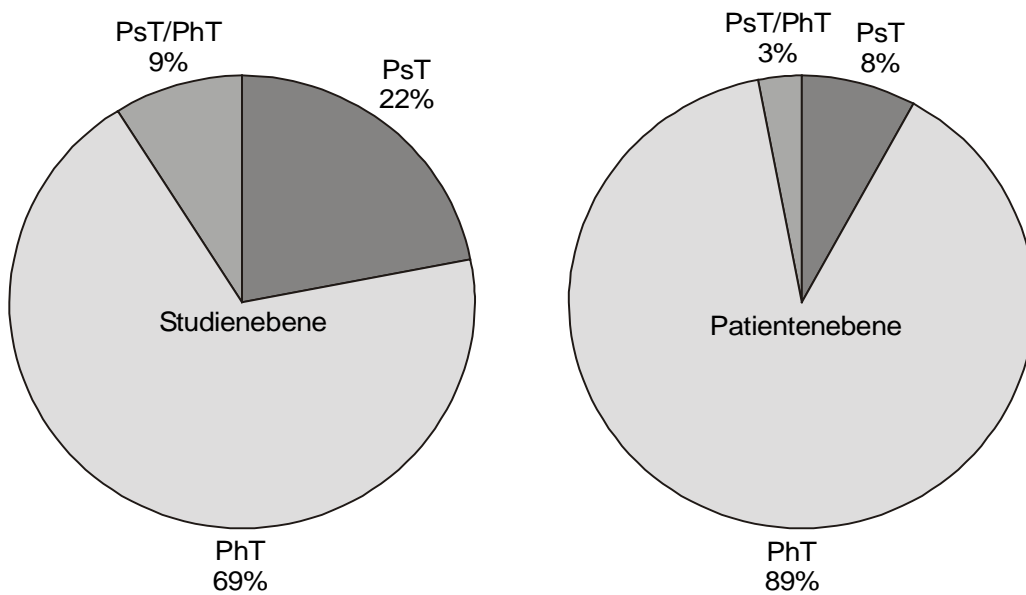


Abbildung 4: Prozentuale Aufteilung der Studien bzw. Patienten auf die untersuchten Therapien

In den folgenden Kapiteln sind die Ergebnisse der Meta-Analyse für die einzelnen Angststörungen dargestellt. Nach einer deskriptiven Beschreibung der Primärstudien folgt ein Vergleich der Prä-Werte im FEM-Modell. Anschließend sind die Ergebnisse für die Psychotherapien getrennt nach Art der Kontrollgruppe aufgeführt. Die Ergebnisse zur Stabilität und der Akzeptanz der Therapie sind dargestellt, sowie der Einfluss von Moderatoren auf die Effektstärken der Kategorie Angst. Es folgen die Analysen zum Vergleich von konfrontativen und kognitiv-behavioralen Verfahren. Abschließend erfolgt eine Beurteilung der Stabilität der Ergebnisse über Fail-Safe N. Im Anschluss sind die Ergebnisse für die Pharmakotherapie aufgeführt. Auch hier erfolgte eine Gliederung nach Überprüfung der Effektivität (einschließlich möglicher unterschiedlicher Wirksamkeit verschiedener Substanzklassen), Überprüfung der zeitlichen Stabilität, Einschätzung der Akzeptanz, Überprüfung des Einflusses von Moderatoren sowie die Einschätzung der Stabilität der Ergebnisse über Fail-Safe N. Es folgen die Ergebnisse des Vergleiches von Psycho- und Pharmakotherapie sowie von Psychotherapie und Kombinationsbehandlung.

Es sei wiederholt, dass die Analysen im REM bei einer Anzahl an Vergleichen von unter fünf nur approximative Ergebnisse darstellen und die optimale Anzahl erst ab 15 Vergleichen gegeben ist.

Die Ausreißeranalyse befindet sich in Anhang F.

Als Sensitivitätsanalysen sind dargestellt die Ergebnisse des FEM-Modells für alle Symptomkategorien. Für Angst als wichtigste Kategorie werden zudem die Ergebnisse der Modelle bei Ausschluss der Studien präsentiert, die keine Berechnung der Effektstärken über Mittelwerte und Standardabweichungen ermöglichten. Es sollte die Frage geklärt werden, wie sich die Resultate der Arbeit verändern, wenn hinsichtlich der verfügbaren Daten das konservativste Einschlusskriterium gewählt würde. Im Rahmen einer post hoc Analyse aufgrund der Ergebnisse für die Zwangsstörung wurde der Einfluss unterschiedlicher Schätzverfahren für die RE-Varianz auf die Resultate der Meta-Analyse überprüft. Diese Sensitivitätsanalyse ist in Anhang G dargestellt.

### 3.3 Generalisierte Angststörung

#### 3.3.1 Deskriptive Analyse der Primärstudien

Es konnten insgesamt 67 Studien gefunden werden, die den Inklusionskriterien entsprachen: 46 Pharmako- und 15 Psychotherapiestudien. Weitere 6 Studien untersuchten beide Verfahren. Es wurden mehr als zehnmal so viele Patienten in Pharmako- im Vergleich zu Psychotherapiestudien untersucht. Im Vergleich zu Daten aus epidemiologischen Studien war der Frauenanteil in den Studien geringer. Nur jede zehnte Untersuchung gab an, dass mindestens 25 % der Patienten eine komorbide Störung aufwiesen mit einer Gleichverteilung von affektiven Störungen, Abusus oder anderen Angststörungen, ein sehr geringer Prozentsatz bei den vorliegenden Komorbiditätsraten aus epidemiologischen Untersuchungen. Weitere deskriptive Daten finden sich in Tabelle 2.

**Tabelle 2:** Deskriptive Beschreibung der Primärstudien

	PsT	PhT	PsT / PhT
Patienten N <sup>a</sup>	599	7002	263
Durchschnittliche Gruppengröße <sup>a</sup>	17.0	52.3	13.2
n ≤ 20 (%)	76.9	28.3	66.7
n > 20 (%)	23.1	71.7	33.3
Alter	38.4	39.8	34.2
Dauer der Störung (Jahre)	11.0	9.3	1.5
Frauen (%)	66.3	58.1	49.2
Dauer der Therapie (PsT: Stunden, PhT: Wochen)	13.1	7.2	PsT: 9.7 PhT: 7.1
Publikationsjahr			
bis 1985 (%)	26.7	10.9	0
1986 – 1995 (%)	34.8	52.1	66.7
ab 1996 (%)	38.5	37.0	33.3
Intent-to-treat Auswertung (%)	7.7	76.1	16.7
Placebowashout (%)		80.4	33.3
Feste Dosierung (%)		41.3	50.0

<sup>a</sup> Zahlen geben die in die Auswertung aufgenommenen Patienten wider

### 3.3.2 Vergleich der Prä-Werte

In allen Symptomkategorien hatten Patienten der WL eine leicht niedrigere Symptomatik im Vergleich zur PsT. Nur einen geringen Unterschied ergab sich dabei in der Kategorie Angst ( $g = -0.10$ , 95%-CI:  $-0.29 - 0.09$ ,  $\chi^2 = 4.30$ ,  $n = 17$ ). In den Kategorien Depression und Lebensqualität zeigten sich dagegen mittlere ESn kleiner  $g = -0.20$  und bei Depression signifikant von Null verschieden ( $g = -0.29$ , 95%-CI:  $-0.49 - -0.09$ ,  $\chi^2 = 11.56$ ,  $n = 14$ ;  $g = -0.25$ , 95%-CI:  $-0.65 - 0.16$ ,  $\chi^2 = 2.50$ ,  $n = 4$ ). Auch der Ausschluss einer Studie mit Extremwert, definiert als größte Abweichung vom ungewichteten Mittel, veränderte die Ergebnisse für Depression nur unwesentlich ( $g = -0.24$ , 95%-CI:  $-0.45 - -0.03$ ,  $\chi^2 = 9.01$ ,  $n = 13$ ). Aufgrund der Richtung des Effektes wurde keine Modifikation der Studienpopulation zu Post durchgeführt, die Ergebnisse zu Post stellen somit konservative Resultate dar. Für Lebensqualität dagegen wurde eine Studie ausgeschlossen, der Unterschied zwischen WL und PsT zu Prä sank daraufhin auf nahe Null ( $g = -0.06$ , 95%-CI:  $-0.52 - 0.41$ ,  $\chi^2 < 0.01$ ,  $n = 3$ ). Die Studie wurde anschließend auch in den Berechnungen zu Post für diese Symptomkategorie ausgeschlossen.

Es zeigte sich kein bedeutsamer Unterschied in der Symptomatik zwischen PsT und behandelter Kontrollgruppe (PP/TP) zu Prä sowohl für Angst als auch Depression ( $g = 0.06$ , 95%-CI:  $-0.20 - 0.33$ ,  $\chi^2 = 3.12$ ,  $n = 8$ ;  $g = 0.10$ , 95%-CI:  $-0.18 - 0.38$ ,  $\chi^2 = 0.99$ ,  $n = 6$ ).

In Studien, die einen direkten Vergleich zwischen VT und KVT betrachteten, wiesen Patienten, die VT erhielten, zu Prä leicht höhere Werte auf, als Patienten, die mit KVT behandelt wurden sowohl in der Kategorie Angst als auch Depression. Keine der mittleren ESn war größer  $g = 0.20$  ( $g = 0.10$ , 95%-CI:  $-0.10 - 0.31$ ,  $\chi^2 = 9.73$ ,  $n = 11$ ;  $g = 0.13$ , 95%-CI:  $-0.10 - 0.36$ ,  $\chi^2 = 4.59$ ,  $n = 8$ ).

Weder in der Symptomkategorie Angst noch Depression zeigte sich ein signifikanter Unterschied zwischen PhT und Placebo, die mittlere ES lag bei Null ( $g = -0.08$ , 95%-CI:  $-0.16 - 0.80 \times 10^{-3}$ ,  $\chi^2 = 34.14$ ,  $n = 29$ ;  $g = -0.06$ , 95%-CI:  $-0.18 - 0.06$ ,  $\chi^2 = 5.78$ ,  $n = 11$ ).

---

Studien, die sowohl PhT als auch PsT untersuchten, zeigten in der Kategorie Angst eine mittlere ES nahe Null ( $g = 0.08$ , 95%-CI:  $-0.21 - 0.37$ ,  $\chi^2 = 3.82$ ,  $n = 8$ ).

### 3.3.3 Psychotherapie

#### Effektivität: Vergleich zur Wartelistenbedingung

Die Mehrzahl an Studien untersuchte kognitiv-behaviorale Verfahren meist in Verbindung mit konfrontativen Elementen. Ca. ein Fünftel untersuchte angewandte Entspannung. Alle ESn der Studien in den Kategorien waren positiv, d.h. keine der Studien zeigte für Angst, Depression oder Lebensqualität ein schlechteres Ergebnis der PsT im Vergleich zur WL. Es ergaben sich zu Post für alle Symptomkategorien signifikant positive mittlere ESn zwischen  $g = 0.76$  (Depression) und  $g = 1.02$  (Lebensqualität). Die untersuchten psychotherapeutischen Verfahren erscheinen somit sehr effektiv in der Behandlung der GAS. Es ist zu beachten, dass aufgrund der Unterschiede zu Prä für Depression die mittlere ES den wahren Effekt leicht unterschätzt. Die Ergebnisse der Analysen sind in Tabelle 3 dargestellt. Einen Überblick über die einzelnen Effektstärken bietet Abbildung 5.

In allen Kategorien lag Homogenität vor. Der Ausschluss der Studien, in denen keine Mittelwerte und Standardabweichungen vorlagen, veränderte die mittlere ES in der Kategorie Angst kaum (REM:  $g = 0.81$ , 95%-CI:  $0.59 - 1.03$ ,  $\sigma_{\theta}^2 = 0$ ; FEM:  $g = 0.81$ , 95%-CI:  $0.61 - 1.01$ ,  $\chi^2 = 10.38$ ,  $n = 17$ ).

Eine Studie untersuchte einen reinen Selbsthilfeansatz ohne therapeutische Begleitung. Es ergab sich eine ES von  $g = 0.95$  in der Kategorie Angst, die für ein hohes Potential dieses Verfahrens spricht, aber allgemeine Aussagen nicht zulässt.

Tabelle 3: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  zu Post für den Vergleich von behavioralen und kognitiv-behavioralen Verfahren zur Warteliste

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
Post	Angst	REM	0.82	0.09	0.62 – 1.01	0		19
		FEM	0.82	0.09	0.63 – 1.00		10.91	19
	Depression	REM	0.76	0.10	0.55 – 0.98	0		15
		FEM	0.76	0.10	0.57 – 0.96		8.69	15
	Lebensqualität	REM	1.02	0.18	0.52 – 1.51	0		5
		FEM	1.02	0.18	1.09 – 1.37		1.09	5

### Effektivität: Vergleich zur behandelten Kontrollgruppe

Im Vergleich zu einer behandelten Kontrollbedingung (Therapie- bzw. Pillenplacebo) ergaben sich durchgehend positive Studien-ESn, d.h. in jeder Studie waren die untersuchten Verfahren, meist kognitiv-behaviorale Therapie, effektiver als TP/PP. Die mittlere ES war wie zu erwarten kleiner als im Vergleich zur WL und betrug zwischen  $g = 0.52$  in der Kategorie Depression und  $g = 0.98$  für Klinische Signifikanz und war in jeder Kategorie bedeutsam von Null verschieden. Somit geht die Wirksamkeit dieser Behandlungen über eine Realisierung gemeinsamer Faktoren hinaus. Die Ergebnisse der Analysen sind in Tabelle 4 dargestellt. Einen Überblick über die einzelnen Effektstärken bietet Abbildung 5.

Es lag in keiner der Kategorien Heterogenität vor. Bei Ausschluss einer Studie, in der keine Mittelwerte und Standardabweichungen zur Berechnung der ESn vorlagen, stieg die mittlere ES in der Kategorie Angst unwesentlich an (REM:  $g = 0.61$ , 95%-CI: 0.32 – 0.90,  $\sigma_{\theta}^2 = 0$ ; FEM:  $g = 0.61$ , 95%-CI: 0.37 – 0.85,  $\chi^2 = 5.94$ ,  $n = 8$ ).



Tabelle 4: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  zu Post für den Vergleich von behavioralen und kognitiv-behavioralen Verfahren zur Placebobedingung

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
Post	Angst	REM	0.57	0.12	0.30 – 0.85	0		9
		FEM	0.57	0.12	0.34 – 0.80		7.28	9
	Depression	REM	0.52	0.14	0.15 – 0.89	0		6
		FEM	0.52	0.14	0.24 – 0.80		1.06	6
	Klin. Sign.	REM	0.98	0.19	0.38 – 1.57	0		4
		FEM	0.98	0.19	0.61 – 1.35		2.80	4

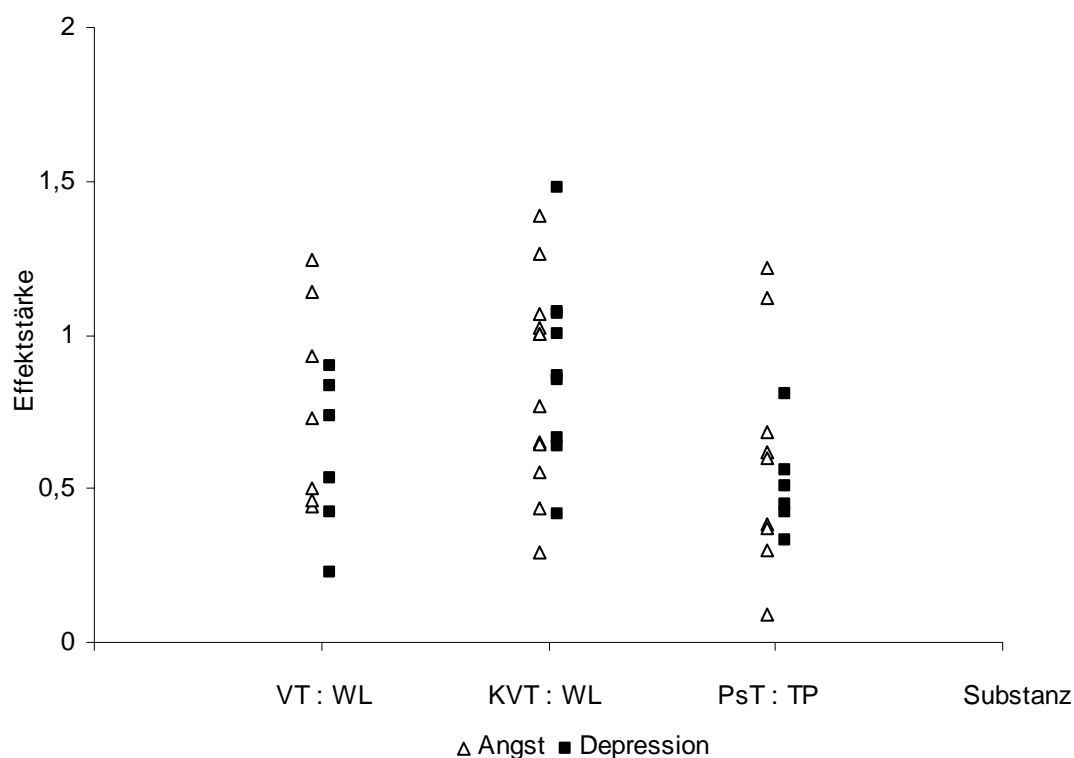


Abbildung 5: Studieneffektstärken für die Symptomkategorien Angst und Depression aufgeteilt nach Art der Kontroll- und Behandlungsgruppe

### Zeitliche Stabilität der Therapie

Aufgrund der Anzahl an Studien, wurde in der Kategorie Angst eine Gliederung der Zeiträume unter sechs bzw. über sechs Monate eingesetzt, für die anderen Symptomkategorien wurden alle Daten zusammen ausgewertet. Die Messungen zu Follow up lassen sich für die Symptomkategorie Angst wie folgt charakterisieren:  $f1 = 1.9$  Monate ( $SD = 1.23$ ) und  $f2 = 8.5$  Monate ( $SD = 3.09$ ). Für die Kategorien Depression und

Klinische Signifikanz betrug der mittlere Katamnesezeitraum 6.4 Monate ( $SD = 4.32$ ) und 8.7 Monate ( $SD = 3.16$ ). Die Wirksamkeit von PsT erwies sich als zeitlich stabil, es gab keine bedeutsame Abnahme der ESn. Für das Langzeit-Follow up zeigte sich, bezogen auf die eingeschlossenen Studien, eine signifikante Zunahme in der Kategorie Angst  $g = 0.20$ . Die bedeutsame Verbesserung der Patienten in der Hauptsymptomatik kann aber noch nicht generalisiert werden. Dies resultierte nicht aus der RE-Varianz, sondern aufgrund der Anzahl an Studien. Die Ergebnisse der Analysen finden sich in Tabelle 5.

Tabelle 5: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  der Veränderung von Post zu Follow up (f) für behaviorale und kognitiv-behaviorale Verfahren

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
f1	Angst	REM	0.07	0.10	-0.15 – 0.30	0		14
		FEM	0.07	0.10	-0.13 – 0.27		3.62	14
f2	Angst	REM	0.20	0.09	-0.01 – 0.40	0		12
		FEM	0.20	0.09	0.02 – 0.38		1.75	12
f1/f2	Depression	REM	0.05	0.10	-0.16 – 0.26	0.05		16
		FEM	0.07	0.08	-0.09 – 0.23		15.19	16
	Klin. Sign.	REM	0.02	0.17	-0.37 – 0.42	0.15		9
		FEM	0.02	0.11	-0.20 – 0.25		18.93 <sup>+</sup>	9

<sup>+</sup>  $p < .10$

### Akzeptanz

Bei behavioraler und kognitiv-behavioraler Therapie brach jeder Zehnte vorzeitig die Studie ab ( $M = 9.9$ ,  $SD = 10.17$ ,  $n = 27$ ). Behaviorale Verfahren zeigten dabei mit 11.1 % ( $SD = 11.24$ ,  $n = 10$ ) eine ähnliche Abbrecherrate im Vergleich mit kognitiv-behavioralen Verfahren mit 9.2 % ( $SD = 9.77$ ,  $n = 17$ ), der Unterschied war nicht signifikant ( $t(25) = 0.47$ , n.s.)<sup>26</sup>. Für die WL-Bedingung betrug die Abbrecherquote 9.9 % ( $SD = 15.29$ ,  $n = 11$ ), für die Placebobedingung 14.0 % ( $SD = 11.98$ ,  $n = 5$ ).

<sup>26</sup> Normalverteilung gegeben.

## Moderatoren

Für die Vergleiche mit WL und mit TP/PP entfielen die multiple Regressionen, da die RE-Varianz Null war. Es lagen für WL leicht höhere ESn basierend auf Fremdbeurteilungsverfahren vor, der Unterschied zur Selbstbeurteilung war aber nicht bedeutsam ( $\beta = -0.14$ , n.s.,  $R^2 = .00$ ;  $n = 25$ ). Dagegen zeigte sich eine signifikant höhere Effektstärke für Fremd- vs. Selbstbeurteilungen im Vergleich der PsT zu TP/PP ( $\beta = -0.54$ ,  $p < .01$ ,  $R^2 = 1.00$ ;  $n = 17$ ). Dieses Ergebnis blieb stabil bei Ausschluss der Studie, in der nur Selbstbeurteilung eingesetzt wurde ( $\beta = -0.51$ ,  $p < .01$ ,  $R^2 = 1.00$ ;  $n = 16$ ) und tendenziell, bzgl. einer möglichen Konfundierung der Ergebnisse mit einer unterschiedlichen Änderungsrate verschiedener Symptomkategorien, auch bei alleiniger Betrachtung der Instrumente der Kategorie Angst ( $\beta = -0.32$ ,  $p < .10$ ,  $R^2 = 1.00$ ;  $n = 16$ ). Die bedeutsam höheren Effektstärken der Fremdbeurteilungsverfahren fanden sich in den Prä-Post-ESn nicht für die Placebobedingung ( $\beta = -0.35$ , n.s.,  $R^2 = 0.65$ ;  $n = 10$ ), aber sehr deutlich für die Behandlungsgruppen ( $\beta = -1.07$ ,  $p < .05$ ,  $R^2 = 0.38$ ;  $n = 16$ ). Somit können als Ursache weder eine potentiell unterschiedliche Änderungssensitivität noch Versuchsleitereffekte ausgeschlossen werden. Zudem ist zu beachten, dass dies die einzige Analyse war, in der es für diese Angststörung einen bedeutsamen Unterschied zwischen Fremd- und Selbstbeurteilung gab. Eine weitere Aufklärung der gefundenen Ergebnisse erscheint aufgrund der Vielfältigkeit der eingesetzten Instrumente zwischen den Studien mit gleichzeitig nur geringer Gesamtanzahl an Untersuchungen nicht möglich.

## Vergleich konfrontative und kognitiv-behaviorale Verfahren

Bei Betrachtung der ES aus dem WL-BG-Vergleich ergaben sich, wie aufgrund der RE-Varianz von Null zu erwarten, keine Unterschiede zwischen den Verfahren in der Kategorie Angst, weder in der einfachen Regression ( $\beta = 0.04$ , n.s.,  $R^2 = .00$ ;  $n = 19$ ), noch bei Kontrolle von Stichprobengröße, Abbrecher, Diagnostischen System und Therapiedauer ( $\beta_{VT:KVT} = 0.05$ , n.s.,  $\beta_N = -0.11 \times 10^{-1}$ , n.s.,  $\beta_{Dropout} = -0.69 \times 10^{-2}$ , n.s.,  $\beta_{Diagnose} = 0.28$ , n.s.,  $\beta_{Dauer} = -0.12 \times 10^{-3}$ , n.s.,  $R^2 = .00$ ;  $n = 19$ ). In der Kategorie Depression zeigten kognitiv-behaviorale Verfahren eine leicht höhere mittlere ES als rein konfrontative Verfahren, dieser Unterschied war allerdings weder in der einfachen ( $\beta = 0.26$ , n.s.,  $R^2 = .00$ ;  $n = 15$ ) noch in der multiplen Regression signifikant ( $\beta_{VT:KVT} = 0.37$ , n.s.,  $\beta_N = -0.26 \times 10^{-2}$ , n.s.,  $\beta_{Dropout} = 0.57 \times 10^{-2}$ , n.s.,  $\beta_{Diagnose} = 0.25$ , n.s.,

---

$\beta_{\text{Dauer}} = -0.15 \times 10^{-2}$ , n.s.,  $R^2 = .00$ ;  $n = 15$ ). Ein Vergleich über die Placebothherapie entfiel aufgrund der geringen Anzahl an Studien mit konfrontativen Verfahren für diese Berechnungen.

Meist wurden für den direkten Vergleich Angewandte Entspannung mit einer kognitiv-behavioralen Therapie in den Studien untersucht. In allen Kategorien ergaben einige der Studien bessere Ergebnisse für konfrontative Verfahren, andere Studien dagegen bessere Ergebnisse für kognitiv-behaviorale Verfahren. Die Variabilität zwischen den Untersuchungen war aber nicht bedeutsam. Es zeigte sich in keiner der Kategorien eine signifikante mittlere ES, beide Verfahren sind somit in der Behandlung der GAD einschließlich der depressiven Symptomatik als gleich effektiv einzuschätzen. Die Ergebnisse der Analysen sind in Tabelle 6 dargestellt. Einen Überblick über die einzelnen Effektstärken bietet Abbildung 6.

Es lag in keiner der Kategorien bedeutsame Heterogenität vor. Betrachtete man nur die ES<sub>n</sub>, für die eine Berechnung über Mittelwerte und Standardabweichungen möglich war, veränderte sich die mittlere ES für die Kategorie Angst kaum (REM:  $g = 0.07$ , 95%-CI:  $-0.17 - 0.30$ ,  $\sigma_{\theta}^2 = 0$ ; FEM:  $g = 0.07$ , 95%-CI:  $-0.14 - 0.27$ ,  $\chi^2 = 6.78$ ,  $n = 11$ ). Es zeigte sich kein signifikanter Unterschied zwischen Fremd- und Selbstbeurteilungsinstrumenten ( $\beta = -0.16$ , n.s.,  $R^2 = .00$ ;  $n = 18$ ).

Die Ergebnisse blieben im Follow up stabil, es zeigte sich kein bedeutsamer Unterschied zwischen den Verfahren. Der durchschnittliche Katamnesezeitraum betrug dabei für die Kategorie Angst 5.9 Monate ( $SD = 4.04$ ).

Tabelle 6: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  zu Post und Follow up (f) für den Vergleich von konfrontativen und kognitiv-behavioralen Verfahren

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
Post	Angst	REM	0.08	0.11	-0.16 – 0.31	0		11
		FEM	0.08	0.11	-0.13 – 0.29		5.84	11
	Depression	REM	0.16	0.12	-0.12 – 0.43	0		8
		FEM	0.16	0.12	-0.07 – 0.38		4.79	8
	Klin. Sign.	REM	-0.07	0.15	-0.47 – 0.33	0.02		6
		FEM	-0.07	0.14	-0.35 – 0.21		5.92	6
f	Angst	REM	0.15	0.12	-0.13 – 0.42	0		9
		FEM	0.15	0.12	-0.09 – 0.38		4.44	9
	Depression	REM	0.03	0.14	-0.32 – 0.39	0.01		6
		FEM	0.03	0.14	-0.24 – 0.30		5.02	6
	Klin. Sign.	REM	0.37	0.31	-0.63 – 1.37	0.27		4
		FEM	0.34	0.18	-0.01 – 0.68		8.58 <sup>+</sup>	4

<sup>+</sup>  $p < .10$

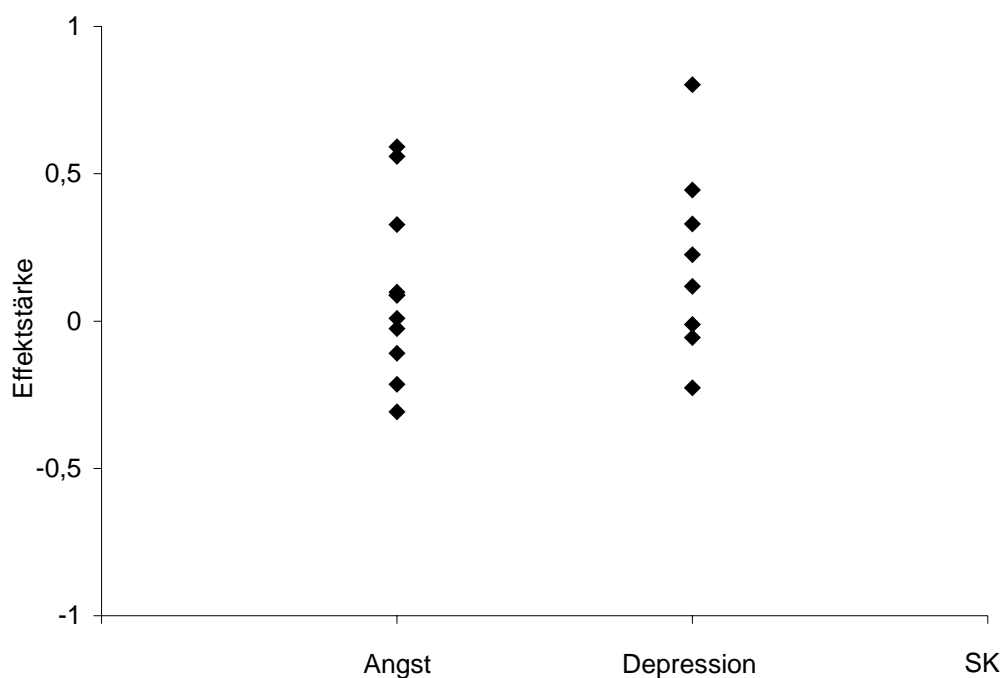


Abbildung 6: Effektstärken für die Symptomkategorien (SK) Angst und Depression der einzelnen Studien für den direkten Vergleich von konfrontativen und kognitiv-behavioralen Verfahren

### **Fail Safe-N**

Für eine Reduzierung der beobachteten Effekte des Vergleiches PsT zu WL auf eine kleine ES von  $g = 0.20$  sind 59 Gruppenvergleiche mit  $g = 0$  notwendig, ca. die Hälfte der von Rosenthal vorgeschlagenen Grenze für stabile Ergebnisse. Das Fail-Safe N ist zumindest fast dreimal so hoch, wie die Anzahl an eingeschlossenen Vergleichen. Für einen Vergleich mit einer behandelten Kontrollgruppe werden nur 18 Studien benötigt, sowie 27 Studien mit leicht höheren Effektivität für die Placebobedingung mit einer Effektstärke von  $g = -0.20$  für einen mittleren Effekt von  $g = 0$ . Obwohl auch hier die Grenze nach Rosenthal mit 55 Studien unterschritten wird, sind zumindest zwei- bis dreimal so viele weitere Studien für eine bedeutsame Änderung des mittleren Effektes notwendig. Für den direkten Vergleich von konfrontativen mit kognitiv-behavioralen Therapien wären 10 bzw. 4 Studien notwendig mit einer jeweils mittleren Effektstärke von  $g = 0.50$  für eines der Verfahren, um einen geringen Effekt von  $g = 0.20$  für konfrontative oder für kognitiv-behaviorale Behandlungen zu erhalten. Die Grenze liegt hier bei 65, es werden zur Veränderung der mittleren Effektstärke weniger Studien benötigt als eingeschlossen wurden. Die Ergebnisse können nicht als stabil gelten.

### **3.3.4 Pharmakotherapie**

#### **Effektivität: Vergleich zur Placebothherapie**

Die am häufigsten untersuchten Substanzklassen waren mit ca. der Hälfte der Studien BZD und mit ca. einem Viertel Azapirone (Serotoninagonist, z.B. Buspiron). Die Mehrzahl der ESn war positiv, nur 2 % der Studieneffektstärken in den Hauptkategorien ergaben dagegen ein besseres Ergebnis für die PP-Bedingung mit ESn bis zu  $g = -0.40$ . In allen Symptomkategorien zeigte sich zu Post eine signifikant positive mittlere ES zwischen  $g = 0.31$  (Depression) und  $g = 0.53$  (Klinische Signifikanz). Die Ergebnisse der Analysen sind in Tabelle 7 dargestellt. Einen Überblick über die einzelnen Effektstärken bietet Abbildung 7.

Besonders für diese Störung mussten für die Sensitivitätsanalyse eine Vielzahl an Studien ausgeschlossen werden, betrachtete man nur die ESn, die aus Mittelwerten und Standardabweichungen berechnet werden konnten. Die mittlere ES stieg in der Kategorie Angst an (REM:  $g = 0.48$ , 95%-CI: 0.32 – 0.64,  $\sigma_{\theta}^2 = 0.04$ ; FEM:  $g = 0.44$ , 95%-CI:

0.34 – 0.55,  $\chi^2 = 24.28$ ,  $n = 19$ ). Es stellt sich allerdings die Frage, ob die restlichen Untersuchungen eine repräsentative Stichprobe darstellen.

Tabelle 7: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_\theta^2$  zu Post für den Vergleich einer medikamentösen zur Placebobedingung

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_\theta^2$	$\chi^2$	n
Post	Angst	REM	0.32	0.03	0.25 – 0.38	0.03 <sup>a</sup>		68
		FEM	0.30	0.02	0.25 – 0.35		74.89	68
	Depression	REM	0.31	0.05	0.20 – 0.42	0.03		26
		FEM	0.30	0.04	0.22 – 0.38		27.77	26
	Klin. Sign.	REM	0.52	0.06	0.40 – 0.65	0.10		44
		FEM	0.45	0.03	0.40 – 0.51		121.22 <sup>+</sup>	44

<sup>a</sup>  $\sigma_\theta^2 > 0$  auch bei Ausschluss der Studie mit extremster ES definiert über die größte Abweichung vom ungewichteten Mittel

<sup>+</sup>  $p < .10$

Für die einzelnen Substanzklassen ergab sich für die Symptomkategorie Angst eine mittlere ES von  $g = 0.32$  für BZD und  $g = 0.30$  für Azapirone. Die multiple Regression mit Kontrolle der Variablen Diagnosesystem, Dosierung, PP Runin, Dropouts, Analyseart und Stichprobengröße ergab keinen signifikanten Unterschied für Azapirone vs. BZD ( $\beta = 0.05$ , n.s.,  $R^2 = .22$ ,  $n = 48$ ). Die am häufigsten untersuchten Substanzklassen unterschieden sich in ihrer Effektivität für die Angstsymptomatik nicht bedeutsam voneinander. Äquivalente Ergebnisse zeigten sich in der Kategorie Depression. Hier ergab sich eine mittlere ES von  $g = 0.22$  für Azapirone und  $g = 0.33$  für BZD. Die multiple Regression<sup>27</sup> ergab keinen Unterschied in der Effektivität ( $\beta = 0.08$ , n.s.,  $R^2 = .74$ ,  $n = 17$ ).

<sup>27</sup> PP Runin nicht als Prädiktor aufgenommen, da alle Studien diesem Design entsprachen.

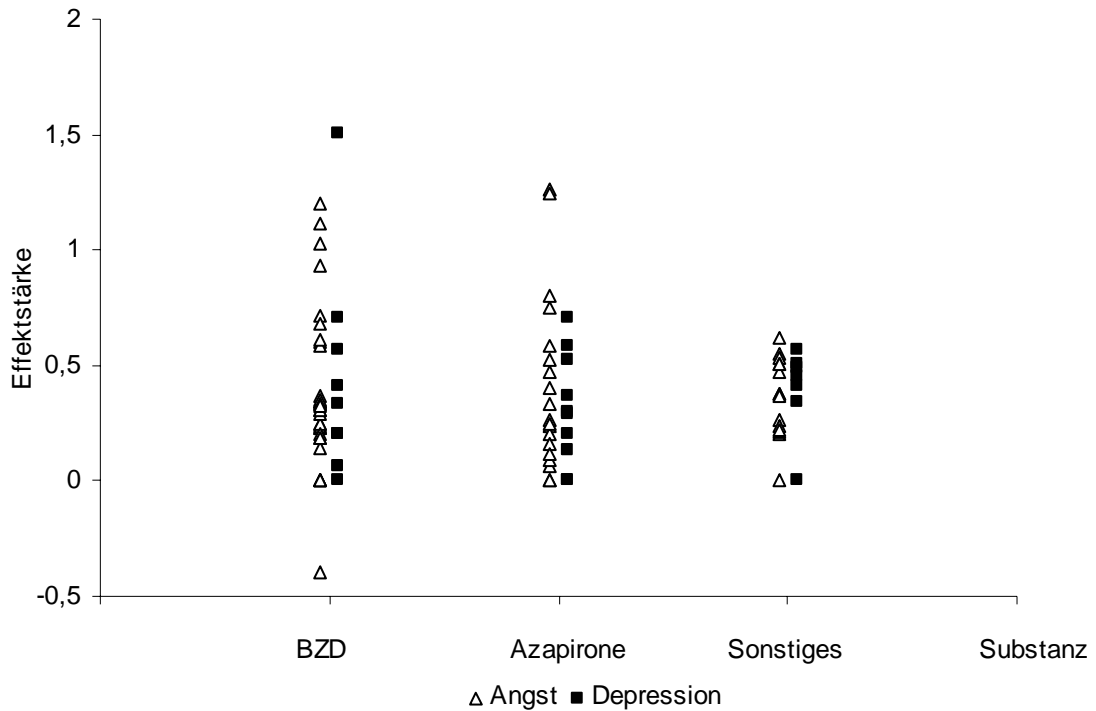


Abbildung 7: Effektstärken der einzelnen Studien für die Symptomkategorien Angst und Depression aufgeteilt nach Substanzklasse (BZD = Benzodiazepine)

### Zeitliche Stabilität der Therapie

Zu Follow up konnte keine mittlere ES berechnet werden. Es lagen in drei Studien Daten für die Kategorie Angst vor, alle Studien betrachteten BZD. In einer Studie wurde die aktive Substanz graduiert ersetzt mit Pillenplacebo. Es zeigte sich nach 1 Monat keine bedeutsame Veränderung ( $g = 0.07$ ). In zwei weiteren Studien wurde das Medikament abgesetzt. Es ergab sich in Studie 1 nach einem Monat eine deutliche Verschlechterung mit  $g = -0.77$  und in Studie 2 nach 4 Monaten keine Veränderung mit  $g = -0.01$ . Auch für die anderen Kategorien lagen nur unzureichend Daten vor. Generelle Aussagen sind nicht möglich.



### Akzeptanz

Die mittlere Dropout-Rate in den BGn betrug 24.4 % ( $SD$ : 15.57,  $n = 65$ ), d.h. jeder Vierte brach vorzeitig die Behandlung ab. Die Akzeptanz erscheint insgesamt somit eher gering. Für Benzodiazepine betrug die mittlere Abbrecherquote 20.5 % ( $SD = 14.30$ ,  $n = 35$ ) und für Azapirone 30.7 % ( $SD = 17.96$ ,  $n = 18$ ). Der Unterschied war signifikant ( $t(51) = -2.25$ ,  $p < .05$ )<sup>28</sup>. Bei PP verließ ca. jeder Dritte vorzeitig die Studie ( $M = 30.2$ ,  $SD = 22.32$ ,  $n = 40$ ).

### Moderatoren

Die multiple Regression ergab, dass ESn tendenziell größer waren, wenn die Stichprobengröße kleiner war. Weder das diagnostische System, die Art der Dosierung, die Anzahl an Abbrechern, die Analyseart noch ob ein PP Runin durchgeführt wurde zeigte einen bedeutsamen Zusammenhang zur ES. Die Ergebnisse sind in Tabelle 8 dargestellt. Es zeigte sich kein bedeutsamer Unterschied in den ESn zwischen Fremd- und Selbstrating ( $\beta = -0.08$ , n.s.,  $R^2 = .01$ ;  $n = 120$ ).

Tabelle 8: Ergebnisse der multiplen Regression zum Einfluss verschiedener Variablen auf die Effektstärke zwischen Pharmakotherapie und Pillenplacebo

Variable	B
Intercept	0.95*
Diagnosestellung	-0.05
Dosierung	-0.04
PP-Washout	-0.20
Dropouts	$-0.22 \times 10^{-2}$
Analyse	-0.04
Stichprobengröße	$-0.96^+ \times 10^{-3}$
Benzodiazepine	-0.13
Azapirone	-0.09

$R^2 = .23$ ,  $n = 60$ , \*  $p < .05$ , +  $p < .10$

<sup>28</sup> Normalverteilung gegeben

## Fail Safe-N

Für eine Reduzierung der beobachteten Effekte auf eine kleine ES von  $g = 0.20$  sind 42 Gruppenvergleiche mit  $g = 0$  notwendig, ca. ein Zehntel der von Rosenthal vorgeschlagenen Grenze für stabile Ergebnisse. Das Fail-Safe N beträgt circa die Hälfte der eingeschlossenen Vergleiche.

### 3.3.5 Vergleich Psycho- und Pharmakotherapie

#### Indirekter Vergleich

Es zeigte sich in den Prä-Post-ESn für die Kategorie Angst eine Differenz von  $g = 0.94$  zwischen PP und WL und für die Kategorie Depression ein Unterschied von  $g = 0.85$  zwischen PP und WL mit jeweils höheren ESn für PP. Es ist allerdings zu beachten, dass, wie schon in der Sensitivitätsanalyse bzgl. der Berechnung der ES aus Mittelwerten und Standardabweichungen im Rahmen der PhT erwähnt, nur eine geringe Anzahl an PhT-Studien eine Berechnung der Prä-Post-ESn ermöglichte ( $n = 20$  für Angst und  $n = 7$  für Depression), so dass die Repräsentativität dieser Schätzungen und darauf aufbauender Analysen möglicherweise eingeschränkt ist. Der Vergleich zwischen TP und PP wurde aufgrund mangelnder Daten nicht durchgeführt. In der Kategorie Angst zeigte sich nach Korrektur der Post-ESn um diese Werte eine signifikant leicht höhere Wirksamkeit von PhT im Vergleich zu PsT ( $\beta = -0.31, p < .01, R^2 = .44, n = 91$ ). Auch nach Kontrolle von Diagnosesystem, Abbrecherrate und Stichprobengröße<sup>29</sup> in der multiplen Regression blieb der Unterschied bestehen und stieg für PhT an ( $\beta_{\text{PhT:PsT}} = -0.44, p < .01, \beta_{\text{Diagnose}} = -0.04, \text{n.s.}, \beta_{\text{N}} = -0.26 \times 10^{-2}, \text{n.s.}, \beta_{\text{N}} = -0.68 \times 10^{-3}, \text{n.s.}, R^2 = .49; n = 83$ ). Dieses Ergebnis blieb stabil bei Ausschluss weniger etablierter Behandlungen (definiert als bei PsT nicht zu konfrontativen bzw. kognitiv-behavioralen Verfahren und bei PhT nicht zu den am häufigsten untersuchten Substanzklassen BZD oder Azapirone gehörend;  $\beta_{\text{PhT:PsT}} = -0.45, p < .01, \beta_{\text{Diagnose}} = -0.07, \text{n.s.}, \beta_{\text{N}} = -0.12 \times 10^{-2}, \text{n.s.}, \beta_{\text{N}} = -0.27 \times 10^{-2}, \text{n.s.}, R^2 = .49; n = 70$ ). Nach Korrektur zeigte sich für die Kategorie Depression eine signifikant kleinere ES für PsT im Vergleich zur PhT ( $\beta = -0.34, \text{n.s.}, R^2 = 1.00, n = 42$ ). Der Unterschied vergrößerte sich mit Kontrolle von Diagnosesystem, Abbrecherrate und Stichprobengröße als auch

<sup>29</sup> Analyseart wurde aufgrund der Verteilungseigenschaften nicht einbezogen.

nach Ausschluss weniger etablierter Behandlungen ( $\beta_{\text{PhT:PsT}} = -0.54$ ,  $p < .01$ ,  $\beta_{\text{Diagnose}} = 0.02$ , n.s.,  $\beta_{\text{Dropout}} = -0.50 \times 10^{-2}$ , n.s.,  $\beta_{\text{N}} = -0.14 \times 10^{-2}$ , n.s.,  $R^2 = 1.00$ ;  $n = 39$ ;  $\beta_{\text{PhTvsPsT}} = -0.49$ ,  $p < .05$ ,  $\beta_{\text{Diagnose}} = 0.07$ , n.s.,  $\beta_{\text{Dropout}} = -0.42 \times 10^{-2}$ , n.s.,  $\beta_{\text{N}} = -0.13 \times 10^{-2}$ , n.s.,  $R^2 = 1.00$ ;  $n = 34$ ).

### Direkter Vergleich

Es lagen 6 Studien vor, die PhT und PsT direkt verglichen, meist mit BZD. Die ESn der Studien schwankten, so lagen in der Kategorie Angst sowohl Studien vor, die ein besseres Ergebnis der PhT zeigten (bis  $g = -0.52$ ), als auch Studien, in denen PsT wirksamer war (bis  $g = 0.85$ ). Die mittlere ES war leicht positiv und im FEM signifikant von Null verschieden, nicht aber im REM (REM:  $g = 0.33$ , 95%-CI:  $-0.02 - 0.67$ ,  $\sigma_{\theta}^2 = 0$ ; FEM:  $g = 0.33$ , 95%-CI:  $0.04 - 0.61$ ,  $\chi^2 = 6.80$ ,  $n = 8$ ). Die eingeschlossenen Studien zeigen somit im Gegensatz zu dem indirekten Vergleich eine bedeutend höhere Effektivität der PsT im Vergleich zur PhT. Diese Ergebnisse können allerdings noch nicht generalisiert werden. Zudem sind die Resultate abhängig von den Entscheidungen des Wissenschaftlers: bei Ausschluss der Studien, die keine Berechnung der ESn über Mittelwerte und Standardabweichungen erlaubten, sank die mittlere ES leicht und war weder im REM, noch im FEM bedeutsam von Null verschieden (REM:  $g = 0.28$ , 95%-CI:  $-0.10 - 0.66$ ,  $\sigma_{\theta}^2 = 0$ ; FEM:  $g = 0.28$ , 95%-CI:  $-0.02 - 0.59$ ,  $\chi^2 = 6.15$ ,  $n = 7$ ). Daten für andere Kategorien lagen nur unzureichend vor. Aufgrund der geringen Anzahl an Studien wird auf eine graphische Darstellung der Effektstärken verzichtet.

In drei Studien lagen Daten für Follow up von 1 bis 4 Monaten vor, es zeigten sich ausschließlich positive ES, d.h. bessere Ergebnisse für PsT im Bereich von  $g = 0.32$  bis  $g = 0.88$ . Weitere Untersuchungen sind allerdings zu fordern.

Aufgrund der RE-Varianz von Null entfiel eine Berechnung der Regressionen. Des weiteren zeigte sich kein signifikanter Unterschied zwischen Fremd- und Selbstrating ( $\beta = 0.08$ , n.s.,  $R^2 = .00$ ;  $n = 9$ ).

In einer Studie wurde PsT verglichen mit einer Kombination aus PsT und PhT. Es zeigte sich hier eine ES von  $g = 0.46$  für die Kategorie Angst mit höherer Effektivität für das Kombinationsverfahren.

### **Akzeptanz**

Es zeigte sich über alle Studien eine signifikant geringere Abbrecherrate für Psycho- im Vergleich zu Pharmakotherapie ( $t(73.1) = 5.30, p < .01$ )<sup>30</sup>. Die Akzeptanz für behaviorale und kognitiv-behaviorale Therapien ist somit von Patienten mit GAD bedeutsam höher.

### **Fail Safe-N**

Wie aufgrund der geringen Anzahl an Vergleichen zu erwarten, zeigen sich die Ergebnisse der Analysen als eher instabil. So sind nur fünf Studien mit einer Effektstärke von  $g = 0$  notwendig für einen mittleren Effekt von  $g = 0.20$ . Ebenfalls fünf Studien werden benötigt für einen mittleren Effekt von  $g = 0.50$  ausgehend von Untersuchungen, die eine hohe Effektivität der Psychotherapie zu Pharmakotherapie mit  $g = 0.80$  zeigen. Immerhin fast doppelt soviel Vergleiche wie vorhanden mit  $g = -0.50$  sind notwendig für einen kleinen Effekt für Pharmakotherapie mit  $g = -0.20$ . Die Grenze nach Rosenthal liegt bei jeweils 50 Studien.

---

<sup>30</sup> Normalverteilung gegeben

### 3.3.6 Diskussion

Die Ergebnisse der Meta-Analyse können folgendermaßen zusammengefasst werden. Die untersuchten psychotherapeutischen Verfahren, insbesondere die kognitiv-behaviorale Behandlung und Angewandte Entspannung waren sehr wirksam bei Patienten mit Generalisierter Angststörung. Die gefundenen hohen Effektstärken zeigten sich dabei in allen Symptomkategorien, wenn auch für Lebensqualität nur eine eingeschränkte Generalisierbarkeit der Ergebnisse aufgrund der geringen Anzahl an Studien gegeben ist. Es zeigte sich eine bedeutend höhere Effektivität behavioraler und kognitiv-behavioraler Verfahren im Vergleich zu einer behandelten Kontrollgruppe, die Wirksamkeit der Therapien geht somit über eine Realisierung gemeinsamer Faktoren hinaus. Obwohl auch im REM die mittlere Effektstärke bedeutsam war, scheinen für eine Generalisierung der Ergebnisse weitere Untersuchungen angemessen. Ca. drei Viertel der Effektivität der untersuchten Psychotherapien ist auf spezifische Faktoren zurückzuführen.

Die Akzeptanz der Therapie ist als sehr gut einzuschätzen. Nur jeder Zehnte brach vorzeitig die Behandlung ab.

Nach Ende der Behandlung verschlechterte sich die Symptomatik der Patienten nicht, in den vorliegenden Studien ergab sich sogar eine bedeutsame Verbesserung nach mehr als sechs Monaten. Generell zeigen die Daten, dass die gefundenen Effekte stabil über die Zeit sind.

Weder der indirekte noch der direkte Vergleich ergab einen bedeutsamen Unterschied zwischen konfrontativen und kognitiv-behavioralen Verfahren, die Ergebnisse deuten auf eine vergleichbare Effektivität in der Behandlung der GAS hin. Brewin (1996) schlug vor, dass kognitive Techniken besonders bei weniger eng umschriebenen Problemen, wie bei der GAS vorhanden, eingesetzt werden sollten. Die zugrunde liegenden unbewussten kognitiven Strukturen seien nur schwer situational zugänglich und darum konfrontative Verfahren nicht die Methode der Wahl. Die vorliegenden Befunde unterstützen nicht diese Annahme.

In einigen der Evaluationsstudien erfolgte keine Diagnose der GAS nach einem diagnostischen System. Aufgrund der gefundenen geringen Variabilität der Ergebnisse zwischen den Untersuchungen, scheinen aber Einschränkungen in der Generalisierbarkeit gering.

Die untersuchten pharmakologischen Therapie, vor allem BZD und Azapirone, waren effektiv in der Behandlung, es ergab sich eine Effektstärke im kleinen Bereich. Zwischen den beiden Substanzklassen zeigten sich keine bedeutsamen Unterschiede, die Befunde sprechen für eine vergleichbare Wirksamkeit sowohl für die Hauptsymptomatik Angst als auch für Depression. Die Akzeptanz der medikamentösen Verfahren war allerdings eher gering: insgesamt jeder vierte Patient brach vorzeitig die Untersuchung ab. Es zeigte sich eine geringere Abbrecherrate für BZD im Vergleich zu Azapironen. Beendete bei BZD jeder Fünfte vorzeitig die Behandlung, war es bei Azapironen jeder Dritte. Bei gleicher Effektivität der Medikamente sprechen die vorliegenden Daten somit eher für einen Einsatz der BZD bei Patienten mit GAS.

Nur ein Viertel der Wirksamkeit war auf spezifische Faktoren zurückführbar. Es zeigte sich bei dieser Störung eine sehr starke Placeboantwort mit einem gleichzeitigen Retest-Effekt nahe Null. Lagen für andere Angststörungen die Prä-Post-Effektstärken um  $g = 0.50$  für Pillenplacebo, ergab sich für GAS eine Effektstärke von  $g = 0.94$ .

Es ergaben sich sehr divergente Ergebnisse hinsichtlich des Vergleiches der Effektivität von Psycho- und Pharmakotherapie. Im indirekten Vergleich ergab sich eine bedeutsam höhere Wirksamkeit für eine medikamentöse Behandlung, der Unterschied war im mittleren Bereich. In Studien, die beide Behandlungen direkt verglichen, ergab sich dagegen eine bedeutsam höhere Effektstärke für Psychotherapieverfahren. Aufgrund der geringen Anzahl an Studien kann diese allerdings nicht verallgemeinert werden, zudem können schon wenige weitere Untersuchungen das Ergebnis substanziell verändern. Für eine Diskussion der jeweiligen Vor- und Nachteile der beiden Berechnungen sei auf das Kapitel zur Panikstörung verwiesen. Ein weiterer Aspekt mit besonderem Stellenwert für GAS wird im Hauptdiskussionsteil erörtert.

Die Akzeptanz war für psychologische Therapien im Vergleich zu Psychopharmaka bedeutend höher. Eine abschließende Empfehlung für eines der Verfahren erscheint aber aufgrund der stark divergenten Befunde verfrüht.

Wie auch bisherige Meta-Analysen zeigte die vorliegende Arbeit die Effektivität von Psycho- und Pharmakotherapie, sowie die zeitliche Stabilität der Effekte bei psychologischen Behandlungen. Die eingesetzten Methoden erlauben dabei eine Generalisierbarkeit der Ergebnisse. Die gefundene Vergleichbarkeit der Abbrecherrate zwischen Pharmako- und Psychotherapie bei Gould & Otto et al. (1997) konnte nicht bestätigt werden, allerdings konnten in die aktuellen Meta-Analyse fast doppelt so viele Studien eingeschlossen werden. Die Autoren fanden keinen bedeutsamen Unterschied zwischen Pharmako- und Psychotherapie, der Vergleich erfolgte aber unabhängig von der eingesetzten Kontrollgruppe und erscheint somit fehlerbehaftet.

### 3.4 Panikstörung mit und ohne Agoraphobie, Agoraphobie ohne Panikstörung in der Vorgeschichte

#### 3.4.1 Deskriptive Analyse der Primärstudien

Es konnten 53 Pharmako- und 47 Psychotherapiestudien gefunden werden, die den Inklusionskriterien entsprachen. In weiteren 24 Studien wurden beide Verfahren zusammen untersucht. Die Mehrzahl an Patienten nahm an medikamentösen Studien teil, auch die Gruppengröße war hier am höchsten. Ca. zwei Drittel bis vier Fünftel der Patienten waren weiblich. Ca. jede zehnte Studie gab an, dass mindestens 25 % der Patienten eine komorbide Störung aufwiesen, meist eine andere Angststörung oder Depression. Die Charakteristika der Studien sind in Tabelle 9 dargestellt.

Tabelle 9: Deskriptive Beschreibung der Primärstudien

	PsT	PhT	PsT / PhT
Patienten N <sup>a</sup>	1811	7725	1747
Durchschnittliche Gruppengröße	15.9	53.5	23.5
n ≤ 20 (%)	78.7	39.6	50.0
n > 20 (%)	21.3	60.4	50.0
Alter	36.7	35.6	36.5
Dauer der Störung (Jahre)	9.0	8.3	8.4
Frauen (%)	74.2	65.1	76.6
Dauer der Therapie (PsT: Stunden, PhT: Wochen)	16.2	8.3	PsT: 13.9 PhT: 12.2
Publikationsjahr			
bis 1985 (%)	14.9	0	25.0
1986 – 1995 (%)	53.2	51.9	41.7
ab 1996 (%)	31.9	49.1	33.3
Intent-to-treat Auswertung (%)	0	67.9	25.0
Placebowashout (%)		49.1	20.8
Feste Dosierung (%)		32.1	4.2

<sup>a</sup> Zahlen geben die in die Auswertung aufgenommenen Patienten wider



### 3.4.2 Vergleich der Prä-Werte

Im Vergleich zwischen PsT und WL zeigte sich weder in den Symptomkategorien Angst, Depressionen noch Lebensqualität ein bedeutsamer Unterschied, alle mittleren ESn lagen um Null ( $g = -0.04$ , 95%-CI:  $-0.17 - 0.10$ ,  $\chi^2 = 12.18$ ,  $n = 29$ ;  $g = -0.09$ , 95%-CI:  $-0.24 - 0.06$ ,  $\chi^2 = 20.71$ ,  $n = 21$ ;  $g = -0.01$ , 95%-CI:  $-0.24 - 0.21$ ,  $\chi^2 = 17.02$ ,  $n = 8$ ). In der Kategorie Lebensqualität zeigte sich somit Heterogenität, die mittlere ES nahe Null resultierte aus einer sehr hohen Variabilität zwischen den Effekten von  $g = -0.90$  bis  $g = 0.31$ . Für den Vergleich von Selbsthilfeansätzen zur WL ergab sich für die Symptomkategorie Angst kein bedeutsamer Unterschied zu Prä, für die Kategorie Depression eine mittlere ES kleiner als  $g = -0.20$  und Heterogenität zwischen den ESn ( $g = -0.07$ , 95%-CI:  $-0.36 - 0.22$ ,  $\chi^2 = 3.78$ ,  $n = 7$ ;  $g = -0.23$ , 95%-CI:  $-0.56 - 0.10$ ,  $\chi^2 = 21.86$ ,  $n = 6$ ). Bei Ausschluss einer Studie mit Extremwert ( $g = -2.85$ ) zeigte sich kein Unterschied mehr zwischen den Bedingungen ( $g = -0.01$ , 95%-CI:  $-0.36 - 0.33$ ,  $\chi^2 = 2.17$ ,  $n = 5$ ). Diese Studie wurde auch in der weiteren Analyse für die Symptomkategorie Depression ausgeschlossen.

In allen Symptomkategorien (Angst, Depression, Lebensqualität) zeigten sich im Vergleich zwischen PsT und TP schwach positive mittlere ESn, d.h. dass die Patienten in der Kontrollbedingung eine leicht schwerere Symptomatik aufwiesen ( $g = 0.10$ , 95%-CI:  $-0.07 - 0.27$ ,  $\chi^2 = 4.00$ ,  $n = 13$ ;  $g = 0.20$ , 95%-CI:  $-0.01 - 0.41$ ,  $\chi^2 = 9.76$ ,  $n = 8$ ;  $g = 0.11$ , 95%-CI:  $-0.11 - 0.32$ ,  $\chi^2 = 10.31$ ,  $n = 7$ ). Die Unterschiede waren nicht signifikant, aber für Depression war die mittlere ES größer  $g = 0.20$ . Nach Ausschluss einer Studie mit der höchsten Differenz zu Prä ( $g = 1.01$ ) betrug die mittlere ES  $g = 0.14$  (95%-CI:  $-0.08 - 0.36$ ,  $\chi^2 = 5.61$ ,  $n = 7$ ).

In Studien, die einen direkten Vergleich zwischen VT und KVT betrachteten, wiesen Patienten, die VT erhielten, zu Prä leicht höhere Werte auf, als Patienten, die mit KVT behandelt wurden sowohl in der Kategorie Angst, als auch Depression und Lebensqualität. Keine der mittleren ESn war größer  $g = 0.20$  ( $g = 0.13$ , 95%-CI:  $-0.04 - 0.30$ ,  $\chi^2 = 5.75$ ,  $n = 15$ ;  $g = 0.13$ , 95%-CI:  $-0.04 - 0.30$ ,  $\chi^2 = 8.45$ ,  $n = 14$ ;  $g = 0.13$ , 95%-CI:  $-0.14 - 0.40$ ,  $\chi^2 = 2.45$ ,  $n = 4$ ).

Weder in der Symptomkategorie Angst, Depression noch Lebensqualität zeigte sich ein signifikanter Unterschied zwischen PhT und Placebo, die mittlere ES lag bei Null ( $g = 0.04$ , 95%-CI:  $-0.02 - 0.10$ ,  $\chi^2 = 33.05$ ,  $n = 57$ ;  $g = 0.02$ , 95%-CI:  $-0.07 - 0.11$ ,  $\chi^2 = 26.84$ ,  $n = 25$ ;  $g = -0.01$ , 95%-CI:  $-0.12 - 0.09$ ,  $\chi^2 = 20.09$ ,  $n = 21$ ).

Studien, die sowohl PhT als auch PsT untersuchten, zeigten sowohl in der Kategorie Angst als auch Depression eine mittlere ES nahe Null ( $g = 0.05$ , 95%-CI:  $-0.12 - 0.22$ ,  $\chi^2 = 1.57$ ,  $n = 10$ ;  $g = 0.05$ , 95%-CI:  $-0.16 - 0.27$ ,  $\chi^2 = 2.87$ ,  $n = 8$ ).

Weder in der Symptomkategorie Angst, Depression noch Lebensqualität zeigte sich ein signifikanter Unterschied zwischen PsT und PsT+PhT, die mittlere ES lag jeweils bei Null ( $g = 0.02$ , 95%-CI:  $-0.13 - 0.17$ ,  $\chi^2 = 4.24$ ,  $n = 18$ ;  $g = -0.03$ , 95%-CI:  $-0.23 - 0.18$ ,  $\chi^2 = 10.29$ ,  $n = 9$ ;  $g = -0.10$ , 95%-CI:  $-0.36 - 0.15$ ,  $\chi^2 = 4.85$ ,  $n = 5$ ). Die Gruppen waren zu Prä vergleichbar.

### 3.4.3 Psychotherapie

#### Effektivität: Vergleich zur Wartelistenbedingung

Ca. zwei Drittel der Studien untersuchten kognitiv-behaviorale Verfahren, meist mit zusätzlichen konfrontativen Verfahren. Die restlichen Studien untersuchten vorwiegend Exposition. Die Mehrzahl an ESN in den Hauptkategorien war positiv, nur eine Studie zeigte eine leicht negative ES mit  $g = -0.02$  (Angst) bzw.  $g = -0.23$  (Depression). Es zeigte sich eine hohe Effektivität von VT und KVT: in allen Symptomkategorien ergaben sich zu post signifikant positive mittlere ESN im Vergleich zur WL im Bereich von  $g = 0.72$  (Depression) bis  $g = 1.36$  (Klinische Signifikanz). PsT scheint dabei auf alle Dimensionen der Angst eine ähnliche Wirksamkeit zu haben. Die Ergebnisse der Analysen sind in Tabelle 10 dargestellt. Einen Überblick über die Studieneffektstärken bietet Abbildung 8.

Bei Ausschluss der Studien, in denen keine Mittelwerte und Standardabweichungen vorlagen, stieg die mittlere ES in der Kategorie Angst leicht an (REM:  $g = 0.92$ , 95%-CI:  $0.73 - 1.10$ ,  $\sigma_{\theta}^2 = 0.04$ ; FEM:  $g = 0.92$ , 95%-CI:  $0.77 - 1.08$ ,  $\chi^2 = 34.02$ ).

Tabelle 10: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  zu Post für den Vergleich von behavioralen und kognitiv-behavioralen Verfahren zur Warteliste

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
Post	Angst	REM	0.87	0.08	0.71 – 1.03	0.04 <sup>a</sup>		32
		FEM	0.87	0.07	0.73 – 1.01		41.11	32
	Vermeidung	REM	0.92	0.13	0.65 – 1.18	0.20		24
		FEM	0.88	0.08	0.72 – 1.04		41.66 <sup>+</sup>	24
	Kognition	REM	0.82	0.11	0.57 – 1.07	0.08		19
		FEM	0.81	0.10	0.62 – 1.00		23.44	19
	Erregung	REM	0.82	0.08	0.63 – 1.01	0.05		28
		FEM	0.82	0.09	0.66 – 0.98		32.23	28
	Depression	REM	0.72	0.09	0.54 – 0.90	0.01		23
		FEM	0.72	0.08	0.56 – 0.88		23.01	23
	Klin. Sign.	REM	1.36	0.12	1.10 – 1.62	0.13		21
		FEM	1.35	0.09	1.18 – 1.54		35.47 <sup>+</sup>	21
	Lebensqualität	REM	0.85	0.15	0.48 – 1.21	0.06		8
		FEM	0.84	0.13	0.59 – 1.09		11.79	8

<sup>a</sup>  $\sigma_{\theta}^2 > 0$  bei Ausschluss der Studie mit extremster ES definiert über die größte Abweichung vom ungewichteten Mittel

<sup>+</sup>  $p < .10$

Acht Studien untersuchten die Wirksamkeit eines Selbsthilfeansatzes im Vergleich zu einer Wartelistenbedingung, in der Hälfte der Studien wurde dabei das Buch Coping with Panic (Clum, 1989) eingesetzt. In den Hauptkategorien Angst und Klinische Signifikanz ergaben sich nur positive ESn in den Primärstudien, in der Kategorie Depression eine leicht negative ESn mit  $g = -0.02$ . Die mittleren ESn waren kleiner als die therapeutengeführten Behandlungen im Bereich von  $g = 0.62$  (Depression) bis  $g = 1.13$  (Kognition) und mit Ausnahme bei Depression in jeder Kategorie signifikant größer als Null. Auch Selbsthilfverfahren scheinen somit ein positives Potential für die Behandlung der PDA zu haben, im Besonderen für eine Verringerung der Angstsymptomatik. Die Ergebnisse deuten darauf hin, dass diese Ansätze speziell einen Einfluss auf die kognitive Dimension der Angst haben. Allerdings sind folgende Einschränkungen zu beachten: zum einen ist die Anzahl an Studien eher gering, zum anderen lag in den meisten Kategorien Heterogenität vor, Selbsthilfverfahren sind somit nicht als gleich effektiv einzuschätzen, die Konfidenzintervalle sind sehr breit. Die Ergebnisse der

Analysen sind in Tabelle 11 dargestellt. In allen Studien lagen zur Berechnung der ESn Mittelwerte und Standardabweichungen vor.

Tabelle 11: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  zu Post für den Vergleich von Selbsthilfeverfahren zur Warteliste

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
Post	Angst	REM	0.80	0.21	0.29 – 1.30	0.16 <sup>a</sup>		8
		FEM	0.70	0.15	0.73 – 1.01		15.37 <sup>+</sup>	8
	Vermeidung	REM	0.65	0.24	0.06 – 1.25	0.21		7
		FEM	0.57	0.16	0.26 – 0.88		11.43 <sup>+</sup>	7
	Kognition	REM	1.13	0.32	0.32 – 1.94	0.38		6
		FEM	1.01	0.18	0.65 – 1.38		16.87 <sup>+</sup>	6
	Erregung	REM	0.87	0.25	0.27 – 1.47	0.31		8
		FEM	0.73	0.15	0.43 – 1.03		18.56 <sup>+</sup>	8
	Depression	REM	0.62	0.21	0.03 – 1.21	0.03		5
		FEM	0.58	0.18	0.22 – 0.96		8.07 <sup>+</sup>	5
	Klin. Sign.	REM	0.98	0.26	0.25 – 1.71	0.16		5
		FEM	0.95	0.19	0.57 – 1.32		6.84	5

<sup>a</sup>  $\sigma_{\theta}^2 > 0$  bei Ausschluss der Studie mit extremster ES definiert über die größte Abweichung vom ungewichteten Mittel

<sup>+</sup>  $p < .10$

### Effektivität: Vergleich zur Placebobedingung

Im Vergleich zu einer behandelten Kontrollbedingung (Therapie- bzw. Pillenplacebo) ergaben sich in der Mehrzahl an Studien positive ESn, 16 % der ESn waren leicht negativ bis maximal  $g = -0.24$ . Nur 15 % der Studien untersuchten rein behaviorale Verfahren, die Mehrzahl betrachtete zusätzliche kognitive Elemente. Die mittlere ES war wie zu erwarten kleiner als im Vergleich zur WL und betrug zwischen  $g = 0.26$  in der Unterkategorie Kognition und  $g = 0.58$  für Klinische Signifikanz. In den störungsspezifischen Hauptkategorien Angst und Klinische Signifikanz zeigten sich signifikant von Null verschiedene mittlere ESn, d.h. VT und KVT zeigen eine Effektivität, die über eine Realisierung gemeinsamer Faktoren hinausgeht. Dagegen zeigten die Verfahren in den Kategorien Depression und Lebenszufriedenheit keinen generellen bedeutsamen Vorteil gegenüber einer Placebothherapie. Nur im FEM, das heißt für die eingeschlossenen Studien, zeigte sich hier eine signifikante mittlere ES. Innerhalb der Unterkate-

gorien der Angst zeigten sich die größten Unterschiede für Vermeidung und Erregung. Es ist zu beachten, dass diese ermittelten ESn den wahren Effekt leicht überschätzen aufgrund der Unterschiede zu Prä. Die Ergebnisse der Analysen sind in Tabelle 12 dargestellt. Einen Überblick über die Studieneffektstärken bietet Abbildung 8.

Bei Ausschluss einer Studie, in der keine Mittelwerte und Standardabweichungen zur Berechnung der ESn vorlagen, sank die mittlere ES leicht, war aber weiterhin bedeutsam von Null verschieden (REM:  $g = 0.47$ , 95%-CI: 0.26 – 0.68,  $\sigma_{\theta}^2 = 0$ ; FEM:  $g = 0.47$ , 95%-CI: 0.29 – 0.66,  $\chi^2 = 10.01$ ,  $n = 12$ ).

Tabelle 12: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  zu Post für den Vergleich von behavioralen und kognitiv-behavioralen Verfahren zur Placebobedingung

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
Post	Angst	REM	0.51	0.10	0.30 – 0.72	0.01 <sup>a</sup>		13
		FEM	0.51	0.09	0.33 – 0.69		13.13	13
	Vermeidung	REM	0.47	0.16	0.09 – 0.84	0.08		8
		FEM	0.44	0.12	0.20 – 0.68		12.20 <sup>+</sup>	8
	Kognition	REM	0.26	0.22	-0.35 – 0.87	0.11		5
		FEM	0.25	0.16	-0.06 – 0.55		6.92	5
	Erregung	REM	0.50	0.16	0.13 – 0.86	0.09		9
		FEM	0.46	0.12	0.23 – 0.69		13.91 <sup>+</sup>	9
	Depression	REM	0.27	0.12	-0.02 – 0.56	0.02		8
		FEM	0.27	0.11	0.05 – 0.49		9.60	8
	Klin. Sign.	REM	0.58	0.15	0.25 – 0.92	0.11		10
		FEM	0.55	0.10	0.35 – 0.75		18.73 <sup>+</sup>	10
	Lebensqualität	REM	0.42	0.20	-0.11 – 0.94	0.15		6
		FEM	0.38	0.13	0.14 – 0.63		12.64 <sup>+</sup>	6

<sup>a</sup>  $\sigma_{\theta}^2 = 0$  bei Ausschluss der Studie mit extremster ES definiert über die größte Abweichung vom ungewichteten Mittel

<sup>+</sup>  $p < .10$

### Zeitliche Stabilität der Therapie

Die Messungen zu Follow up lassen sich für die Symptomkategorie Angst wie folgt charakterisieren:  $f_1 = 2.7$  Monate ( $SD = 0.78$ ),  $f_2 = 6.0$  Monate ( $SD = 0.00$ ) und  $f_3 = 14.8$  Monate ( $SD = 4.63$ ). Die Wirksamkeit von PsT erwies sich als zeitlich stabil,

es gab keine bedeutsame Abnahme der ESn. Für das Langzeit-Follow up zeigte sich eine signifikante Zunahme in der Kategorie Angst und bezogen auf die eingeschlossenen Studien in der Kategorie Klinische Signifikanz mit einer mittleren ES von  $g = 0.17$  bzw.  $g = 0.15$ . Patienten mit PsT verbesserten sich somit in der Hauptsymptomatik auch nach Abschluss der Behandlung leicht und zeigten in den anderen Kategorien keine Verschlechterung. Die Ergebnisse sind in Tabelle 13 dargestellt.

Tabelle 13: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  der Veränderung von Post zu Follow up (f1: bis sechs Monate, f2: 6 bis 12 Monate, f3: ab 12 Monate) für behaviorale und kognitiv-behaviorale Verfahren

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
f1	Angst	REM	0.06	0.08	-0.10 – 0.22	0		22
		FEM	0.06	0.08	-0.09 – 0.21		5.92	22
	Depression	REM	-0.13	0.09	-0.32 – 0.05	0		15
		FEM	-0.13	0.09	-0.30 – 0.03		4.90	15
	Klin. Sign.	REM	-0.06	0.08	-0.24 – 0.11	0		16
		FEM	-0.06	0.08	-0.22 – 0.10		9.31	16
	Lebensqualität	REM	0.05	0.11	-0.22 – 0.33	0		7
		FEM	0.05	0.11	-0.17 – 0.27		2.75	7
f2	Angst	REM	0.11	0.08	-0.05 – 0.28	0		22
		FEM	0.11	0.08	-0.04 – 0.27		15.63	22
	Depression	REM	0.01	0.08	-0.17 – 0.18	0		18
		FEM	0.01	0.08	-0.15 – 0.17		12.18	18
	Klin. Sign.	REM	-0.03	0.08	-0.31 – 0.25	0.16		17
		FEM	0.03	0.13	-0.13 – 0.19		39.71 <sup>+</sup>	17
	Lebensqualität	REM	0.11	0.10	-0.12 – 0.33	0.01		12
		FEM	0.11	0.10	-0.08 – 0.30		11.42	12
f3	Angst	REM	0.17	0.06	0.04 – 0.30	0		27
		FEM	0.17	0.06	0.05 – 0.30		18.13	27
	Depression	REM	0.03	0.08	-0.13 – 0.19	0.04		26
		FEM	0.03	0.07	-0.10 – 0.17		28.15	26
	Klin. Sign.	REM	0.15	0.07	-0.01 – 0.30	0		21
		FEM	0.15	0.07	0.01 – 0.30		19.50	21
	Lebensqualität	REM	0.24	0.17	-0.15 – 0.63	0.12		9
		FEM	0.20	0.12	-0.04 – 0.43		11.59	9

<sup>+</sup>  $p < .10$

## Akzeptanz

Bei behavioraler und kognitiv-behavioraler Therapie brachen 15.1 % der Patienten vorzeitig die Studie ab ( $SD = 12.83$ ,  $n = 81$ ). Behaviorale Verfahren zeigten dabei mit 18.3 % ( $SD = 14.53$ ,  $n = 35$ ) eine höhere Abbrecherrate im Vergleich mit kognitiv-behavioralen Verfahren mit 12.7 % ( $SD = 10.92$ ,  $n = 46$ ), der Unterschied war marginal signifikant ( $t(79) = 1.98$ ,  $p < .10$ )<sup>31</sup>. Für die WL-Bedingung betrug die Abbrecherquote 7.1 % ( $SD = 9.98$ ,  $n = 13$ ), für die Placebobedingung 17.3 % ( $SD = 12.76$ ,  $n = 10$ ).

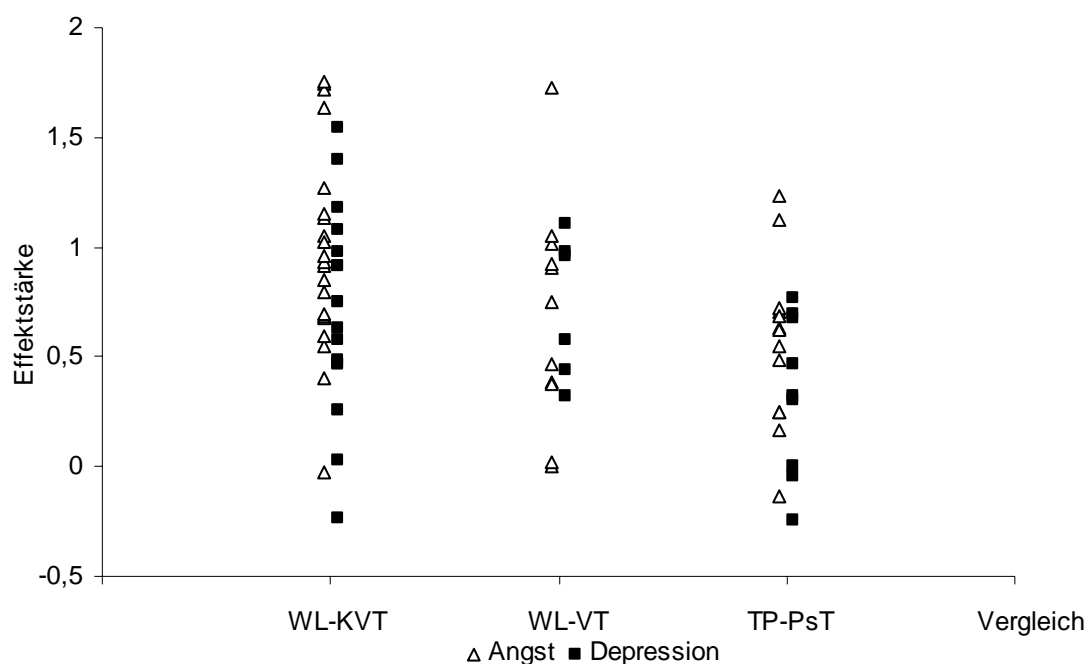


Abbildung 8: Effektstärken für die Symptomkategorien Angst und Depression für die Vergleiche psychotherapeutischer Verfahren mit verschiedenen Kontrollgruppen

## Moderatoren

Die multiple Regression ergab einen signifikanten Einfluss des Diagnosesystems nicht aber der Therapiedauer und der Stichprobengröße<sup>32</sup>. Die ES<sub>n</sub> waren umso höher, wenn die Störung nach einem neueren diagnostischen System diagnostiziert wurde, die mittlere ES<sub>n</sub> der Kategorie Angst aus Tabelle 10 stellen somit angewandt auf das

<sup>31</sup> Normalverteilung gegeben.

<sup>32</sup> Die Abbrecherrate wurde nicht als Prädiktor eingesetzt, da ansonsten ca. die Hälfte der Vergleiche ausgeschlossen würde.

aktuelle Diagnosesystem eher konservative Schätzungen dar. Die anderen Variablen zeigten dagegen keinen bedeutsamen Einfluss. Die Ergebnisse der multiplen Regression sind in Tabelle 14 dargestellt. Für Fremdbeurteilungsinstrumente ergaben sich tendenziell höhere ES<sub>n</sub> im Vergleich zur Selbstbeurteilung ( $\beta = -0.30$ ,  $p < .10$ ,  $R^2 = .24$ ;  $n = 41$ ). Nach Ausschluss der Untersuchungen, in denen nur eine Art des Ratings eingesetzt wurden, in allen Fällen Selbstbeurteilung, sank der Unterschied leicht und war nicht mehr bedeutsam von Null verschieden ( $\beta = -0.27$ , n.s.,  $R^2 = .03$ ;  $n = 26$ ).

Tabelle 14: Ergebnisse der multiplen Regression zum Einfluss verschiedener Variablen auf die Effektstärke zwischen VT/KVT und Warteliste

Variable	B
Intercept	0.06
Diagnosestellung	0.33*
Therapiedauer	$-0.53 \times 10^{-3}$
Stichprobengröße	0.01
VT vs. KVTa	0.20

<sup>a</sup>Entspannungsverfahren in VT eingeschlossen  
 $R^2 = .75$ ; \*  $p < .05$

Für den Vergleich mit einer Placebothherapie entfiel die multiple Regression, da im Rahmen der Sensitivitätsanalyse die RE-Varianz Null war. Es zeigte sich zudem kein bedeutsamer Unterschied zwischen Fremd- und Selbstbeurteilungsverfahren ( $\beta = 0.02$ , n.s.,  $R^2 = .00$ ;  $n = 21$ ). Trotz Heterogenität wurde aufgrund der geringen Studienanzahl für Selbsthilfverfahren auf eine Aufklärung der RE-Varianz sowie der Fremd- vs. Selbstbeurteilung verzichtet.

### Vergleich konfrontative und kognitiv-behaviorale Verfahren

Bei Betrachtung der ES aus dem WL-BG-Vergleich zeigten kognitiv-behaviorale Verfahren eine leicht höhere mittlere ES als rein konfrontative Verfahren in der Kategorie Angst, dieser Unterschied war allerdings nicht signifikant ( $\beta = 0.29$ , n.s.,  $R^2 = .00$ ;  $n = 29$ ). Auch in der nachfolgenden multiplen Regression mit Kontrolle von Diagnosesystem und Stichprobengröße ergab sich kein bedeutsamer Unterschied zwischen den Verfahren ( $\beta_{VT:KVT} = 0.13$ , n.s.,  $\beta_{Diagnose} = 0.19$ , n.s.,  $\beta_N = 0.01$ , n.s.,  $R^2 = .00$ ;  $n = 29$ ). Eine Betrachtung der anderen Hauptkategorien sowie ein Vergleich über die Placebothherapie entfielen aufgrund der geringen Anzahl an Studien mit konfrontativen Verfah-



---

ren für diese Berechnungen. Es zeigte sich kein signifikanter Unterschied zwischen Fremd- und Selbstbeurteilungsinstrumenten ( $\beta = -0.01$ , n.s.,  $R^2 = .00$ ;  $n = 35$ ).

Es lagen ausreichend Studien vor, die beide Verfahren direkt miteinander verglichen. In allen Kategorien ergaben einige der Studien bessere Ergebnisse für konfrontative Verfahren, andere Studien dagegen bessere Ergebnisse für kognitiv-behaviorale Verfahren. Es zeigte sich weder in den Kategorien Angst, Klinische Signifikanz noch Lebensqualität ein bedeutsamer Unterschied zwischen den Verfahren, die mittleren ESn lagen nahe Null. In der Kategorie Depression ergab sich dagegen eine bedeutsam positive ES mit  $g = 0.18$ , kognitiv-behaviorale Verfahren zeigten somit für depressive Symptomatik eine leicht erhöhte Effektivität im Vergleich zu rein konfrontativen Verfahren. Die Ergebnisse der Analysen sind in Tabelle 15 dargestellt. Eine Übersicht über die Studieneffektstärken bietet Abbildung 9. In der Symptomkategorie Angst war die RE-Varianz Null, eine bedeutsame Variabilität zwischen den Studien konnte nicht gefunden werden. Auf die Durchführung einer Regressionsanalyse wurde somit verzichtet.

Der Ausschluss von Studien, in denen die ESn nicht über Mittelwerte und Standardabweichungen erfolgte, führte zu einer Erhöhung der mittleren ES (REM:  $g = 0.18$ , 95%-CI:  $-0.01 - 0.37$ ,  $\sigma_{\theta}^2 = 0$ ; FEM:  $g = 0.18$ , 95%-CI:  $0.01 - 0.35$ ,  $\chi^2 = 13.08$ ,  $n = 17$ ). Im FEM, d.h. bzgl. der eingeschlossenen Studien, ergab sich eine signifikant höhere Effektivität für KVT im Vergleich zu konfrontativen Verfahren. Die Ergebnisse sind somit abhängig von den Entscheidungen des Meta-Analysten hinsichtlich der eingesetzten Selektionskriterien und müssen als nicht stabil gelten.

Die Ergebnisse blieben bei Betrachtung von Follow-up relativ stabil, f1 bezog sich dabei auf einen Zeitraum von 2.3 Monaten ( $SD = 1.20$ ), f2 auf 6 Monate ( $SD = 0.00$ ) und f3 auf 13.3 Monate ( $SD = 2.65$ ). In den Kategorien Angst und Klinische Signifikanz zeigte sich für die einzelnen Zeiträume kein bedeutsamer Unterschied zwischen den Verfahren. Für Depression war ab dem 6-Monats-Follow-up der Unterschied nicht mehr signifikant. Für f2 und f3 lagen divergente Ergebnisse für REM und FEM vor, im FEM ergab sich eine bedeutend höhere ES für KVT bei Klinischer Signifikanz. Allerdings lag Heterogenität vor. Die Ergebnisse der Analysen sind ebenfalls in Tabelle 15 dargestellt.

Tabelle 15: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  zu Post und Follow up (f1: bis sechs Monate, f2: 6 bis 12 Monate, f3: ab 12 Monate) für den Vergleich von konfrontativen und kognitiv-behavioralen Verfahren

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
Post	Angst	REM	0.09	0.07	-0.07 – 0.24	0		26
		FEM	0.09	0.07	-0.06 – 0.23		19.61	26
	Vermeidung	REM	0.10	0.08	-0.06 – 0.27	0		23
		FEM	0.10	0.08	-0.05 – 0.26		20.50	23
	Kognition	REM	0.01	0.08	-0.16 – 0.19	0		20
		FEM	0.01	0.08	-0.15 – 0.18		13.84	20
	Erregung	REM	0.04	0.09	-0.14 – 0.21	0.04		25
		FEM	0.05	0.07	-0.10 – 0.19		34.48 <sup>+</sup>	25
	Depression	REM	0.18	0.08	0.01 – 0.35	0		22
		FEM	0.18	0.08	0.02 – 0.34		17.29	22
	Klin. Sign.	REM	0.13	0.12	-0.13 – 0.39	0.09		15
		FEM	0.15	0.09	-0.02 – 0.33		26.71 <sup>+</sup>	15
	Lebensqualität	REM	-0.11	0.16	-0.63 – 0.42	0.01		4
		FEM	-0.10	0.15	-0.40 – 0.20		3.05	4
f1	Angst	REM	0.15	0.10	-0.08 – 0.38	0		12
		FEM	0.15	0.10	-0.05 – 0.35		10.71	12
	Depression	REM	0.31	0.12	0.05 – 0.57	0		10
		FEM	0.31	0.12	0.08 – 0.53		7.85	10
	Klin. Sign.	REM	0.23	0.18	-0.16 – 0.62	0.22		11
		FEM	0.24	0.10	0.04 – 0.44		29.68 <sup>+</sup>	11
f2	Angst	REM	-0.03	0.13	-0.34 – 0.28	0		7
		FEM	-0.03	0.13	-0.28 – 0.22		2.14	7
	Depression	REM	0.15	0.15	-0.28 – 0.57	0		5
		FEM	0.15	0.15	-0.15 – 0.45		3.30	5
	Klin. Sign.	REM	0.23	0.26	-0.50 – 0.97	0.23		5
		FEM	0.34	0.14	0.06 – 0.62		10.70 <sup>+</sup>	5
f3	Angst	REM	0.11	0.11	-0.15 – 0.37	0		9
		FEM	0.11	0.11	-0.11 – 0.33		6.27	9
	Depression	REM	0.32	0.17	-0.07 – 0.72	0.15		9
		FEM	0.36	0.11	0.13 – 0.58		18.23 <sup>+</sup>	9
	Klin. Sign.	REM	0.43	0.21	-0.08 – 0.94	0.20		7
		FEM	0.48	0.12	0.23 – 0.72		15.43 <sup>+</sup>	7

<sup>+</sup>  $p < .10$

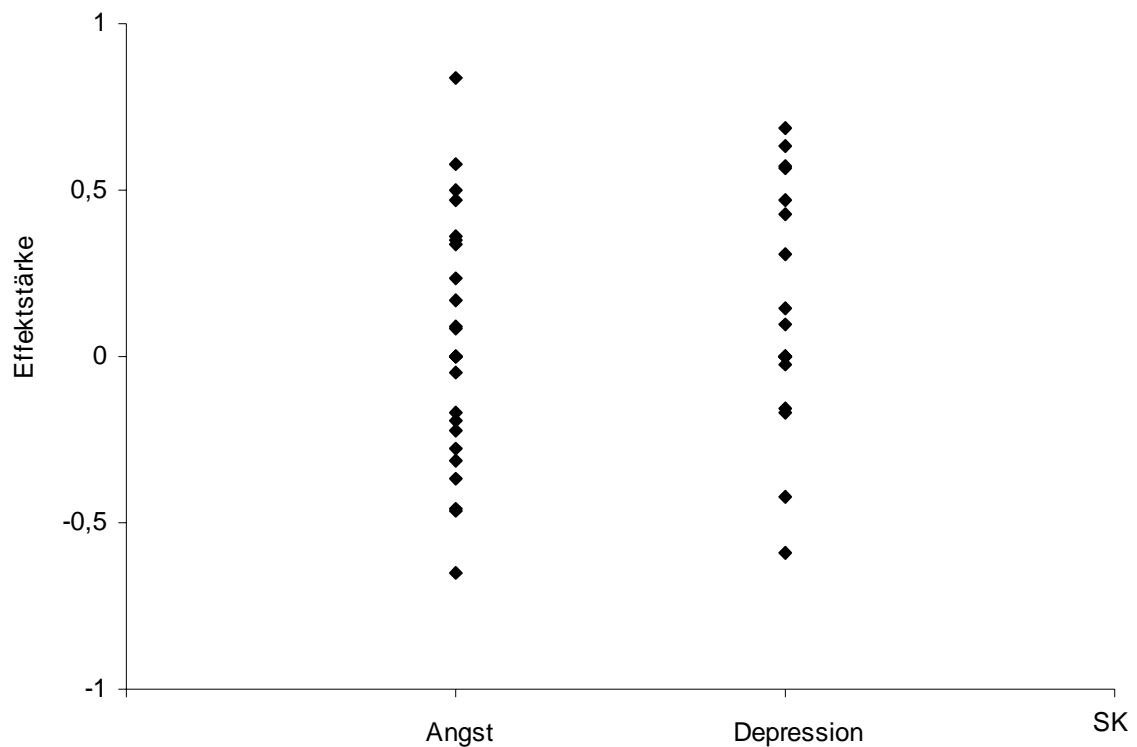


Abbildung 9: Effektstärken für die Symptomkategorien (SK) Angst und Depression der einzelnen Studien für den direkten Vergleich von konfrontativen und kognitiv-behavioralen Verfahren

### Fail-Safe N

Für eine Reduzierung der beobachteten Effekte des Vergleiches PsT zu WL auf eine kleine ES von  $g = 0.2$  sind 107 Gruppenvergleiche mit  $g = 0$  notwendig, ca. zwei Drittel der von Rosenthal vorgeschlagenen Grenze für stabile Ergebnisse. Das Fail-Safe N ist zumindest über dreimal so hoch, wie die Anzahl an eingeschlossenen Vergleichen. Für einen Vergleich mit einer behandelten Kontrollgruppe werden 20 Gruppenvergleiche benötigt, sowie 33 Gruppenvergleiche mit einer Effektstärke von  $g = -0.2$  für einen mittleren Effekt von  $g = 0$ . Die geringe Stabilität der Ergebnisse wird deutlich: die Grenze nach Rosenthal beträgt 75 Studien, der Faktor hinsichtlich der eingeschlossenen Studien beträgt 1.6 bzw. 2.6. Für den direkten Vergleich von konfrontativen mit kognitiv-behavioralen Therapien wären 24 bzw. 9 Studien notwendig mit einer jeweils mittleren Effektstärke für eines der Verfahren, um einen geringen Effekt für konfrontative oder für kognitiv-behaviorale Behandlungen zu erhalten. Die Grenze liegt hier bei 135, der Faktor bezüglich der eingeschlossenen Vergleiche beträgt 1.0 bis 0.4, auch hier können die Ergebnisse noch nicht als stabil gelten.

### 3.4.4 Pharmakotherapie

#### **Effektivität: Vergleich zur behandelten Kontrollgruppe**

Die am häufigsten untersuchten Substanzklassen waren mit jeweils ca. einem Drittel der Studien BZD und TCA und mit ca. einem Fünftel SSRI. Die Mehrzahl der ES<sub>n</sub> war positiv, nur 6 % der Studieneffektstärken in den Hauptkategorien ergaben dagegen ein besseres Ergebnis für die PP-Bedingung mit ES<sub>n</sub> bis zu  $g = -0.54$ . In allen Symptomkategorien zeigte sich zu Post eine signifikant positive mittlere ES zwischen  $g = 0.30$  (Vermeidung/Kognition) und  $g = 0.53$  (Klinische Signifikanz). Meist lagen die mittleren ES<sub>n</sub> im kleinen Bereich. Pharmakotherapie scheint dabei auf alle Bereiche der Angstsymptomatik eine ähnliche Effektivität zu haben. In jeder der Symptomkategorien lag Heterogenität vor, es gibt eine bedeutsame Variabilität zwischen den Untersuchungen. Die Ergebnisse sind in Tabelle 16 dargestellt. Einen Überblick über die Studieneffektstärken bietet Abbildung 10.

Wurden nur diejenigen Studien betrachtet, in denen die Berechnung der ES<sub>n</sub> über Mittelwerte und Standardabweichungen gegeben war, erhöhte sich die mittlere ES leicht (REM:  $g = 0.51$ , 95%-CI: 0.40 – 0.61,  $\sigma_{\theta}^2 = 0.05$ ; FEM:  $g = 0.47$ , 95%-CI: 0.40 – 0.54,  $\chi^2 = 73.19$ ,  $n = 52$ ).

Tabelle 16: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  zu Post für den Vergleich einer medikamentösen zur Placebo-bedingung

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
Post	Angst	REM	0.38	0.03	0.31 – 0.45	0.03 <sup>a</sup>		78
		FEM	0.32	0.02	0.27 – 0.36		104.89 <sup>+</sup>	78
	Vermeidung	REM	0.30	0.05	0.20 – 0.40	0.05		45
		FEM	0.23	0.03	0.18 – 0.29		79.23 <sup>+</sup>	45
	Kognition	REM	0.30	0.06	0.17 – 0.42	0.05		26
		FEM	0.19	0.03	0.13 – 0.25		54.98 <sup>+</sup>	26
	Erregung	REM	0.38	0.04	0.30 – 0.46	0.05		67
		FEM	0.31	0.02	0.26 – 0.36		105.10 <sup>+</sup>	67
	Depression	REM	0.34	0.06	0.21 – 0.47	0.09		42
		FEM	0.29	0.03	0.23 – 0.35		63.06 <sup>+</sup>	42
	Klin. Sign.	REM	0.51	0.05	0.40 – 0.61	0.09		57
		FEM	0.49	0.02	0.44 – 0.53		101.89 <sup>+</sup>	57
	Lebensqualität	REM	0.35	0.06	0.22 – 0.48	0.08		36
		FEM	0.33	0.03	0.26 – 0.40		66.06 <sup>+</sup>	36

<sup>a</sup>  $\sigma_{\theta}^2 > 0$  bei Ausschluss der Studie mit extremster ES definiert über die größte Abweichung vom ungewichteten Mittel

<sup>+</sup>  $p < .10$

Für die einzelnen Substanzklassen ergab sich für die Symptomkategorie Angst eine mittlere ES von  $g = 0.40$  für BZD,  $g = 0.39$  für TCA und  $g = 0.40$  für SSRI. Die multiple Regression mit Kontrolle der Variablen Diagnosesystem, Dosierung, PP Runin, Dropouts, Analyseart und Stichprobengröße ergab keinen signifikanten Unterschied für BZD vs. TCA ( $\beta = -0.01$ , n.s.,  $R^2 = .00$ ,  $n = 48$ ), für BZD vs. SSRI ( $\beta = -0.002$ , n.s.,  $R^2 = .23$ ,  $n = 43$ ) und für TCA vs. SSRI ( $\beta = 0.02$ , n.s.,  $R^2 = .44$ ,  $n = 38$ ). Die am häufigsten untersuchten Substanzklassen unterschieden sich in ihrer Effektivität für die Angstsymptomatik nicht bedeutsam voneinander. Anders war dies in der Kategorie Depression. Hier ergab sich eine mittlere ES von  $g = 0.30$  für BZD,  $g = 0.43$  für TCA und  $g = 0.49$  für SSRI. Die multiple Regression ergab keinen bedeutsamen Unterschied zwischen BZD und SSRI sowie zwischen TCA und SSRI (BZD vs. SSRI:  $\beta = 0.03$ , n.s.,  $R^2 = .15$ ,  $n = 24$ ; TCA vs. SSRI:  $\beta = 0.04$ , n.s.,  $R^2 = 1.00$ ,  $n = 22$ ). Dagegen zeigten BZD tendenziell signifikant geringere ESn im Vergleich zu TCA

( $\beta = 0.42$ ,  $p < .10$ ,  $R^2 = .29$ ,  $n = 22$ ), nach Alphafehlerkorrektur war der Unterschied nicht bedeutsam.

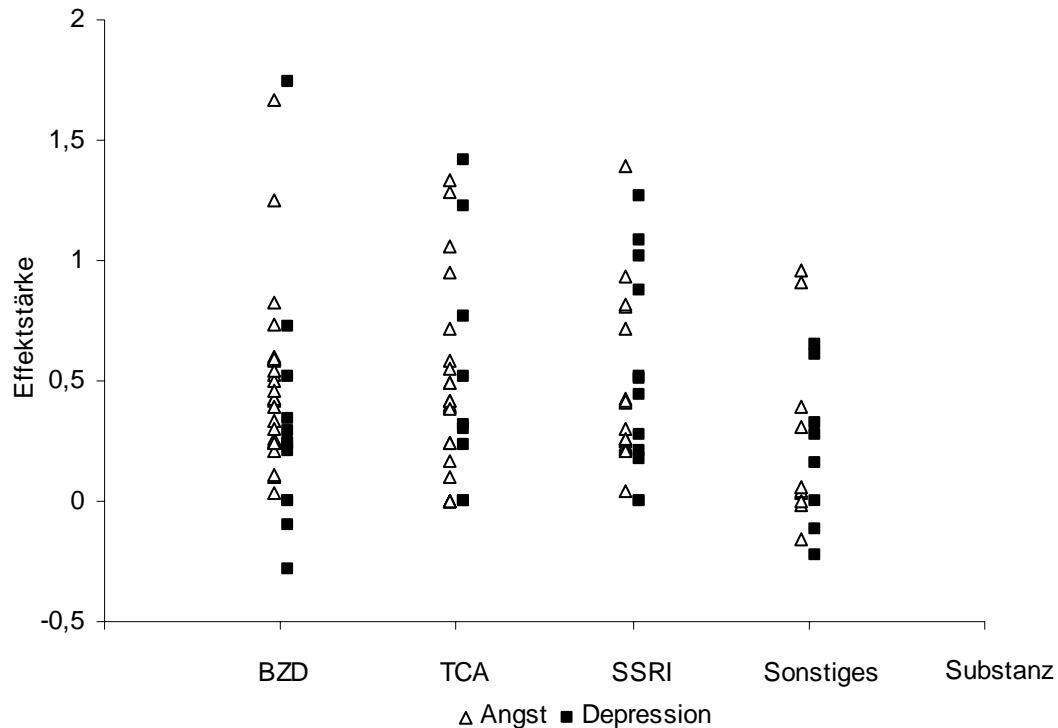


Abbildung 10: Effektstärken der einzelnen Studien für die Symptomkategorien Angst und Depression aufgeteilt nach Substanzklasse

### Zeitliche Stabilität der Therapie

Nur in wenigen Studien war in der Kategorie Angst die Berechnung einer Effektstärke für das Follow up möglich. Es zeigte sich in zwei Studien eine Verschlechterung der Symptomatik bis  $g = -0.39$  zu 12 Monaten, in einer Studie eine Verbesserung mit  $g = 0.29$  zu 6 Monaten.

### Akzeptanz

Die mittlere Dropout-Rate in den BGn betrug 20.4 % ( $SD: 15.27$ ,  $n = 75$ ), d.h. jeder Fünfte brach vorzeitig die Behandlung ab. Für Benzodiazepine betrug die mittlere Abbrecherquote 17.7 % ( $SD = 13.63$ ,  $n = 27$ ), für TCA 23.5 % ( $SD = 17.04$ ,  $n = 21$ ) und für SSRI 23.1 % ( $SD = 15.26$ ,  $n = 18$ ). Diese Werte unterschieden sich nicht signifikant voneinander (BZD vs. TCA:  $t(46) = -1.31$ , n.s.; BZD vs. SSRI:  $t(43) = -1.25$ , n.s.; TCA

vs. SSRI:  $t(37) = 0.07$ , n.s.)<sup>33</sup>. Bei PP verließ ca. jeder Dritte vorzeitig die Studie ( $M = 30.1$ ,  $SD = 19.16$ ,  $n = 56$ ).

### Moderatoren

Die multiple Regression ergab, dass ES signifikant größer waren, wenn die Stichprobengröße kleiner war. Weder das diagnostische System, die Art der Dosierung, die Abbrecherrate, die Analyseart noch ob ein PP Runin durchgeführt wurde zeigte einen bedeutsamen Zusammenhang zur ES. Die Ergebnisse sind in Tabelle 17 dargestellt. Es ergab sich kein bedeutsamer Einfluss der Art der Erhebung, Fremdbeurteilung zeigte vergleichbare ESn wie Selbstbeurteilung ( $\beta = -0.05$ , n.s.,  $R^2 = .00$ ;  $n = 151$ ).

Tabelle 17: Ergebnisse der multiplen Regression zum Einfluss verschiedener Variablen auf die Effektstärke zwischen Pharmakotherapie und Pillenplacebo

Variable	B
Intercept	0.42*
Diagnosestellung	0.02
Dosierung	-0.01
PP-Washout	-0.08
Dropouts	-0.01
Analyse	-0.11
Stichprobengröße	-0.45* x 10 <sup>-3</sup>
Benzodiazepine	0.24
TCA	0.23
SSRI	0.26

\*  $p < .05$ ,  $R^2 = .21$ ,  $n = 73$

### Fail-Safe N

Für eine Reduktion der mittleren Effektstärke auf  $g = 0.20$  sind 70 Studien notwendig mit einem Nulleffekt. Die Grenze nach Rosenthal beträgt jeweils 400 Studien, eine sehr hohe Zahl, die bei weitem unterschritten wird.

<sup>33</sup> Normalverteilung gegeben

### 3.4.5 Vergleich Psycho- und Pharmakotherapie

#### Indirekter Vergleich

Es zeigte sich in den Prä-Post-ESn für die Kategorie Angst eine Differenz von  $g = 0.30$  zwischen PP und WL mit höheren ES für PP und ein Unterschied von  $g = 0.20$  zwischen PP und TP, wobei TP höhere ESn aufwies als PP. Es zeigte sich nach Korrektur der Post-ESn um diese Werte eine signifikant leicht höhere Wirksamkeit von PsT im Vergleich zu PhT ( $\beta = 0.20, p < .05, R^2 = .14, n = 123$ ). Auch mit Kontrolle von Diagnosesystem und Stichprobengröße<sup>34</sup> in der multiplen Regression war der Unterschied signifikant; zudem ergab die Regression, dass Studien mit mehr Patienten signifikant geringere ESn zeigten ( $\beta_{PhT:PsT} = 0.16, p < .05, \beta_{Diagnose} = 0.01, n.s., \beta_N = -0.45 \times 10^{-3}, p < .01, R^2 = .29; n = 120$ ). Schloss man Studien mit weniger etablierten Behandlungen aus (definiert als bei PsT nicht zu konfrontativen bzw. kognitiv-behavioralen Verfahren und bei PhT nicht zu den am häufigsten untersuchten Substanzklassen TCA, SSRI oder BZD gehörend), wurde die Differenz zwischen beiden Behandlungsformen kleiner, der Unterschied war tendenziell signifikant; es zeigte sich weiterhin ein signifikanter Einfluss der Stichprobengröße mit geringeren ESn bei höherer Patientenzahl ( $\beta_{PhT:PsT} = 0.15, p < .10, \beta_{Diagnose} = 0.02, n.s., \beta_N = -0.49 \times 10^{-3}, p < .01, R^2 = .38; n = 110$ ). Für die Kategorie Depression ergab sich in den Prä-Post-ESn eine Differenz von  $g = 0.35$  im Vergleich zur WL mit höheren ESn für PP und von  $g = 0.26$  im Vergleich zu TP mit geringeren ESn für PP. Nach Korrektur zeigte sich kein bedeutsamer Unterschied zwischen PsT und PhT ( $\beta = 0.05, n.s., R^2 = .00, n = 74$ ). Dieses Ergebnis blieb stabil mit Kontrolle von Diagnosesystem und Stichprobengröße als auch nach Ausschluss weniger etablierter Behandlungen ( $\beta_{PhT:PsT} = 0.03, n.s., \beta_{Diagnose} = -0.01, n.s., \beta_N = -0.24 \times 10^{-3}, n.s., R^2 = .00; n = 73; \beta_{PhT:PsT} = 0.01, n.s., \beta_{Diagnose} = 0.03, n.s., \beta_N = -0.26 \times 10^{-3}, n.s., R^2 = .00; n = 64$ ). Insgesamt scheinen somit behaviorale und kognitiv-behaviorale Therapien mindestens so effektiv zu sein wie medikamentöse Therapien.

---

<sup>34</sup> Analyseart und Abbrecherrate wurden aufgrund der Verteilungseigenschaften bzw. hohen Ausschlussrate für PsT-Studien nicht einbezogen.



### Direkter Vergleich

Es lagen 11 Studien vor, die PhT und PsT direkt verglichen. Ein Drittel untersuchte rein behaviorale Verfahren. Die meist untersuchten Substanzklassen waren TCA und SSRI. Die ES<sub>n</sub> der Studien schwankten stark, so lagen in der Kategorie Angst sowohl Studien vor, die ein besseres Ergebnis der PhT zeigten (bis  $g = -0.56$ ), als auch Studien, in denen PsT wirksamer war (bis  $g = 1.50$ ). Die mittleren ES<sub>n</sub> waren leicht positiv, aber nicht signifikant von Null verschieden, in keiner der Kategorien lag somit ein bedeutsamer Unterschied zwischen den beiden Verfahren vor. Hinsichtlich der Angstdimensionen zeigten sich der geringste Unterschied in der Subkategorie Erregung. Es lag in allen Hauptkategorien Heterogenität vor, d.h. es gibt eine bedeutsame Variabilität zwischen den Studien. Die Ergebnisse der Analysen sind in Tabelle 18 dargestellt, einen Überblick über die einzelnen Effektstärken bietet Abbildung 11.

In allen Studien erfolgte die Berechnung der Effektstärken über Mittelwerte und Standardabweichungen.

Es lagen Daten für spätere Messzeitpunkte vor, das durchschnittliche Follow up betrug 10.8 Monate ( $SD = 2.4$ ). Die Ergebnisse bleiben stabil, es zeigt sich auch zu Follow up kein bedeutsamer Unterschied zwischen den Verfahren. Die Ergebnisse sind in Tabelle 18 dargestellt. Bei Anwendung der strengeren Kriterien für Pharmakotherapie, d.h. Beendigung nach Studienende und Betrachtung der kompletten Stichprobe statt nur der Responder, blieben die Ergebnisse stabil. Die mittlere ES betrug dann in der Kategorie Angst  $g = 0.29$  (95%-CI:  $-0.29 - 0.87$ ,  $\sigma_{\theta}^2 = 0$ ; FEM:  $g = 0.29$ , 95%-CI:  $-0.07 - 0.64$ ,  $\chi^2 = 2.84$ ,  $n = 4$ ). Es ist allerdings die nur geringe Gesamtanzahl an Studien zu beachten.

Tabelle 18: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  zu Post und Follow up (f) für den Vergleich von Pharmako- zu Psychotherapie

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
Post	Angst	REM	0.27	0.16	-0.07 – 0.62	0.21 <sup>a</sup>		13
		FEM	0.13	0.08	-0.03 – 0.30		37.06 <sup>+</sup>	13
	Vermeidung	REM	0.24	0.14	-0.09 – 0.56	0.06		9
		FEM	0.21	0.11	-0.01 – 0.42		13.50 <sup>+</sup>	9
	Kognition	REM	0.33	0.39	-0.75 – 1.40	0.63		5
		FEM	0.17	0.15	-0.12 – 0.45		27.96 <sup>+</sup>	5
	Erregung	REM	0.03	0.16	-0.36 – 0.42	0.07		7
		FEM	0.02	0.12	-0.22 – 0.25		13.50 <sup>+</sup>	7
	Depression	REM	0.21	0.24	-0.34 – 0.75	0.45		10
		FEM	0.10	0.10	-0.11 – 0.30		49.49 <sup>+</sup>	10
	Klin. Sign.	REM	0.09	0.20	-0.38 – 0.56	0.22		8
		FEM	0.08	0.09	-0.11 – 0.26		19.62 <sup>+</sup>	8
f	Angst	REM	0.20	0.12	-0.13 – 0.54	0		5
		FEM	0.20	0.12	-0.04 – 0.44		3.23	5
	Klin. Sign.	REM	0.21	0.25	-0.44 – 0.85	0.27		6
		FEM	0.26	0.11	0.04 – 0.48		14.35 <sup>+</sup>	6

<sup>a</sup>  $\sigma_{\theta}^2 > 0$  bei Ausschluss der Studie mit extremster ES definiert über die größte Abweichung vom ungewichteten Mittel

<sup>+</sup>  $p < .10$

Aufgrund der geringen Anzahl an Studien mit gleichzeitig hoher Anzahl an Prädiktoren wurden zur Stabilisierung der Ergebnisse einfache Regressionen berechnet. Keiner der Prädiktoren zeigte einen bedeutsamen Zusammenhang zur ES. Die Ergebnisse sind in Tabelle 19 dargestellt. Des weiteren zeigte sich kein signifikanter Unterschied zwischen Fremd- und Selbstrating ( $\beta = 0.13$ , *n.s.*,  $R^2 = .00$ ;  $n = 21$ ).

Tabelle 19: Ergebnisse der einfachen Regressionen zum Einfluss verschiedener Variablen auf die Effektstärke zwischen Pharmako- und Psychotherapie

Variable	B	R <sup>2</sup>
VT vs. KVT <sup>a</sup>	0.01	.00
Stichprobengröße <sup>a</sup>	-0.01	.24
Completer vs. ITT <sup>a</sup>	0.23x10 <sup>-3</sup>	.00
TCA vs. SSRI <sup>b</sup>	0.02	.00

<sup>a</sup>  $n = 13$ , <sup>b</sup>  $n = 10$

**Fail-Safe N**

Für eine Reduktion der mittleren Effektstärke auf  $g = 0$  sind 18 Studien notwendig mit einem kleinen Effekt für Pharmakotherapie von  $g = -0.20$  und 7 Studien mit einem Effekt von  $g = -0.50$ . 10 Studien mit einer hohen Effektstärke für Psychotherapie von  $g = 0.80$  sind notwendig für einen mittleren Effekt von  $g = 0.50$ . Schon eine geringe Anzahl weiterer Untersuchungen könnte zu substantiell anderen Ergebnissen führen.

**Akzeptanz**

Über alle Studien ergab sich eine bedeutsam geringere Akzeptanz für medikamentöse Therapie im Vergleich zur psychologischen Behandlung ( $t(154) = 2.37, p < .05$ ).

**3.4.6 Vergleich Psychotherapie mit Kombination aus Psycho- und Pharmakotherapie****Effektivität**

20 Studien untersuchten die Effektivität von PsT im Vergleich zu einer Kombinationsbehandlung meist mit TCA oder SSRI. Jeweils die Hälfte der Studien untersuchte rein konfrontative oder kognitiv-behaviorale Verfahren. Für zwei der Untersuchungen lagen nur Daten für Follow up vor.

16 % der ESn zu Post in den Hauptkategorien waren negativ, d.h. zeigten ein besseres Ergebnis der PsT. Es ergab sich in den mittleren ESn ein signifikant leicht höhere ES für die Kombinationsbehandlung in den Kategorien Angst, Depression und Klinische Signifikanz im Bereich von  $g = 0.23$  (Angst) bis  $g = 0.40$  (Klinische Signifikanz). In keiner der Unterkategorien der Angst zeigte sich eine signifikante mittlere ES, möglicherweise auf die im Vergleich zur Hauptkategorie geringere Studienanzahl zurückzuführen. Zudem ist zu beachten, dass Messinstrumente zur Erfassung der allgemeinen Ängstlichkeit bzw. PDA-Symptomatik insgesamt in keiner der Unterkategorien einbezogen wurden, auch dies könnte vielleicht diese Ergebnisse teilweise erklären, wenn man von einer unterschiedlichen Änderungssensitivität ausgeht. Die Ergebnisse sind in Tabelle 20 dargestellt, die einzelnen Effektstärken finden sich in Abbildung 11.

Bei Ausschluss der Studien, in denen zur Berechnung der ESn keine Mittelwerte und Standardabweichungen vorlagen, stieg die mittlere ES leicht und betrug in der Kategorie Angst  $g = 0.29$  (95%-CI: 0.12 – 0.46,  $\sigma_{\theta}^2 = 0$ ; FEM:  $g = 0.29$ , 95%-CI: 0.14 – 0.44,  $\chi^2 = 7.73$ ,  $n = 15$ ).

Aufgrund der Anzahl an Studien wurden die Zeiträume des Follow up zusammengefasst und der jeweils letzte Messzeitpunkt einer Studie in die Analyse eingeschlossen. Im Mittel betrug das Follow up 16.8 Monate ( $SD = 11.06$ ). Ca. zwei Drittel der ESn der Studien in den Hauptkategorien waren negativ, d.h. sie zeigten eine höhere Wirksamkeit für die reine Psychotherapiebedingung. In keiner der Kategorien zeigte sich eine bedeutsam von Null verschiedene mittlere ES, d.h. beide Verfahren sind langfristig gesehen gleich effektiv. Die Ergebnisse der Analysen sind in Tabelle 20 dargestellt. Bei Anwendung der strengeren Kriterien für Pharmakotherapie, d.h. Beendigung nach Studienende und Betrachtung der kompletten Stichprobe statt nur der Responder, blieben die Ergebnisse stabil. Die mittlere ES betrug dann in der Kategorie Angst  $g = -0.08$  (95%-CI: -0.48 – 0.32,  $\sigma_{\theta}^2 = 0$ ; FEM:  $g = -0.08$ , 95%-CI: 0.37 – 0.21,  $\chi^2 = 1.32$ ,  $n = 5$ ).

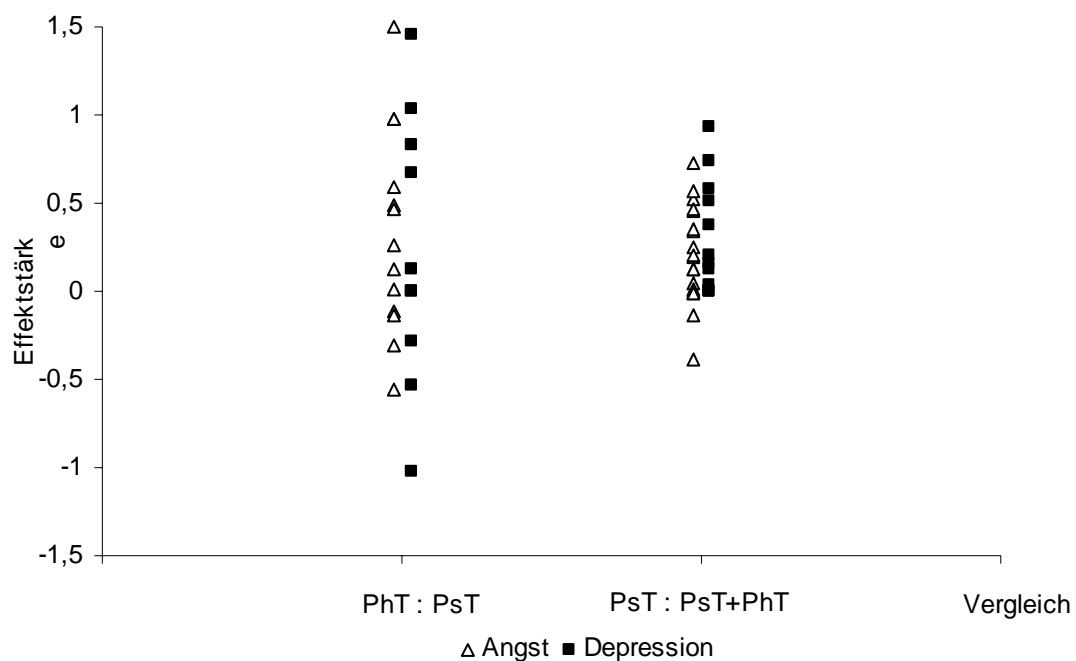


Abbildung 11: Effektstärken für die Symptomkategorien Angst und Depression für den Vergleich von Pharmako- mit Psychotherapie bzw. Psychotherapie mit einem Kombinationsverfahren

Tabelle 20: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  zu Post und Follow up (f) für den Vergleich von Psychotherapie zu einem Kombinationsverfahren

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
Post	Angst	REM	0.23	0.07	0.09 – 0.37	0		19
		FEM	0.23	0.07	0.10 – 0.36		11.48	19
	Vermeidung	REM	0.20	0.09	-0.01 – 0.41	0		12
		FEM	0.20	0.09	0.02 – 0.38		8.78	12
	Kognition	REM	0.09	0.15	-0.40 – 0.57	0		4
		FEM	0.09	0.15	-0.21 – 0.38		1.54	4
	Erregung	REM	0.14	0.14	-0.17 – 0.45	0.11		11
		FEM	0.16	0.09	-0.02 – 0.33		15.79	11
	Depression	REM	0.29	0.09	0.09 – 0.49	0		12
		FEM	0.29	0.09	0.11 – 0.46		10.51	12
	Klin. Sign.	REM	0.40	0.08	0.23 – 0.56	0.01		13
		FEM	0.40	0.08	0.25 – 0.55		12.81	13
	Lebensqualität	REM	0.25	0.15	-0.18 – 0.68	0.03		5
		FEM	0.26	0.13	-0.01 – 0.53		4.76	5
F	Angst	REM	-0.16	0.09	-0.39 – 0.06	0		8
		FEM	-0.16	0.09	-0.35 – 0.02		4.68	8
	Depression	REM	-0.07	0.20	-0.59 – 0.46	0.15		6
		FEM	-0.11	0.12	-0.34 – 0.11		21.06 <sup>+</sup>	6
	Klin. Sign.	REM	0.03	0.24	-0.53 – 0.59	0.30		8
		FEM	-0.06	0.11	-0.28 – 0.15		14.08 <sup>+</sup>	8

<sup>+</sup>  $p < .10$

### Akzeptanz

In der Kombinationsbedingung brach ca. jeder Fünfte vorzeitig die Behandlung ab ( $M = 20.5$ ,  $SD = 11.18$ ,  $n = 15$ ). Es zeigte sich zu den PsT-Bedingungen dieser Studien mit einer Abbrecherrate von 18.9 % ( $SD = 8.66$ ,  $n = 15$ ) kein bedeutsamer Unterschied ( $t(28) = -0.42$ , n.s.)<sup>35</sup>.

<sup>35</sup> Normalverteilung gegeben.

**Moderatoren**

Aufgrund der RE-Varianz von Null wurde keine Regressionsanalyse durchgeführt. Es zeigte sich zudem kein signifikanter Unterschied zwischen Fremd- und Selbstbeurteilungsverfahren ( $\beta = -0.11$ , *n.s.*,  $R^2 = .00$ ,  $n = 33$ ).

**Fail-Safe N**

Für eine Reduktion der mittleren Effektstärke auf  $g = 0$  sind 22 Studien notwendig mit einem kleinen Effekt für Psychotherapie von  $g = -0.20$  und 9 Studien mit einem Effekt von  $g = -0.50$ . 17 Studien mit einer hohen Effektstärke für das Kombinationsverfahren von  $g = 0.80$  sind dagegen notwendig für einen mittleren Effekt von  $g = 0.50$ . Die Anzahl an Untersuchungen in Schubladen entspricht somit jeweils circa der Anzahl eingeschlossener Untersuchungen.

### 3.4.7 Diskussion

Die Daten der vorliegenden Meta-Analyse zeigen, dass die untersuchten Psychotherapien, insbesondere kognitiv-behaviorale Verfahren und Exposition, sehr effektiv in der Behandlung der PDA sind. Es ergaben sich im Vergleich zu einer unbehandelten Gruppe bedeutsame hohe Effektstärken über alle Symptomdimensionen. Die Wirksamkeit der betrachteten Psychotherapie geht dabei für die Angstsymptomatik über eine Realisierung gemeinsamer Faktoren hinaus. Dieses Ergebnis zeigte sich in den vorliegenden Studien auch für Depression und Lebensqualität, es kann aber nicht generalisiert werden. Die Befunde deuten darauf hin, dass die Vorteile der meist kognitiv-behavioralen Behandlung zu einem Pillen- oder Therapieplacebo eher in den Subkategorien Vermeidung und Erregung und weniger für Kognitionen ergeben waren. Dies ist allerdings aufgrund der geringen Anzahl an Studien mit Vorsicht zu interpretieren. Somit scheinen auch ohne spezielle kognitive Techniken Veränderungen in Kognitionen gegeben. Mehr als die Hälfte der Effektivität behavioraler und kognitiv-behavioraler Therapien scheint auf die Wirkung spezifischer Faktoren zurückführbar. Insgesamt ist die Akzeptanz als gut einzuschätzen: nur jeder Siebente brach vorzeitig die Therapie ab.

Die gefundenen positiven Ergebnisse blieben auch im Follow up stabil. Für die Langzeit, d.h. ab 12 Monate, zeigten Patienten sogar eine weitere Verbesserung in ihrer Symptomatik. Eine Übertragbarkeit der gelernten Techniken in die Alltagssituation scheint somit gegeben.

Sind kognitiv-behaviorale Verfahren effektiver als rein konfrontative? Im indirekten Vergleich ergab sich kein bedeutsamer Unterschied zwischen den Behandlungen für die Angstsymptomatik. Dieses Ergebnis wurde stabilisiert anhand von Studien, die beide Therapien direkt verglichen. Konfrontative Verfahren waren in einer Reduktion der Angstsymptomatik ebenso effektiv wie kognitiv-behaviorale Verfahren. Beide Therapien scheinen dabei auf alle Dimensionen der Angst eine vergleichbare Wirksamkeit zu haben. Auch hier zeigten sich keine besonderen Vorteile der speziellen kognitiven Techniken für eine Modifikation angstrelevanter Kognitionen. Dagegen ergab sich in der Behandlung der depressiven Symptomatik von Patienten mit PDA ein bedeutsamer Vorteil für kognitiv-behaviorale Therapien. Die erhöhte Effektivität kann dabei auch generalisiert werden. Langfristig scheint, zumindest für die eingeschlossenen

Studien, dieser Unterschied stabil. Zudem zeigten Patienten für kognitiv-behaviorale Verfahren eine leicht höhere Akzeptanz als für rein behaviorale Verfahren. Kognitiv-behaviorale Therapien erscheinen somit für die Behandlung von Patienten mit PDA empfehlenswert.

Auch Selbsthilfeverfahren scheinen wirksam in der Behandlung. Allerdings ergab sich eine hohe Variabilität in der Effektivität, das teilweise gleiche Material zeigte in verschiedenen Studien sehr unterschiedliche Effektstärken. Weitere Untersuchungen sollten somit unterschiedliche Realisierungen und z.B. den Zusammenhang zum Aufbau einer hohen Compliance überprüfen.

Auch Pharmakotherapie, insbesondere BZD, TCA und SSRI, ist im Vergleich zu einem Pillenplacebo über alle Symptomdimensionen effektiv. Die Wirksamkeit lag dabei im kleinen Bereich. Die vorliegenden Ergebnisse sprechen für eine annähernde Gleichverteilung von gemeinsamen und spezifischen Faktoren. Die Unterschiede zwischen den Substanzklassen erscheinen gering sowohl hinsichtlich Effektivität als auch Akzeptanz.

Die untersuchten psychologischen Therapien waren im indirekten Vergleich mit Kontrolle methodischer Variablen und nach Ausschluss weniger etablierter Behandlung leicht effektiver in der Reduktion der Angstsymptomatik. Keine Unterschiede lagen hingegen für die depressive Symptomatik vor. In Studien, die beide Verfahren direkt miteinander verglichen, ergab sich dagegen kein bedeutsamer Vorteil für eine Behandlung. Allerdings zeigen die Ergebnisse eine hohe Variabilität zwischen den Untersuchungen, d.h. es liegen Studien vor, die eine höhere Effektivität für Pharmakotherapie zeigten, als auch Studien mit höherer Effektivität der Psychotherapie. Dabei scheinen die jeweiligen Untergruppen der untersuchten Therapien (also VT vs. KVT, TCA vs. SSRI), wie auch schon aus den Vergleichen zu den Kontrollbedingungen zu erwarten, keinen bedeutsamen Einfluss auf die gefundenen Unterschiede zu haben. Dies konnte allerdings aufgrund der geringen Studienanzahl nur in univariaten Analysen überprüft werden. Weitere Untersuchungen erscheinen notwendig. Diese erhöhen zum einen die Stabilität der Ergebnisse, zum anderen kann dann eine Überprüfung potentiell relevanter Variablen im Rahmen multivariater Verfahren durchgeführt werden.



Eine Einschätzung, welche Analyse bei divergenten Ergebnissen stärker gewichtet werden sollte, ist schwer zu treffen. Im indirekten Vergleich ergeben sich als Vorteile die höhere Studienanzahl und eine optimierte Realisierung der Therapie. So erfolgt eine Durchführung durch Wissenschaftler, die der jeweiligen Behandlung loyal gegenüber stehen. Allerdings kann nicht ausgeschlossen werden, dass dritte Variablen, die in der vorliegenden Arbeit nicht kontrolliert wurden, mit den Treatment-Bedingungen korrelieren und die gefundenen Effekte auf diese zurückzuführen sind. Gibt es mögliche systematische Unterschiede in der Stichprobe zwischen isolierten und vergleichenden Evaluationsstudien, die die divergenten Ergebnisse verursachen könnten? Dieser Punkt kann nicht ausgeschlossen werden, einige Aspekte sprechen aber eher dagegen. Zum einen sind die Selektionskriterien der Studien sehr ähnlich<sup>36</sup>. Auch Unterschiede zwischen Gruppen an Patienten mit unterschiedlicher Präferenz für eine Therapie scheinen gering. So fanden Bakker und Kollegen (2000) bei Patienten mit PDA, die eine randomisierte Aufteilung auf Psycho- oder Pharmakotherapie ablehnten und sich stattdessen für die Psychotherapie entschieden, keinen Unterschied zu Patienten, die diese Aufteilung akzeptierten, weder in demographischen Variablen, Symptomschwere vor der Behandlung noch Effektivität der Behandlung.

Behaviorale und kognitiv-behaviorale Verfahren erscheinen zusammengefasst mindestens so effektiv wie eine medikamentöse Therapie in der Behandlung der PDA. Bedeutsame Vorteile zeigten sich für die Psychotherapie in der Akzeptanz der Behandlung. Die Abbrecherrate war in pharmakologischen Studien bedeutend höher.

Die Kombinationsbehandlung aus Pharmako- und Psychotherapie, insbesondere mit einem SSRI oder TCA, zeigte eine bedeutsam leicht höhere Wirksamkeit bei Patienten mit PDA als Psychotherapie allein. Dieses Ergebnis wurde in allen Symptomkategorien gefunden mit Ausnahme der Lebensqualität. Allerdings war die Anzahl an Studien hier auch nur sehr gering. Betrachtete man die Daten nach Absetzen der Medikation, so zeigte sich in den vorliegenden Studien langfristig kein bedeutsamer Unterschied zwischen den Verfahren. Allerdings wurde in nicht einmal der Hälfte an Untersuchungen ein Follow up durchgeführt. Einschränkungen in der Repräsentativität der Ergeb-

---

<sup>36</sup> Der Punkt der zusätzlichen Medikation bei Psychotherapiestudien, die oft nicht als Ausschlusskriterium gilt, wird später im Hauptdiskussionsteil der Arbeit erörtert.

---

nisse scheinen gegeben. Deskriptiv zeigte sich ein möglicherweise sehr wichtiger Aspekt, der weitere Studien dringend erfordert. War zu Post nur jede sechste Studieneffektstärke negativ, d.h. zeigte eine geringere Symptomatik von Patienten der Psychotherapiebedingung, waren es im Follow up zwei von Drei. Eine Zusammenstellung theoretischer Annahmen zu potentiell negativen Wirkungen der Kombinationsbehandlung findet sich bei Westra und Stewart (1998), sowohl mögliche negative als auch positive Aspekte werden von Angendt, Frommberger und Berger (1998) diskutiert. Ein möglicher negativer Einfluss der Kombinationsbehandlung erscheint zu Post nicht gegeben, aufgrund der vorliegenden Daten scheint diese Annahme für die langfristige Effektivität noch nicht angemessen überprüft. Generelle Empfehlungen können somit nur eingeschränkt gegeben werden. Ein wichtiger Kritikpunkt im Design dieser Studien ist der in der Hälfte der Untersuchungen parallele Einsatz von Psychotherapie und Pillenplacebo in einer Gruppe. Mögliche Implikationen für die Validität der Studien werden im Hauptdiskussionsteil erläutert.

Es zeigte sich, wie auch schon von einigen anderen Autoren kritisiert, in einer Vielzahl an Studien eine Besonderheit der Definition der Klinischen Signifikanz. Oft wurde Endstate operationalisiert über Panikfreiheit. Vermeidung als wichtige Dimension der Störung wird dabei vollständig ignoriert. Eine vollständige Panikfreiheit kann sich aber auch ergeben über ein ausgeprägtes Vermeidungsverhalten.

Wie in bisherigen Meta-Analysen auch zeigen die Ergebnisse die Effektivität psychotherapeutischer und pharmakologischer Interventionen in der Behandlung der PDA. Auch die zeitliche Stabilität der Effekte bei Psychotherapie konnte bestätigt werden. Die aktuelle Arbeit ermöglicht dabei eine Generalisierbarkeit der Ergebnisse. Die von Ruhmland (2001c) gefundene höhere Effektivität von Exposition in vivo im Vergleich zu KVT konnte nicht gezeigt werden, allerdings wurde Exposition in der vorliegenden Arbeit mit anderen konfrontativen Verfahren in einer Kategorie zusammen untersucht. Die in einigen Meta-Analysen gefundene höhere Effektivität der Psychotherapie im indirekten Vergleich konnte in der vorliegenden Arbeit auch nach Anpassung der Kontrollgruppe und Kontrolle methodischer Variablen bestätigt werden, allerdings war der Unterschied nur marginal signifikant. Der bedeutsame Vorteil von SSRI wie bei Boyer (1995) konnte in der vorliegenden Arbeit nicht gezeigt werden.

## 3.5 Posttraumatische Belastungsstörung

### 3.5.1 Deskriptive Analyse der Primärstudien

Es konnten insgesamt 26 Studien gefunden werden, die den Inklusionskriterien entsprachen: 14 Pharmako- und 12 Psychotherapiestudien. In keiner der Studien wurden beide Verfahren zusammen überprüft. Es wurden fast dreimal so viele Patienten in Pharmako- im Vergleich zu Psychotherapiestudien untersucht. Vor allem Psychotherapiestudien wurden erst in den letzten Jahren publiziert. Die Hälfte der Studien, meist Pharmakotherapie, untersuchte eine Stichprobe mit gemischten Traumata, gefolgt von Kriegserlebnissen, sexuellem bzw. nicht-sexuellem Überfall sowie Verkehrsunfällen. In ca. einem Drittel der Studien waren mehr als 25 % der Patienten komorbid depressiv. Weitere deskriptive Daten sind in Tabelle 21 dargestellt.

Tabelle 21: Deskriptive Beschreibung der Primärstudien

	PsT	PhT
Patienten N <sup>a</sup>	686	1905
Durchschnittliche Gruppengröße <sup>a</sup>	19.7	60.8
N ≤ 20 (%)	66.7	42.9
N > 20 (%)	33.3	57.1
Alter	38.0	40.2
Dauer der Störung (Jahre)	7.4	10.5
Frauen (%)	54.5	40.1
Dauer der Therapie (PsT: Stunden, PhT: Wochen)	17.0	9.7
Publikationsjahr		
bis 1985 (%)	0	0
1986 – 1995 (%)	16.7	42.9
ab 1996 (%)	83.3	57.1
Intent-to-treat Auswertung (%)	16.7	71.4
Placebowashout (%)		28.6
Feste Dosierung (%)		7.1

<sup>a</sup> Zahlen geben die in die Auswertung aufgenommenen Patienten wieder

### 3.5.2 Vergleich der Prä-Werte

In allen Symptomkategorien (Angst, Depression, Lebensqualität) zeigte sich im Vergleich zwischen PsT und WL kein signifikanter Unterschied ( $g = 0.10$ , 95%-CI:  $-0.07 - 0.28$ ,  $\chi^2 = 4.79$ ,  $n = 12$ ;  $g = 0.10$ , 95%-CI:  $-0.09 - 0.29$ ,  $\chi^2 = 4.90$ ,  $n = 9$ ;  $g = 0.05$ , 95%-CI:  $-0.24 - 0.34$ ,  $\chi^2 = 1.09$ ,  $n = 5$ ).

Bei Betrachtung der Prä-ESn zwischen PsT und TP zeigte sich in der Symptomkategorie Angst keine signifikante mittlere ES ( $g = -0.13$ , 95%-CI:  $-0.51 - 0.25$ ,  $\chi^2 = 0.29$ ,  $n = 4$ ), in den anderen Symptomkategorien lagen zu wenig Daten vor.

In Studien, die einen direkten Vergleich zwischen VT und KVT untersuchten, ergab sich zu Prä eine nicht signifikante mittlere ES von  $g = -0.12$  in der Kategorie Angst (95%-CI:  $-0.32 - 0.08$ ,  $\chi^2 = 12.33$ ,  $n = 8$ ), in der Kategorie Depression von  $g = -0.14$  (95%-CI:  $-0.37 - 0.09$ ,  $\chi^2 = 2.41$ ,  $n = 6$ ). Gruppen mit behavioraler Behandlung haben bei Beginn der Untersuchung eine etwas geringere Symptomatik gegenüber kognitiv-behavioralen Gruppen.

In keiner der Symptomkategorien Angst, Depression und Lebensqualität ergab sich zu Prä ein signifikanter Unterschied zwischen PhT und PP, die mittleren ESn sind nahe Null ( $g = 0.01$ , 95%-CI:  $-0.10 - 0.12$ ,  $n = 12$ ;  $g = -0.03$ , 95%-CI:  $-0.16 - 0.09$ ,  $n = 7$ ;  $g = 0.05$ , 95%-CI:  $-0.10 - 0.19$ ,  $n = 5$ ).

### 3.5.3 Psychotherapie

#### Effektivität: Vergleich zur Warteliste

Alle Studien mit behavioralen Verfahren und Warteliste untersuchten die Wirksamkeit von Exposition. Auch bei kognitiv-behavioralen Verfahren war Exposition meist ein Bestandteil der Therapie. Alle ESn der Primärstudien in den Hauptkategorien waren positiv, d.h. keine Studie zeigte ein schlechteres Ergebnis der BG im Vergleich zur unbehandelten KG. In den Hauptkategorien zeigte sich zu post eine signifikant positive mittlere ES im Vergleich zur WL im Bereich von  $g = 0.80$  (Lebensqualität) bis  $g = 1.36$  (Klinische Signifikanz). Die untersuchten psychotherapeutischen Verfahren erscheinen

somit sehr effektiv in der Behandlung der PTB. Die Wirksamkeit zeigte sich dabei nicht nur in der Kernsymptomatik, sondern auch für Depression und Lebensqualität. Die Ergebnisse der Analysen finden sich in Tabelle 22, die einzelnen ESn in Abbildung 12.

Betrachtete man nur Studien, die eine Berechnung der ES über Mittelwerte und Standardabweichungen ermöglichten, veränderte sich die mittlere ES in der Kategorie Angst kaum (REM:  $g = 0.95$ , 95%-CI: 0.67 – 1.23,  $\sigma_{\theta}^2 = 0.06$ ; FEM:  $g = 0.92$ , 95%-CI: 0.73 – 1.11,  $\chi^2 = 14.53$ ,  $n = 11$ ). Es zeigten sich trotz einzelner Heterogenität für die Interpretation der Ergebnisse keine bedeutsamen Unterschiede zwischen den beiden Modellen; die Konfidenzintervalle enthielten nie Null.

Eine Studie untersuchte ein Selbsthilfeprogramm, es ergab sich hier in der Kategorie Angst eine ES von  $g = 0.21$ , deutlich geringer als die mittlere ES für eine Behandlung mit Therapeuten.

Tabelle 22: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  zu Post für den Vergleich von behavioralen und kognitiv-behavioralen Verfahren zur Warteliste

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
Post	Angst	REM	0.91	0.13	0.62 – 1.20	0.09 <sup>a</sup>		12
		FEM	0.88	0.09	0.70 – 1.07		17.48 <sup>+</sup>	12
	Kognition	REM	0.83	0.27	0.07 – 1.59	0.24		5
		FEM	0.70	0.13	0.45 – 0.96		9.49 <sup>+</sup>	5
	Depression	REM	0.84	0.13	0.55 – 1.14	0.05		10
		FEM	0.85	0.10	0.65 – 1.04		11.53	10
	Klin. Sign.	REM	1.36	0.17	0.94 – 1.77	0.04		7
		FEM	1.36	0.15	1.06 – 1.65		6.61	7
	Lebensqualität	REM	0.80	0.24	0.12 – 1.47	0.17		5
		FEM	0.81	0.16	0.50 – 1.11		7.80 <sup>+</sup>	5

<sup>a</sup>  $\sigma_{\theta}^2 > 0$  auch bei Ausschluss der Studie mit extremster ES definiert über die größte Abweichung vom ungewichteten Mittel

<sup>+</sup>  $p < .10$

### Effektivität: Vergleich zur behandelten Kontrollgruppe

Aufgrund der geringen Gruppenvergleiche mit TP ( $n = 4$ ), wurde nur die Kategorie Angst berechnet. Auch hier ergaben sich nur positive ESn in den Primärstudien. Die mittlere ES war wie zu erwarten kleiner als im Vergleich zur WL und betrug  $g = 0.46$ , diese war im REM nicht signifikant (95%-CI: -0.16 - 1.08,  $\sigma_{\theta}^2 = 0$ ), aber im FEM (95%-CI: 0.08 - 0.84,  $\chi^2 = 1.29$ ). In den vorliegenden Studien zeigen somit behaviorale und kognitiv-behaviorale spezifische Effekte, die Wirksamkeit dieser Verfahren ist in den eingeschlossenen Untersuchungen nicht allein auf gemeinsame Faktoren zurückzuführen.

### Zeitliche Stabilität der Therapie

Die Messungen zu Follow-up lassen sich für die Symptomkategorie Angst wie folgt charakterisieren:  $f1 = 3.0$  Monate ( $SD = 0.00$ ) und  $f2 = 8.7$  Monate ( $SD = 2.63$ ) einschließlich einer Studie mit Ergebnissen zur Langzeiteffektivität (ab 12 Monate). Es zeigte sich weder zu  $f1$  noch zu  $f2$  eine bedeutsame Veränderung der mittleren ES in den Symptomkategorien. In Tabelle 23 sind die Ergebnisse der Analysen dargestellt. Es lag in keiner der Symptomkategorien Heterogenität vor. In beiden statistischen Modellen REM und FEM wurden äquivalente Ergebnisse berechnet.

Tabelle 23: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  der Veränderung von Post zu Follow up (f) für behaviorale und kognitiv-behaviorale Verfahren

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
f1	Angst	REM	-0.04	0.10	-0.26 - 0.18	0		10
		FEM	-0.04	0.10	-0.22 - 0.15		1.83	10
	Depression	REM	-0.09	0.10	-0.31 - 0.13	0		9
		FEM	-0.09	0.10	-0.28 - 0.10		3.82	9
f2	Angst	REM	0.01	0.09	-0.20 - 0.21	0		10
		FEM	0.01	0.09	-0.17 - 0.18		8.38	10
	Depression	REM	-0.06	0.09	-0.28 - 0.16	0		9
		FEM	-0.06	0.09	-0.24 - 0.13		1.76	9
f1/f2	Klin. Sign.	REM	0.03	0.09	-0.19 - 0.26	0		8
		FEM	0.03	0.09	-0.15 - 0.22		3.94	8

## Akzeptanz

Ca. jeder Fünfte brach vorzeitig die PsT ab ( $M = 18.0$ ,  $SD = 10.33$ ,  $n = 21$ ). Die Akzeptanz erscheint somit gut. Für behaviorale Verfahren ergab sich mittlere Abbrecherrate von 16.6 ( $SD = 8.33$ ,  $n = 8$ ), für kognitiv-behaviorale Verfahren von 18.8 ( $SD = 11.63$ ,  $n = 8$ ). Der Unterschied war nicht signifikant ( $t(19) = 0.47$ , n.s.)<sup>37</sup>. Die Abbrecherrate in der Wartelistenbedingung war mit 5 % niedrig ( $SD = 5.72$ ,  $n = 7$ ).

## Moderatoren

Alle Analysen wurden aufgrund der Studienanzahl an den ESn im Vergleich zur WL durchgeführt. Es ergab sich ein signifikanter Unterschied in der Art der Erhebung: Fremdratings zeigten signifikant höhere ESn als Selbstratings, trotz meist formaler Blindheit des Erhebers ( $\beta = -0.53$ ,  $p < .05$ ,  $R^2 = .69$ ;  $n = 20$ ). Auch bei Einschluss von Studien, die eine Erhebung über beide Arten von Instrumenten durchführten, ergab sich der Unterschied ( $\beta = -0.53$ ,  $p < .05$ ,  $R^2 = .77$ ;  $n = 18$ ). Um den Einfluss der Änderungssensitivität verschiedener Symptomkategorien auszuschließen, wurde in der nachfolgenden Analyse nur die Kategorie Angst betrachtet. Weiterhin ergab sich eine höhere ES für das Fremdrating, der Unterschied war allerdings nicht mehr signifikant ( $\beta = -0.18$ , n.s.,  $R^2 = .00$ ;  $n = 19$ ). Es zeigte sich kein signifikanter Einfluss der Stichprobengröße auf die ES, allerdings bzgl. der Therapiedauer in Stunden wider Erwarten eine signifikante Abnahme der ES bei längerer Therapie ( $\beta_N = -0.32 \times 10^{-2}$ , n.s.,  $\beta_{\text{Dauer}} = -0.06$ ,  $p < .10$ ,  $\beta_{VT:KVT} = 0.30$ , n.s.,  $R^2 = .88$ ;  $n = 12$ ). Betrachtete man die untersuchte Population zeigte sich, dass die beiden längsten Therapien die geringsten ESn aufwiesen und ausschließlich Vietnamveteranen mit langer Chronizität der Erkrankung untersucht wurden. Bei Ausschluss dieser beiden Studien ergab sich kein bedeutsamer Zusammenhang zwischen Therapiedauer und Effektivität ( $\beta_N = -0.36 \times 10^{-2}$ , n.s.,  $\beta_{\text{Dauer}} = -0.03$ , n.s.,  $\beta_{VT:KVT} = 0.22$ , n.s.,  $R^2 = .00$ ;  $n = 10$ ). Der Einschluss anderer Prädiktoren wurde aufgrund von Verteilungseigenschaften nicht durchgeführt.

---

<sup>37</sup> Normalverteilung gegeben

### Vergleich behaviorale und kognitiv-behaviorale Verfahren

Bei Betrachtung der ES aus dem WL-BG-Vergleich zeigten kognitiv-behaviorale Verfahren bei Kontrolle der Therapiedauer und Stichprobengröße eine höhere ES als rein behaviorale Verfahren, dieser Unterschied war nicht signifikant; die Ergebnisse der Regression sind im Abschnitt Moderatoren aufgeführt.

Betrachtete man nur Studien, in denen die beiden Verfahren direkt miteinander verglichen wurden, zeigten sich in allen Symptomkategorien mittlere ESn nahe Null. Diese Werte unterschätzen bei zusätzlicher Betrachtung der leicht niedrigeren Werte der VT-Gruppen zu Prä etwas den Vorteil von KVT. Die ESn der einzelnen Studie lagen dabei zwischen  $g = -0.41$  und  $g = 0.37$ . Es liegen somit sowohl Studien vor, die ein besseres Ergebnisse für VT im Vergleich zu KVT zeigen, als auch Studien, in denen KVT effektiver war. Die Daten sprechen allerdings nicht dafür, dass die Unterschiede zwischen den Untersuchungen in der Kategorie Angst bedeutsam sind. Bei Depression ergab sich dagegen bedeutsame Heterogenität. Die Ergebnisse der Analysen sind in Tabelle 24 dargestellt.

Alle ESn wurden über Mittelwerte und Standardabweichungen berechnet. Eine Regressionsanalyse wurde aufgrund der RE-Varianz von Null nicht durchgeführt: es wurde keine bedeutsame Variabilität zwischen den Studien gefunden. Des weiteren ergab sich kein bedeutsamer Unterschied zwischen Fremd- und Selbstrating ( $\beta = -0.42 \times 10^{-4}$ , n.s.,  $R^2 = .00$ ;  $n = 16$ ).

Tabelle 24: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  zu Post und Follow up (f) für den Vergleich von konfrontativen und kognitiv-behavioralen Verfahren

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	N
Post	Angst	REM	0.03	0.10	-0.21 – 0.28	0		8
		FEM	0.03	0.10	-0.17 – 0.24		8.39	8
	Depression	REM	-0.10	0.17	-0.53 – 0.33	0.07		6
		FEM	-0.04	0.12	-0.27 – 0.18		11.34 <sup>+</sup>	6
	Klin. Sign.	REM	-0.09	0.12	-0.38 – 0.19	0.01		7
		FEM	-0.09	0.10	-0.30 – 0.11		4.31	7

<sup>+</sup>  $p < .10$



### **Fail-Safe N**

Um die beobachteten Effekte im Vergleich von PsT zu WL auf einen kleinen Effekte von  $g = 0.2$  zu reduzieren, würden 43 Gruppenvergleiche mit  $g = 0$  benötigt. Dies übersteigt nicht die von Rosenthal vorgeschlagene Grenze, erscheint aber ausreichend hoch, um von einem relativ stabilen Effekt sprechen zu können. Im Vergleich zu einem TP/PP wären nur 5 Studien mit einer Effektstärke von 0 notwendig zur Reduzierung des mittleren Effektes auf  $g = 0.2$  sowie 9 Studien mit einer Effektstärke von  $-0.2$  für einen mittleren Effekt von  $g = 0$ . Das Ergebnis kann somit nicht als stabil gelten. Wie aufbauend auf der Anzahl der vorliegenden Studien zu erwarten, sind für den direkten Vergleich von VT und KVT nur wenige Studien notwendig, um ein abweichendes Bild zu erhalten. So werden 6 Untersuchungen mit einer ES von  $g = -0.5$  benötigt, um einen mittleren Effekt von  $g = -0.2$  zugunsten VT zu erhalten, demgegenüber stehen 5 Studien mit einer ES von  $g = 0.5$  für einen mittleren Effekt von  $g = 0.2$  zugunsten KVT.

### **3.5.4 Pharmakotherapie**

#### **Effektivität: Vergleich zur Placebobedingung**

Mehr als die Hälfte an Studien untersuchte SSRI, die restlichen Untersuchungen waren verschiedene Substanzklassen. In einer Studie zeigten Patienten mit PP zu Post eine geringere Symptomatik als die Medikamentenbedingung ( $g = -0.26$ ), in allen anderen Studien war PhT effektiver als PP. In allen Symptomkategorien ergab sich zu post im Vergleich zur KG eine signifikant positive mittlere ES. Die kleinste mittlere ES zeigte sich dabei in den Unterkategorien Kognition und Erregung mit  $g = 0.24$ , die größte mittlere ES in der Hauptkategorie Klinische Signifikanz mit  $g = 0.44$ . In allen Symptomkategorien waren somit die untersuchten Psychopharmaka, insbesondere SSRI, effektiv in der Behandlung von Patienten mit PTB. Die Ergebnisse der Analysen sind in Tabelle 25 dargestellt, die einzelnen ESn in Abbildung 12.

Bei Ausschluss der Studien, in denen eine Berechnung nicht über Mittelwerte und Standardabweichungen erfolgte, stieg die Effektstärke stark an, es wurden allerdings 50% der Effektstärken ausgeschlossen, möglicherweise eine Einschränkung der Repräsentativität (REM:  $g = 0.52$ , 95%-CI:  $0.15 - 0.89$ ,  $\sigma_{\theta}^2 = 0.05$ ; FEM:  $g = 0.46$ , 95%-CI:  $0.25 - 0.67$ ,  $\chi^2 = 9.93$ ,  $n = 7$ ).

Es lag in keiner der Symptomkategorien Heterogenität vor, so dass in beiden statistischen Modellen REM und FEM äquivalente Ergebnisse berechnet wurden.

Tabelle 25: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  zu Post für den Vergleich einer medikamentösen zur Placebobedingung

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
Post	Angst	REM	0.28	0.05	0.18 – 0.39	0		15
		FEM	0.28	0.05	0.19 – 0.38		15.76	15
	Vermeidung	REM	0.27	0.05	0.15 – 0.39	0		10
		FEM	0.27	0.05	0.16 – 0.37		6.64	10
	Kognition	REM	0.24	0.06	0.10 – 0.37	0.01		10
		FEM	0.23	0.05	0.13 – 0.33		9.11	10
	Erregung	REM	0.24	0.05	0.11 – 0.37	0		6
		FEM	0.24	0.05	0.13 – 0.34		0.98	6
	Depression	REM	0.29	0.07	0.14 – 0.45	0.01		10
		FEM	0.28	0.05	0.18 – 0.38		10.54	10
	Klin. Sign.	REM	0.44	0.07	0.28 – 0.59	0.01		10
		FEM	0.45	0.05	0.35 – 0.55		11.87	10
	Lebensqualität	REM	0.27	0.06	0.12 – 0.43	0		6
		FEM	0.27	0.06	0.15 – 0.39		4.20	6

### Zeitliche Stabilität der Therapie

Zu Follow up konnte keine mittlere ES berechnet werden, da nur aus einer Studie angemessene Daten vorlagen. In dieser Studie wurde Alprazolam überprüft und es zeigte sich eine Abnahme der ES im Rahmen einer Crossover-Studie mit Pillenplacebo von  $g = -1.00$  zu 1.25 Monaten, allerdings betrug die Stichprobengröße nur vier Personen.

### Akzeptanz

Die mittlere Dropout-Rate in den BGN betrug 33.7 % ( $SD: 11.62, n = 14$ ). Das heißt, jeder Dritte brach vorzeitig die Behandlung ab, die Akzeptanz durch Patienten mit PTB ist nicht zufrieden stellend. Auch in der KG zeigte sich eine hohe Dropout-Rate von 32.8 % ( $SD = 22.80, n = 12$ ).

## Moderatoren

Aufgrund der RE-Varianz von Null entfiel eine Berechnung der Regressionen. Es zeigte sich keine bedeutsame Variabilität zwischen den Studien. Es ergab sich zudem kein bedeutsamer Unterschied zwischen Fremd- und Selbstrating ( $\beta = -0.05$ , n.s.,  $R^2 = .00$ ;  $n = 26$ ).

## Fail-Safe N

Um die beobachteten Effekte im Vergleich von PhT zu PP auf einen kleinen Effekte von  $g = 0.2$  zu reduzieren, würden, da die mittlere ES nur gering war, 6 Gruppenvergleiche mit  $g = 0$  benötigt. Insgesamt zeigen sich die Ergebnisse, im Besonderen für Pharmakotherapie als auch für den Vergleich von behavioralen und kognitiv-behavioralen Verfahren als wenig stabil gegenüber potentiellen Studien in Schubladen.

### 3.5.5 Vergleich Psycho- und Pharmakotherapie

In keiner der vorliegenden Studien wurden die beiden Verfahren direkt miteinander verglichen. Da nur eine Studie ausschließlich TP einsetzte, wurde nur eine Schätzung der Effektivität von WL vs. PP durchgeführt. Es zeigte sich hier in der Regression ein Unterschied in den Prä-Post-ES von  $g = 0.26$ , wobei PP höhere ESn aufwies als die WL-Bedingung. Nach Korrektur der Post-ESn um diesen Wert ergaben sich für PsT signifikant höhere ESn im Vergleich zur PhT ( $\beta = 0.30$ ,  $p < .05$ ,  $R^2 = 0.51$ ;  $n = 27$ ). Betrachtete man in einer multiplen Regression zusätzlich das Ausmaß an Abbrechern und die Stichprobengröße<sup>38</sup> zeigte PsT weiterhin leicht höhere Effekte, die beiden Verfahren waren allerdings nicht mehr bedeutsam voneinander verschieden; es ergab sich kein Einfluss der Stichprobengröße auf ES, aber tendenziell geringere Effekte für Therapien mit einer hohen Abbrecherquote ( $\beta_{\text{PhT:PsT}} = 0.14$ , n.s.,  $\beta_{\text{N}} = -0.15 * 10^{-3}$ , n.s.,  $\beta_{\text{Dropout}} = -0.02$ ,  $p < .10$ ,  $R^2 = 1$ ;  $n = 24$ ). Aufgrund der hohen Unterschiede für die Ergebnisse der Effektivität der Pharmakotherapie betrachtete man nur Studien mit Berechnung der Effektstärken aus Mittelwerten und Standardabweichungen, erfolgte im Rahmen einer Sensitivitätsanalyse eine Regression ohne entsprechende Untersuchun-

---

<sup>38</sup> Aufgrund der Verteilungseigenschaften wurde zur Stabilisierung der Ergebnisse auf die Prädiktoren Analyseart und Diagnosesystem verzichtet.

gen. Die einfache Regression ergab keinen bedeutsamen Unterschied zwischen den Verfahren ( $\beta = 0.13$ , n.s.,  $R^2 = 0.03$ ;  $n = 18$ ), die Ergebnisse sind somit abhängig von den Entscheidungen des Meta-Analysten. In der multiplen Regression waren beide Verfahren gleich effektiv und es zeigte sich eine tendenziell geringere Effektstärke bei Therapien mit hoher Abbrecherrate ( $\beta_{\text{PhT:PsT}} = 0.02$ , n.s.,  $\beta_{\text{N}} = -0.25 \cdot 10^{-2}$ , n.s.,  $\beta_{\text{Drop-out}} = -0.02$ ,  $p < .10$ ,  $R^2 = .57$ ;  $n = 17$ ).

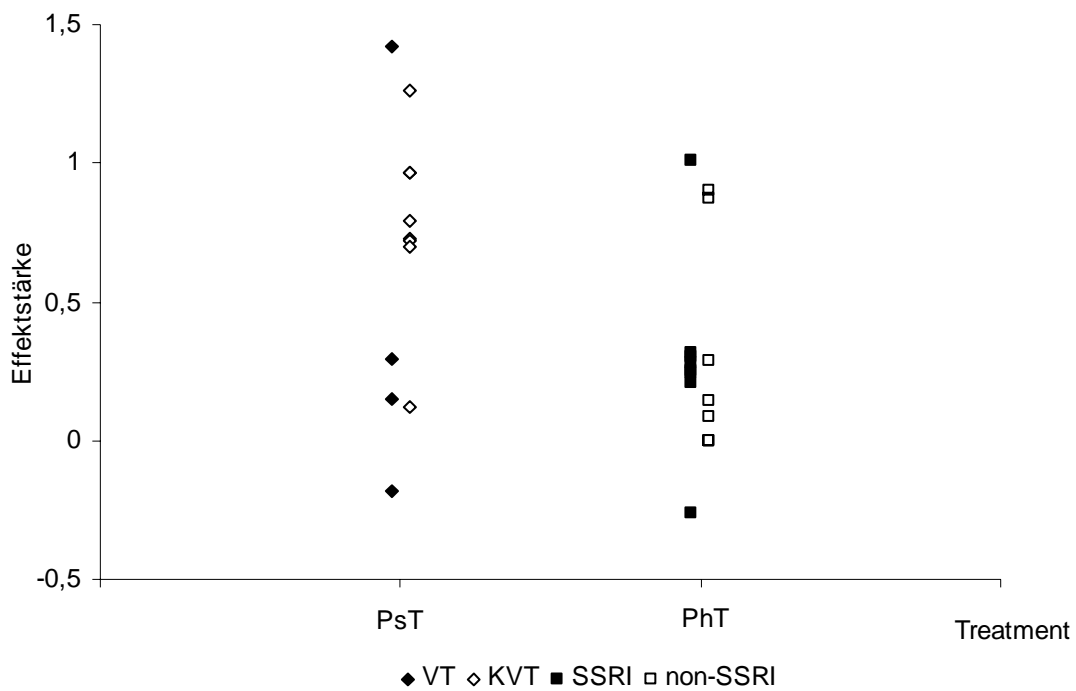


Abbildung 12: Studieneffektstärken für die Symptomkategorie Angst für die einzelnen Therapien vom Vergleich Psychotherapie zur Warteliste und Pharmakotherapie zu Pillenplacebo

Patienten aus PsT-Studien brachen signifikant weniger häufig die Behandlung ab als Patienten aus PhT-Studien ( $t(33) = 4.19$ ,  $p < .001$ )<sup>39</sup>. Die Akzeptanz für Psychotherapie ist somit bedeutend höher.

<sup>39</sup> Normalverteilung gegeben

### 3.5.6 Diskussion

Die Ergebnisse der Meta-Analyse zeigen, dass die untersuchten behavioralen und kognitiv-behavioralen Verfahren, im Besonderen mit Exposition als Element oder eigenständige Therapie, sehr effektiv in der Behandlung der PTB sind und Patienten eine bedeutend geringere Symptomatik nach Abschluss der Therapie im Vergleich zu einer unbehandelten Gruppen aufweisen. In den eingeschlossenen Studien ergab sich zudem, dass die Wirksamkeit nicht allein auf gemeinsame Faktoren zurückzuführen ist, sondern spezifische Faktoren der untersuchten Therapien eine zusätzliche effektive Komponente darstellen. Spezifische und gemeinsame Faktoren scheinen dabei annähernd gleichverteilt, wenn auch auf unterschiedliche Studien basierend und mit einer nur geringen Anzahl an Untersuchungen mit Therapieplacebo. Die Ergebnisse erscheinen stabil über die Zeit. Auch nach Abschluss der Behandlung erlebten die Patienten keine Verschlechterung der Symptomatik.

Kognitiv-behaviorale Verfahren scheinen ebenso wirksam zu sein wie rein konfrontative Verfahren. Weder im indirekten, noch im direkten Vergleich ergaben sich bedeutende Unterschiede. Nach Brewin (1996) sollte eine kognitiv-behaviorale Behandlung insbesondere eine Veränderung des verbal zugänglichen Wissens ermöglichen. Somit könnten vor allem Patienten profitieren mit starken sekundären Emotionen, wie Schuld oder Trauer, oder mit Befürchtungen sehr komplexer Natur, wie z.B. das Erleben der gesamten Welt als gefährlich. Die vorliegenden Daten ermöglichten keine Beurteilung dieser Annahme, mögliche Unterschiede könnten sich schon auf Studienebene ausmitteln. Eine direkte Überprüfung in den Primärstudien wäre wünschenswert. Die Akzeptanz war für beide Verfahren gleich.

Die medikamentöse Behandlung der PTB, insbesondere mit SSRI, ist im Vergleich zu einer Pillenplacebobedingung effektiv. Die Höhe der Wirksamkeit war allerdings abhängig von den eingeschlossenen Studien. Betrachtete man sämtliche vorliegende Daten, ergab sich ein kleiner Effekt, schloss man dagegen nur Studien ein, die eine Berechnung über Mittelwerte und Standardabweichungen ermöglichten, zeigte sich ein mittlerer Effekt. Allerdings erfolgte die Berechnung dann nur für ca. die Hälfte an Untersuchungen, die Generalisierbarkeit erscheint aufgrund der geringen Anzahl an verbleibenden Studien gering. Die vorliegenden Befunde sprechen dafür, dass sich die

Wirksamkeit der medikamentösen Therapie über alle Pharmakotherapiestudien zu gleichen Teilen aus spezifischen und gemeinsamen Faktoren zu ergeben scheint. Eine deutliche Erhöhung des Anteils spezifischer Faktoren ergibt sich dagegen bei Betrachtung der Untersuchungen mit Berechnung der Effektstärken aus Mittelwerten und Standardabweichungen. Auch hier ergibt sich wieder die Einschränkung einer Schätzung über verschiedene Untersuchungen: der Anteil des Retest-Effektes wurde über die Wartelistebedingungen der Psychotherapiestudien geschätzt. Auffällig ist die eher geringe Akzeptanz für Pharmakotherapie. Jeder Dritte brach vorzeitig die Untersuchung ab.

Es zeigte sich im indirekten Vergleich von Psycho- und Pharmakotherapie nach Kontrolle methodischer Variablen eine ähnliche Wirksamkeit der beiden Verfahren. Allerdings war Akzeptanz für die Psychotherapie bedeutend höher als für die Pharmakotherapie. Betrachtet man zudem die Stabilität der Ergebnisse der Psychotherapien über die Zeit, während für Pharmakotherapien dementsprechende Aussagen nicht möglich waren, scheinen die vorliegenden Daten dafür zu sprechen, Patienten mit PTB eher mit behavioralen oder kognitiv-behavioralen Therapien denn mit medikamentösen Therapien zu behandeln.

Die Ergebnisse der vorliegenden Meta-Analyse zeigen, dass behaviorale, kognitiv-behaviorale und medikamentöse Therapie (bes. SSRI) effektiv in der Behandlung der PTB sind und die gefundenen positiven Resultate für Psychotherapie zeitlich stabil sind. Dieses Ergebnis wurde auch in den bisherigen Arbeiten gefunden, die aktuelle Arbeit ermöglicht allerdings die Generalisierbarkeit der Ergebnisse. Wie auch bei van Etten et al. (1998) ergab sich eine bedeutsam höhere Akzeptanz für Psycho- im Vergleich zu Pharmakotherapie. Die Autoren fanden zudem eine höhere Effektivität psychologischer Therapie (einschließlich EMDR, Hypnose, Psychodynamische Therapie) im Vergleich zu pharmakologischer Therapie für die PTB-Symptomatik bei Selbstbeurteilung, VT (einschließlich KVT) und SSRI (als wirksamste medikamentöse Therapie) waren aber gleich effektiv. Für die Fremdbeurteilung fanden sie eine leicht höhere Effektstärke für VT auch im Vergleich zu SSRI, allerdings nicht bedeutsam. In der aktuellen Arbeit ergaben sich gemittelt über alle Messinstrumente höhere Effektstärken für VT und KVT im Vergleich zu Pharmakotherapie, nach Kontrolle methodischer Variablen war der Unterschied allerdings nicht mehr bedeutsam.

Insgesamt ergibt sich aufgrund der geringen Anzahl an Untersuchungen nur eine eingeschränkte Interpretierbarkeit und Generalisierbarkeit der Ergebnisse. Änderungen in der Auswahl an Studien verändern die Resultate zum Teil drastisch. Auch die Stabilität der Ergebnisse berechnet über Fail-Safe N erscheint partiell gering. Weitere Untersuchungen sind somit zu fordern vor allem mit paralleler Überprüfung von Psycho- und Pharmakotherapie.

## 3.6 Soziale Phobie

### 3.6.1 Deskriptive Analyse der Primärstudien

Es konnten insgesamt 50 Studien gefunden werden, die den Inklusionskriterien entsprachen: 20 Pharmako- und 22 Psychotherapiestudien. In weiteren acht Studien wurden beide Verfahren zusammen untersucht einschließlich drei Studien, in denen zum Teil zusätzlich eine Kombinationsbehandlung untersucht wurde. Die Mehrzahl an Psychotherapiestudien wurde vor 1995 publiziert, während Pharmakotherapiestudien erst in den letzten Jahren veröffentlicht wurden. Knapp mehr als die Hälfte der untersuchten Patienten waren männlich, dies entspricht nur annähernd den gefundenen Daten in epidemiologischen Untersuchungen. In ca. einem Drittel der Studien hatten mehr als 25 % der Patienten eine komorbide Störung, meist eine selbstunsichere Persönlichkeitsstörung. In nur 8 % der Studien lagen komorbide Angststörungen oder Affektive Störungen vor. Weitere deskriptive Daten finden sich in Tabelle 26.

**Tabelle 26:** Deskriptive Beschreibung der Primärstudien

	PsT	PhT	PhT / PsT
Patienten Gesamt <sup>a</sup>	1006	2235	772
Durchschnittliche Gruppengröße <sup>a</sup>	17.9	49.9	27.5
n ≤ 20 (%)	68.2	25.0	37.5
n > 20 (%)	31.8	75.0	62.5
Alter	33.2	36.7	37.4
Dauer der Störung (Jahre)	16.7	19.1	22.0
Frauen (%)	48.0	43.3	50.7
Dauer der Therapie (PsT: Stunden, PhT: Wochen)	15.6	11.4	PsT: 17.5 PhT: 12.2
Publikationsjahr			
bis 1985 (%)	22.7	0	12.5
1986 – 1995 (%)	54.6	30.0	37.5
ab 1996 (%)	22.7	70.0	50.0
Intent-to-treat Auswertung (%)	0	80.0	37.5
Placebowashout (%)		40.0	12.5
Feste Dosierung (%)		25.0	0

<sup>a</sup> Zahlen geben die in die Auswertung aufgenommenen Patienten wider



### 3.6.2 Vergleich der Prä-Werte

In den Symptomkategorien Angst und Depression zeigte sich im Vergleich zwischen PsT und WL kein signifikanter Unterschied, alle mittleren ESn lagen um Null ( $g = 0.02$ , 95%-CI:  $-0.15 - 0.19$ ,  $\chi^2 = 3.72$ ,  $n = 19$ ;  $g = 0.05$ , 95%-CI:  $-0.20 - 0.30$ ,  $\chi^2 = 3.24$ ,  $n = 9$ ).

Bei Betrachtung der Prä-ESn zwischen PsT und TP/PP zeigte sich sowohl in den Symptomkategorien Angst als auch Depression eine leicht positive mittlere ES, diese war aber nicht signifikant ( $g = 0.16$ , 95%-CI:  $-0.05 - 0.38$ ,  $\chi^2 = 5.32$ ,  $n = 7$ ;  $g = 0.11$ , 95%-CI:  $-0.11 - 0.32$ ,  $\chi^2 = 2.56$ ,  $n = 6$ ). Personen in der BG hatten leicht geringere Prä-Werte als Personen in der KG.

Studien mit einem direkten Vergleich von VT und KVT zeigten zu Prä eine mittlere ES nahe Null in der Kategorie Angst ( $g = -0.01$ , 95%-CI:  $-0.20 - 0.18$ ,  $\chi^2 = 10.76$ ,  $n = 13$ ). In der Kategorie Depression ergab sich dagegen eine mittlere ES größer  $g = 0.20$  ( $g = 0.23$ , 95%-CI:  $-0.09 - 0.56$ ,  $\chi^2 = 6.08$ ,  $n = 6$ ), der Ausschluss einer Studie mit höchstem positiven Effekt reduzierte die mittlere ES auf  $g = 0.10$ , diese Studie wurde auch in der Analyse zu Post ausgeschlossen.

Weder in der Symptomkategorie Angst noch Depression zeigte sich ein signifikanter Unterschied zwischen PhT und PP, die mittlere ES lag bei Null ( $g < 0.01$ , 95%-CI:  $-0.09 - 0.10$ ,  $\chi^2 = 5.54$ ,  $n = 22$ ;  $g = 0.03$ , 95%-CI:  $-0.07 - 0.14$ ,  $\chi^2 = 13.69$ ,  $n = 18$ ).

In Studien, die einen direkten Vergleich von PhT und PsT untersuchten, ergab sich zu Prä eine mittlere ES nahe Null in der Kategorie Angst ( $g = 0.07$ , 95%-CI:  $-0.19 - 0.33$ ,  $\chi^2 = 0.93$ ,  $n = 6$ ).

### 3.6.3 Psychotherapie

#### **Effektivität: Vergleich zur Warteliste**

Zwei Drittel der Vergleiche zur unbehandelten Kontrollgruppe beinhaltete behaviorale Verfahren. Als häufigstes Verfahren wurde Exposition untersucht, auch mit zusätzlichen kognitiven Elementen. Ca. ein Drittel der Behandlungsgruppen war Training Sozialer Kompetenz meist ohne kognitive Restrukturierung. Die Mehrzahl der ES<sub>n</sub> der Primärstudien in den Hauptkategorien war positiv, zwei Studien zeigten in der Kategorie Depression ein leicht schlechteres Ergebnis der BG im Vergleich zur unbehandelten KG. In allen Symptomkategorien ergaben sich zu Post signifikant positive mittlere ES<sub>n</sub> im Vergleich zur WL im Bereich von  $g = 0.53$  (Depression) bis  $g = 1.38$  (Klinische Signifikanz). Die RE-Varianz war Null, es gab keine bedeutsame Variabilität zwischen den Studien.

In der Kategorie Depression lag Heterogenität vor, auch im REM enthielt das Konfidenzintervall nicht Null. In den übrigen Kategorien war Homogenität gegeben, so dass in beiden statistischen Modellen REM und FEM äquivalente Ergebnisse berechnet wurden. Die Ergebnisse der Analysen sind in Tabelle 27 dargestellt. Einen Überblick über die einzelnen ES<sub>n</sub> aufgeteilt nach behavioralen und kognitiv-behavioralen Verfahren findet sich in Abbildung 13.

Bei Ausschluss der Studien, in denen zur Berechnung der ES<sub>n</sub> keine Mittelwerte und Standardabweichungen vorlagen, sank die mittlere ES leicht und betrug in der Kategorie Angst  $g = 0.70$  (95%-CI: 0.48 – 0.92,  $\sigma_{\theta}^2 = 0$ ; FEM:  $g = 0.70$ , 95%-CI: 0.50 – 0.90,  $\chi^2 = 7.78$ ,  $n = 16$ ).

Tabelle 27: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  zu Post für den Vergleich von behavioralen und kognitiv-behavioralen Verfahren zur Warteliste

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
Post	Angst	REM	0.75	0.09	0.56 – 0.94	0		21
		FEM	0.75	0.09	0.57 – 0.92		9.95	21
	Vermeidung	REM	0.76	0.14	0.44 – 1.08	0		8
		FEM	0.76	0.14	0.49 – 1.02		6.73	8
	Kognition	REM	0.76	0.11	0.52 – 1.00	0		13
		FEM	0.76	0.11	0.54 – 0.98		9.88	13
	Depression	REM	0.55	0.18	0.12 – 0.99	0.11		8
		FEM	0.55	0.14	0.28 – 0.82		13.48 <sup>+</sup>	8
	Klin. Sign.	REM	1.38	0.16	1.00 – 1.75	0		8
		FEM	1.38	0.16	1.07 – 1.69		3.92	8

<sup>+</sup>  $p < .10$

### Effektivität: Vergleich zur Placebothherapie

Im Vergleich zu einer Placebothherapie waren die Mehrzahl an untersuchten Therapieformen kognitiv-behaviorale Verfahren. Mit Ausnahme einer ES in der Kategorie Depression waren alle ES positiv. Die mittlere ES war wie zu erwarten kleiner als im Vergleich zur WL, war aber für die Hauptkategorien Angst und Klinische Signifikanz bedeutsam von Null verschieden. Das heißt, behaviorale und kognitiv-behaviorale Verfahren zeigen für die Angstsymptomatik eine Wirksamkeit, die über eine Realisierung gemeinsamer Faktoren hinausgeht. Veränderungen zeigten sich dabei vor allem in der Unterkategorie Erregung, weniger für Kognitionen und Vermeidung. In der Kategorie Lebensqualität ergab sich eine leicht positive ES, diese war nicht signifikant im REM, aber im FEM. Es zeigte sich dagegen in den vorliegenden Studien kein bedeutsamer Unterschied zwischen PsT und TP/PP in der Kategorie Depression. Es ist zu beachten, dass diese ermittelten ESn den wahren Effekt leicht überschätzen aufgrund der Unterschiede zu Prä. Die Ergebnisse der Analysen sind in Tabelle 28 dargestellt, die einzelnen ESn in Abbildung 13.

Bei Ausschluss der Studien, in denen zur Berechnung der ESn keine Mittelwerte und Standardabweichungen vorlagen, stieg die mittlere ES leicht an und betrug in der

Kategorie Angst  $g = 0.29$  (95%-CI: 0.03 – 0.54,  $\sigma_{\theta}^2 = 0$ ; FEM:  $g = 0.29$ , 95%-CI: 0.08 – 0.50,  $\chi^2 = 6.49$ ,  $n = 8$ ).

Tabelle 28: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  zu Post für den Vergleich von behavioralen und kognitiv-behavioralen Verfahren zur Placebobedingung

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
Post	Angst	REM	0.25	0.08	0.07 – 0.43	0		11
		FEM	0.25	0.08	0.09 – 0.41		8.64	11
	Vermeidung	REM	0.30	0.12	-0.01 – 0.60	0.04		7
		FEM	0.26	0.09	0.08 – 0.43		9.53	7
	Kognition	REM	0.15	0.09	-0.10 – 0.41	0.03		9
		FEM	0.13	0.11	-0.04 – 0.30		10.98	9
	Erregung	REM	0.57	0.13	0.20 – 0.94	0		5
		FEM	0.57	0.13	0.31 – 0.83		2.18	5
	Depression	REM	0.16	0.18	-0.29 – 0.61	0.10		6
		FEM	0.14	0.12	-0.09 – 0.37		11.30 <sup>+</sup>	6
	Klin. Sign.	REM	0.67	0.24	0.02 – 1.33	0.18		5
		FEM	0.64	0.14	0.37 – 0.92		10.61 <sup>+</sup>	5
	Lebensqualität	REM	0.33	0.13	-0.01 – 0.67	0.04		6
		FEM	0.32	0.10	0.13 – 0.51		5.81	6

<sup>+</sup>  $p < .10$

### Zeitliche Stabilität der Therapie

Die Messungen zu Follow-up lassen sich für die Symptomkategorie Angst wie folgt charakterisieren: f1 = 2.7 Monate ( $SD = 0.84$ ), f2 = 6.0 Monate ( $SD = 0.00$ ) und f3 = 14.5 Monate ( $SD = 2.95$ ). Die mittleren ESn erwiesen sich als zeitlich stabil, es gab zu keinem der Messzeitpunkte eine signifikante Zu- oder Abnahme der Effekte.

Es zeigten sich zwischen den Modellen keine substanziellen Unterschiede in der Interpretation der Bedeutsamkeit der mittleren ES. Die Ergebnisse der Analysen sind in Tabelle 29 dargestellt.

Tabelle 29: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  der Veränderung von Post zu Follow up (f1: bis sechs Monate, f2: 6 bis 12 Monate, f3: ab 12 Monate) für behaviorale und kognitiv-behaviorale Verfahren

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
f1	Angst	REM	0.13	0.10	-0.09 – 0.35	0.08	11.53	20
		FEM	0.12	0.08	-0.04 – 0.28			
	Depression	REM	-0.19	0.17	-0.60 – 0.21	0.10	7.65	8
		FEM	-0.15	0.12	-0.39 – 0.09			
f2	Angst	REM	0.07	0.13	-0.23 – 0.37	0	2.26	8
		FEM	0.07	0.13	-0.18 – 0.32			
	Depression	REM	0.08	0.14	-0.28 – 0.45	0	2.50	6
		FEM	0.08	0.14	-0.19 – 0.36			
f3	Angst	REM	0.26	0.14	-0.10 – 0.63	0	1.26	6
		FEM	0.26	0.14	-0.01 – 0.54			
	Depression	REM	0.08	0.16	-0.43 – 0.59	0.01	2.94	4
		FEM	0.08	0.16	-0.23 – 0.38			

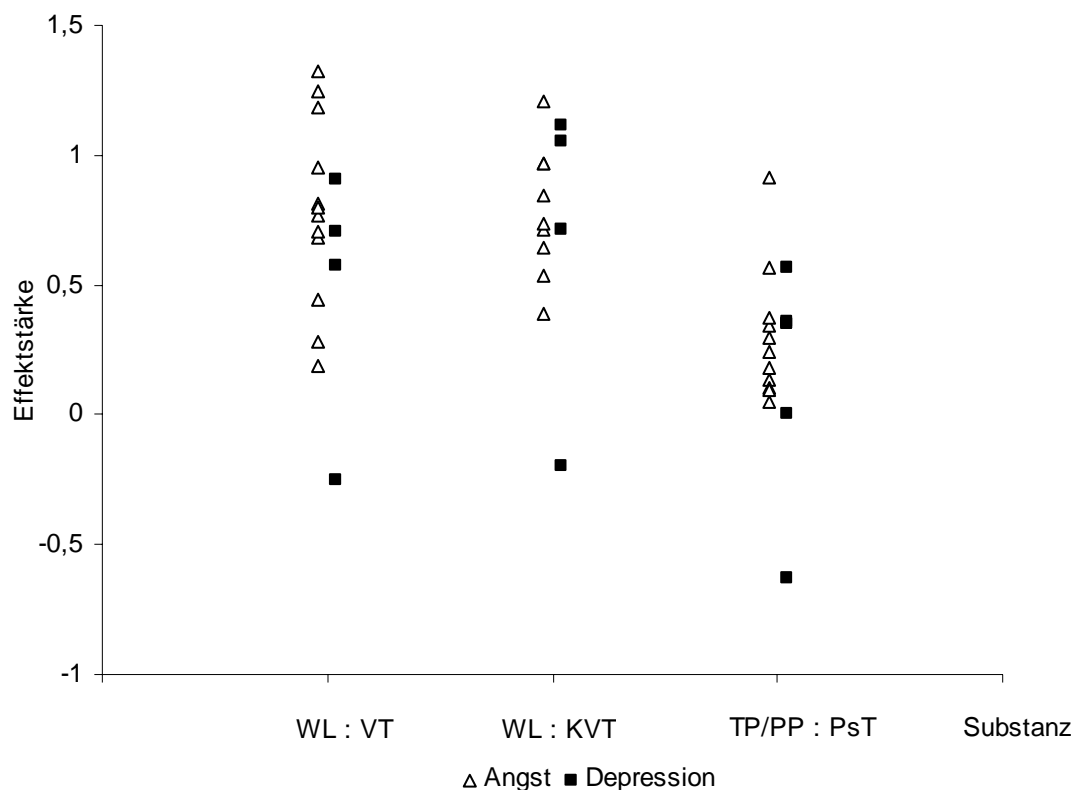


Abbildung 13: Studieneffektstärken für die Symptomkategorien Angst und Depression aufgeteilt nach Art der Kontroll- und Behandlungsgruppe

### **Akzeptanz**

Die mittlere Dropout-Rate betrug für die Behandlungsgruppen insgesamt 15.8 % ( $SD = 9.04$ ,  $n = 39$ ), für behaviorale Verfahren 14.4 % ( $SD = 10.41$ ,  $n = 18$ ) und für kognitiv-behaviorale Verfahren 16.8 % ( $SD = 8.16$ ,  $n = 19$ ). Die beiden Therapien unterschieden sich nicht in ihrer Akzeptanz ( $t(35) = 0.77$ , n.s.). Nur 4.7 % der Patienten der Wartelistenbedingung verließen vorzeitig die Studie ( $SD = 5.56$ ,  $n = 10$ ).

### **Moderatoren**

Aufgrund der RE-Varianz von 0 im Rahmen der Sensitivitätsanalyse entfiel eine Berechnung der Regressionen sowohl für den Vergleich mit einer WL als auch mit TP/PP. Im Vergleich zur WL entfiel zusätzlich die Betrachtung von Fremd- vs. Selbstrating, da meist letzteres eingesetzt wurde. Für den Vergleich zu einer Placebotherapie ergab sich kein bedeutsamer Unterschied zwischen den beiden Arten der Beurteilung ( $\beta = -0.11$ , n.s.,  $R^2 = 0$ ;  $n = 17$ ).

### **Vergleich konfrontative<sup>40</sup> und kognitiv-behaviorale Verfahren**

Der Vergleich über die WL-ES zeigte im Rahmen einer multiplen Regression, wie aufgrund der RE-Varianz von Null zu erwarten, bei Kontrolle von Abbrecherrate, Stichprobengröße, Diagnosesystem und Therapiedauer keinen bedeutsamen Unterschied zwischen den Verfahren ( $\beta_{VT:KVT} = 0.07$ , n.s.,  $\beta_{Dropout} = 0.93 \times 10^{-2}$ , n.s.,  $\beta_N = 0.42 \times 10^{-3}$ , n.s.,  $\beta_{Diagnose} = -0.07$ , n.s.,  $\beta_{Dauer} = 0.01$ , n.s.,  $R^2 = 0$ ;  $n = 17$ ). Eine Untersuchung der anderen Symptomkategorien wurde aufgrund der Studienanzahl nicht durchgeführt.

Der direkte Vergleich ergab keinen signifikanten Unterschied zwischen den Verfahren in den Kategorien Angst und Depression. In der Kategorie Klinische Signifikanz zeigten die vorliegenden Studien eine im FEM bedeutsam signifikante mittlere ES mit höherer Effektivität für kognitiv-behaviorale Verfahren. Allerdings war die Studienanzahl in dieser Kategorie sehr klein. Die Ergebnisse der Analysen sind in Tabelle 30 dargestellt. Bei Ausschluss der Studien, in denen zur Berechnung der ESn keine Mittelwerte und Standardabweichungen vorlagen, veränderte sich die mittlere ES kaum

<sup>40</sup> Einschließlich TSK, mit Ausnahme es wurden zusätzliche kognitive Elemente eingesetzt.

und betrug in der Kategorie Angst  $g = 0.05$  (95%-CI:  $-0.19 - 0.28$ ,  $\sigma_{\theta}^2 = 0$ ; FEM:  $g = 0.05$ , 95%-CI:  $-0.16 - 0.26$ ,  $\chi^2 = 6.07$ ,  $n = 13$ ). Die einzelnen ESn der Kategorien Angst und Depression sind in Abbildung 14 dargestellt.

Aufgrund der geringen Anzahl an Studien, die Fremdbeurteilungsverfahren einsetzten, wurde auf eine Überprüfung potentieller Unterschiede zwischen Fremd- und Selbst-rating verzichtet.

Tabelle 30: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  zu Post für den Vergleich von konfrontativen und kognitiv-behavioralen Verfahren

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
Post	Angst	REM	0.04	0.10	-0.17 – 0.25	0		15
		FEM	0.04	0.10	-0.15 – 0.23		6.11	15
	Vermeidung	REM	0.05	0.15	-0.33 – 0.44	0		6
		FEM	0.05	0.15	-0.24 – 0.35		0.90	6
	Kognition	REM	0.06	0.10	-0.18 – 0.29	0		12
		FEM	0.06	0.10	-0.15 – 0.26		8.37	12
	Depression	REM	0.28	0.28	-0.49 – 1.05	0.21		5
		FEM	0.28	0.18	-0.08 – 0.64		8.90 <sup>+</sup>	5
Klin. Sign.	REM	0.43	0.22	-0.27 – 1.13	0.03		4	
	FEM	0.42	0.20	0.03 – 0.81		4.05	4	

<sup>+</sup>p < .10

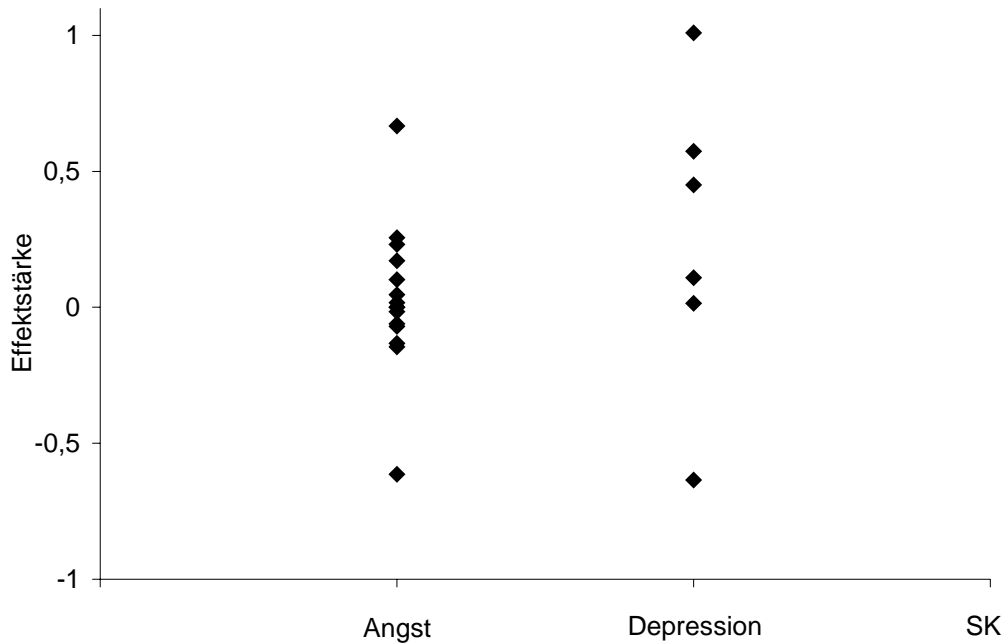


Abbildung 14: Effektstärken für die Symptomkategorien (SK) Angst und Depression der einzelnen Studien für den direkten Vergleich von konfrontativen und kognitiv-behavioralen Verfahren

### Fail Safe-N

Für eine Reduzierung der beobachteten Effekte des Vergleiches PsT zu WL auf eine kleine ES von  $g = 0.2$  sind 57 Gruppenvergleiche mit  $g = 0$  notwendig. Dies ist ca. die Hälfte der von Rosenthal vorgeschlagenen Grenze für stabile Ergebnisse. Das Fail-Safe N ist zumindest fast dreimal so hoch, wie die Anzahl an eingeschlossenen Vergleichen. Für einen Vergleich mit einer behandelten Kontrollgruppe werden nur 3 Gruppenvergleiche benötigt, sowie 14 Gruppenvergleiche mit einer Effektstärke von  $g = -0.2$  für einen mittleren Effekt von  $g = 0$ . Die geringe Stabilität der Ergebnisse wird deutlich: die Grenze nach Rosenthal beträgt 65 Studien. Für den direkten Vergleich von konfrontativen mit kognitiv-behavioralen Therapien wären 12 bzw. 8 Studien notwendig mit jeweils einer mittleren Effektstärke für eines der Verfahren, um einen geringen Effekt für konfrontative oder für kognitiv-behaviorale Behandlungen zu erhalten. Die Grenze liegt hier bei 85, der Faktor bezüglich der eingeschlossenen Vergleiche beträgt 0.8 bis 0.5, auch hier können die Ergebnisse noch nicht als stabil gelten.



### 3.6.4 Pharmakotherapie

#### Effektivität: Vergleich zur Placebobedingung

In ca. jeweils einem Drittel der Studien wurden SSRI oder Monoaminoxidase-Inhibitoren (MAOI; einschließlich reversibler und selektiver Inhibitoren der Monoaminoxidase-A, z.B. Moclobemid) untersucht. Die Studien-ESn der Hauptkategorien waren meist positiv, 11 % der ESn zeigten dagegen ein besseres Ergebnis für PP bis  $g = -0.33$ . In allen Symptomkategorien ergab sich zu Post im Vergleich zur KG eine signifikant positive mittlere ES. Die kleinste mittlere ES der Hauptkategorien zeigte sich bei Lebensqualität mit  $g = 0.36$ , die größte mittlere ES für Klinische Signifikanz mit  $g = 0.76$ . Es ergaben sich Unterschiede in der Unterkategorien der Angst: deutliche Verbesserungen zeigten sich für Erregung ( $g = 0.60$ ), geringere dagegen für Vermeidung und Kognition ( $g = 0.46$ ;  $g = 0.36$ ). Bei Ausschluss der Studien, die keine Berechnung der ES über Mittelwerte und Standardabweichungen ermöglichten, stieg die mittlere ES in der Kategorie Angst auf  $g = 0.54$  (95%-CI: 0.31 – 0.77,  $\sigma_{\theta}^2 = 0.16$ ; FEM:  $g = 0.48$ , 95%-CI: 0.38 – 0.58,  $\chi^2 = 56.34$ ,  $n = 21$ ).

In jeder der Symptomkategorien lag Heterogenität vor. Die Ergebnisse der Analysen sind in Tabelle 31 dargestellt. Einen Überblick über die einzelnen Studien-ESn für die Symptomkategorien Angst und Depression bietet Abbildung 15.

Zur differentiellen Wirksamkeit verschiedener Substanzklassen zeigte sich in der Symptomkategorie Angst für MAOI eine mittlere ES von  $g = 0.56$  und für SSRI von  $g = 0.39$ . Beide waren bedeutsam wirksamer als PP. Der Unterschied zwischen den Gruppen war in einer multiplen Regression mit Kontrolle von diagnostischem System, Analyseart, Dosierung, Placebo Run-in, Dropouts und Stichprobengröße nicht signifikant ( $-\beta = 0.06$ , *n.s.*,  $R^2 = .36$ ;  $n = 19$ ). Eine Berechnung in der Kategorie Depression entfiel aufgrund der geringen Zahl an Studien mit SSRI, in denen Daten für diese Symptomkategorie vorlagen.

Tabelle 31: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  zu Post für den Vergleich einer medikamentösen zur Placebobedingung

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
Post	Angst	REM	0.46	0.08	0.29 – 0.63	0.12 <sup>a</sup>		29
		FEM	0.40	0.04	0.32 – 0.48		65.58 <sup>+</sup>	29
	Vermeidung	REM	0.46	0.09	0.27 – 0.65	0.10		21
		FEM	0.40	0.05	0.31 – 0.50		49.95 <sup>+</sup>	21
	Kognition	REM	0.36	0.11	0.13 – 0.60	0.12		16
		FEM	0.33	0.05	0.23 – 0.43		35.23 <sup>+</sup>	16
	Erregung	REM	0.60	0.10	0.38 – 0.81	0.14		20
		FEM	0.49	0.05	0.39 – 0.58		61.44 <sup>+</sup>	20
	Depression	REM	0.40	0.11	0.16 – 0.64	0.16		19
		FEM	0.34	0.05	0.24 – 0.44		53.92 <sup>+</sup>	19
	Klin. Sign.	REM	0.76	0.09	0.58 – 0.95	0.10		22
		FEM	0.67	0.05	0.58 – 0.76		63.00 <sup>+</sup>	22
	Lebensqualität	REM	0.34	0.09	0.14 – 0.53	0.10		18
		FEM	0.30	0.04	0.21 – 0.39		32.61 <sup>+</sup>	18

<sup>a</sup>  $\sigma_{\theta}^2 > 0$  auch bei Ausschluss der Studie mit extremster ES definiert über die größte Abweichung vom ungewichteten Mittel

<sup>+</sup>  $p < .10$

### Zeitliche Stabilität der Therapie

In sechs Studien konnten ESn zu Follow up berechnet werden, meist handelte es sich um einen nur kurzen Zeitraum von ein oder zwei Monaten nach Beendigung der Hauptdosierung, eine Studie untersuchte dagegen ein 24-Monats-Follow up. Der Mittelwert betrug 5.7 Monate ( $SD = 8.22$ ). Es zeigte sich auch nach Absetzen der Medikation keine bedeutsame Zunahme der Symptomatik, die Effekte scheinen stabil ( $g = -0.12$ , 95%-CI:  $-0.45 - 0.21$ ,  $\sigma_{\theta}^2 = 0$ ; FEM:  $g = -0.12$ , 95%-CI:  $-0.38 - 0.14$ ,  $\chi^2 = 2.37$ ,  $n = 7$ ).

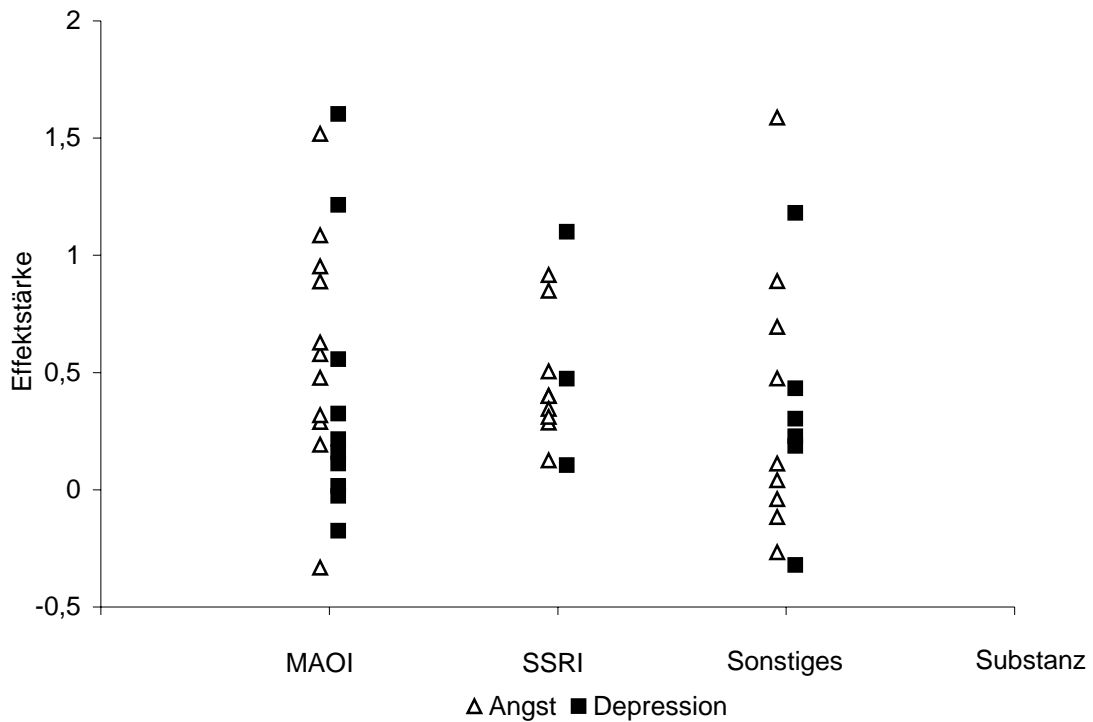


Abbildung 15: Effektstärken der einzelnen Studien für die Symptomkategorien Angst und Depression aufgeteilt nach Substanzklasse

### Akzeptanz

Durchschnittlich jeder Fünfte brach vorzeitig die Behandlung mit einer aktiven Substanz ab ( $M = 18.8$ ,  $SD = 12.03$ ,  $n = 28$ ). Für SSRI betrug die mittlere Abbrecherrate 25.4 % ( $SD = 10.54$ ,  $n = 8$ ), für MAOI dagegen nur 14.4 % ( $SD = 10.15$ ,  $n = 11$ ). Der Unterschied war signifikant ( $t(17) = 2.30$ ,  $p < .05$ )<sup>41</sup>. Die mittlere Dropout-Rate in der PP-Bedingung betrug 19.4 % ( $SD = 12.39$ ,  $n = 25$ ).

### Moderatoren

Keine der untersuchten Variablen zeigte einen bedeutsamen Einfluss auf die Effektstärke. Die Ergebnisse der multiplen Regression sind in Tabelle 32 dargestellt. Auch nach Kontrolle der untersuchten Variablen war weiterhin ungeklärte Varianz vorhanden ( $\sigma_{\theta}^2 = .13$ ). Es zeigte sich kein bedeutsamer Unterschied zwischen Selbst- und Fremdbeurteilung ( $\beta = -0.09$ , n.s.,  $R^2 = 0$ ;  $n = 54$ ).

<sup>41</sup> Normalverteilung gegeben

Tabelle 32: Ergebnisse der multiplen Regression zum Einfluss verschiedener Variablen auf die Effektstärke zwischen Pharmakotherapie und Pillenplacebo

Variable	B
Intercept	0.66
Diagnosestellung	-0.12
Dosierung	0.08
PP-Washout	-0.23
Dropouts	-0.01
Analyse	0.21
Stichprobengröße	$0.30 \times 10^{-3}$
SSRI	0.14
MAOI	0.08

$R^2 = .00, n = 28$

### Fail Safe-N

Für eine Reduzierung der beobachteten Effekte auf eine kleine ES von  $g = 0.2$  sind 38 Gruppenvergleiche mit  $g = 0$  notwendig, ca. ein Viertel der von Rosenthal vorgeschlagenen Grenze für stabile Ergebnisse. Das Fail Safe-N ist etwas höher als die Anzahl an eingeschlossenen Effektstärken.

### 3.6.5 Vergleich Psycho- und Pharmakotherapie

Es zeigte sich ein mittlerer Unterschied in den Prä-Post-ES von  $g = 0.45$  mit höheren ESn für PP im Vergleich zu WL und von  $g = -0.12$  mit niedrigeren ESn für PP im Vergleich zu TP. Nach Korrektur der Post-ESn um diese Werte ergab sich kein bedeutsamer Unterschied zwischen den Verfahren ( $\beta = 0.09$ , n.s.,  $R^2 = .10$ ;  $n = 60$ ). Die multiple Regression mit Kontrolle von Abbrecherrate, Stichprobengröße und Diagnosesystem ergab dagegen einen tendenziell bedeutsamen Unterschied zwischen den Verfahren mit leicht höheren ESn für die medikamentöse Therapie sowie niedrigeren ESn bei einer höheren Abbrecherrate ( $\beta_{\text{PhT:PsT}} = -0.18$ ,  $p < .10$ ,  $\beta_{\text{Dropout}} = -0.01$ ,  $p < .05$ ,  $\beta_{\text{N}} = -0.65 \times 10^{-3}$ , n.s.,  $\beta_{\text{Diagnose}} = 0.05$ , n.s.,  $R^2 = .03$ ;  $n = 56$ ). Bei Ausschluss weniger etablierter Behandlungen (PsT: nicht konfrontativ einschließlich TSK oder kognitiv-behavioral, PhT: nicht MAOI oder SSRI) ergab sich ein bedeutsamer Unterschied zwischen PsT und PhT; die multiple Regression zeigte zudem signifikant geringere ESn

bei hoher Abbrecherrate und tendenziell geringere ESn bei einer höheren Stichprobengröße ( $\beta_{\text{PhT:PsT}} = -0.20$ ,  $p < .05$ ,  $\beta_{\text{Dropout}} = -0.01$ ,  $p < .01$ ,  $\beta_{\text{N}} = -0.81 \times 10^{-3}$ ,  $p < .10$ ,  $\beta_{\text{Diagnose}} = 0.10$ , n.s.,  $R^2 = .77$ ;  $n = 46$ ). Eine Berechnung für die Kategorie Depression wurde nicht durchgeführt aufgrund der geringen Anzahl an Studien, die eine Berechnung der Prä-Post-ESn für WL und TP ermöglichten (jeweils  $n = 4$ ).

Sieben Studien untersuchten beide Verfahren direkt miteinander. Es wurden dabei meist kognitiv-behaviorale Therapien betrachtet, die Substanzklassen waren dagegen sehr gemischt. Es lagen sowohl Studien vor, in denen PhT eine höhere Effektivität zeigte, als auch Studien, in den PsT wirksamer war. Es zeigte sich in keiner der Symptomkategorien ein bedeutsamer Unterschied zwischen den Verfahren, die mittleren ESn lagen nahe Null. Es ist zu beachten, dass in der Mehrzahl an Kategorien Heterogenität vorliegt, aufgrund der geringen Studienanzahl, auch hinsichtlich der sehr unterschiedlichen Substanzklassen, wird aber auf eine Aufklärung der RE-Varianz verzichtet. Bei Ausschluss der Studien, in denen zur Berechnung der ESn keine Mittelwerte und Standardabweichungen vorlagen, stieg die mittlere ES leicht und betrug in der Kategorie Angst  $g = 0.09$  (95%-CI:  $-0.40 - 0.57$ ,  $\sigma_{\theta}^2 = 0.03$ ; FEM:  $g = 0.09$ , 95%-CI:  $-0.21 - 0.40$ ,  $\chi^2 = 5.02$ ,  $n = 5$ ). Die Ergebnisse der Analysen sind in Tabelle 33 dargestellt. Auf eine graphische Darstellung wird aufgrund der geringen Anzahl an Studien-ESn verzichtet.

Die Art der Erhebung zeigte keinen bedeutsamen Einfluss auf die ESn ( $\beta = 0.07$ , n.s.,  $R^2 = .00$ ;  $n = 15$ ).

Im Follow up nach durchschnittlich 2.6 Monaten ( $SD = 1.95$ ) zeigte sich ebenfalls kein bedeutsamer Unterschied zwischen den Verfahren (REM:  $g = 0.19$ , 95%-CI:  $-0.32 - 0.71$ ,  $\sigma_{\theta}^2 = 0.03$ , FEM:  $g = 0.20$ , 95%-CI:  $-0.13 - 0.52$ ,  $\chi^2 = 4.48$ ,  $n = 5$ ).

Tabelle 33: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  zu Post für den Vergleich von Pharmako- zu Psychotherapie

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
Post	Angst	REM	-0.06	0.14	-0.39 – 0.27	0.06 <sup>a</sup>		8
		FEM	-0.06	0.10	-0.25 – 0.13		10.91	8
	Vermeidung	REM	-0.02	0.20	-0.58 – 0.55	0.14		5
		FEM	0.01	0.10	-0.20 – 0.21		10.64 <sup>+</sup>	5
	Kognition	REM	-0.01	0.27	-0.70 – 0.68	0.33		6
		FEM	-0.09	0.10	-0.29 – 0.11		10.67 <sup>+</sup>	6
	Lebensqualität	REM	-0.01	0.12	-0.38 – 0.36	0		4
		FEM	-0.01	0.12	-0.24 – 0.22		1.94	4
	Klin. Sign.	REM	-0.12	0.28	-0.89 – 0.65	0.27		5
		FEM	-0.05	0.15	-0.34 – 0.23		11.83 <sup>+</sup>	5

<sup>a</sup>  $\sigma_{\theta}^2 > 0$  auch bei Ausschluss der Studie mit extremster ES definiert über die größte Abweichung vom ungewichteten Mittel

<sup>+</sup>  $p < .10$

Es zeigte sich über alle Studien kein bedeutsamer Unterschied zwischen Pharmako- und Psychotherapie in der Anzahl an Abbrechern ( $t(49.9) = 1.30$ , n.s.). Beide Verfahren scheinen somit von den Patienten gleich gut akzeptiert zu werden.

Drei Studien untersuchten den Vergleich von PsT mit einer Kombination aus PsT+PhT. Es zeigten sich dabei ESn für die Kategorie Angst von  $g = 0.12$  (Kombination mit Sertraline),  $g = 0$  (Kombination mit Propanolol) bzw.  $g = -0.16$  (Kombination mit Bupirone). Weitere Untersuchungen sind somit notwendig.

### Fail Safe-N

Hinsichtlich des direkten Vergleiches von Psycho- und Pharmakotherapie wären nur jeweils zwei Studien notwendig mit einer jeweils mittelhohen Effektstärke von  $g = 0.50$  für eines der beiden Verfahren für einen mittleren geringen Effekt von  $g = 0.20$ . Die Ergebnisse sind somit äußerst instabil, weitere Untersuchungen werden dringend benötigt.

### 3.6.6 Diskussion

Die Ergebnisse der Meta-Analyse können wie folgt zusammengefasst werden. Erstens zeigen die hier untersuchten Psychotherapieverfahren, insbesondere Exposition mit und ohne kognitive Elemente sowie ein Training Sozialer Kompetenz, eine mittlere bis hohe Effektivität in der Behandlung der SoP. Patienten der Therapiegruppen zeigen am Behandlungsende eine bedeutsam geringere Angstsymptomatik als unbehandelte Patienten. Aufgrund der gefundenen Homogenität ist nicht von bedeutsamen Unterschieden zwischen den Studien auszugehen. Auch für depressive Symptome zeigte sich – wenn auch im geringeren Maße – eine Effektivität der Psychotherapien, allerdings sind aufgrund der geringen Studienzahl die Resultate nur eingeschränkt generalisierbar. Der Einsatz von Instrumenten zur Erfassung dieses Aspektes, aber auch zur Erfassung der Lebensqualität sollte von künftigen Studien gefordert werden. Für letzteres kann keine Aussage hinsichtlich der Effektivität der Therapie getroffen werden, da nur mangelnde Daten vorlagen.

Auch für den Vergleich mit einer behandelten Kontrollgruppe (Pillen- oder Therapie-placebo) zeigten die untersuchten Psychotherapien, meist kognitiv-behaviorale Verfahren, eine bedeutsam geringere Angstsymptomatik nach der Behandlung. Es zeigte sich in den vorliegenden Untersuchungen, wie anhand des Rational der KVT zu erwarten, die Wirksamkeit dabei nicht im Bereich der angstbezogenen Kognitionen, sondern für Erregung. Veränderungen in Kognitionen scheinen in den Primärstudien somit auch ohne spezielle Techniken der kognitiven Verhaltenstherapie aufgetreten zu sein. Hinsichtlich Lebensqualität und Klinischer Signifikanz ergaben die Untersuchungen bedeutsame Effekte, nicht aber für Depression. Hier ergab sich kein Unterschied zwischen den Bedingungen. Insgesamt sind aber die Möglichkeiten zur Generalisierung der Ergebnisse aufgrund der Studienzahl eingeschränkt. Für die Angstsymptomatik, mit der größten Anzahl an eingeschlossenen Untersuchungen, sprechen die Ergebnisse für eine Wirksamkeit, die über gemeinsame Faktoren hinausgeht. Unter der Annahme einer additiven Zusammensetzung von Therapieeffekten mit einer Realisierung gemeinsamer Faktoren in behandelten Kontrollgruppen, nicht aber in Wartelisten, war allerdings nur ein Drittel der Effektivität auf spezifische Faktoren zurückzuführen.

---

Die positiven Effekte der untersuchten Verfahren blieben auch über die Zeit stabil. Es zeigte sich keine bedeutsame Verschlechterung in der Symptomatik.

Hinsichtlich der differentiellen Wirksamkeit von konfrontativen und kognitiv-behavioralen Verfahren weisen die Ergebnisse der Meta-Analyse auf eine gleiche Effektivität der Therapien hin. Weder der Vergleich über verschiedene Studien noch der direkte Vergleich innerhalb von Studien zeigte bedeutsame Unterschiede. Im letzteren liegen Hinweise vor, die allerdings aufgrund der geringen Anzahl an Untersuchungen weitere Forschung benötigen, dass eine Differenzierung der Verfahren vielleicht über die Klinische Signifikanz, mit höheren Effektstärken für KVT, möglich ist. Auch für diese Studien zeigte sich, dass eine Änderung in Kognitionen keine spezifischen kognitiven Techniken zu benötigen scheint. Die Akzeptanz der Patienten scheint für beide Therapien vergleichbar.

Auch die untersuchten medikamentösen Therapien zeigten über alle Symptomdimensionen eine bedeutsame Effektivität in der Behandlung der Sozialen Phobie. Die Effektstärken lagen dabei im kleinen bis mittleren Bereich. Es wurden insbesondere SSRIs und MAOIs untersucht. Die Ergebnisse der Meta-Analyse sprechen dafür, dass sich beide nicht in ihrer Wirksamkeit für die Angstsymptomatik zu unterscheiden scheinen, für MAOIs aber, definiert über die Abbrecherrate, eine höhere Akzeptanz besteht. Die gefundenen starken Unterschiede zwischen den Studien konnten durch keine der hier untersuchten Variablen aufgeklärt werden. Inwieweit Eigenschaften der Patienten einen moderierenden Einfluss haben, sollte in weiteren Untersuchungen speziell überprüft werden. Ein möglicher Aspekt könnte die Unterteilung in SoP des generalisierten oder spezifischen Typus sein, häufig erfolgte keine genaue Angabe in den bisherigen Studien. Betrachtet man die gefundenen Differenzen in den Prä-Post-ESn von WL und PP ergibt sich für die Pharmakotherapie der Sozialen Phobie eine annähernde Gleichverteilung spezifischer und gemeinsamer Faktoren für die Wirksamkeit. Die wenigen Daten mit Aussagen zur Stabilität der Ergebnisse nach Absetzen der Medikation sprechen in den eingeschlossenen Studien gegen eine bedeutsame Verschlechterung, Generalisierungen erscheinen aber verfrüht.



Vergleicht man die Effektivität von Psycho- und Pharmakotherapie, ergaben sich divergente Befunde in Abhängigkeit von der betrachteten Studienpopulation. Im indirekten Vergleich zeigte sich nach Kontrolle methodischer Variablen eine leicht höhere Effektstärke für Pharmakotherapie für die Angstsymptomatik. Studien, die beide Verfahren überprüften, ergaben dagegen keinen bedeutsamen Unterschied in der Effektivität. Allerdings war in einigen der Symptomkategorien Heterogenität gegeben, es scheint somit bestimmte Bedingungen zu geben, in denen Psychotherapie angemessener ist und andere, in denen Pharmakotherapie die Methode der Wahl darstellt. Welche Bedingungen allerdings bedeutsam sind, seien es Eigenschaften der Patienten oder der Methode konnte aufgrund der geringen Anzahl an Studien nicht geklärt werden. Zudem können schon sehr wenige weitere Studien die Ergebnisse substantziell verändern. Eine generelle Empfehlung für eine der Behandlungsarten erscheint aufgrund der vorliegenden Datenlage nur schwer möglich. Weitere Untersuchungen sind somit wünschenswert. Anmerkungen hinsichtlich des Aspektes, welche Analyse bei divergenten Ergebnissen stärker zu gewichten sei, finden sich im Diskussionsteil der PDA.

Die Akzeptanz war für beide Therapien vergleichbar.

Auch bisherige Meta-Analysen zeigten eine bedeutsame Effektivität von behavioralen, kognitiv-behavioralen und pharmakologischen Therapien mit einer zeitlichen Stabilität der Ergebnisse für Psychotherapien. Die bei Federoff et al. (2001) gefundene höhere Wirksamkeit von BZD konnte in der vorliegenden Arbeit nicht überprüft werden aufgrund der geringen Anzahl an kontrollierten Studien, die BZD untersuchten.

## 3.7 Spezifische Phobie

### 3.7.1 Deskriptive Analyse der Primärstudien

Es konnten insgesamt 29 Studien gefunden werden, die den Inklusionskriterien entsprachen. 23 Studien untersuchten Psychotherapie, in 5 Studien die Gabe eines Medikamentes bei Konfrontation und in einer Studien eine Langzeitpharmakotherapie (diese wurde in Tabelle 34 der Kategorie PsT/PhT zugeordnet). Ein Drittel der eingeschlossenen Studien untersuchte Zahnarztphobiker, gefolgt von Tier-, Höhen- und Flugphobikern. Die in epidemiologische Studien gefundenen Geschlechterunterschiede hinsichtlich eines häufigeren Auftretens bei Frauen spiegeln sich auch in der Zusammensetzung der Stichprobe in den vorliegenden Studien wider. Viele der eingeschlossenen Studien sind älteren Publikationsdatums. In keiner der Studien war angegeben, dass mindestens 25 % der Patienten eine komorbide Störung haben. Die durchschnittliche Anzahl an Patienten ist eher gering. Weitere deskriptive Daten sind in Tabelle 34 dargestellt.

Tabelle 34: Deskriptive Beschreibung der Primärstudien

	PsT	PhT / PsT
Patienten N <sup>a</sup>	889	184
Durchschnittliche Gruppengröße <sup>a</sup>	13.5	11.7
n < 20 (%)	91.3	100
n > 20 (%)	8.7	0
Alter	33.6	38.8
Dauer der Störung (Jahre)	15.6	9.7
Frauen (%)	75.8	77.6
Dauer der Therapie (Stunden)	6.4	1.4
Publikationsjahr		
bis 1985 (%)	60.9	16.7
1986 – 1995 (%)	21.7	33.3
ab 1996 (%)	17.4	50.0
Intent-to-treat Auswertung (%)	0	0

<sup>a</sup> Zahlen geben die in die Auswertung aufgenommenen Patienten wider

### 3.7.2 Vergleich der Prä-Werte

Es zeigte sich kein signifikanter Unterschied zwischen PsT und WL in der Symptomkategorie Angst, die mittlere ES lag nahe Null ( $g = -0.03$ , 95%-CI:  $-0.22 - 0.15$ ,  $\chi^2 = 8.61$ ,  $n = 17$ ).

### 3.7.3 Psychotherapie

#### Effektivität: Vergleich zur Warteliste

Die am häufigsten untersuchten Verfahren mit jeweils einem Drittel der Gruppen waren Systematische Desensibilisierung und Exposition. Ca. ein Zehntel der Vergleiche beziehen sich auf computergestützte Therapien einschließlich virtueller Realität. Alle ESn der Primärstudien in den Hauptkategorien waren positiv, d.h. keine Studie zeigte ein schlechteres Ergebnis der BG im Vergleich zur unbehandelten KG. Aufgrund der vorliegenden Daten werden nur die Ergebnisse für die Hauptkategorie Angst und die Unterkategorie Vermeidung präsentiert. Es zeigte sich zu post eine signifikant positive mittlere ES im Vergleich zur WL sowohl in der Kategorie Angst, als auch in der Unterkategorie Vermeidungsverhalten.

Bei Ausschluss der Studien, in denen zur Berechnung der ESn keine Mittelwerte und Standardabweichungen vorlagen, stieg die mittlere ES leicht an und betrug in der Kategorie Angst  $g = 1.22$  (95%-CI:  $0.93 - 1.51$ ,  $\sigma_{\theta}^2 = 0.20$ ; FEM:  $g = 1.36$ , 95%-CI:  $1.18 - 1.55$ ,  $\chi^2 = 61.23$ ,  $n = 21$ ).

Tabelle 35: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  zu Post für den Vergleich von behavioralen und kognitiv-behavioralen Verfahren zur Warteliste

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
Post	Angst	REM	1.08	0.10	0.87 – 1.29	0.15 <sup>a</sup>		33
		FEM	1.12	0.07	0.97 – 1.26		93.78 <sup>+</sup>	33
	Vermeidung	REM	1.01	0.20	0.59 – 1.43	0.37		15
		FEM	0.91	0.11	0.69 – 1.13		39.73 <sup>+</sup>	15

<sup>a</sup>  $\sigma_{\theta}^2 = 0$  nach Ausschluss eines Extremwertes definiert als größte mittlere Abweichung vom ungewichteten Mittel

<sup>+</sup>  $p < .10$

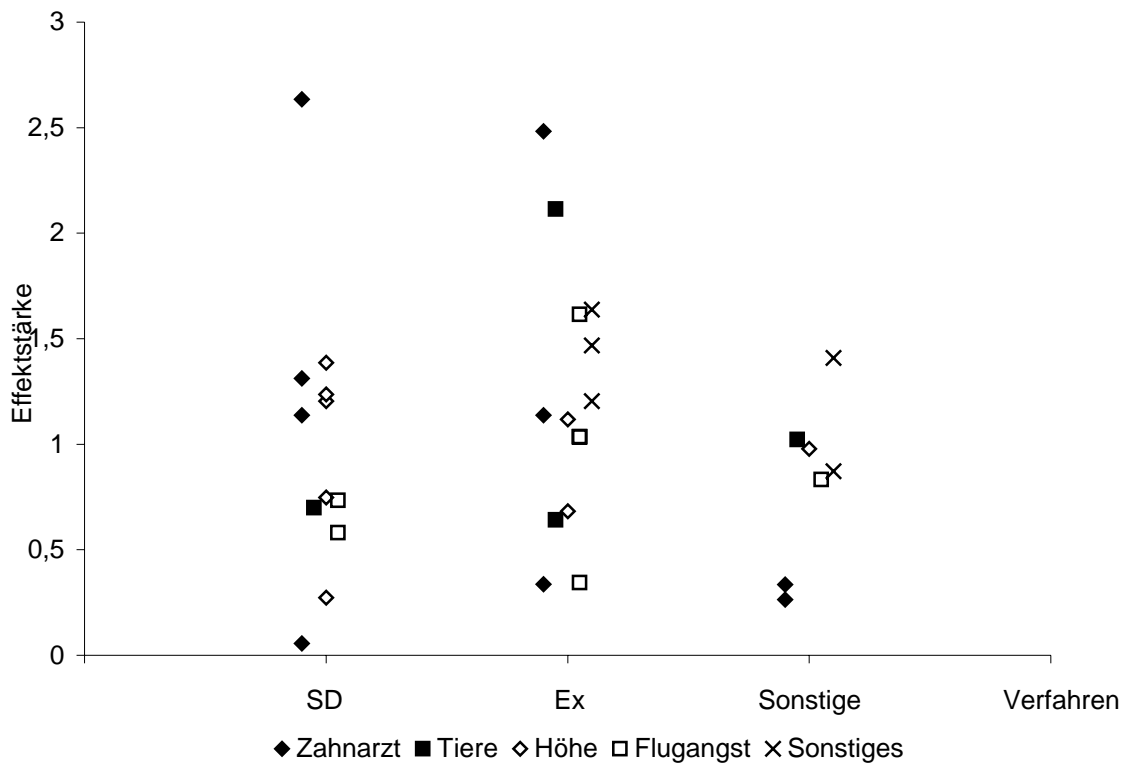


Abbildung 16: Studieneffektstärken für die Symptomkategorie Angst für den Vergleich von Psychotherapie zur Warteliste aufgeteilt nach Verfahren und gefürchtetem Stimulus

### Effektivität: Vergleich zu einer Placebobedingung

Einen Vergleich zu TP ermöglichten sechs Gruppenvergleiche aus drei Studien, die Interpretation der Befunde ist somit stark eingeschränkt. Die mittlere ES war entsprechend den Erwartungen niedriger als im Vergleich zur WL und betrug  $g = 0.32$  (95%-CI:  $-0.19 - 0.83$ ,  $\sigma_{\theta}^2 = 0$ ; FEM:  $g = 0.32$ , 95%-CI:  $-0.07 - 0.70$ ,  $\chi^2 = 1.19$ ,  $n = 6$ ).

### Zeitliche Stabilität der Therapie

Weder für die Messungen zu f1 ( $M = 1.3$  Monate,  $SD = 0.52$ ) noch zu f2 ( $M = 1.3$  Monate,  $SD = 0.52$ ) ergab sich eine signifikante Zu- oder Abnahme der Effektstärke ( $g = 0.09$ , 95%-CI:  $-0.32 - 0.55$ ,  $\sigma_{\theta}^2 = 0$ ; FEM:  $g = 0.09$ , 95%-CI:  $-0.22 - 0.41$ ,  $\chi^2 = 1.52$ ,  $n = 6$ ;  $g = 0.29$ , 95%-CI:  $-0.11 - 0.70$ ,  $\sigma_{\theta}^2 = 0$ ; FEM:  $g = 0.29$ , 95%-CI:  $-0.03 - 0.62$ ,  $\chi^2 = 2.21$ ,  $n = 7$ ). Die Behandlungsergebnisse sind somit über die Zeit stabil.

**Akzeptanz**

Es zeigten sich im Besonderen bei dieser Störung Schwierigkeiten in einer genauen Bestimmung der Dropout-Rate. In einer Vielzahl an Studien wurde nicht explizit mitgeteilt, ob Patienten die Behandlung vorzeitig abbrachen. Es wird deshalb im Folgenden nur eine approximative mittlere Abbrecherquote angegeben, die sich ergibt aus Studien, die mehr als eine Therapiesitzung durchführten und die genaue Dropout-Rate angaben. Im Schnitt brach ca. jeder Zehnte die Behandlung vorzeitig ab ( $M = 9.6$ ,  $SD = 7.67$ ,  $n = 18$ ). Eine sehr geringe Dropout-Rate ergab sich für die unbehandelte Kontrollgruppe: Patienten der Wartelistenbedingung beendeten nur zu 3.6 % die Studie vorzeitig ( $SD = 6.62$ ,  $n = 7$ ).

**Moderatoren**

Aufgrund der RE-Varianz von 0 im Rahmen der Sensitivitätsanalyse entfiel eine Berechnung der Regressionen. Auch bzgl. des Ratings wurde keine Analyse durchgeführt, da kaum Fremdbeurteilungen eingesetzt wurden.

**Vergleich behaviorale und kognitiv-behaviorale Verfahren**

Vier Studien erlaubten einen direkten Vergleich zwischen VT und KVT. Es zeigte sich in keiner der Studien ein Vorteil der kognitiv-behavioralen Verfahren. Die mittlere ES betrug  $g = -0.20$ , diese war nicht signifikant von Null verschieden (95%-CI:  $-0.90 - 0.51$ ,  $\sigma_{\theta}^2 = 0$ ; FEM:  $g = -0.20$ , 95%-CI:  $-0.63 - 0.23$ ,  $\chi^2 = 1.24$ ;  $n = 4$ ).

**Fail-Safe N**

Um die beobachteten Effekte im Vergleich von PsT zu WL auf einen kleinen Effekte von  $g = 0.2$  zu reduzieren, würden 146 Gruppenvergleiche mit  $g = 0$  benötigt. Dies übersteigt nicht die von Rosenthal vorgeschlagene Grenze, erscheint aber ausreichend hoch, um von einem relativ stabilen Effekt sprechen zu können. Wie aufbauend auf der Anzahl der vorliegenden Studien zu erwarten, sind für den direkten Vergleich von VT und KVT nur wenige Studien notwendig, um ein vollständig anderes Bild zu erhalten. So werden nur vier Untersuchungen mit einer ES von  $g = -0.8$  benötigt, um einen

mittleren Effekt von  $g = -0.5$  zugunsten VT zu erhalten, demgegenüber stehen 10 Studien mit einer ES von  $g = 0.8$  für einen mittleren Effekt von  $g = 0.5$  zugunsten KVT.

#### **3.7.4 Pharmakotherapie**

Sechs Studien untersuchten die Wirksamkeit eines Medikamentes bei spezifischer Phobie. Es wurde auf die Berechnung einer Effektstärke aufgrund der unterschiedlichen Designs verzichtet: fünf Studien untersuchten die Effektivität während bzw. kurz nach einer einmaligen Exposition unter Wirkung der Medikation. Drei dieser Studien betrachteten zusätzlich einem späteren Zeitpunkt ohne Medikation. In den fünf Studien zeigten sich positive Effekte unter Wirkung des Medikamentes während bzw. kurz nach Exposition im Vergleich zu Exposition mit PP bzw. zur WL. Bei Betrachtung des zweiten Messzeitpunktes, d.h. bei erneuter Konfrontation mit dem phobischen Stimulus ohne Medikation, zeigten sich für drei Studien im Vergleich zur PP-Bedingung bzw. reiner Psychotherapie negative Effekte bzw. keine bedeutsamen Unterschiede. Eine Studie untersuchte die Langzeitwirkung von Paroxetin, die Messung erfolgte im Gegensatz zu den anderen Studien nicht während oder kurz nach Exposition. Es zeigte sich hier nach einem Monat noch unter Wirkung des Medikamentes eine signifikant positive ES im Vergleich zu PP.

### 3.7.5 Diskussion

Die Ergebnisse der Meta-Analyse zeigen, dass die untersuchten behavioralen und kognitiv-behavioralen Verfahren, im Besonderen Systematische Desensibilisierung und Exposition, sehr effektiv in der Behandlung der spezifischen Phobie sind. Dies war unabhängig vom gefürchteten Stimulus. Die hohe Effektivität kann dabei schon mit einer relativ geringen Therapiedauer erreicht werden. Zusätzliche kognitive Elemente scheinen für die spezifische Phobie zu keiner weiteren Steigerung der Wirksamkeit der Therapie zu führen, mehr Studien mit einem direkten Vergleich sind aber notwendig für generalisierbare Aussagen. Die Akzeptanz der Therapie kann als sehr hoch eingeschätzt werden. Zudem sind die gefundenen positiven Effekte der Behandlung stabil über die Zeit. Dieses Ergebnis konnte auch in der Arbeit von Ruhmland (2001a) gefunden werden, allerdings ohne die Möglichkeit der Generalisierbarkeit der Ergebnisse. Ob die untersuchten Therapien auch im Vergleich zu einem Therapieplacebo bedeutsame Vorteile zeigen, ist ein zu überprüfendes Forschungsdesiderat. Bisherige Studien weisen den Erwartungen entsprechend auf eine im Vergleich zu den Effekten aus dem Wartelistendesign geringeren Überlegenheit der behavioralen bzw. kognitiv-behavioralen Therapie hin.

Eine zusätzliche Medikation scheint aufbauend auf den vorliegenden Befunden nicht Erfolg versprechend. Bei erneuter Konfrontation mit dem Stimulus ohne aktive Substanz zeigte keine der Studien positive Effekte. Beachtet man zudem mögliche Nebenwirkungen erscheint Pharmakotherapie bei spezifischer Phobie nicht angemessen. Die bisherigen Studien untersuchten hierbei meist eine einmalige Dosis.

Hinsichtlich des Studiendesigns gab es Besonderheiten für diese Angststörung. Nur selten wurden Maße außerhalb der Angstsymptomatik erhoben. Dies könnte im Zusammenhang stehen mit einer eher seltenen Komorbidität mit Störungen außerhalb des Angstspektrums, wenn spezifische Phobie die Hauptsymptomatik ist (Antony et al., 2002). Hinsichtlich der Art der Erhebung wurden überwiegend Selbstbeurteilungsinstrumente eingesetzt. Nur in 40 % der Studien erfolgte eine Diagnose nach einem Diagnostischen System. Dies ist zu verstehen in einem relativ frühen Publikationsdatum vieler Studien, die aufgrund der genauen Benennung der Störung trotzdem in die Auswertung eingeschlossen werden konnten. Im Zusammenhang mit dem Alter der

Studien steht vermutlich auch die nur geringe Integritätsüberprüfung der durchgeführten Therapien. Nur in 17 % der Psychotherapiestudien wurde die Durchführung überwacht.



## 3.8 Zwangsstörung

### 3.8.1 Deskriptive Analyse der Primärstudien

Es konnten insgesamt 41 Studien gefunden werden, die den Inklusionskriterien entsprachen: 20 Pharmako- und 9 Psychotherapiestudien. In weiteren 12 Studien wurden beide Verfahren zusammen untersucht meist mit einer Kombinationsbehandlung. In Pharmakotherapiestudien wurden fast zehnmal so viele Patienten untersucht wie in Psychotherapiestudien. Der Frauenanteil erscheint mit den gefundenen epidemiologischen Daten vergleichbar. Ca. ein Fünftel der Studien gab an, dass mindestens 25 % der Patienten eine komorbide Störung, meist Depression, aufwiesen. Weitere deskriptive Daten finden sich in Tabelle 36.

Tabelle 36: Deskriptive Beschreibung der Primärstudien

	PsT	PhT	PhT / PsT
Patienten Gesamt <sup>a</sup>	287	2820	482
Durchschnittliche Gruppengröße <sup>a</sup>	15.9	44.7	16.0
n ≤ 20 (%)	66.7	55.0	66.7
n > 20 (%)	33.3	45.0	33.3
Alter	34.2	36.6	35.1
Dauer der Störung (Jahre)	10.1	13.4	13.3
Frauen (%)	57.5	51.3	54.5
Dauer der Therapie (PsT: Stunden, PhT: Wochen)	23.1	9.6	PsT: 25.9 PhT: 17.1
Publikationsjahr			
bis 1985 (%)	11.1	10.0	8.3
1986 – 1995 (%)	44.4	55.0	58.4
ab 1996 (%)	44.4	35.0	33.3
Intent-to-treat Auswertung (%)	11.1	45.0	8.3
Placebowashout (%)		60.0	16.7
Feste Dosierung (%)		45.0	8.3

<sup>a</sup> Zahlen geben die in die Auswertung aufgenommenen Patienten wider

### 3.8.2 Vergleich der Prä-Werte

Es zeigte sich im Vergleich von PsT zur WL kein signifikanter Unterschied in der Kategorie Angst, die mittlere ES war nahe Null ( $g < 0.01$ , 95%-CI: -0.27 – 0.27,  $\chi^2 = 3.07$ ,  $n = 6$ ).

Für Studien, die einen direkten Vergleich zwischen VT und KVT untersuchten, ergab sich kein bedeutsamer Unterschied zwischen den Gruppen, die mittleren ESn lagen in den Kategorien Angst und Depression bei Null ( $g = 0.05$ , 95%-CI: -0.18 – 0.28,  $\chi^2 = 0.87$ ,  $n = 7$ ;  $g = 0.01$ , 95%-CI: -0.23 – 0.25,  $\chi^2 = 5.77$ ,  $n = 6$ ).

Es zeigte sich für PhT im Vergleich zu PP weder in der Symptomkategorie Angst noch Depression ein signifikanter Unterschied zwischen den Gruppen, die mittlere ES lag bei Null ( $g = 0.06$ , 95%-CI: -0.04 – 0.16,  $\chi^2 = 4.49$ ,  $n = 19$ ;  $g = -0.06$ , 95%-CI: -0.22 – 0.10,  $\chi^2 = 6.10$ ,  $n = 9$ ).

Für Studien mit einem Vergleich von PsT zu einer Kombinationsbehandlung ergaben die Analysen keinen bedeutsamen Unterschied zu Prä, in der Kategorie Depression zeigte die reine Psychotherapie eine leicht geringere Symptomatik ( $g = 0.05$ , 95%-CI: -0.19 – 0.28,  $\chi^2 = 2.46$ ,  $n = 8$ ;  $g = -0.17$ , 95%-CI: -0.42 – 0.08,  $\chi^2 = 5.69$ ,  $n = 7$ ).

### 3.8.3 Psychotherapie

#### Effektivität: Vergleich zur Wartelistenbedingung

Insgesamt ist die geringe Anzahl an Studien mit Kontrollgruppendesign auffallend. Zwei Drittel der Vergleiche zur Warteliste beinhalteten Exposition mit Reaktionsverhinderung. Es ergab sich für die Kategorie Angst eine signifikant positive mittlere Effektstärke im Vergleich zur WL mit  $g = 0.94$  (95%-CI: 0.56 – 1.32,  $\sigma_{\theta}^2 = 0$ ; FEM:  $g = 0.94$ , 95%-CI: 0.65 – 1.23,  $\chi^2 = 5.28$ ,  $n = 6$ ). Für andere Symptomkategorien sowie hinsichtlich eines Vergleiches zur Placebobedingung konnte aufgrund mangelnder Daten keine Effektstärke berechnet werden. Alle Effektstärken wurden über Mittelwerte und Standardabweichungen berechnet. Auf eine graphische Darstellung der Effektstärken wird verzichtet.

### Zeitliche Stabilität der Therapie

Patienten zeigten auch im Mittel 4 Monate nach Beendigung der Therapie ( $SD = 2.27$ ) keine Veränderung in ihrer Symptomatik, die positive Effekte der Psychotherapie scheinen somit stabil. Die Ergebnisse der Analysen sind in Tabelle 37 dargestellt.

Tabelle 37: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  der Veränderung von Post zu Follow up (f) für behaviorale und kognitiv-behaviorale Verfahren

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
f	Angst	REM	0.05	0.13	-0.24 – 0.35	0		8
		FEM	0.05	0.13	-0.19 – 0.30		1.04	8
	Depression	REM	0.08	0.13	-0.22 – 0.37	0		8
		FEM	0.08	0.13	-0.17 – 0.32		1.90	8
	Klin.Sign.	REM	0.01	0.17	-0.39 – 0.40	0		5
		FEM	0.01	0.12	-0.28 – 0.28		0.91	5

### Akzeptanz

Patienten der Behandlungsgruppen brachen zu 16.9 % vorzeitig die Studie ab ( $SD = 9.70$ ,  $n = 25$ ), d.h. jeder Sechste verließ die Untersuchung. Es zeigte sich kein bedeutsamer Unterschied zwischen behavioralen und kognitiv-behavioralen Verfahren (VT:  $M = 17.0$ ,  $SD = 10.87$ ,  $n = 14$ , KVT:  $M = 15.8$ ,  $SD = 8.33$ ,  $n = 10$ ,  $t(22) = 0.29$ , n.s.<sup>42</sup>). Die mittlere Dropout-Rate in der WL-Bedingung betrug 7.0 % ( $SD = 6.08$ ,  $n = 3$ ).

### Moderatoren

Aufgrund der RE-Varianz von Null sowie der insgesamt sehr geringen Anzahl an Studien wurde auf eine Berechnung der Regressionen einschließlich der Überprüfung des Einflusses der Erhebungsart verzichtet.

<sup>42</sup> Normalverteilung gegeben

### Vergleich konfrontative und kognitiv-behaviorale Verfahren

Der Vergleich über die WL-ESn entfiel aufgrund der geringen Anzahl an Studien.

Im direkten Vergleich waren 72 % der Studien-ESn über alle Symptomkategorien positiv, d.h. sie zeigten ein besseres Ergebnis der kognitiv-behavioralen Verfahren. Die mittlere ES war in keiner der Kategorien bedeutsam von Null verschieden. Beide Verfahren scheinen somit gleich effektiv in der Behandlung von Patienten mit OCD zu sein, auch hier ist wieder die geringe Anzahl an Studien zu beachten. Schloss man die Studien aus, in denen keine Berechnung über Mittelwerte und Standardabweichungen gegeben war, veränderte sich die mittlere ES kaum ( $g = 0.12$ , 95%-CI:  $-0.30 - 0.54$ ,  $\sigma_{\theta}^2 = 0$ ; FEM:  $g = 0.12$ , 95%-CI:  $-0.18 - 0.42$ ,  $\chi^2 = 2.89$ ,  $n = 5$ ). Auch im Follow up nach durchschnittlich 5 Monaten ( $SD = 3.08$ ) waren beide Verfahren gleich effektiv. Die Ergebnisse der Analysen sind in Tabelle 38 dargestellt. Auf eine graphische Darstellung wird aufgrund der geringen Anzahl an Studien verzichtet.

Es ergab sich kein bedeutsamer Unterschied zwischen Fremd- und Selbstrating ( $\beta = 0.15$ , *n.s.*,  $R^2 = 0$ ;  $n = 14$ ).

Tabelle 38: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  zu Post und Follow up (f) für den Vergleich von konfrontativen und kognitiv-behavioralen Verfahren

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
Post	Angst	REM	0.09	0.13	-0.22 – 0.41	0		7
		FEM	0.09	0.13	-0.16 – 0.34		3.04	7
	Depression	REM	0.15	0.13	-0.20 – 0.49	0		6
		FEM	0.15	0.13	-0.11 – 0.41		5.97	6
	Klin. Sign.	REM	-0.04	0.26	-0.76 – 0.68	0.22		5
		FEM	-0.04	0.14	-0.32 – 0.23		12.37 <sup>+</sup>	5
f	Angst	REM	0.02	0.16	-0.42 – 0.47	0		5
		FEM	0.02	0.16	-0.29 – 0.34		1.24	5
	Depression	REM	-0.04	0.16	-0.48 – 0.40	0		5
		FEM	-0.04	0.16	-0.35 – 0.27		1.25	5

<sup>+</sup>  $p < .10$

### **Fail Safe-N**

Für eine Reduzierung der beobachteten Effekte des Vergleiches PsT zu WL auf eine kleine ES von  $g = 0.20$  sind 22 Gruppenvergleiche mit  $g = 0$  notwendig, ca. die Hälfte der von Rosenthal vorgeschlagenen Grenze für stabile Ergebnisse. Das Fail-Safe N beträgt fast viermal soviel wie die eingeschlossenen Vergleiche. Für den direkten Vergleich von konfrontativen mit kognitiv-behavioralen Therapien sind 7 bzw. 3 Studien notwendig mit einer jeweils mittleren Effektstärke von  $g = 0.50$  für eines der Verfahren, um einen geringen Effekt von  $g = 0.20$  für konfrontative oder für kognitiv-behaviorale Behandlungen zu erhalten. Die Grenze liegt hier bei 45, schon eine geringe Anzahl weiterer Untersuchungen kann die mittlere Effektstärke verändern. Die Ergebnisse können nicht als stabil gelten.

### **3.8.4 Pharmakotherapie**

#### **Effektivität: Vergleich zur Placebobedingung**

In ca. der Hälfte der Studien wurden SSRI untersucht, in weiteren 37 % TCA. Nur eine Studien-ES war negativ, ansonsten ergab sich eine geringere Symptomatik für Patienten der aktiven Substanzgruppe am Ende der Behandlung. In allen Symptomkategorien ergab sich zu post im Vergleich zur KG eine signifikant positive mittlere ES. Die kleinste mittlere ES der Hauptkategorien zeigte sich bei Depression mit  $g = 0.21$ , die größte mittlere ES für Klinische Signifikanz mit  $g = 0.56$ . Die Ergebnisse der Analysen sind Tabelle 39 dargestellt. Eine Darstellung der einzelnen Effektstärken findet sich in Abbildung 17.

Bei Ausschluss der Studien, die keine Berechnung der ES über Mittelwerte und Standardabweichungen ermöglichten, stieg die mittlere ES in der Kategorie Angst auf  $g = 0.96$  (95%-CI: 0.76 – 1.16,  $\sigma_{\theta}^2 = 0.01$ ; FEM:  $g = 1.02$ , 95%-CI: 0.86 – 1.18,  $\chi^2 = 37.00$ ,  $n = 11$ ). Die beiden Berechnungen unterschieden sich bedeutsam voneinander, Entscheidungen des Meta-Analysten haben einen starken Einfluss auf die Ergebnisse der Analysen. Aussagen, welches die „korrekte“ Effektstärke ist, sind nicht möglich. Obwohl eine Berechnung der Effektstärken über Alternativen im aktuellen Fall zu einer Unterschätzung des wahren Effektes zu führen scheint, ist andererseits zu beachten,

---

dass 50 % der Vergleiche ausgeschlossen wurden und somit die Repräsentativität des letzteren Ergebnisses eingeschränkt ist.

Es zeigte sich in den vorliegenden Daten ein divergentes Ergebnis hinsichtlich der RE-Varianz. Es wurde zum einen eine RE-Varianz von Null berechnet, im Gegensatz dazu aber war der Homogenitätstest signifikant (d.h. die beobachtete Varianz übersteigt die aufgrund des Stichprobenfehlers zu erwartende Varianz: die RE-Varianz ist bedeutsam größer Null). Dies könnte Unterschiede in den Ergebnissen bezogen auf Population bzw. Stichprobe widerspiegeln. Gilt in der Population eine RE-Varianz von Null, d.h. ist Homogenität gegeben, ist das Ergebnis erklärbar durch einen  $\alpha$ -Fehler des Homogenitätstestes. Gilt in der Population eine RE-Varianz größer Null, d.h. es ist Heterogenität gegeben, zeigt der Schätzer der RE-Varianz einen  $\beta$ -Fehler. Die Schätzer könnten auch verschieden sensibel gegenüber Stichprobenfehler reagieren. Statistische Simulationsstudien sollten die Angemessenheit der beiden Schätzer in Abhängigkeit von Parameterkonstellationen evaluieren<sup>43</sup>.

Zur differentiellen Wirksamkeit verschiedener Substanzklassen zeigte sich in der Symptomkategorie Angst für SSRI eine mittlere ES von  $g = 0.32$ , für TCA von  $g = 0.61$ . Der Unterschied zwischen den Gruppen war in einer multiplen Regression mit Kontrolle von diagnostischem System, Analyseart, Dropouts und Stichprobengröße marginal signifikant, zudem zeigte sich eine signifikant geringere Effektstärke bei Studien mit einer höheren Abbrecherrate ( $\beta_{\text{SSRI:TCA}} = 0.23, p < .10, \beta_{\text{Dropout}} = -0.02, p < .05, \beta_{\text{N}} = -0.69 \times 10^{-4}, \text{n.s.}, \beta_{\text{Diagnose}} = -0.02, \text{n.s.}, R^2 = .00; n = 18$ ). Eine Berechnung der multiplen Regression in der Kategorie Depression entfiel aufgrund der geringen Gesamtanzahl an Studien.

---

<sup>43</sup> Zur Evaluation der Stabilität der Ergebnisse der Modellanalysen in Abhängigkeit vom eingesetzten Berechnungsalgorithmus erfolgte post hoc eine Sensitivitätsanalyse. Die Ergebnisse sind in Anhang G dargestellt.

Tabelle 39: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  zu Post für den Vergleich einer medikamentösen zur Placebobedingung

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
Post	Angst	REM	0.42	0.04	0.33 – 0.51	0		22
		FEM	0.42	0.04	0.33 – 0.50		35.63 <sup>+</sup>	22
	OCD-Gedanken	REM	0.31	0.06	0.16 – 0.45	0		7
		FEM	0.31	0.06	0.19 – 0.42		2.46	7
	OCD-Handlungen	REM	0.27	0.06	0.13 – 0.42	0		7
		FEM	0.27	0.06	0.16 – 0.39		5.61	7
	Depression	REM	0.21	0.06	0.07 – 0.35	0		10
		FEM	0.21	0.06	0.09 – 0.33		2.41	10
	Klin. Sign.	REM	0.56	0.05	0.44 – 0.67	0		13
		FEM	0.56	0.05	0.46 – 0.66		21.74 <sup>+</sup>	13

<sup>+</sup>  $p < .10$

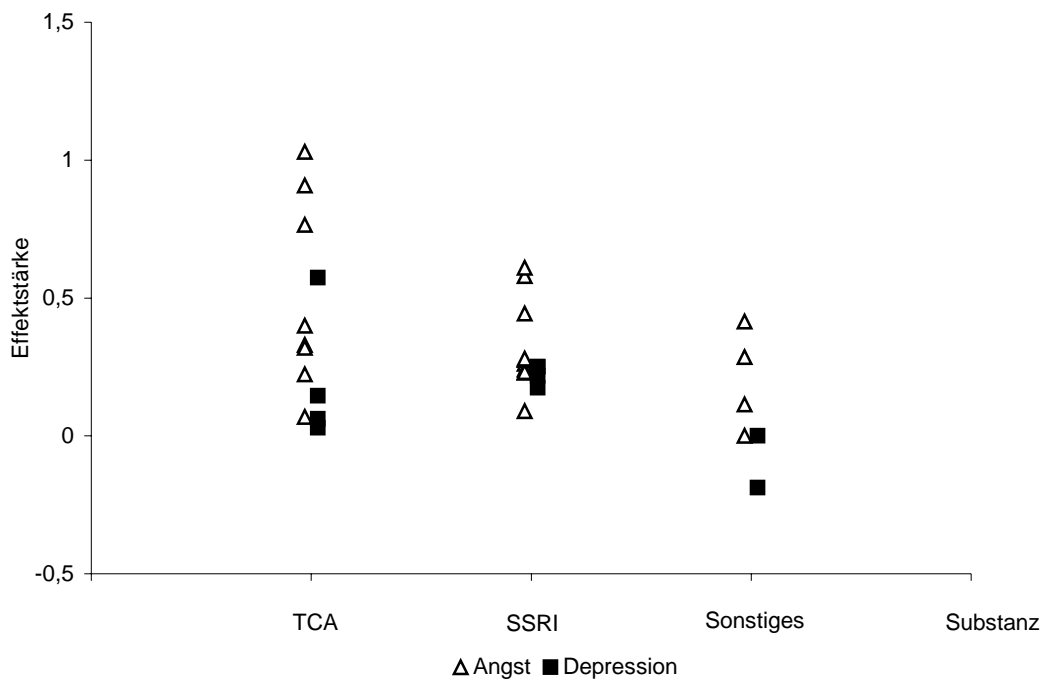


Abbildung 17: Effektstärken der einzelnen Studien für die Symptomkategorien Angst und Depression aufgeteilt nach Substanzklasse

### **Zeitliche Stabilität der Therapie**

Nur eine Studie ermöglichte im Rahmen eines Crossover eine Berechnung einer ES für Follow up, es handelte sich um einen SSRI. Es zeigte sich Abnahme von  $g = -0.15$  nach 1.5 Monaten Placebosubstitution, die Gruppengröße betrug allerdings nur 7 Personen.

### **Akzeptanz**

Durchschnittlich jeder Sechste brach vorzeitig die Behandlung mit einer aktiven Substanz ab ( $M = 17.1\%$ ,  $SD = 9.25$ ,  $n = 24$ ). Für SSRI betrug die mittlere Abbrecher-rate  $20.5\%$  ( $SD = 7.79$ ,  $n = 10$ ) und für TCA  $18.3\%$  ( $SD = 10.86$ ,  $n = 8$ ). Der Unterschied war nicht signifikant ( $t(16) = 0.51$ , n.s.)<sup>44</sup>. Die mittlere Dropout-Rate in der PP-Bedingung betrug  $17.1\%$  ( $SD = 12.77$ ,  $n = 20$ ).

### **Moderatoren**

Aufgrund der RE-Varianz von Null wird keine multiple Regression zur Überprüfung methodischer Variablen berechnet. Zwischen Selbst- und Fremdbeurteilung ergab sich kein bedeutsamer Unterschied ( $\beta = -0.14$ , n.s.,  $R^2 = 0$ ;  $n = 39$ ).

### **Fail Safe-N**

Für eine Reduzierung der beobachteten Effekte auf eine kleine ES von  $g = 0.20$  sind 24 Gruppenvergleiche mit  $g = 0$  notwendig, ca. ein Fünftel der von Rosenthal vorgeschlagenen Grenze für stabile Ergebnisse. Das Fail-Safe N entspricht annähernd der Anzahl eingeschlossener Vergleiche.

### **3.8.5 Vergleich Psycho- und Pharmakotherapie**

Aufgrund der geringen Anzahl an Effektstärken, die die Grundlage für den Effektivitätsschätzer der WL bilden würden ( $n = 4$ ), wurde kein indirekter Vergleich durchgeführt.

---

<sup>44</sup> Normalverteilung gegeben



Es zeigte sich über alle Studien kein bedeutsamer Unterschied in der Abbrecherrate für Psycho- im Vergleich zu Pharmakotherapie ( $t(47) = 0.09$ , n.s.).

Zwei Studien untersuchten beide Verfahren direkt miteinander. In der Kategorie Angst waren beide ESn negativ, d.h. zeigten ein besseres Ergebnis für die Pharmakotherapie von  $g = -0.14$  und  $g = -0.32$ . Auf die Berechnung einer mittleren Effektstärke wurde verzichtet.

### 3.8.6 Vergleich Psychotherapie mit einer Kombinationsbehandlung

In Studien mit Kombinationsbehandlung erfolgte in ca. zwei Drittel der Studien eine Kombination mit einem SSRI, in den restlichen Studien mit einem TCA. Ca. jeweils die Hälfte der Studien untersuchten Exposition mit oder ohne zusätzliche kognitive Elemente, in einer Studie Entspannung. Da letztere nicht als etablierte Behandlung für OCD gilt, wird diese Untersuchung im Rahmen einer Sensitivitätsanalyse ausgeschlossen. Nur 12 % der Studien-ESn waren negativ, d.h. meist zeigte das Kombinationsverfahren eine höhere Effektivität. Sowohl im FEM als auch REM ergab sich, dass Kombinationsverfahren eine bedeutsam höhere Effektivität als reine Psychotherapien in der Behandlung der Angst- aber auch der Depressionssymptomatik haben. Die Höhe der ESn war dabei im kleinen Bereich. Die Ergebnisse der Analysen sind in Tabelle 40 dargestellt. Einen Überblick über die Effektstärken bietet Abbildung 18.

Die mittlere ES blieb stabil nach Ausschluss der Studien mit Entspannung und in denen keine Mittelwerte und Standardabweichungen angegeben waren ( $g = 0.31$ , 95%-CI: 0.03 – 0.58,  $\sigma_{\theta}^2 = 0$ ; FEM:  $g = 0.31$ , 95%-CI: 0.08 – 0.54,  $\chi^2 = 6.80$ ,  $n = 8$ ).

Die Art der Erhebung zeigte keinen bedeutsamen Einfluss auf die ESn ( $\beta = 0.11$ , n.s.,  $R^2 = .00$ ;  $n = 20$ ). Für das Follow up lagen keine ausreichenden Daten vor, in drei Studien ergaben sich nach jeweils 12 Monaten ESn von  $g = 0.02$  bis  $g = 0.52$ . Aussagen sind aufgrund der Anzahl an Untersuchungen nicht möglich.

Tabelle 40: Statistische Kennwerte der mittleren Effektstärken der Symptomkategorien (SK) für Random und Fixed Effects-Modell (REM, FEM) einschließlich der Random Effects-Varianz  $\sigma_{\theta}^2$  zu Post für den Vergleich von Psychotherapie zu einer Kombinationsbehandlung

Zeit	SK	Modell	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
Post	Angst	REM	0.28	0.10	0.06 – 0.49	0		12
		FEM	0.28	0.10	0.09 – 0.47		10.31	12
	Depression	REM	0.44	0.14	0.12 – 0.75	0.08		10
		FEM	0.43	0.11	0.22 – 0.63		17.65 <sup>+</sup>	10
	Klin. Sign.	REM	0.62	0.19	0.01 – 1.23	0.04		4
		FEM	0.60	0.16	0.28 – 0.91		4.57	4

<sup>+</sup>  $p < .10$

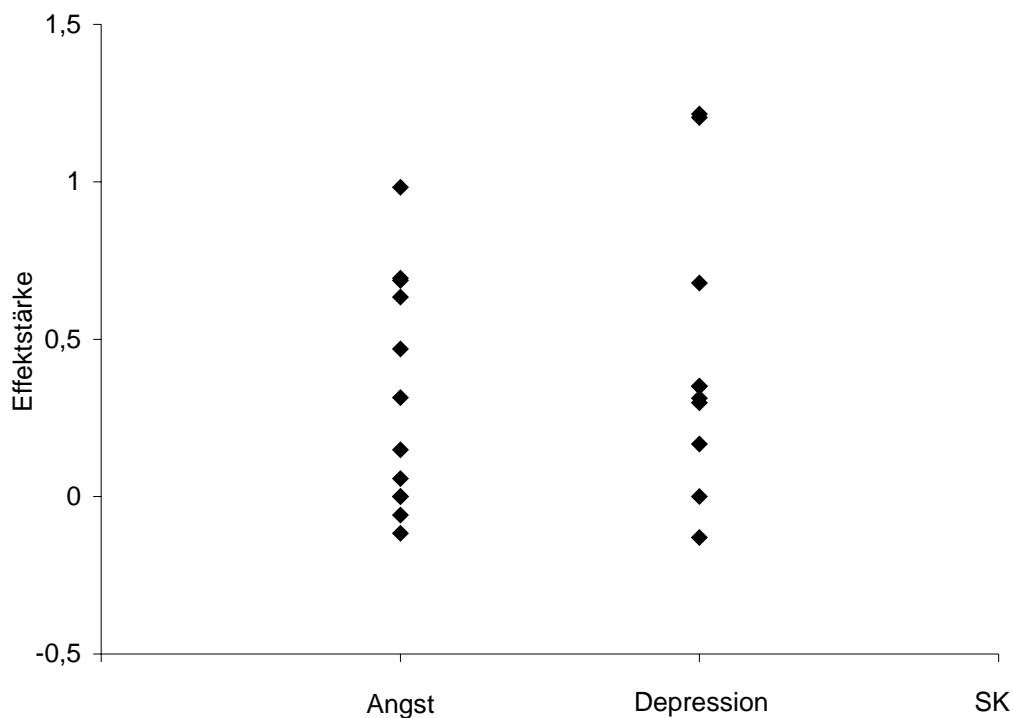


Abbildung 18: Effektstärken für die Symptomkategorien (SK) Angst und Depression der einzelnen Studien für den Vergleich von Psychotherapie mit einer Kombinationsbehandlung

In den Studien, die Psychotherapie mit einer Kombinationsbehandlung verglichen, zeigte sich kein bedeutsamer Unterschied in der Abbrecherrate zwischen den Verfahren (PsT:  $M = 15.7$ ,  $SD = 8.65$ ,  $n = 10$ , PsT+PhT:  $M = 17.8$ ,  $SD = 12.79$ ,  $n = 10$ ,  $t(18) = -0.43$ , n.s.<sup>45</sup>).

### **Fail Safe-N**

Wie aufgrund der geringen Anzahl an Vergleichen zu erwarten, zeigen sich die Ergebnisse der Analysen als eher instabil. So sind nur fünf Studien mit einer Effektstärke von  $g = 0$  notwendig für einen mittleren Effekt von  $g = 0.20$ . Neun Studien werden benötigt für einen mittleren Effekt von  $g = 0.50$  ausgehend von Untersuchungen, die eine hohe Effektivität der Kombinationsbehandlung im Vergleich zu Psychotherapie mit  $g = 0.80$  zeigen. Für einen kleinen Effekt für Psychotherapie mit  $g = -0.20$  werden 19 weitere Studien benötigt mit  $g = -0.50$ . Die Grenze nach Rosenthal liegt bei jeweils 70 Studien.

---

<sup>45</sup> Normalverteilung gegeben

### 3.8.7 Diskussion

Es ergaben sich in der Meta-Analyse folgende Ergebnisse. Die vorliegenden Studien zeigten eine sehr hohe Effektivität der untersuchten psychotherapeutischen Verfahren in der Behandlung der Zwangsstörung im Vergleich zu einer unbehandelten Kontrollgruppe. Die am häufigsten untersuchte Therapieform war dabei Exposition mit Reaktionsverhinderung. Allerdings können die Ergebnisse nicht generalisiert werden, da nur eine geringe Anzahl an Studien vorlag. Die Frage, ob die Wirksamkeit behavioraler und kognitiv-behavioraler Therapien über eine Realisierung gemeinsamer Faktoren hinausgeht, kann nicht beantwortet werden. Weitere kontrollierte Studien sind wünschenswert. Jeder sechste Patient verließ vorzeitig die Psychotherapie, die Akzeptanz erscheint gut.

Betrachtet man die Ergebnisse der Psychotherapien zu Follow up, so ergab sich weder eine bedeutsame Verschlechterung noch Verbesserung in der Symptomatik. Die Befunde sprechen für eine zeitliche Stabilität der Effekte.

Es zeigte sich kein bedeutsamer Unterschied zwischen konfrontativen und kognitiv-behavioralen Verfahren weder in der Effektivität der Therapien noch in der Akzeptanz durch die Patienten. Beide Verfahren scheinen somit gleich angemessen in der Behandlung der Zwangsstörung. Auch hier war die Anzahl an Studien allerdings gering.

Medikamentöse Therapien, insbesondere SSRI und TCA waren effektiv im Vergleich zu Pillenplacebo in der Behandlung der Zwangsstörung. Die Höhe der Wirksamkeit war allerdings abhängig von Entscheidungen des Meta-Analysten. So ergab sich eine geringe Effektivität über alle Studien, dagegen ein hoher Effekt, betrachtete man nur Effektstärken aus Studien, die Mittelwerte und Standardabweichungen angaben. Allerdings wurde dabei auch die Hälfte der Studien ausgeschlossen, eine Einschränkung der Repräsentativität scheint gegeben. TCA zeigten eine leicht höhere Wirksamkeit in der Behandlung der Angstsymptomatik im Vergleich zu SSRI. Auch dieses Ergebnis erscheint abhängig von den methodischen Kriterien der vorliegenden Meta-Analyse. Von acht Effektstärken mit TCA konnten sechs über Mittelwerte und Standardabweichungen berechnet werden, von 10 Effektstärken mit SSRI dagegen nur zwei. Eine Einschätzung, wie die gefundenen Effekte zu interpretieren und zu generalisieren sind,

ist schwierig. Spiegelt die bedeutsame Zunahme der Effektstärke bei Ausschluss von Studien mit Mittelwerten und Standardabweichungen die geringere Effektivität der SSRI wider oder ergibt sich die geringere Effektivität der SSRI aufgrund der alternativen Berechnung der Effektstärken? Nur weitere Studien mit ausführlicher Dokumentation der Ergebnisse können diese Frage klären.

Aufgrund der geringen Anzahl an Psychotherapiestudien bzw. Studien mit einem direkten Vergleich von psychologischer mit medikamentöser Therapie kann nicht eingeschätzt werden, welche Therapieform eine höhere Effektivität zeigt. Die Akzeptanz der Patienten scheint für beide Verfahren vergleichbar.

Dagegen ergab die vorliegende Meta-Analyse, dass bei gleicher Akzeptanz eine Kombination aus Psycho- und Pharmakotherapie, besonders mit einem SSRI oder TCA, bedeutend wirksamer in der Behandlung von Patienten mit einer Zwangsstörung ist, als Psychotherapie allein. Dies galt sowohl für die Angst- als auch für die Depressions-symptomatik. Die beiden Verfahren unterschieden sich dabei um eine Effektstärke im kleinen Bereich. Inwieweit allerdings der Vorteil der Kombinationsbehandlung auch nach Absetzen der Medikation gegeben ist, kann aufgrund der vorliegenden Daten nicht eingeschätzt werden. Zudem gelten vergleichbare Einschränkungen wie für Studien mit Kombinationsbehandlung bei PDA. So untersuchten annähernd zwei Drittel der eingeschlossenen Studien nicht Psychotherapie sondern Psychotherapie mit Pillen-placebo in einer Behandlungsgruppe. Mögliche Schwierigkeiten dieses Designs werden später diskutiert.

Die vorliegende Meta-Analyse zeigte, wie in bisherigen Arbeiten auch, dass behaviorale und kognitiv-behaviorale, aber auch medikamentöse Therapien (im speziellen TCA und SSRI) effektiv in der Behandlung von Patienten mit Zwangssymptomatik sind und die Ergebnisse für Psychotherapie zeitlich stabil erscheinen. Die in der aktuellen Arbeit gefundene leicht höhere Effektivität für TCA im Vergleich zu SSRI wurde in bisherigen Meta-Analysen nicht gezeigt. Allerdings werden wichtige Unterschiede in der Kategorisierung der Effektstärken deutlich. In der vorliegenden Arbeit wurde Clomipramin der Klasse der TCA zugeordnet (vgl. z.B. auch Möller, Müller, Volz, 2000). Aufgrund der starken Beeinflussung des serotonergen Systems und als klassisches Medikament in der

Behandlung von Zwangserkrankungen, wurde es in vielen bisherigen Arbeiten extra gewertet (z.B. Piccinelli et al., 1995) oder zusammen mit SSRI ausgewertet (Kobak et al., 1998; Stein et al., 1995; van Balkom et al., 1994). In den bisherigen Meta-Analysen zeigten sich deskriptiv höhere Effektstärken für Clomipramin, in der vorliegenden Arbeit untersuchten 6 von 8 Studien aus dem Bereich der TCA im speziellen dieses Medikament, eine Erklärung für die unterschiedlichen Resultate. Die Ergebnisse einer zum Teil gefundenen geringeren Effektstärke für Fremdbeurteilungsverfahren in anderen Meta-Analysen konnten in der vorliegenden Arbeit nicht repliziert werden.

Insgesamt ist festzustellen, dass die Ergebnisse der vorliegenden Meta-Analyse nicht als stabil gelten können. Sowohl eine meist nur geringe Anzahl von Studien in Schubladen als auch andere methodische Entscheidungen des Meta-Analysten führen zu substantiell anderen Resultaten, als hier gefunden wurde. Die Zwangsstörung muss somit als eher ungenügend untersuchte Störung angesehen werden, für die noch keine sicheren generalisierbaren Aussagen getroffen werden können.

## 4 Diskussion

Das Kapitel umfasst die folgenden Bereiche: zunächst werden die Ergebnisse der Meta-Analyse zusammenfassend über alle Angststörungen<sup>46</sup> dargestellt. Es folgt eine Diskussion über potentielle Einschränkungen der Interpretierbarkeit der vorliegenden Resultate. Hier geht es zum einen um methodische Besonderheiten der eingeschlossenen Primärstudien einschließlich eines Ausblicks auf zukünftige Forschungsaspekte für Evaluationsstudien. Außerdem erfolgt eine Diskussion über Probleme der aktuellen Meta-Analyse bzw. der Methode der Meta-Analyse im Allgemeinen.

### 4.1 Zusammenfassende Betrachtung der Ergebnisse

#### 4.1.1 Effektivität der behavioralen und kognitiv-behavioralen Verfahren

Es zeigte sich in der vorliegenden Arbeit, wie auch in bisherigen Meta-Analysen, dass die untersuchten psychotherapeutischen Verfahren sehr effektiv in der Behandlung der Angststörungen sind. Sowohl im Vergleich zu einer Warteliste als auch zu einer behandelten Kontrollgruppe mit Therapie- oder Pillenplacebo ergaben sich bedeutsame Effektstärken für die Angstsymptomatik. Die Wirksamkeit von behavioralen und kognitiv-behavioralen Therapien geht somit über eine Realisierung gemeinsamer Faktoren hinaus. Konfrontative und kognitiv-behaviorale Verfahren scheinen sich in ihrer Wirksamkeit nicht zu unterscheiden. Dies ergaben sowohl die Daten aus den direkten Vergleichen, d.h. wenn beide Verfahren in einer Studie zusammen untersucht wurden, als auch die indirekten Vergleiche über die Effektstärken aus den Kontrollgruppenberechnungen.

Für die depressive Symptomatik zeigte sich im Vergleich zu einer unbehandelten Kontrollgruppe eine bedeutsame Effektivität der hier untersuchten Psychotherapien. Im Vergleich zu einem Therapie- oder Pillenplacebo lagen dagegen über die Störungen divergente Befunde vor: es zeigte sich für die Generalisierte Angststörung eine generelle bedeutsame Wirksamkeit, für Panikstörung zeigten die vorliegenden Studien

---

<sup>46</sup> Zur Vereinfachung wird im Folgenden von Panikstörung gesprochen, es beinhaltet aber alle Störungen, die unter PDA zusammengefasst wurden.

---

eine Wirksamkeit, die aber nicht generalisiert werden kann, während es für Soziale Phobie keinen bedeutsamen Unterschied in der Effektivität gab. Es ergab sich deskriptiv dagegen konsistent, sowohl im Vergleich zu einer unbehandelten als auch behandelten Kontrollgruppe, dass die Effektstärken für die depressive Symptomatik geringer waren als für die Angstsymptomatik. Der Unterschied zwischen den Symptomgruppen war aber nicht bedeutsam. Hinsichtlich einer differentiellen Wirksamkeit von konfrontativen und kognitiv-behavioralen Therapien zeigten sich für Studien, die beide Verfahren direkt miteinander verglichen, größere Differenzen für die depressive denn für die Angstsymptomatik. Kognitiv-behaviorale Verfahren führten zumindest für Panikstörung zu einer bedeutsam höheren Reduktion der depressiven Symptomatik als rein konfrontative Verfahren.

Die vorliegenden Befunde sprechen dafür, dass behaviorale und kognitiv-behaviorale Therapien die Lebensqualität der Patienten sowohl gegenüber einer unbehandelten Kontrollgruppe als auch in den eingeschlossenen Studien gegenüber einem Therapie- oder Pillenplacebo bedeutsam erhöhen. Allerdings wird deutlich, dass eine Erfassung dieses Aspektes nur in einer geringen Anzahl an Untersuchungen erfolgte und somit, trotz bedeutsamen Effektes für den Vergleich mit einer Wartelistenbedingung auch im Random Effects-Modell, eine Generalisierbarkeit eingeschränkt ist. Eine systematische Erfassung dieses Aspektes ist von weiteren Evaluationsstudien zu fordern.

Über alle Störungen und Vergleiche der Therapieverfahren mit einer Kontrollbedingung lagen die höchsten Effektstärken in der Kategorie Klinische Signifikanz vor. Hinsichtlich der praktischen Relevanz waren die untersuchten Verfahren somit sehr wirksam. Dies spricht für eine hohe Änderungssensitivität des Konzeptes. Jedoch könnte die Inflation der Effektstärke auch gegeben sein durch die Berechnung über Häufigkeiten und nicht über Mittelwerte. Hasselblad et al. (1995) sprechen von einer Vergleichbarkeit bei gleichen Varianzen und einer logistischen (Normal-)Verteilung der zugrunde liegenden kontinuierlichen Daten, Voraussetzungen, die in der Regel vom Meta-Analysten nicht überprüft werden können. Bei Verletzung dieser Annahmen liegen Abweichungen vor, die die Unterschiede bedingen könnten<sup>47</sup>. Die vorliegenden Ergeb-

---

<sup>47</sup> Es sei allerdings anzumerken, dass diese Annahmen bei Hedges  $g$  auch aus anderen Berechnungsalternativen implizit als erfüllt betrachtet werden. Kritik findet sich z.B. bei Grissom und Kim (2001).



nisse implizieren, dass eine Sensitivitätsanalyse angemessen erscheint, wenn man Effektstärken, die über Häufigkeiten berechnet wurden, mit anderen Effektstärkemaßen in einer Symptomkategorie integriert. Es sollte überprüft werden, ob dieses Vorgehen zu einer Inflation des mittleren Effektes führt. Es ist zudem zu beachten, dass die Methoden zur Bestimmung der Klinischen Signifikanz zwischen den Studien differierten, möglicherweise Ursache für die teilweise gefundene Heterogenität der Effektstärken. Die Klassifikation der Patienten ist abhängig von Entscheidungen des Wissenschaftlers: welche Instrumente bilden die Grundlage (es ergibt sich eine Abhängigkeit von psychometrischen Eigenschaften der Verfahren) und welche Kriterien werden eingesetzt (z.B. mindestens 25 % oder 50 % Verbesserung in der Symptomatik oder kein Erfüllen der diagnostischen Kriterien). So fanden z.B. Ogles, Lambert und Sawyer (1995) Unterschiede in der Klassifikation von depressiven Patienten abhängig vom eingesetzten Messinstrument (BDI, HAMD, Hopkins Symptom Checklist). Eine Überprüfung der Angemessenheit verschiedener Varianten sowie darauf aufbauend eine Standardisierung der Erhebung ist ein Forschungsdesiderat.

Für behaviorale und kognitiv-behaviorale Verfahren zeichnete sich eine Stabilität der Ergebnisse ab. Auch in einem längeren Zeitraum nach Abschluss der Psychotherapie zeigten Patienten keine bedeutsame Verschlechterung in ihrer Symptomatik.

#### **4.1.2 Akzeptanz behavioraler und kognitiv-behavioraler Verfahren**

Es zeigte sich über alle Störungen eine sehr gute bis gute Akzeptanz der Patienten für die untersuchten Psychotherapien. Die Abbrecherraten lagen zwischen 10 % (Generalisierte Angststörung, Spezifische Phobie) und 18 % (Posttraumatische Belastungsstörung).

#### **4.1.3 Effektivität der medikamentösen Therapien**

Auch die untersuchten Pharmakotherapien waren wirksam in der Behandlung der Angststörungen im Vergleich zu einer Pillenplacebobedingung. Die mittleren Effektstärken waren meist im kleinen Bereich, d.h. unter 0.5. Im Gegensatz zur Psychotherapie zeigten medikamentöse Verfahren über Angst, Depression und Lebensqualität ähnliche Effektstärken mit Ausnahme bei Zwangsstörung mit leicht geringeren Effekten

bei Depression. Die Effektstärken für die Klinische Signifikanz waren ebenfalls höher. Hinsichtlich der Stabilität der Ergebnisse ist allerdings zu beachten, dass nur eine geringe Anzahl an Studien mit geeignetem Design vorliegt. Die Repräsentativität der Daten ist somit stark eingeschränkt. Aber nur mit ausreichend Untersuchungen sind generalisierbare Aussagen möglich, ob die gefundenen positiven Effekte zeitlich auch nach Absetzen der Medikation stabil sind oder ob – im negativen Fall – eine Verschlechterung in der Symptomatik eine Langzeitmedikation implizieren würde.

#### **4.1.4 Akzeptanz medikamentöser Therapien**

Die Akzeptanz der Patienten für die Pharmakotherapie war gut bis zufrieden stellend. Die Abbrecherraten lagen zwischen 17 % (Zwangsstörung) und 24 % (Generalisierte Angststörung). Bei einer medikamentösen Behandlung der Posttraumatischen Belastungsstörung brach dagegen jeder Dritte vorzeitig die Untersuchung ab, dies kann nicht als befriedigend angesehen werden.

#### **4.1.5 Vergleichende Effektivität von Psycho- und Pharmakotherapie**

In Studien, die Psycho- und Pharmakotherapie direkt miteinander verglichen, ergab sich für keine der untersuchten Angststörungen ein bedeutsamer Unterschied zwischen den Verfahren. Für die Generalisierte Angststörung ergab sich zwar für die eingeschlossenen Studien eine bedeutsam höhere Wirksamkeit für Psychotherapie, diese kann aber nicht generalisiert werden. Es wurde deutlich, dass sowohl Untersuchungen existieren, die ein besseres Ergebnis für die psychologische Therapie zeigen, als auch Untersuchungen mit einer höheren Effektivität für die medikamentöse Therapie. Aufgrund der RE-Varianz ungleich Null bei Sozialer Phobie und Panikstörung ist nicht von einer gemeinsamen Populationseffektstärke auszugehen, wegen der geringen Studienzahl wurden aber keine Regressionsanalysen durchgeführt. Es gibt somit bedeutsame Unterschiede zwischen den Untersuchungen, die aber aufgrund der Datenlage nicht aufgeklärt werden können. Weitere Studien sind notwendig für Aussagen, welche Therapie für welchen Patienten angemessen ist bzw. welche methodischen Eigenschaften der Untersuchungen die Ergebnisse beeinflussen.

Im indirekten Vergleich ergaben sich dagegen divergente Befunde: bei Sozialer Phobie und Generalisierter Angststörung zeigte Pharmakotherapie eine erhöhte Effektivität, bei Posttraumatischer Belastungsstörung wurde kein Unterschied zwischen den Verfahren gefunden und bei Panikstörung war Psychotherapie etwas wirksamer. Es wurde deutlich, dass eine Kontrolle methodischer Variablen für einen Vergleich von Therapien über verschiedene Studien wichtig ist. In den einfachen Regressionen ergab sich bei den beiden letztgenannten Störungen eine bedeutend höhere Effektivität für Psychotherapie.

Betrachtet man zusätzlich die Akzeptanz der Behandlung zeigte sich mindestens eine gleich hohe Akzeptanz für Psychotherapie, bei Generalisierter Angststörung, Panikstörung und Posttraumatischer Belastungsstörung eine bedeutsam höhere. Dies bezieht sich auf Daten, nachdem sich ein Patient entschieden hat, mit der Behandlung zu beginnen. Diese Unterschiede scheinen sich noch zu verstärken, betrachtet man die gesamte Patientenpopulation. Es liegen Befunde vor, dass Patienten bei Entscheidung für oder gegen eine bestimmte Therapie eher Pharmako- als Psychotherapie ablehnen, d.h. gar nicht mit der Behandlung beginnen (z.B. Bakker et al., 2000; Hofmann et al., 1998).

#### **4.1.6 Vergleichende Effektivität von Psychotherapie mit einem integrativen Ansatz aus Psycho- und Pharmakotherapie**

Hinsichtlich des Vergleiches von Psychotherapie mit einem Kombinationsverfahren lagen nur für Panikstörung und Zwangsstörung ausreichend Daten vor. In beiden zeigte die Kombinationsbehandlung eine bedeutsame höhere Effektivität im kleinen Bereich unter Wirkung der Substanz. Allerdings ergab sich für die Panikstörung nach Absetzen der Medikation im Follow up mit einer allerdings nur kleinen Anzahl an Studien kein bedeutsamer Unterschied zwischen den Verfahren, für die Zwangsstörung lagen hierzu keine ausreichenden Daten vor.

#### **4.1.7 Zusammenfassung der Ergebnisse zur Effektivität der Therapien**

Insgesamt verdeutlichen die vorliegenden Befunde entsprechend Marks et al. (2000), dass viele Wege ins Rom der Angstreduktion führen, nicht nur hinsichtlich behavioraler oder kognitiv-behavioraler Verfahren sondern auch hinsichtlich Pharmakotherapie. Die

---

untersuchten psychologischen Therapien stellen mindestens eine gleichrangige Alternative in der Behandlung der Angststörungen dar. Psychotherapie wird dabei oft von Patienten besser akzeptiert als Pharmakotherapie. Abschließend sei noch auf die unterschiedlichen Kosten der Behandlung verwiesen, wenn auch im Rahmen dieser Arbeit nur am Rande diskutiert. So fanden Gould et al. (1995) bei Panikstörung bzw. Gould, Buckminster et al. (1997) bei Sozialer Phobie, dass die kognitiv-behaviorale Einzeltherapie nach neun bzw. sechs Monaten kostengünstiger war als eine Langzeitmedikation mit einem SSRI (Flouxetin bzw. Fluvoxamin) sowie hochdosierten Alprazolam und die kognitiv-behaviorale Gruppentherapie schon nach einem bzw. drei Monaten kostengünstiger war.

#### **4.1.8 Einschätzung der Verteilung gemeinsamer und spezifischer Faktoren für die Wirksamkeit der untersuchten Therapien**

Für die Aufteilung des Gesamteffektes in spezifische und gemeinsame Faktoren ergab sich für Pharmakotherapie für Soziale Phobie, Panikstörung und Posttraumatische Belastungsstörung eine Gleichverteilung der beiden Aspekte. Für die Generalisierte Angststörung ergab sich dagegen ein sehr hoher Anteil gemeinsamer Faktoren<sup>48</sup>. In einer Meta-Analyse von Kirsch und Sapirstein (1998) in der Behandlung der Depression fanden die Autoren einen zweimal so großen Anteil für gemeinsame Faktoren denn für spezifische Faktoren für die medikamentöse Behandlung. Auch in dieser Meta-Analyse erfolgte die Berechnung über verschiedene Studien. Allerdings wird die Angemessenheit der eingeschlossenen Untersuchungen angezweifelt (Klein, 1998), da trotz des Umfangs an wissenschaftlichen Untersuchungen nur 19 Pharmako- und 19 Psychotherapiestudien zur Schätzung des Retest-Effektes in die Analyse einbezogen wurden. Keine der Studien wurde nach 1994 publiziert. Möglicherweise ist somit die Repräsentativität der Ergebnisse eingeschränkt. Inwieweit sich Unterschiede auch aufgrund eines anderen Störungsbildes ergeben und somit ob eine Abhängigkeit des Verhältnisses der verschiedenen Anteile zur untersuchten Störung gegeben ist, sollte im Rahmen einer gemeinsamen Meta-Analyse betrachtet werden. Dies ermöglicht es, den Einfluss von

---

<sup>48</sup> Wie später in der Diskussion ausgeführt, erscheint aber besonders für diese Störung eine starke Einschränkung gegeben über die Berechnung aus verschiedenen Studien.

---

Unterschieden in der meta-analytischen Methode (z.B. Suchstrategien, Inklusionskriterien, Berechnungen) auszuschließen.

Für Psychotherapie ergaben sich dagegen divergente Befunde: eine Gleichverteilung bei Posttraumatischer Belastungsstörung und Panikstörung, ein höherer Anteil gemeinsamer Faktoren für Soziale Phobie und ein höherer Anteil spezifischer Faktoren für Generalisierte Angststörung. Stevens, Hynan und Allen (2000) fanden in einer aktuellen Meta-Analyse, durchgeführt über eine Vielzahl an Störungsbildern und Psychotherapien, einen fast doppelt so hohen Anteil spezifischer Faktoren im Vergleich zu gemeinsamen Faktoren. Im Gegensatz zu dem Vorgehen in der aktuellen Arbeit schlossen die Autoren nur Studien ein, in denen sowohl eine Wartelistenbedingung als auch eine behandelte Kontrollgruppe vorlag, so dass die Anteile direkt berechnet werden konnten. Dies hat einen bedeutenden Vorteil, da Unterschiede zwischen Studien, wie später beschrieben, einen starken Einfluss auf die Ergebnisse haben können. Stevens et al. (2000) fanden einen höheren Anteil gemeinsamer Faktoren bei eng umschriebenen Störungsbildern. Möglicherweise könnte dies die gefundenen Unterschiede zwischen den Angststörungen erklären, gilt insbesondere die Generalisierte Angststörung als wenig umschriebene Störung. Entsprechend der Argumentation von Stevens et al. (2000) scheinen Patienten mit Generalisierter Angststörung weniger von gemeinsamen Faktoren der Therapie zu profitieren. Eine Interpretation wird allerdings erschwert durch die gefundenen starken Veränderungen der Symptomatik bei Pillenplacebo.

Sowohl für Psychotherapie als auch in noch stärkerem Maße für Pharmakotherapie sei allerdings hinsichtlich der Interpretation der Ergebnisse betont, dass eine Schätzung des Anteils der verschiedenen Faktoren einschließlich des Retest-Effektes über verschiedene Studien erfolgte, mögliche Einschränkungen werden später diskutiert.

#### **4.1.9 Einfluss methodischer Variablen**

Hinsichtlich der externen Validität der Ergebnisse ergab sich für die untersuchten Variablen in der Mehrzahl an Analysen kein bedeutsamer Zusammenhang zur Effektstärke. Es zeigte sich mit Ausnahme der Panikstörung kein bedeutsamer Einfluss des diagnostischen Systems, die Ergebnisse älterer Studien scheinen somit auch auf Patienten diagnostiziert mit einem aktuellen System anwendbar zu sein, die mittlere Effekt-

stärke erscheint generalisierbar. Für Panikstörung zeigten sich höhere Effektstärken für neuere Diagnosesysteme, die mittlere Effektstärke erscheint somit konservativ. Die Wirksamkeit der medikamentösen Therapie war entsprechend der Intention unabhängig von fester vs. flexibler Dosierung, das Vorgehen des Einschlusses der effektivsten Behandlungsgruppe bei mehrfacher Stufung der Dosierung erscheint somit angemessen.

Es zeigte sich wider Erwarten in keiner der Berechnungen ein bedeutsamer Unterschied zwischen Completer- und Intent to treat-Analyse für Pharmakotherapie. Effektstärken aus ITT-Daten waren somit nicht bedeutsam geringer als Effektstärken aus Completer-Daten. Gould, Otto et al. (1997) fanden in einer Meta-Analyse zur Behandlung der Generalisierten Angststörung ebenfalls keinen Unterschied zwischen den Analysearten. Möglicherweise steht dies im Zusammenhang mit der Nichtbeachtung der Unterschiede in den Abbrecherraten von Kontroll- und Behandlungsgruppe im Rahmen eines Interaktionsterms. Beide Gruppen gehen aber in die Berechnung der Effektstärke ein. Bei Abbruch eines höheren Prozentsatzes von Patienten in der Placebobedingung könnte sich in der ITT-Analyse eine höhere Effektstärke ergeben als bei einem höheren Prozentsatz von Patienten in der Behandlungsgruppe, Unterschiede könnten sich somit ausmitteln. Zudem erfolgte ein Vergleich der beiden Analysearten über verschiedene Untersuchungen und nicht an Studien, die beide Methoden einsetzten. Gould et al. (1995) fanden bei vier Studien, die sowohl Daten einer Completer- als auch ITT-Analyse präsentierten, jeweils eine höhere Effektstärke für die ITT-Analyse: in jeder dieser Untersuchungen war die Abbrecherrate für die Placebobedingung höher.

Nicht in jeder Analyse zeigte sich ein Zusammenhang von Abbrecherrate der Behandlungsgruppe auf die Effektstärke. Gab es einen Einfluss, war dieser negativ, d.h. in Studien, in denen mehr Patienten die Behandlung abbrachen, ergaben sich geringere Effektstärken. Möglicherweise basiert dies darauf, dass Patienten mit weniger effektiven Behandlungen eher bereit sind, eine Studie zu verlassen bzw. damit im Zusammenhang in geringerem Ausmaß dazu bereit sind, die verbundenen Nebenwirkungen zu akzeptieren. In Verbindung mit der internen Validität einer Studie ergaben sich somit teilweise geringere Effektstärken bei weniger repräsentativen Studien. Dieses Ergebnis liegt aber nicht konsistent vor. Die Abbrecherrate kann somit, muss aber nicht, in Zusammenhang mit der Effektstärke stehen, dies entspricht auch den von Matt et al. (1997) gefundenen Ergebnissen.

In der Mehrzahl an Analysen ergab sich kein bedeutsamer Unterschied zwischen Fremd- und Selbstrating. Lagen Differenzen vor, waren Effektstärken von Fremdbeurteilungen höher als von Selbstbeurteilungen. Die Bedeutsamkeit der Unterschiede erschien aber abhängig von den eingeschlossenen Studien (bei Sozialer Phobie war der Unterschied nicht mehr bedeutsam nach Ausschluss von Untersuchungen, die nur eine Art der Erhebung durchführten) und unter Beachtung der Änderungssensitivität der Symptomkategorien (bei Posttraumatischer Belastungsstörung war der Unterschied nicht mehr bedeutsam betrachtete man nur die Kategorie Angst). Nur im Vergleich von Psychotherapie mit Therapieplacebo ergab sich für die Generalisierte Angststörung auch nach Kontrolle dieser Aspekte eine bedeutsam höhere Effektstärke für Fremdbeurteilungen. Ob dies auf eine erhöhte Änderungssensitivität oder einen Versuchsleiterereffekt zurückzuführen ist, kann nicht beurteilt werden.

Hinsichtlich des Zusammenhanges von Stichprobengröße und Effektstärke ergab sich ein ähnliches Bild wie bei der Abbrecherrate. Es zeigte sich nicht in jeder Analyse ein Einfluss, war er bedeutsam, so ergaben sich wie zu erwarten geringere Effektstärken in Studien mit höherer Stichprobengröße. Dies steht vermutlich in Abhängigkeit zu Aspekten der Power einer Untersuchung und somit der Signifikanz und des Publikationsstatus. Interessant wäre die Überprüfung der Stabilität dieser Ergebnisse bei zusätzlichem Einschluss unveröffentlichter Studien und somit des vermuteten Zusammenhanges. Hier sind die Grenzen der Meta-Analyse zu erkennen.

Es zeigte sich in keiner der Analysen ein bedeutsamer Unterschied zwischen Studien, die ein Pillenplacebo Run-in einsetzten oder direkt mit der Medikation begannen. Das Vorgehen erscheint somit nicht geeignet, die Reaktion auf Placebo zu reduzieren bzw. führt zu einer äquivalenten Reduktion in der Behandlungsgruppe. Die Ergebnisse stimmen damit mit den bisherigen Arbeiten von Oosterbaan (2001) und Trivedi et al. (1994) überein. Bezieht man den Kostenaspekt einer Untersuchung ein, so ist die Durchführung eines Placebo Run-in nicht zu befürworten.

Nur für Posttraumatische Belastungsstörung ergab sich ein signifikanter Einfluss der Therapiedauer: Studien mit längeren Therapien zeigten geringere Effektstärken. Dieses Ergebnis war allerdings abhängig von den eingeschlossenen Studien und spiegelt die für

alle untersuchten Variablen geltende Einschränkung der korrelativen Natur der Daten einer Meta-Analyse wider. Dies wird später diskutiert für den indirekten Vergleich von Psycho- und Pharmakotherapie.

## **4.2 Einschränkungen der Arbeit**

### **4.2.1 Besonderheiten der eingeschlossenen Primärstudien**

#### **Psychotherapiestudien**

Sind die Ergebnisse der Psychotherapiestudien möglicherweise zu progressiv? In vielen Psychotherapiestudien nimmt zumindest ein Teil der Patienten zusätzliche Medikation ein. So kritisierten z.B. Power und Sharp (1995) das Design:

We ... would question whether CBT [kognitiv-behaviorale Therapie] is effective when used in the absence of concurrent psychotropic medication as is claimed on the basis of these predominantly flawed studies. ... We believe that patients taking part in trials of psychological treatments should certainly not keep taking the tablets! (S. 72)

Diese Sichtweise erscheint allerdings aufgrund der vorliegenden Daten zu negativ. Es zeigte sich für behaviorale und kognitiv-behaviorale Verfahren auch zur behandelten Kontrollgruppe (Therapie- oder Pillenplacebo) eine bedeutsame Wirksamkeit der Therapie im Besonderen für die Angstsymptomatik. Die RE-Varianz war hierbei für jede Störung Null, so dass von einer gemeinsamen Effektstärke ausgegangen wird. Das bedeutet, dass es keinen Unterschied in der Effektivität der Psychotherapie gab, unabhängig ob eine Berechnung im Vergleich zu einem Therapieplacebo (d.h. zusätzliche Medikation möglich) oder zu einem Pillenplacebo (d.h. zusätzliche Medikation nicht erlaubt) erfolgte. Die Einschränkung der internen Validität ist möglicherweise gering. So wird meist als Einschlusskriterium eine stabile Medikation gefordert und trotzdem erfüllen die Patienten zu Prä die Diagnosekriterien. Des Weiteren führt die meist eingesetzte Randomisierung zu einer Gleichverteilung auf die verschiedenen Gruppen. Zudem ist zu beachten, dass dieses Problem nicht nur in Psycho- sondern auch in Pharmakotherapiestudien zu finden ist. So fanden Clark et al. (1990) in einer placebokontrollierten Studie zur Effektivität von Alprazolam und Imipramin bei 25 %



von 79 Patienten bei einer Serumanalyse verbotene Substanzen während der substanzfreien Anfangszeit und/oder während der aktiven Studie. Dies war nicht nur in der Placebogruppe gegeben sondern auch in den aktiven Substanzgruppen. Somit könnten sowohl in Psycho- als auch in Pharmakostudien die Ergebnisse beeinflusst sein durch zusätzliche Medikation; dies wird auch in künftigen Evaluationsstudien kaum zu vermeiden sein. Bei regelmäßiger Überprüfung des Serums und Ausschluss dieser Patienten ergibt sich eine erhöhte interne Validität allerdings mit gleichzeitigen Einbußen für die externe Validität. Zudem ist zu beachten, dass auch eine Psychotherapie in Studien zur Untersuchung einer pharmakologischen Substanz nicht stets ein Ausschlusskriterium darstellt bzw. dies überprüft wird. Somit sollte die Kritik nicht einseitig erfolgen.

Als Problem hinsichtlich der Interpretation der Ergebnisse zur Stabilität der Behandlung ist zu beachten, dass in der Mehrzahl an Studien entweder zusätzliche Therapien im Follow up-Zeitraum von den Patienten in Anspruch genommen bzw. keine Aussagen darüber in den Publikationen getroffen wurden. Es ist zu vermuten, dass vor allem Patienten, die eine Verschlechterung erleben oder auch nach der untersuchten Behandlung noch eine hohe Symptomatik aufweisen, eine zusätzliche Therapie wünschen, so dass die in der vorliegenden Arbeit gefundenen Ergebnisse eher progressiver Natur sind. Einfluss kann dies auch auf die Ergebnisse zum Vergleich verschiedener Therapien im Follow up haben. Es ergibt sich eine Einschränkung der internen Validität der Resultate. Dass Patienten andere Behandlungen nutzen, ist auch in künftigen Evaluationsstudien nicht zu vermeiden. Welche Möglichkeiten könnten eingesetzt werden, um eine Interpretation zu erleichtern? Zum einen sollte die Informationsvermittlung in den Publikationen verbessert werden. Es sollte mitgeteilt werden, wie viele Patienten eine zusätzliche Therapie durchführten, welche Therapie in Anspruch genommen wurde und ob sich diese Patienten von denen ohne weitere Behandlung zu Post unterschieden haben. Diese Variablen könnten dann als Prädiktoren in meta-analytische Regressionsanalysen eingesetzt werden. Zum anderen könnte eine Änderung des methodischen Designs von Therapiestudien erfolgen. Nach Erarbeitung einer hohen Compliance des Patienten könnte eine Vereinbarung getroffen werden, dass wenn sich der Patient entschließt, eine weitere Therapie zu beginnen, vor Beginn der Behandlung eine Erhebung der Symptomatik erfolgt, im einfachen Fall mit Hilfe von Selbstbeurteilungsinstrumenten, und diese Informationen der Forschergruppe bereitgestellt werden.

Im Sinne einer Intent to treat-Analyse könnten diese Daten in die Berechnung der Ergebnisse zu Follow up einfließen. Eine Überprüfung der Durchführung dieses Konzeptes wäre bei künftigen Untersuchungen zu empfehlen.

Bislang liegen nur sehr wenige Psychotherapiestudien vor, die eine Analyse der Ergebnisse mit einem Intent to treat-Ansatz durchführten. Auch wenn in der vorliegenden Arbeit sich bei keiner der Störungen ein bedeutsamer Unterschied zwischen Completer- und ITT-Analyse für die Pharmakotherapie zeigte, sollte diese Überprüfung auch für Psychotherapiestudien erfolgen. Eine Übertragbarkeit der Ergebnisse der Pharmako- auf Psychotherapiestudien erscheint eingeschränkt. Wie bereits erwähnt könnte eine Ursache für die gleichen Effektstärken von Completer- und ITT-Analyse eine Ausmittelung der Effektstärken bei ITT sein, basierend auf einer unterschiedlichen Abbrecherrate von Kontroll- und Behandlungsgruppe. Wenn in der Kontrollgruppe im Vergleich zur Behandlungsgruppe ein hoher Prozentsatz an Patienten die Studie abbricht, könnten sich höhere Effektstärken ergeben, bei einem höheren Prozentsatz in der Behandlungsgruppe dagegen niedrigere. In Pharmakotherapiestudien lagen sowohl Studien vor, in denen die Dropout-Rate höher war für die Kontroll- als auch für die Behandlungsgruppe. In Psychotherapiestudien mit Wartelistendesign zeigte sich dagegen überwiegend eine höhere Abbrecherquote für die Behandlungsgruppe. Möglicherweise kann dies zu geringeren Effektstärken für die ITT-Analyse führen. Die in den letzten Jahren häufiger werdende ITT-Auswertung auch bei Psychotherapiestudien ist zu unterstützen.

### **Pharmakotherapiestudien**

Es existiert möglicherweise das Problem, dass auch die Ergebnisse der Pharmakotherapiestudien zu progressiv sind. Ein großes Problem bei der Durchführung medikamentöser Untersuchungen ist der Aspekt, inwieweit die intendierte Doppelblindheit tatsächlich aufrechterhalten wurde. Es liegen mehrere Studien vor, die zeigten, dass sowohl Patienten als auch Rater bzw. Ärzte mit einer Häufigkeit, die über den Zufall hinausgeht, einschätzen können, ob der Patient eine aktive Substanz oder Placebo erhalten hatte (z.B. Bakker, Spinhoven, van Balkom, Matser & van Dyck, 1999; Basoglu, Marks, Livanou & Swinson, 1997; Margraf et al., 1991). Mögliche Einschränkungen der internen Validität ergeben sich aus einer reduzierten positiven Erwartungs-

---

haltung der Placebogruppe<sup>49</sup>, die somit eine unangemessene Realisierung gemeinsamer Faktoren darstellt sowie einem Versuchsleitereffekt, so dass insgesamt die Effektivität der Behandlung potentiell überschätzt wird. Möglichkeiten im Umgang mit diesem Problem diskutieren Even, Siobud-Dorocant und Dardennes (2000), z.B. eine unabhängige Erhebung der Zielkriterien von den Nebenwirkungen (in den hier eingeschlossenen Studien nur von einem geringem Prozentsatz erfüllt) sowie eine systematische Überprüfung der Blindheit. Letzteres könnte im Rahmen von Meta-Analysen die Grundlage für Prädiktoren bilden.

Ein weiterer Aspekt, der nicht in die Berechnung der Effektstärken eingegangen ist, sondern nur indirekt über die Abbrecherrate in der vorliegenden Arbeit betrachtet wurde, sind Quantität und Qualität der Nebenwirkungen der Medikation. Diese stellen allerdings einen wichtigen Punkt im Bereich der Lebensqualität dar und würden bei Bewertung diese Effektstärke reduzieren<sup>50</sup>. Eine Berechnung einer Effektstärke für Nebenwirkungen setzt eine Standardisierung von Erhebung und Präsentation voraus; sind z.B. nur die Symptome angegeben, die Patienten berichteten, würde die Effektstärke unangemessen berechnet werden, da potentielle Beeinträchtigungen, die z.B. andere Substanzklassen hervorrufen, unbeachtet blieben und nicht mit Null eingingen. Zudem wird in den Publikationen deutlich, dass vorrangig das Vorhandensein eines Symptoms überprüft wird, nicht aber dessen Schwere. Aber nur beide Dimensionen ermöglichen angemessene Aussagen über eine Beeinträchtigung. Insgesamt ergibt sich die Frage, ob Nebenwirkungen medikamentöser Therapie in Kauf genommen werden sollten, wenn die hier untersuchten psychotherapeutischen Verfahren ähnliche Effekte in der Behandlung der Angststörungen zeigen.

---

<sup>49</sup> Der Einfluss dieses Aspektes scheint allerdings weniger häufig gegeben. So erfolgte zwar bei Bakker et al. (1999) eine korrekte Zuordnung der Patienten, wenn dieser eine aktive Substanz erhalten hatten, aber sowohl die Einschätzung von Arzt als auch Patient waren nur entsprechend dem Zufall korrekt, wenn der Patient der Placebogruppe zugehörte. Bei Basoglu et al. (1997) zeigte die Gruppenzugehörigkeit dagegen keinen Unterschied.

<sup>50</sup> Eine Einschränkung der Lebensqualität ist allerdings potentiell auch für Psychotherapie gegeben, z.B. hinsichtlich des Zeitaufwandes incl. Hausaufgaben. Eine systematische Erhebung möglicher als negativ erlebter Einschränkungen wäre wünschenswert. Es stellt sich die Frage, inwieweit diese Aspekte mit den Nebenwirkungen einer Medikation zu vergleichen sind.

### **Kombinationsbehandlung**

Es ergaben sich spezielle methodische Probleme für Studien, die Psychotherapie und Kombinationsbehandlung direkt miteinander verglichen. So wurde nur selten die Integrität der Psychotherapie von Dritten geprüft, im Besonderen aber für diesen Forschungsbereich wichtig, um den Einfluss einer möglichen Loyalität des Wissenschaftlers gegenüber einer Behandlung zu verringern. Oft wurden Psychotherapie und Pillenplacebo zusammen als Vergleichsgruppe untersucht. Es ergeben sich zu einer Einschränkung in der externen Validität als auch in der internen Validität, da nicht alle potentiellen negativen Effekte der Kombinationsbehandlung (vgl. z.B. Westra et al., 1998) kontrolliert werden können. Hierzu zählen zum einen potentiell negative Einflüsse von Attributionsprozessen. Der erlebte Therapieerfolg könne auf die verabreichte Substanz attribuiert werden und es können sich in diesem Zusammenhang mögliche Einschränkungen in der Selbstwirksamkeit ergeben. Auch das Erleben eines aktiven Problemlöseprozesses könne eingeschränkt sein. Speziell bei Panikstörung, die charakterisiert ist durch eine Katastrophisierung körperlicher Empfindungen, könne während der Gabe des Pillenplacebo eine Attribution physiologischer Symptome als Nebenwirkung erfolgen und ein Lernprozess, dass diese nichts gefährliches darstellen, verhindert werden. Des Weiteren ist zu beachten, dass die verabreichte Substanz eine Art des Sicherheitsverhaltens darstellt. Der negative Einfluss dieser Aspekte im Vergleich zu reiner Psychotherapie könnte sich vor allem im Follow up zeigen. Weitere Studien sind somit zu fordern.

### **Generalisierbarkeit der Ergebnisse**

Ein Problem der vorliegenden Arbeit ist eine möglicherweise eingeschränkte Generalisierbarkeit auf die therapeutische Situation in der Praxis. Evaluationsstudien lassen sich in ein Kontinuum von Efficacy- bis Effectiveness-Studien einordnen. Seligman (1995) definierte Efficacy-Studien über eine randomisierte Zuteilung, kontrolliertes Design, Manualisierung mit Überprüfung der Integrität, eine feste Dauer der Therapie, operationalisierte Zielkriterien, blinde Rater, Patienten ohne komorbide Störung sowie das Vorliegen eines Katamnesezeitraums. Das Gegenstück sind Effectiveness-Studien, in denen die Therapie überprüft wird, wie sie im klinischen Setting durchgeführt wird. Dazu liegen keine exakten Kriterien vor. Während erstere somit die Gewichtung auf die interne Validität legen, ist bei letzteren die externe Validität erhöht. In der vorliegenden

Arbeit sind, aufgrund der eingesetzten Selektionskriterien, fast alle eingeschlossenen Studien sehr stark dem Pol der Efficacy-Studien zuzuordnen. Ist trotzdem eine Generalisierung der Ergebnisse angemessen?

In mehreren Meta-Analysen wurde der Einfluss einzelner Variablen oder Variablenkonstellationen, die wichtig für die Unterscheidung der beiden Konzepte sind, auf die Wirksamkeit einer Therapie überprüft. Matt et al. (1997) fanden bei acht Meta-Analysen, die den Einfluss des Settings (universitär vs. klinisch) untersuchten, bei fünf keinen bedeutsamen Zusammenhang, bei zwei eine signifikant höhere Effektstärke für das klinische Setting und bei einer Meta-Analyse eine signifikant höhere Effektstärke für das universitäre Setting. Die Erfahrung der Therapeuten zeigte nur in einer von 15 Meta-Analysen einen bedeutsamen (positiven) Zusammenhang zur Effektstärke. In keiner von acht Meta-Analysen ergab sich ein bedeutsamer Einfluss der Schwere der Störung. Shadish et al. (1997) erhielten nach Sekundäranalyse von Daten aus 15 anderen Meta-Analysen eine leicht geringere Effektstärke für klinisch repräsentative Therapien<sup>51</sup> von 0.52 im Vergleich zu einer Mittelung über alle Therapiestudien mit 0.59. Neben einer sehr starken Heterogenität der zugrunde liegenden Studien waren die Ergebnisse allerdings auch abhängig von methodischen Eigenschaften der beteiligten Meta-Analysen, wie die unterschiedliche Berechnung der Effektstärken. In einer methodisch sehr anspruchsvollen Meta-Analyse untersuchten Shadish und Kollegen (Shadish et al., 2000) Unterschiede in der Wirksamkeit von Efficacy- und Effectiveness-Studien an 90 Untersuchungen mit einer psychotherapeutischen Intervention bei psychologischen, behavioralen oder emotionalen Problemen. Sie fanden in einer multiplen Regression unter Kontrolle verschiedener methodischer Variablen (z.B. Therapiedauer, Art der Aufteilung der Patienten) keinen Einfluss der klinischen Reprä-

---

<sup>51</sup> Klinisch repräsentative Therapien waren definiert als: nicht-universitäres Setting, Patienten nicht durch Forscher angesprochen, erfahrene und professionelle Therapeuten, keine Manualisierung der Therapie und keine Überwachung der therapeutischen Integrität. In der Planung der Studie wurde eine weitere Stufe eingeführt, für die zusätzlich Patienten mit heterogenen Charakteristika (z.B. Herkunft, Alter) und Problemen sowie Therapeuten ohne spezielles Training gefordert wurde. Allerdings lag keine (!) randomisierte Studie vor.

---

sentativität<sup>52</sup> auf die Effektstärke. Dies spricht – mit aller Vorsicht – für die Möglichkeit einer Generalisierbarkeit<sup>53</sup> der vorliegenden Ergebnisse auch auf die therapeutische Situation in der Praxis, eine Annahme allerdings, die nicht anhand der eingeschlossenen Studien überprüft werden kann.

Studien direkt aus dem Bereich der Angststörungen, die spezifische Aspekte der Differenzierung von Efficacy- und Effectiveness-Studien untersuchten, fanden, dass es Unterschiede in der Effektivität geben kann aber nicht muss. Eine wichtige Variable sind Eigenschaften der Stichprobe, Kritiker von Efficacy-Studien argumentieren, dass die Ergebnisse nicht übertragbar seien, da es sich hier um bedeutend weniger schwer gestörte Patienten handele, da komorbide Störungen meist als Exklusionskriterium dienen. Welcher Zusammenhang besteht zwischen Komorbidität und Therapieerfolg? Nur ein geringer Prozentsatz der in die vorliegende Arbeit eingeschlossenen Studien untersuchte Patienten mit komorbiden Störungen, eine direkte Überprüfung war nicht möglich. In einem Review für Panikstörung fanden Mennin und Heimberg (2000) sowohl Studien, in denen die Diagnose einer Major Depression einen (negativen) Einfluss auf den Status eines Patienten zu Post hatte, als auch Studien, in denen sich keine differentielle Wirksamkeit zeigte, die Ergebnisse sind somit nicht eindeutig, möglicherweise aber auch aufgrund methodischer Unterschiede in den Studien. Die Autoren sehen dagegen einen negativen Einfluss von Persönlichkeitsstörungen auf die Effektivität einer Behandlung, wobei auch hier einige der Studien keinen Zusammenhang zu einem schlechteren Ergebnis zeigten. Dagegen schlussfolgern Dreessen und Arntz (1998) in einem Review an 15 Studien, dass kein genereller negativer Einfluss einer Persönlichkeitsstörung auf den Therapieerfolg bei Angststörungen bestätigt werden kann. Es ist festzustellen, dass der vermutete negative Zusammenhang zwischen Komorbidität und Therapieerfolg aufgrund der vorliegenden Daten nicht eindeutig gestützt werden kann. Die Ergebnisse der vorliegenden Meta-Analyse scheinen somit,

---

<sup>52</sup> Klinische Repräsentativität wurde operationalisiert über 10 Kriterien und war umso höher, wenn z.B. die Therapeuten praktizierende Kliniker bzw. Professionelle waren, keine Überprüfung der Durchführung der Behandlung stattfand oder eine flexible Anzahl an Sitzungen eingesetzt wurde.

<sup>53</sup> Generalisierbarkeit steht hier für Transfer und meint eine Extrapolation von Beobachtungen, die in einer bestimmten Bedingung erfolgten auf eine andere Bedingung (Matt, in press).

allerdings nicht anhand der eingeschlossenen Studien überprüft, auch auf Patienten mit komorbiden Störungen generalisierbar zu sein.

#### **4.2.2 Meta-Analytische Besonderheiten**

In der vorliegenden Arbeit erfolgte eine Anwendung aktueller statistischer Methoden meta-analytischer Verfahren<sup>54</sup>. Dazu zählte insbesondere im Vergleich zu bisherigen Meta-Analysen im Bereich der Therapieforschung der Einsatz des Random Effects-Modell zur Generalisierbarkeit der Ergebnisse, der Einsatz multivariater Verfahren einschließlich einer Gewichtung mit der Stichprobengröße sowie eine Anwendung von Sensitivitätsanalysen zur Überprüfung der Stabilität der Ergebnisse. Es wurde eine Kontrolle der Vergleichbarkeit der Gruppen zu Prä durchgeführt und ein Vergleich verschiedener Therapieformen über eine Anpassung der „Effektivität“ der Kontrollgruppe. Neben den bereits diskutierten Einschränkungen aufgrund Charakteristika der Primärstudien zeigten sich auch methodische Grenzen der Arbeit.

Als Problem ergab sich der Umgang mit fehlenden Informationen der Primärstudien. Hinsichtlich fehlender Daten zur Berechnung der Effektstärken wurde der Einfluss von Berechnungsalternativen auf den mittleren Effekt über eine Sensitivitätsanalyse geprüft. Ray et al. (1996) konnten zeigen, dass für gleiche Datensätze verschiedene Berechnungsalternativen zu teilweise sehr unterschiedlichen Effektstärken führen können. In der vorliegenden Arbeit ergaben sich meist nur geringe Veränderungen der Ergebnisse bei Ausschluss von Studien, in denen die Berechnung nicht über Mittelwerte und Standardabweichungen erfolgte. Die Ergebnisse scheinen somit größtenteils stabil hinsichtlich der eingeschlossenen Studienpopulation. Dagegen waren bei Panikstörung für den Vergleich von konfrontativen mit kognitiv-behavioralen Verfahren die Ergebnisse abhängig von den eingeschlossenen Studien: betrachtete man nur Untersuchungen, die Mittelwerte und Standardabweichungen präsentierten, zeigte sich für die vorliegenden Studien eine bedeutsam höhere Effektivität für kognitiv-behaviorale Verfahren,

---

<sup>54</sup> Es sei erinnert, dass einige der aktuellen Techniken nicht eingesetzt wurden, da die notwendigen Informationen zur Berechnung nicht gegeben waren. Dazu zählte eine Korrektur der Effektstärke von systematischen Artefakten nach Hunter et al. (1990) sowie Verfahren zum Umgang mit stochastisch abhängigen Werten (z.B. Gleser & Olkin, 1994).

nicht dagegen, wenn sämtliche Untersuchungen eingeschlossen wurden. Fünf der ausgeschlossenen Vergleiche waren Effektstärken von Null aufgrund nicht-signifikanter Ergebnisse in den Primärstudien. Es wird die Unsicherheit der Ergebnisse von Meta-Analysen deutlich, dies spricht für den Einsatz des Random Effects-Modells, welches konservativer ist. Im Random Effects-Modell waren in beiden Fällen die Effektstärken nicht bedeutsam voneinander verschieden. Für die medikamentöse Behandlung der Zwangsstörung ergab sich dagegen eine bedeutsam höhere Effektstärke nach Ausschluss von Studien, in denen die Berechnung nicht über Mittelwerte und Standardabweichungen möglich waren – auch im Random Effects-Modell. Es zeigte sich zudem eine Konfundierung mit der untersuchten Substanzklasse: vor allem Studien mit SSRI mussten in dieser Sensitivitätsanalyse ausgeschlossen werden. Insgesamt muss festgestellt werden, dass eine Antwort, welches die „richtigen“ Ergebnisse sind, nicht gegeben werden kann. Soll man die Genauigkeit der Effektstärkenberechnungen oder die Repräsentativität der Stichprobe stärker gewichten? Zudem ist zu beachten, dass möglicherweise auch ein Zusammenhang bestehen kann zwischen Ergebnis einer Studie und Umfang der Publikation und somit Raum für Mittelwertstabellen. Die zum Teil beobachteten Differenzen könnten sich also ergeben aus künstlichen Unterschieden aufgrund verschiedener Berechnungen sowie wahren Unterschieden. Die Notwendigkeit von Sensitivitätsanalysen wird deutlich.

Eine weitere Einschränkung ergibt sich aus fehlenden Werten für Prädiktoren. Entsprechende Studien wurden aus den Analysen ausgeschlossen. Es ist keine potentielle Verzerrung gegeben, geht man davon aus, dass diese Studien eine zufällige Stichprobe aus der Gesamtheit an Untersuchungen darstellen (Pigott, 2001). Insbesondere hinsichtlich der Exklusion von Studien mit fehlender Abbrecherrate scheint aber diese Annahme verletzt, es ist möglicherweise gegeben, dass in diesen Untersuchungen entweder kein Patient oder sehr viele Patienten die Behandlung vorzeitig verließen. Pigott (2001) stellt statistische Möglichkeiten zur Schätzung fehlender Werte vor, z.B. Maximum Likelihood Methoden. Sie verdeutlicht das positive Potential dieser Verfahren, zeigt aber auch Einschränkungen auf. So sind Zusatzannahmen hinsichtlich der Verteilung der Daten notwendig, spezielle meta-analytische Aspekte, wie der Einbezug der Gewichtung mit der Genauigkeit der Ergebnisse, bleiben unberücksichtigt und es sind ausreichend große Datensätze erforderlich. Weitere Simulationsstudien erscheinen somit notwendig zur Einschätzung der Angemessenheit der Verfahren für die Meta-



Analyse. Aufgrund der Anzahl an Studien insbesondere für die Analysen der Psychotherapie hätte zudem kein kontinuierlicher Einsatz der Methoden in der vorliegenden Arbeit erfolgen können.

Die Anzahl an Studien für eine stabile Schätzung des Random Effects-Modell war in einigen Analysen gering. Auch die Power der durchgeführten Regressionen erscheint teilweise gering. Die Anzahl an Studien im Verhältnis zur Anzahl an Prädiktoren war in einigen Analysen sehr klein.

Trotz vielfältiger Bemühungen für unpublizierte Studien war die Anzahl in der vorliegenden Arbeit sehr gering. Schloss man zudem Material aus, welches in naher Zukunft erscheinen soll, konnten nur vier Untersuchungen lokalisiert werden, die den Selektionskriterien entsprachen: zwei Dissertationen, ein Kongressbeitrag sowie ein unveröffentlichter Forschungsbericht. Die Ergebnisse könnten somit zu progressiv sein beachtet man Daten bisheriger Meta-Analysen (z.B. Lipsey et al., 1993). Eine Evaluation des Publikationsbias erfolgte über Fail-Safe N. Vor allem hinsichtlich der Vergleiche zu einer Kontrollgruppe scheinen die Ergebnisse stabil. Für direkte Vergleiche zwischen mehreren Therapien war die Anzahl an benötigten Studien zur substantiellen Veränderung der Effektstärke dagegen gering. Es ergibt sich für diese Methode das Problem, dass die Definition einer Grenze bzgl. der Stabilität der Ergebnisse immer subjektiv ist. Aufgrund der hier eingesetzten Selektionskriterien ist aber aufgrund der relativ hohen Kosten für entsprechende Studien möglicherweise von einer relativ geringen Anzahl an Untersuchungen in Schubladen aus der Praxis auszugehen. Eine wie in vielen Kliniken übliche standardisierte Erhebung der Symptomatik zur Evaluation der durchgeführten Therapie würde in dieser Meta-Analyse ausgeschlossen, da keine Kontrollgruppen oder Vergleichstherapien realisiert werden. Auch hinsichtlich eines Vergleiches von Psychotherapie mit Pharmakotherapie oder Kombinationsbehandlung scheint eine Vielzahl dieser Daten nicht für die vorliegende Meta-Analyse angemessen, da oft während der Behandlungszeit eine Adaption der Strategie erfolgt, wie eine Veränderung des Medikationsstatus (z.B. Althaus, Zaudig, Hauke, Röper & Butollo, 2000). Andererseits können natürlich Untersuchungen in ungewissem Umfang existieren, die, entgegen guter wissenschaftlicher Praxis, auch auf Anfrage zurückgehalten werden aufgrund negativer, einschließlich nichtsignifikanter, Ergebnisse. Bemühungen für eine internationale Datenbank, in der Studien bei Beginn der Erhebung und somit unabhängig vom

---

Ergebnis registriert werden, sollten forciert werden. Auch die Entwicklung statistischer Modelle zur Kompensation des Publikationsfehlers sollte weiter erfolgen (z.B. Vevea & Hedges, 1995), sowie deren Überprüfung von unabhängigen Forschergruppen.

### **Vergleich verschiedener Therapien zwischen Studien**

In der vorliegenden Meta-Analyse erfolgte ein Vergleich von Psycho- und Pharmakotherapien aus verschiedenen Studien. Grundlage bildete eine Anpassung der Effektstärken aufgrund unterschiedlicher Kontrollgruppen über die Prä-Post-Effektstärken. Eine empirische Exploration ergab, dass dieses Vorgehen angemessen erscheint unter der zusätzlichen Annahme eines gleichen Retest-Effektes. Einschränkungen in der Interpretation der Ergebnisse ergeben sich aufgrund eines grundlegenden Problems des Verfahrens der Meta-Analyse: „The conditions under which those studies were conducted ... were not assigned to studies at random. Therefore, those choices are probably confounded with each other in mostly unknowable but likely systematic ways.” (Shadish, 1996, S. 60)<sup>55</sup>. Lipsey und Wilson (2001) sprechen dabei auch von der Persönlichkeit von Evaluationsstudien. Über eine multiple Regression wurden potentielle konfundierende Variable einbezogen, allerdings war die Auswahl dieser Variablen eingeschränkt. Das eingesetzte Random Effects-Modell berücksichtigt zwar diesbezügliche Unsicherheiten. Einer der nach Ansicht der Autorin problematischsten Aspekte ist aber ein möglicher systematischer Unterschied in den Messinstrumenten<sup>56</sup>, besonders deutlich und im Folgenden dargestellt am Beispiel der Generalisierten Angststörung. Zum einen sind verschiedene Instrumente zur Messung des gleichen Konstruktes unterschiedlich änderungssensitiv: „For now, however, it is clear that dependent measures are not equivalent in their tendencies to reflect change and that meta-analysis, as it is typically used to combine different measures, cannot overcome the differences between measures“ (Hervorhebung entfernt d. Verf., Lambert & Hill, 1994, S. 84). So wurden in einem meta-analytischen Vergleich höhere Effektstärken für

---

<sup>55</sup> Die korrelative Natur der Daten in Meta-Analysen ist eine generelle Einschränkung des Verfahrens.

<sup>56</sup> Dagegen erscheinen Unterschiede in der Stichprobe weniger bedeutsam. Potentielle Differenzen in der Störung wurden kontrolliert über das diagnostische System. Die Anzahl an Studien mit Patienten, die komorbide Störungen aufwiesen, war insgesamt eher gering. Die zusätzliche Medikation bei Psychotherapiestudien wurde schon diskutiert. Die in den Primärstudien eingesetzten Selektionskriterien sind ansonsten sehr ähnlich.

---

die HAMD im Gegensatz zum BDI für die Therapie der Depression gefunden (1.57 vs. 1.16; n. Lambert et al., 1994). In den eingeschlossenen Studien bei Generalisierter Angststörung setzten ca. die Hälfte der Psychotherapiestudien das BDI ein, aber keine der Pharmakotherapiestudien, hier wurde überwiegend die HAMD verwandt. Auch bei gleicher Effektivität der Therapien könnten somit Studien mit BDI (und somit Psychotherapiestudien) geringere Effektstärken zeigen. Der Einfluss einer unterschiedlichen Änderungssensitivität der eingesetzten Instrumente kann auch für die Kategorie Angst nicht ausgeschlossen werden. Zudem wurden die verschiedenen Aspekte des Konstruktes Angst mit unterschiedlich starker Gewichtung gemessen. In Pharmakotherapiestudien erfolgt eine Erfassung der Störung meist über die HAMA. Obwohl diese sowohl somatische als auch kognitive Charakteristika misst, werden letztere, entsprechend der Anzahl an Items, weniger stark gewichtet als den Kriterien des DSM-III-R bzw. DSM-IV entsprechen würde. In Psychotherapiestudien werden hierzu zum Teil spezielle Instrumente zur Erfassung des Sich Sorgens eingesetzt (z.B. Penn State Worry Questionnaire, Meyer et al., 1990), dies ist künftig auch von medikamentösen Untersuchungen zu fordern. Unterschiedliche Effektstärken zwischen den Verfahren können sich somit nur aufgrund einer unterschiedlichen Auswahl von Messinstrumenten ergeben. Die Ergebnisse sind somit beeinflusst vom „Apfel und Birnen“-Problem: es werden Effektstärken miteinander verglichen, die auf verschiedenen Konzepten basieren. Eine Lösung innerhalb von Meta-Analysen erscheint nicht möglich. Wählt man als Selektionskriterium den Einsatz einer speziellen Instrumentenbatterie sinkt die Anzahl an Studien drastisch, die Repräsentativität erscheint verletzt. Es sind Bemühungen zu unterstützen für einen einheitlichen Standard für Erhebungsinstrumente in Therapiestudien.

Insbesondere von Klein (1996, 1997) wurde gefordert, dass sowohl Psychotherapiestudien als auch Psycho- mit Pharmakotherapie vergleichende Untersuchungen zukünftig eine Pillenplacebobedingung einsetzen sollten. Dies basiere auf der Notwendigkeit zu zeigen, dass die untersuchte Stichprobe „was indeed a medication-responsive sample“ (Klein, 1996, S. 83). Zudem seien Studien, die Psychotherapie nur mit einer Wartelistenbedingung verglichen „uninterpretably biased“ (Klein, 1996, S. 86)<sup>57</sup>. Dies

---

<sup>57</sup> Weitere Vorteile einer Pillenplacebobedingung im Rahmen von Vergleichsstudien nennen Jacobson und Hollon (1996): eine mögliche Steigerung der internen Validität aufgrund der Überprüfbarkeit einer möglichen geringen Änderungssensitivität der eingesetzten Messinstrumente bzw. bei nicht-signifikanten

resultiere aus der Annahme, dass unspezifische Faktoren, wie Erwartungen oder Reaktionen auf die Platzierung in die verschiedenen Gruppen, nicht im Wartelistendesign wohl aber bei Einsatz von Pillenplacebo angemessen kontrolliert würden. Es folgte eine leidenschaftliche Diskussion über diese Forderung von Klein (1996, 1997), die hier nur fragmentarisch dargestellt werden soll. Die wichtigsten Kritikpunkte betrafen:

- a) McNally (1996) legte dar, dass ein fehlender Beleg für eine Effektivität der Pharmakotherapie in einer Studie nicht unbedingt auf eine mangelnde Repräsentativität der Stichprobe zurückzuführen sei, sondern auf eine fehlende Wirksamkeit der Substanz. Könnte Psychotherapie zudem eine Wirksamkeit unabhängig von der untersuchten Stichprobe zeigen, wäre eine effektivere Behandlung gefunden (Heimberg, 1997). D.h. Psychotherapie ist nicht nur dann der Pharmakotherapie überlegen, wenn eine höhere Effektivität bei einer auf Medikation ansprechenden Stichprobe gezeigt wurde, sondern auch selbst dann, wenn es bei dieser Stichprobe keinen bedeutsamen Unterschied gibt, aber Psychotherapie auch bei einer Stichprobe wirkt, die nicht auf Medikation anspricht. Hier zeigt sich der Vorteil von Meta-Analysen bzgl. einer Ergebniszusammenfassung über verschiedene Stichproben.
- b) Die angemessene Kontrolle unspezifischer Faktoren in einer Pillenplacebo-bedingung erscheint problematisch. Es stellt sich das Problem, inwieweit Patienten dieser Gruppe tatsächlich gegenüber der eingenommenen Substanz blind seien (Heimberg, 1997; vgl. Ausführungen oben) und sich dementsprechend Erwartungen bzgl. einer Verbesserung zwischen aktiver und inaktiver Substanz vergleichen lassen.
- c) Auch im Rahmen von Therapieplacebobedingungen könne eine angemessene Kontrolle unspezifischer Faktoren realisiert werden (Crits-Christoph, 1997; Heimberg, 1997). So fanden z.B. Bowers und Clum (1988) in einer Subgruppe von 16 Studien aus ihrer Meta-Analyse, die die Glaubwürdigkeit von Behandlung und Therapieplacebo überprüften, dass in 88 % der Studien beide Gruppen gleich glaubwürdig eingeschätzt wurden und somit Aspekte wie positive Erwartungshaltung in der Kontrollgruppe nicht eingeschränkt sind. In der

---

Unterschieden zwischen Pharmako- und Psychotherapie ist die Klärung der Frage, ob beide Verfahren effektiv oder ineffektiv waren.

vorliegenden Arbeit zeigten sich über die Prä-Post-ESn von Pillen- und Therapieplacebo sowohl bei Sozialer Phobie als auch bei Panikstörung vergleichbare Veränderungen in beiden Bedingungen mit leicht höheren Effektstärken für das Therapieplacebo. Obwohl die bereits diskutierten Einschränkungen zum Vergleich von Daten aus verschiedenen Untersuchungen auch hier gelten, scheint somit innerhalb des eingesetzten Therapieplacebos eine vergleichbare Realisierung gemeinsamer Faktoren wie bei Pillenplacebo zu gelten.

Von Seiten der Meta-Analyse ist durchaus ein *zusätzlicher* Einschluss einer Pillenplacebobedingung im Rahmen von Psychotherapiestudien (neben einer Warteliste) wünschenswert, allerdings im gleichen Ausmaß die *zusätzliche* Betrachtung einer Wartelistenbedingung bei placebokontrollierten Pharmakotherapiestudien. Dies basiert im Vergleich zu Klein (1996, 1997) nicht auf der Annahme, dass Studien ohne diese Gruppen nicht interpretierbar wären, sondern auf der darauf aufbauenden Möglichkeit, besser die unterschiedlichen Effekte von Warteliste und Pillenplacebo schätzen zu können. Würden mehr Studien mit beiden Kontrollgruppen vorliegen, wäre die Aussagekraft zukünftiger Meta-Analysen zum Vergleich von Pharmako- und Psychotherapie aus verschiedenen Studien größer, da die Stärke der unterschiedlichen Kontrollbedingungen direkt gemessen werden könnte. Dennoch sind auch hier spezifische Einschränkungen zu beachten. Psychotherapiestudien müssten eine Änderung der Stichprobenselektion hinsichtlich der bereits diskutierten Einnahme zusätzlicher Medikation durchführen. Würde eine Pillenplacebobedingung eingesetzt, dürften Patienten keine Psychopharmaka einnehmen. Zudem ist die Annahme notwendig, dass die Ergebnisse auf Stichproben früherer Kohorten generalisiert werden können. Es bleibt weiterhin der Vorteil von Studien, die beide Verfahren direkt miteinander vergleichen, dass deren Ergebnisse von Variablen des Designs oder der Stichprobe weniger beeinflusst sind. Dennoch würde das vorgeschlagene Design eine fehlerfreiere Integration der Vielzahl an bisherigen Einzelergebnissen aus isolierten Pharmako- und Psychotherapiestudien ermöglichen.

**Abschließende Bemerkungen**

Meta-Analysen sind ein sehr wertvolles Verfahren. Sie ermöglichen generalisierbare Aussagen und die Untersuchung von Fragestellungen, die in Primärstudien bislang noch nicht oder nur in geringem Maße betrachtet wurden. Allerdings sind die Ergebnisse von Meta-Analysen keinesfalls so sicher, wie sie oft präsentiert werden.

## 5 Zusammenfassung

Angststörungen zählen zu den häufigsten psychischen Störungen, Aussagen zur Wirksamkeit von Therapien haben somit einen besonderen Stellenwert. In der vorliegenden Arbeit wurde die Effektivität von behavioralen, kognitiv-behavioralen und medikamentösen Behandlungen untersucht. Dazu wurden die Ergebnisse bisheriger Evaluationsstudien quantitativ integriert mit der Methode der Meta-Analyse.

*Methode:* Über eine umfassende Literatursuche wurden potentiell relevante kontrollierte Therapiestudien lokalisiert. Es konnten insgesamt 340 Studien mit über 30 000 Patienten in die Meta-Analyse eingeschlossen werden. Als Effektstärke wurde Hedges  $g$  berechnet. Zur Integration der einzelnen Effektstärken wurde das Random Effects-Modell eingesetzt. Es ermöglicht die Betrachtung einer zusätzlichen Varianzkomponente, die ein Maß für die ungeklärte Heterogenität der Primärstudien darstellt. Das Modell erlaubt Generalisierungen über die eingeschlossenen Untersuchungen hinaus. Der Vergleich verschiedener Therapieformen erfolgte zum einen an Studien, die verschiedene Verfahren untersuchten, als auch zwischen Studien über eine Anpassung der „Wirksamkeit“ der Kontrollbedingungen. Dieses Vorgehen ermöglichte zudem eine Einschätzung des Verhältnisses von gemeinsamen und spezifischen Faktoren in der Effektivität der Therapien. Zusätzlich wurden im Vergleich der Therapien verschiedene methodische Variablen kontrolliert. Ein weiterer Aspekt, der untersucht wurde, war die Akzeptanz der Patienten für die Behandlungen, operationalisiert über die Abbrecherrate.

*Ergebnisse:* Behaviorale und kognitiv-behaviorale Therapien sind sehr effektiv in der Behandlung der Angststörungen. Die Wirksamkeit der Verfahren geht dabei, zumindest für die Hauptsymptomatik, über eine Realisierung gemeinsamer Faktoren hinaus. Die hohe Wirksamkeit der Behandlungen war zeitlich stabil. Es ergab sich für keine der Angststörungen eine Verschlechterung in der Symptomatik nach Therapieende. Konfrontative und kognitiv-behaviorale Therapien scheinen in der Behandlung der Angstsymptomatik eine vergleichbare Effektivität zu haben. Unterschiede ergaben sich nur in der Behandlung der depressiven Symptomatik bei Panikpatienten mit höherer Effektivität der kognitiv-behavioralen Verfahren. Die untersuchten psychologischen Therapien wurden von den Patienten gut akzeptiert, bei keiner Angststörung lagen die Abbrecherraten über 20 %. Auch pharmakologische Therapien waren effektiv für

Patienten mit Angststörungen. Die Abbrecherraten betragen bis zu 34 %. Die Wirksamkeit der beiden Behandlungsansätze erscheint vergleichbar, die Akzeptanz der Patienten war allerdings häufig für psychologische Verfahren höher. Kombinationsbehandlungen können kurzfristig zu einer weiteren Steigerung der Effektivität führen.

*Schlussfolgerungen:* Die Ergebnisse zeigen, dass behaviorale und kognitiv-behaviorale Verfahren mindestens so effektiv wie Pharmakotherapie in der Behandlung der Angststörungen sind, aber zum Teil besser von den Patienten akzeptiert werden. Die positiven Wirkungen der Psychotherapie erscheinen zudem zeitlich stabil. Die Daten der vorliegenden Arbeit sprechen dafür, dass Patienten, die das Risiko von Nebenwirkungen nicht eingehen möchten, eine wirksame Alternative in der Therapie der Angststörungen nutzen können. Das Therapieangebot sollte dementsprechend angepasst werden; die gefundene Dominanz medikamentöser Behandlungen erscheint aufgrund der vorliegenden Daten nicht gerechtfertigt. Die aktuelle Arbeit zeigte allerdings auch, dass Meta-Analysten ein verstärktes Bewusstsein für dieses Verfahren entwickeln sollten. Methodische Entscheidungen hatten einen teilweise sehr hohen Einfluss auf die Ergebnisse. Einige der Resultate können somit noch nicht als stabil gelten. Es zeigten sich einige methodische Einschränkungen der bisherigen Evaluationsstudien. Empfehlungen für weitere Untersuchungen wurden gegeben.



## 6 Literatur

- Abramowitz, J.S. (1996). Variants of exposure and response prevention in the treatment of obsessive-compulsive disorder: A meta-analysis. *Behavior Therapy*, 27, 583-600.
- Ackerman, D.L. & Greenland, S. (2002). Multivariate meta-analysis of controlled drug studies for obsessive-compulsive disorder. *Journal of Clinical Psychopharmacology*, 22, 309-317.
- Althaus, D., Zaudig, M., Hauke, W., Röper, G. & Butollo, W. (2000). Wirksamkeit eines spezifisch für Zwangsstörungen entwickelten stationären Gruppentherapiekonzeptes bei gleichzeitiger Verhaltenstherapie und pharmakologischer Behandlung. *Verhaltenstherapie*, 10(1), 16-23.
- American Psychiatric Association (1968). *Diagnostic and Statistical Manual of Mental Disorders* (2<sup>nd</sup> ed.). Washington, DC: Author.
- American Psychiatric Association (1980). *Diagnostic and Statistical Manual of Mental Disorders* (3<sup>rd</sup> ed.). Washington, DC: Author.
- American Psychiatric Association (1987). *Diagnostic and Statistical Manual of Mental Disorders* (3<sup>rd</sup> ed., rev.). Washington, DC: Author.
- American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4<sup>th</sup> ed.). Washington, DC: Author.
- Angenendt, J., Frommberger, U. & Berger, M. (1998). Zum Stellenwert psychotherapeutisch-pharmakologischer Kombinationstherapien bei Angststörungen. *Verhaltenstherapie*, 8, 160-169.
- Antony, M.M. & Barlow, D.H. (2002). Specific phobias. In: D.H. Barlow, *Anxiety and its disorders: The nature and treatment of anxiety and panic*; 2<sup>nd</sup> ed. (S. 380-417). New York: The Guilford Press.
- Arrindell, W.A. (2001). Changes in waiting-list patients over time: Data on some commonly-used measures. Beware! *Behaviour Research and Therapy*, 39, 1227-1247.
- Baker, B.L., Cohen, D.C. & Saunders, J.T. (1973). Self-directed desensitization for acrophobia. *Behaviour Research and Therapy*, 11, 79-89.
- Bakker, A., Spinhoven, P., van Balkom, A.J.L.M., Matser, D. & van Dyck, R. (1999). Double-blindness procedure did not mask giving of medication in panic disorder. *Journal of Affective Disorder*, 54, 189-192.

- 
- Bakker, A., Spinhoven, P., van Balkom, A.J.L.M., Vleugel, L. & van Dyck, R. (2000). Cognitive therapy by allocation versus cognitive therapy by preference in the treatment of panic disorder. *Psychotherapy and Psychosomatics*, *69*, 240-243.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191-215.
- Barlow, D.H. (2002). *Anxiety and its disorders: The nature and treatment of anxiety and panic* (2<sup>nd</sup> ed.). New York: The Guilford Press.
- Basham, R.B. (1986). Scientific and practical advantages of comparative design in psychotherapy outcome research. *Journal of Consulting and Clinical Psychology*, *54*, 88-94.
- Basoglu, M., Marks, I., Livanou, M. & Swinson, R. (1997). Double-blindness procedures, rater blindness, and ratings of outcome. Observations from a controlled trial. *Archives of General Psychiatry*, *54*, 744-748.
- Bebbington, P.E. (1998). Epidemiology of obsessive-compulsive disorder. *British Journal of Psychiatry*, *173*(suppl.35), 2-6.
- Beck, A.T. & Clark, D.A. (1996). An information processing model of anxiety: automatic and strategic processes. *Behaviour Research and Therapy*, *35*(1), 49-58.
- Beck, A.T., Emery, G. & Greenberg, R.L. (1985). *Anxiety disorders and phobias: a cognitive perspective*. New York: Basic Books.
- Beck, A.T., Epstein, N., Brown, G. & Steer, R.A. (1988). An inventory for measuring clinical anxiety. *Journal of Consulting and Clinical Psychology*, *56*, 893-897.
- Becker, B.J. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology*, *41*, 257-278.
- Beelmann, A. & Bliesener, T. (1994). Aktuelle Probleme und Strategien der Metaanalyse. *Psychologische Rundschau*, *45*, 221-233.
- Begg, C.B. (1994): Publication Bias. In: H. Cooper & L.V. Hedges (Eds.), *The handbook of research synthesis* (S.399-410). New York: Russell Sage Foundation.
- Bergin, A.E. & Garfield, S.L. (1994). *Handbook of psychotherapy and behaviour change*. New York: Wiley.
- Bowers, T. & Clum, G. (1988). Relative contributions of specific and nonspecific treatment effects: Meta-analysis of placebo-controlled behavior therapy research. *Psychological Bulletin*, *103*, 315-323.
- Boyer, W. (1995). Serotonin uptake inhibitors are superior to imipramine and alprazolam in alleviating panic attacks: A meta-analysis. *International Clinical*

- 
- Psychopharmacology*, 10, 45-49.
- Brewin, C.R. (1996). Theoretical foundations of cognitive-behavior therapy for anxiety and depression. *Annual Review of Psychology*, 47, 33-57.
- Brown, T.A. & Barlow, D.H. (2002). Classification of anxiety and mood disorder. In: D.H. Barlow, *Anxiety and its disorders: The nature and treatment of anxiety and panic*; 2<sup>nd</sup> ed. (S. 292-327). New York: The Guilford Press.
- Brown, T.A., Campbell, L.A., Lehman, C.L., Grisham, J.R. & Mancill, R.B. (2001). Current and lifetime comorbidity of the DSM-IV anxiety and mood disorders in a large clinical sample. *Journal of Abnormal Psychology*, 110, 49-58.
- Campbell, D.T. & Stanley, J.C. (1966). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Carson, K.P. & Schriesheim, C.A.(1990). The usefulness of the „fail-safe“ statistic in meta-analysis. *Educational & Psychological Measurement*, 50(2), 233-240.
- Chambless, D.L. & Gillis, M.M. (1993). Cognitive therapy of anxiety disorders. *Journal of Consulting and Clinical Psychology*, 61(2), 248-260.
- Clark, D.B., Taylor, C.B., Roth, W.T., Hayward, C., Ehlers, E., Margraf, J. & Agras, W.S. (1990). Surreptitious drug use by patients in a panic disorder study. *American Journal of Psychiatry*, 147, 507-509.
- Clum, G.A., Clum, G.A. & Surls, R. (1993). A meta-analysis of treatments for panic disorder. *Journal of Consulting and Clinical Psychology*, 61(2), 317-326.
- Cohen, J. (1988). *Statistical power analysis for behavioral sciences* (2<sup>nd</sup> ed.). Hillsdale, NJ: Erlbaum.
- Cohen, J. (1994). The earth is round ( $p < .05$ ). *American Psychologist*, 49(12), 997-1003.
- Cook, T.D. & Campbell, D.T. (1979). *Quasi-experimentation: Design & analysis issues for field settings*. Boston: Houghton Mifflin.
- Cox, B.J., Swinson, R.P., Morrison, B. & Lee, P.S. (1993). Clomipramine, fluoxetine, and behavior therapy in the treatment of obsessive-compulsive disorder: A meta-analysis. *Journal of Behavior Therapy and Experimental Psychiatry*, 24(2), 149-153.
- Craske, M.G., Rapee, R.M., Jackel, L. & Barlow, D.H. (1989). Qualitative dimensions of worry in DSM-III-R generalized anxiety disorder subjects and nonanxious controls. *Behaviour Research and Therapy*, 27(4), 397-402.
- Creamer, M. (1993). Recent developments in posttraumatic stress disorder. *Behavior*

- Change*, 10, 219-227.
- Crits-Christoph, P., (1997). Control groups in psychotherapy research revisited. *Treatment*, 1, Comment 1, [http://journals.apa.org/treatment/vol1/97\\_c1.html](http://journals.apa.org/treatment/vol1/97_c1.html)
- Curtis, G.C., Magee, W.J., Eaton, W.W., Wittchen, H.-U., & Kessler, R.C. (1998). Specific fears and phobias: Epidemiology and classification. *British Journal of Psychiatry*, 173, 212-217.
- Davey, G.C.L. (1992). Classical conditioning and the acquisition of human fears and phobias: A review and synthesis of the literature. *Advances in Behaviour Research and Therapy*, 14, 29-66.
- Davidson, J.R.T., Book, S.W., Colket, J.T., Tupler, L.A., Roth, S., David, D., Hertzberg, M., Mellman, T., Beckham, J.C., Smith, R.D., Davison, R.M., Katz, R. & Feldman, M.E. (1997). Assessment of a new self-rating scale for post-traumatic stress disorder. *Psychological Medicine*, 27(1), 153-160.
- Dreessen, L. & Arntz, A. (1998). The impact of personality disorders on treatment outcome of anxiety disorders: Best-evidence synthesis. *Behaviour Research and Therapy*, 36, 483-504.
- Dupuy, J.-B., Beaudoin, S., Rhéaume, J., Ladoceur, R. & Dugas, M.J. (2001). Worry: Daily self-report in clinical and non-clinical populations. *Behaviour Research and Therapy*, 39(10), 1249-1255.
- Edelmann, R.J. & Baker, S.R. (2002). Self-reported and actual physiological responses in social phobia. *British Journal of Clinical Psychology*, 41, 1-14.
- Ehlers, A. & Steil, R. (1995). Maintenance of intrusive memories in Posttraumatic Stress Disorder: A cognitive approach. *Behavioural and Cognitive Psychotherapy*, 23, 217-250.
- Ehlers, A. (1999). *Posttraumatische Belastungsstörung*. Aus: Fortschritte der Psychotherapie, Band 8. Göttingen: Hogrefe.
- Erdfelder, E. (1983). Erzeugung und Verwendung empirischer Daten. In: J. Bredenkamp & H. Feger (Hrsg.), *Enzyklopädie der Psychologie: Themenbereich B Methodologie und Methoden Serie I Forschungsmethoden der Psychologie, Band 1 Methodologische Grundlagen der Psychologie* (S. 47-97). Göttingen: Hogrefe.
- Erez, A., Bloom, M.C. & Wells, M.T. (1995). *On a proper meta-analytic model for correlations* (Working Paper 95-11). Ithaca, NY: Cornell University.
- Even, C., Siobud-Dorocant, E. & Dardennes, R.M. (2000). Critical approach to

- antidepressant trials. *British Journal of Psychiatry*, 177, 47-51.
- Eysenck, H.J. (1978). An exercise in mega-silliness. *American Psychologist*, 33, 517.
- Federoff, I.C. & Taylor, S (2001). Psychological and pharmacological treatments of social phobia: A meta-analysis. *Journal of Clinical Psychopharmacology*, 21, 311-324.
- Feske, U. & Chambless, D.L. (1995). Cognitive behavioral versus exposure only treatment for social phobia: A meta-analysis. *Behavior Therapy*, 26, 695-720.
- Field, A.P. (2001). Meta-analysis of correlation coefficients: A Monte Carlo comparison of fixed and random effects methods. *Psychological Methods*, 6(2), 161-180.
- Fleiss, J.L. (1981). *Statistical methods for rates and proportions* (2<sup>nd</sup> ed.). New York: Wiley.
- Foa, E.B. & Kozak, M.J. (1986). Emotional processing of fear: Exposure to corrective information in rape victims. *Psychological Bulletin*, 99, 20-35.
- Fredrikson, M., Annas, P., Fisher, H. & Wik, G. (1996). Gender and age differences in the prevalence of specific fears and phobias. *Behaviour Research & Therapy*, 26, 241-244.
- Glass, G.V. (1978). In defence of generalization. *Behavioral and Brain Sciences*, 3, 394-395.
- Glass, G.V., McGraw, B. & Smith, M.L. (1981). *Meta-analysis in social research*. Beverly Hills, CA.: Sage.
- Gleser, L.J. & Olkin, I. (1994). Stochastically dependent effect sizes. In: H. Cooper & L.V. Hedges (Eds.). *The handbook of research synthesis* (S. 340-355). New York: Russell Sage Foundation.
- Goisman, R.M., Warshaw, M.G., Steketee, G.S., Fierman, E.J., Rogers, M.P., Goldenberg, I., Weinshenker, N.J., Vasile, R.G. & Keller, M.B. (1995). DSM-IV in the disappearance of agoraphobia without a history of panic disorder: New data on a controversial diagnosis. *American Journal of Psychiatry*, 152, 1438-1443.
- Goodman, W.K., Price, L.H., Rasmussen, S.A., Mazure, C., Delgado, P., Heninger, G.R. & Charney, D.S. (1989). The Yale-Brown Obsessive Compulsive Scale: I. Development, use, and reliability. *Archives of General Psychiatry*, 46, 1006-1011.
- Gould, R.A., Buckminster, S., Otto, M.W., Pollack, M.H. & Yap, L. (1997). Cognitive-behavioral and pharmacological treatment for social phobia: A meta-analysis. *Clinical Psychology: Science and Practice*, 4, 291-306.

- 
- Gould, R.A., Otto, M.W. & Pollack, M.H. (1995). A meta-analysis of treatment outcome for panic disorder. *Clinical Psychology Review, 15*(8), 819-844.
- Gould, R.A., Otto, M.W., Pollack, M.H. & Yap, L. (1997). Cognitive behavioral and pharmacological treatment of generalized anxiety disorder: A preliminary meta-analysis. *Behavior Therapy, 28*, 285-305.
- Grawe, K., Donati, R. & Bernauer, F. (1994). *Psychotherapie im Wandel. Von der Konfession zur Profession*. Göttingen: Hogrefe.
- Greenberg, P.E., Sisitsky, T., Kessler, R.C., Finkelstein, S.N., Berndt, E.R., Davidson, J.R.T., Ballenger, J.C. & Fyer, A.J. (1999). The economic burden of anxiety disorders in the 1990s. *Journal of Clinical Psychiatry, 60*(7), 427-435.
- Grissom, R.J. & Kim, J.J. (2001). Review of assumptions and problems in the appropriate conceptualization of effect size. *Psychological Methods, 6*(2), 135-146.
- Guy, W. (1976). *ECDEU Assessment Manual for Psychotherapy – Revised*. Bethesda, MD: National Institute of Mental Health.
- Haddock, C.K., Rindskopf, D. & Shadish, W.R. (1998). Using odds ratio as effect sizes for meta-analysis of dichotomous data: A primer on methods and issues. *Psychological Methods, 3*, 339-353.
- Hall, J.A., Tickle-Degnen, L., Rosenthal, R. & Mosteller, F. (1994). Hypotheses and problems in research synthesis. In: In: H. Cooper & L.V. Hedges (Eds.). *The handbook of research synthesis* (S. 17-28). New York: Russell Sage Foundation.
- Hand, I. (1996). Expositionsbehandlung. In: M. Linden & M. Hautzinger (Hrsg). *Verhaltenstherapie, 3*. überarb. und erw. Auflage (S. 139-149). Berlin: Springer.
- Hartmann, A. & Herzog, T. (1995). Varianten der Effektstärkenberechnung in Metaanalysen: Kommt es zu variablen Ergebnissen? *Zeitschrift für Klinische Psychologie, 24*(4), 337-343.
- Harwell, M. (1997). An empirical study of Hedges's homogeneity test. *Psychological Methods, 2*(2), 219-231.
- Hasselblad, V. & Hedges, L.V. (1995). Meta-analysis of screening and diagnostic tests. *Psychological Bulletin, 117*, 167-178.
- Hedges, L.V. & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- Hedges, L.V. & Vevea, J.L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods, 3*(4), 486-504.

- 
- Hedges, L.V. (1994). Fixed effects models. In: H. Cooper & L.V. Hedges (Eds.). *The handbook of research synthesis* (S.285-299). New York: Russell Sage Foundation.
- Heimberg, R.G. (1997). The use of pill placebo as a control strategy in the evaluation of psychotherapy: Reply to Klein. *Treatment, 1*, Comment 2, [http://journals.apa.org/treatment/vol1/97\\_c2.html](http://journals.apa.org/treatment/vol1/97_c2.html)
- Heimberg, R.G., Dodge, C.S., Hope, D.A., Kennedy, C.R. & Zollo, L.J. (1990). Cognitive behavioral group treatment for social phobia: comparison with a credible placebo control. *Cognitive Therapy and Research, 14*(1), 1-23.
- Heimberg, R.G., Holt, C.S., Schneier, F.R., Spitzer, R.L. & Liebowitz, M.R. (1993). The issue of subtypes in the diagnosis of social phobia. *Journal of Anxiety Disorders, 7*, 249-269.
- Heimberg, R.G., Müller, G., Holt, C.S., Hope, D.A. & Liebowitz, M.R. (1992). Assessment of anxiety in social interaction and being observed by others: The social interaction anxiety scale and the social phobia scale. *Behavior Therapy, 23*, 53-73.
- Heinsman, D.T. & Shadish, W.R. (1996). Assignment methods in experimentation: when do nonrandomized experiments approximate answers from randomised experiments? *Psychological Methods, 1*(2), 154-169.
- Hirschfeld, R.M.A. (1996). Placebo response in the treatment of panic disorder. *Bulletin of the Menninger Clinic, 60*(2), A76-A86.
- Hodgson, R.J. & Rachman, S. (1977). Obsessional-compulsive complaints. *Behaviour Research and Therapy, 15*, 389-395.
- Hofmann, S.G. & Barlow, D.H. (2002). Social phobia (social anxiety disorder). In: D.H. Barlow. *Anxiety and its disorders: The nature and treatment of anxiety and panic*; 2<sup>nd</sup> ed. (S. 454-476). New York: The Guilford Press.
- Hofmann, S.G., Barlow, D.H., Papp, L.A., Detweiler, M.F., Ray, S.A., Shear, M.K., Woods, S.W. & Gorman, J.G. (1997). Pretreatment attrition in a comparative treatment outcome study on panic disorder. *American Journal of Psychiatry, 155*, 43-47.
- Horowitz, M.J., Wilner, N. & Alvarez, W. (1979). The Impact of Event Scale: A measure of subjective stress. *Psychosomatic Medicine, 41*(3), 209-218.
- Horvath, P. (1988). Placebos and common factors in two decades of psychotherapy research. *Psychological Bulletin, 104*, 214-225.
- Howard, W.A., Murphy, S.M. & Clarke, J.C. (1983). The nature and treatment of fear

- of flying: A controlled investigation. *Behavior Therapy*, 14, 557-567.
- Hsu, L.M. (1989). Random sampling, randomization, and equivalence of contrasted groups in psychotherapy outcome research. *Journal of Consulting and Clinical Psychology*, 57(1), 131-137.
- Hunter, J.E. & Schmidt (F.L. (1990). *Methods of meta-analysis*. Newbury Park: Sage.
- Jacobson, N.S. & Hollon, S.D. (1996). Prospects for future comparisons between drugs and psychotherapy: Lessons from the CBT-versus-pharmacotherapy exchange. *Journal of Consulting and Clinical Psychology*, 64(1), 104-108.
- Kazdin, A. (1999). Overview of research design issues in clinical psychology. In: P.C. Kendall, J.N. Butcher & C.N. Holmbeck (Eds.), *Handbook of research methods in clinical psychology*. New York: Wiley.
- Kazdin, A.E. (1978). Evaluating the generality of findings in analogue therapy research. *Journal of Consulting and Clinical Psychology*, 46(4), 673-686.
- Kessler, R.C., McGonagle, K.A., Zhao, S., Nelson, C.B., Hughes, M., Eshleman, S., Wittchen, H.-U. & Kendler, K.S. (1994). Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: Results from the National Comorbidity Survey, *British Journal of Psychiatry*, 168, 17-30.
- Kessler, R.C., Sonnega, A., Bromet, E., Hughes, M. & Nelson, C.B. (1995). Posttraumatic stress disorder in the national comorbidity survey. *Archives of General Psychiatry*, 52, 1048-1060.
- Kirsch, I. & Sapirstein, G. (1998). Listening to prozac but hearing placebo: A meta-analysis of antidepressant medication. *Prevention and Treatment*, 1, Artikel 0002a, <http://journals.apa.org/prevention/volume1/pre0010002a.html>
- Klein, D.F. (1996). Preventing hung juries about therapy studies. *Journal of Consulting and Clinical Psychology*, 64(1), 81-87.
- Klein, D.F. (1997). Control groups in pharmacotherapy and psychotherapy evaluations. *Treatment*, 1, Artikel 1, [http://journals.apa.org/treatment/vol1/97\\_a1.html](http://journals.apa.org/treatment/vol1/97_a1.html)
- Klein, D.F. (1998). Listening to meta-analysis but hearing bias. *Prevention and Treatment*, 1, Artikel 0006c, <http://journals.apa.org/prevention/volume1/pre0010006c.html>
- Klerman, G.L., Weissman, M.M., Markowitz, J., Glick, I., Wilner, P.J., Mason, B. & Shear, M.K. (1994). Medication and psychotherapy. In: A.E. Bergin, & S.L. Garfield (Eds.). *Handbook of psychotherapy and behaviour change* (S. 734-782). New York: Wiley.



- 
- Kobak, K.A., Greist, J.H., Jefferson, J.W., Katzelnick, D.J. & Henk, H.J. (1998). Behavioral versus pharmacological treatments of obsessive compulsive disorder: A meta-analysis. *Psychopharmacology*, *136*, 205-216.
- Kopta, S.M., Howard, K.I., Lowry, J.L. & Beutler, L.E. (1994). Patterns of symptomatic recovery in psychotherapy. *Journal of Consulting and Clinical Psychology*, *62*(5), 1009-1016.
- Kraemer, H.C. (1981). Coping strategies in psychiatric clinical research. *Journal of Consulting and Clinical Psychology*, *49*, 309-319.
- Lambert, M.J. & Bergin, A.E. (1994). The effectiveness of psychotherapy. In: A.E. Bergin, & S.L. Garfield (Eds.). *Handbook of psychotherapy and behaviour change* (S. 143-189). New York: Wiley.
- Lambert, M.J. & Hill, C. (1994). Assessing psychotherapy outcomes and processes. In: A.E. Bergin, & S.L. Garfield (Eds.). *Handbook of psychotherapy and behaviour change* (S. 72-113). New York: Wiley.
- Lipsey, M.W. & Wilson, D.B. (1993). The efficacy of psychological, educational, and behavioral treatment. *American Psychologist*, *18*(12), 1181-1201.
- Lipsey, M.W. & Wilson, D.B. (2001). The way in which intervention studies have "personality" and why it is important to meta-analysis. *Evaluation and Health Professions*, *24*(4), 236-254.
- Magee, W.J., Eaton, W.W., Wittchen, H.-U., McGonagle, K.A. & Kessler, R.C. (1996). Agoraphobia, simple phobia, and social phobia in the National Comorbidity Survey. *Archives of General Psychiatry*, *53*, 159-168.
- Margraf, J. & Poldrack, A. (2000). Angstsyndrome in Ost- und Westdeutschland: Eine repräsentative Bevölkerungserhebung. *Zeitschrift für Klinische Psychologie und Psychotherapie*, *29*(3), 157-169.
- Margraf, J., Ehlers, E., Roth, W.T., Clark, D.B., Sheikh, J., Agras, W.S. & Taylor, C.B. (1991). How "blind" are double-blind studies? *Journal of Consulting and Clinical Psychology*, *59*(1), 184-187.
- Marks, I. & Dar, R. (2000). Fear reductions by psychotherapies. Recent findings, future directions. *British Journal of Psychiatry*, *176*, 507-511.
- Marks, I.M. & Mathews, A.M. (1979). Brief standard self-rating for phobic patients. *Behaviour Research and Therapy*, *17*, 263-267.
- Matt, G.E. & Navarro, A.M. (1997). What meta-analyses have and have not taught us about psychotherapy effects: A review and future directions. *Clinical Psychology*

- 
- Review, 17(1), 1-32.*
- Matt, G.E. (1989). Decision rules for selecting effect sizes in meta-analysis: a review and reanalysis of psychotherapy outcome studies. *Psychological Bulletin, 105*, 106-115.
- Matt, G.E. (in press). Will it work in Münster? Meta-analysis and the empirical generalization of causal relationships.
- McNally, R.J. (1996). Methodological controversies in the treatment of panic disorder. *Journal of Consulting and Clinical Psychology, 64(1)*, 88-91.
- Mendlowitz, M.V. & Stein, M.B. (2000). Quality of life in individuals with anxiety disorders. *American Journal of Psychiatry, 157*, 669-682.
- Mennin, D.S. & Heimberg, R.G. (2000). The impact of comorbid mood and personality disorders in the cognitive-behavioral treatment of panic disorder. *Clinical Psychology Review, 20(3)*, 339-357.
- Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T.D. (1990). Development and validation of the Penn State Worry Questionnaire. *Behaviour Research & Therapy, 28(6)*, 487-495.
- Möller, H.-J., Müller, W.E. & Volz, H.-P. (2000). *Psychopharmakotherapie: Ein Leitfaden für Klinik und Praxis* (2. überarb. u. erw. Aufl.). Stuttgart: Kohlhammer.
- Oei, T.P.S., Llamas, M. & Devilly, G.J. (1999). The efficacy and cognitive processes of cognitive behaviour therapy in the treatment of panic disorder with agoraphobia. *Behavioural and Cognitive Psychotherapy, 27*, 63-88.
- Ogles, B.M., Lambert, M.J. & Sawyer, J.D. (1995). The clinical significance of the National Institute of Mental Health collaborative depression study data. *Journal of Consulting & Clinical Psychology, 63*, 321-326.
- Oosterbaan, D.B. (2000). *Social Phobia. Cognitive and pharmacological treatment*. Unveröffentlichte Dissertation. Universität Amsterdam.
- Orwin, R.G. (1983). A fail-safe  $N$  for effect size in meta-analysis. *Journal of Educational Statistics, 8(2)*, 157-159.
- Öst, L.-G. (1987). Applied relaxation: Description of a coping technique and review of controlled studies. *Behaviour Research & Therapy, 25*, 397-409.
- Overton, R.C. (1998). A comparison of fixed-effects and mixed (random effects) models for meta-analysis tests of moderator variable effects. *Psychological Methods, 3(3)*, 354-379.
- Peterson, R.A., & Reis, S. (1992). *Anxiety Sensitivity Index - Revised test manual*. Ohio:

---

IDS Publishing Corporation.

- Piccinelli, M., Pini, S. Bellantuono, C. & Wilkinson, G. (1995). Efficacy of drug treatment in obsessive-compulsive disorder. A meta-analytic review. *British Journal of Psychiatry*, *166*, 424-443.
- Pigott, T.D. (2001). Missing predictors in models of effect size. *Evaluation & The Health Professions*, *24*(3), 277-307.
- Power, K.G. & Sharp, D.M. (1995). Keep taking the tablets? Inadequate controls for concurrent psychotropic medications in studies of psychological treatments for panic disorder. *Journal of Psychopharmacology*, *9*(1), 71-72.
- Raudenbush, S.W. (1994). Random effects models. In: H. Cooper & L.V. Hedges (Eds.). *The handbook of research synthesis* (S.301-321). New York: Russell Sage Foundation.
- Ray, J.W. & Shadish, W.R. (1996). How interchangeable are different estimators of effect size? *Journal of Consulting and Clinical Psychology*, *64*(6), 1316-1325.
- Rice, D.P. & Miller, L.S. (1998). Health economics and cost implications of anxiety and other mental disorders in the United States. *British Journal of Psychiatry*, *173* (suppl. 34), 4-9.
- Rosenthal, R. (1994). Parametric measures of effect size. In: H. Cooper & L.V. Hedges (Eds.). *The handbook of research synthesis* (S. 232-234). New York: Russell Sage Foundation.
- Ruhmland, M. & Margraf, J. (2001a). Effektivität psychologischer Therapien von spezifischer Phobie und Zwangsstörung: Meta-Analysen auf Störungsebene. *Verhaltenstherapie*, *11*, 14-26.
- Ruhmland, M. & Margraf, J. (2001b). Effektivität psychologischer Therapien von generalisierter Angststörung und sozialer Phobie: Meta-Analysen auf Störungsebene. *Verhaltenstherapie*, *11*, 27-40.
- Ruhmland, M. & Margraf, J. (2001c). Effektivität psychologischer Therapien von Panik und Agoraphobie: Meta-Analysen auf Störungsebene. *Verhaltenstherapie*, *11*, 41-53.
- Salkovskis, P.M. & Hackman, A. (1997). Agoraphobia. In: G.C.L. Davey, *Phobias. A handbook of theory, research and treatment* (S. 27-61). Chichester: John Wiley & Sons.
- Sanavio, E. (1988). Obsessions and compulsions: The Padua Inventory. *Behaviour Research and Therapy*, *26*, 169-177.

- 
- Seligman, M.E.P. (1995). The effectiveness of psychotherapy. The consumer reports study. *American Psychologist*, *50*(12), 965-974.
- Shadish, W.R. & Haddock, C.K. (1994). Combining estimates of effect size. In: H. Cooper & L.V. Hedges (Eds.). *The handbook of research synthesis* (S.261-281). New York: Russel Sage Foundation.
- Shadish, W.R. & Ragsdale, K. (1996). Random versus non-random assignment in controlled experiments: do you get the same answer? *Journal of Consulting and Clinical Psychology*, *4*(6), 1290-1303.
- Shadish, W.R. (1996). Meta-analysis and the exploration of causal mediating processes: A primer of examples, methods, and issues. *Psychological Methods*, *1*(1), 47-65.
- Shadish, W.R., Matt, G.E., Navarro, A.M. & Phillips, G. (2000). The effects of psychological therapies under clinically representative conditions: A meta-analysis. *Psychological Bulletin*, *126*(4), 512-529.
- Shadish, W.R., Matt, G.E., Navarro, A.M, Siegle, G., Crits-Christoph, P., Hazelrigg, M.D., Jorm, A.F., Lyons, L.C., Nietzel, M.T., Prout, H.T., Robinson, L., Smith, M.L., Svartberg, M. & Weiss, B. (1997). Evidence that therapy works in clinically representtative conditions. *Journal of Consulting and Clinical Psychology*, *65*(3), 355-365.
- Sharpe, D. (1997). Of apples and oranges, file drawers and garbage: Why validity issues in meta-analysis will not go away. *Clinical Psychological Review*, *17*(8), 881-901.
- Shear, M.K. & Maser, J.D. (1994). Standardized assessment for panic disorder research. *Archives of General Psychiatry*, *51*, 346-354.
- Sherman, J.J. (1998). Effects of psychotherapeutic treatments for PTSD: A meta-analysis of controlled clinical trials. *Journal of Traumatic Stress*, *11*(6), 413-435.
- Smith, M.L. & Glass, G.V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, *32*, 752-760.
- Smith, M.L., Glass, G.V. & Miller, T.I. (1980). *The benefits of psychotherapy*. Baltimore, London: John Hopkins University Press.
- Sohn, D. (1996). Publication bias and the evaluation of psychotherapy efficacy in reviews of research literature. *Clinical Psychological Review*, *16*(2), 147-156.
- Spielberger, C.D., Gorsuch, R.L. & Lushene, R.E. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychllogists Press.
- Stangier, U. & Heidenreich, T. (1997). Diagnostik der Sozialen Phobie. *Verhaltenstherapie*, *7*, 107-118.

- 
- Stein, D.J., Spadaccini, E. & Hollander, E. (1995). Meta-analysis of pharmacotherapy trials for obsessive-compulsive disorder. *International Clinical Psychopharmacology*, *10*, 11-18.
- Stein, D.J., Zungu-Dirwayi, N., van der Linden, G.J.H. & Seedat, S. (2001). Pharmacotherapy for posttraumatic stress disorder. *The Cochrane Library*, *2*. Oxford: Update Software.
- Stein, M.B., Torgrud, L.J. & Walker, J.R. (2000). Social phobia symptoms, subtypes, and severity. Findings from a community survey. *Archives of General Psychiatry*, *57*, 1046-1052.
- Stevens, S.E., Hynan, M.T. & Allen, M. (2000). A meta-analysis of common factor and specific treatment effects across the outcome domains of the phase model of psychotherapy. *Clinical Psychology: Science and Practice*, *7*, 273-290.
- Taylor, C.B., Sheikh, J., Agras, W.S., Roth, W.T., Margraf, J., Ehlers, A., Maddock, R.J. & Gossard, D. (1986). Self-report of panic attacks: Agreement with heart rate changes. *American Journal of Psychiatry*, *143*, 478-482.
- Taylor, S. (1996). Meta-analysis of cognitive-behavioral treatments for social phobia. *Journal of Behavior Therapy and Experimental Psychiatry*, *27*(1), 1-9.
- Trivedi, M.H. & Rush, H. (1994). Does a placebo run-in or a placebo treatment cell affect the efficacy of antidepressant medications? *Neuropsychopharmacology*, *11*(1), 33-43.
- Turner, S.M., Beidel, D.C. & Townsley, R.M. (1992). Social phobia: A comparison of specific and generalized subtypes and avoidant personality disorder. *Journal of Abnormal Psychology*, *101*, 326-331.
- Turner, S.M., Beidel, D.C., Dancu, C.V. & Stanley, M.A. (1989). An empirically derived inventory to measure social fears and anxiety: The social phobia and anxiety inventory (SPAI). *Psychological Assessment*, *1*, 35-40.
- Van Balkom, A.J.L.M., Bakker, A., Spinhoven, P., Blauw, B.M.J.W., Smeenk, S. & Ruesink, B. (1997). A meta-analysis of the treatment of panic disorder with or without agoraphobia: A comparison of psychopharmacological, cognitive-behavioral, and combination treatments. *The Journal of Nervous and Mental Disease*, *185*, 510-516.
- Van Balkom, A.J.L.M., van Oppen, P., Vermeulen, A.W.A., van Dyck, R., Nauta, M.C.E. & Vorst, H.C.M. (1994). A meta-analysis on the treatment of obsessive compulsive disorder: A comparison of antidepressants, behavior, and cognitive

- 
- therapy. *Clinical Psychology Review*, 14(5), 359-381.
- Van Etten, M.L. & Taylor, S. (1998). Comparative efficacy of treatments for post-traumatic stress disorder: A meta-analysis. *Clinical Psychology and Psychotherapy*, 5, 126-144.
- Vevea, J.L. & Hedges, L.V. (1995). A general linear model for estimating effect size in the presence of publication bias. *Psychometrika*, 60(3), 419-435.
- Watson, D. & Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of Consulting & Clinical Psychology*, 33, 448-457.
- Weathers, F.W., Blake, D.D., Krinsley, K.E., Haddad, W., Huska, J.A. & Keane, T.M. (1992). *The Clinician Administered PTSD Scale: Reliability and construct validity*. Paper presented at the 26th Annual Convention of the Association for Advancement of Behavior Therapy, Boston.
- Westen, D. & Morrison, K. (2001). A multidimensional meta-analysis of treatments for depression, panic, and generalized anxiety disorder: An empirical examination of the status of empirically supported therapies. *Journal of Consulting and Clinical Psychology*, 69(6), 875-899.
- Westra, H.A. & Stewart, S.H. (1998). Cognitive behavioural therapy and pharmacotherapy: Complementary or contradictory approaches to the treatment of anxiety? *Clinical Psychology Review*, 18(3), 307-340.
- White, K.S. & Barlow, D.H. (2002). Panic disorder and agoraphobia. In: D.H. Barlow. *Anxiety and its disorders: The nature and treatment of anxiety and panic*; 2<sup>nd</sup> ed. (S. 328-379). New York: The Guilford Press.
- Wilson, D.B. & Lipsey, M.W. (2001). The role of method in treatment effectiveness research: Evidence from meta-analysis. *Psychological Methods*, 6(4), 413-429.
- Wittchen, H.-U. & Fehm, L. (2001). Epidemiology, patterns of comorbidity, and associated disabilities of social phobia. *The Psychiatric Clinics of North America*, 24(4), 617-641.
- Wittchen, H.-U. (1991). Der Langzeitverlauf unbehandelter Angststörungen: Wie häufig sind Spontanremissionen? *Verhaltenstherapie*, 1, 273-282.
- Wittchen, H.-U., Zhao, S., Kessler, R.C. & Eaton, W.W. (1994). DSM-III-R generalized anxiety disorder in the National Comorbidity Survey. *Archives of General Psychiatry*, 51, 355-364.
- Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition*. Stanford, CA: Stanford University Press.

- Wortman, P.M. (1994). Judging research quality. In: H. Cooper & L.V. Hedges (Eds.). *The handbook of research synthesis* (S. 97-109). New York: Russell Sage Foundation.
- Yonkers, K.A., Warshaw, M.R., Massion, A.O. & Keller, M.B. (1996). Phenomenology and course of generalised anxiety disorder. *British Journal of Psychiatry*, 168, 308-313.

## 6.1 Eingeschlossene Studien Generalisierte Angststörung

Anmerkung: es wurde nur eine Publikation pro Studie aufgenommen.

- Allgulander, C., Hackett, D. & Salinas, E. (2001). Venlafaxine extended release (ER) in the treatment of generalised anxiety disorder. *British Journal of Psychiatry*, 179, 15-22.
- Arntz, A. (in press). Cognitive therapy versus applied relaxation as treatment of generalized anxiety disorder. *Behaviour Research and Therapy*.
- Ballenger, J.C., McDonald, S., Noyes, R., Rickels, K., Sussman, N., Woods, S., Patin, J. & Singer, J. (1991). The first double-blind, placebo-controlled trial of a partial benzodiazepine agonist abecarnil (ZK 112-119) in generalized anxiety disorder. *Psychopharmacology Bulletin*, 27(2), 171-179.
- Barlow, D.H., Cohen, A.S., Waddell, M.T., Vermilyea, B.B., Klosko, J.S., Blanchard, E.B. & DiNardo, P.A. (1984). Panic and generalized anxiety disorder: Nature and treatment. *Behavior Therapy*, 15, 431-449.
- Barlow, D.H., Rapee, R.M. & Brown, T.A. (1992). Behavioral treatment of generalized anxiety disorder. *Behavior Therapy*, 23, 551-570.
- Biswas, A., Biswas, D. & Chattopadhyay, P.K. (1995). Cognitive behaviour therapy in generalised anxiety disorder. *Indian Journal of Clinical Psychology*, 22(2), 1-10.
- Bjerrum, H., Allerup, P., Thuneborg, K., Jakobsen, K. & Bech, P. (1992). Treatment of generalized anxiety disorder: Comparison of a new beta-blocking drug (CGP361A), low-dose neuroleptic (flupenthixol), and placebo. *Pharmacopsychiatry*, 25, 229-232.
- Blowers, C., Cobb, J. & Mathews, A. (1987). Generalised anxiety: A controlled treatment study. *Behaviour Research and Therapy*, 25(6), 493-502.
- Borison, R.L., Albrecht, J.W. & Diamond, B.I. (1990). Efficacy and safety of a putative anxiolytic agent: Ipsapirone. *Psychopharmacology Bulletin*, 26(2), 207-210.
- Borkovec, T.D. & Costello, E. (1993). Efficacy of applied relaxation and cognitive-behavioral therapy in the treatment of generalized anxiety disorder. *Journal of Consulting and Clinical Psychology*, 61(4), 611-619.
- Bowman, D., Scogin, F., Floyd, M., Patton, E. & Gist, L. (1997). Efficacy of self-examination therapy in the treatment of generalized anxiety disorder. *Journal of Counseling Psychology*, 44(3), 267-273.
- Boyer, W.F. & Feighner, J.P. (1993). A placebo-controlled double-blind multicenter



- trial of two doses of ipsapirone versus diazepam in generalized anxiety disorder. *International Clinical Psychopharmacology*, 8, 173-176.
- Butler, G., Cullington, A., Hibbert, G., Klimes, I. & Gelder, M. (1987). Anxiety Management for persistent generalised anxiety. *British Journal of Psychiatry*, 151, 535-542.
- Butler, G., Fennell, M., Robson, P. & Gelder, M. (1991). Comparison of behavior therapy and cognitive behavior therapy in the treatment of generalized anxiety disorder. *Journal of Consulting and Clinical Psychology*, 59(1), 167-175.
- Casacchia, M., Bolino, F. & Ecari, U. (1990). Etizolam in the treatment of generalized anxiety disorder: A double-blind study versus placebo. *Current Medical Research and Opinion*, 12(4), 215-223.
- Castillo, A., Sotillo, C. & Mariategui, J. (1987). Alprazolam compared to clobazam and placebo in anxious outpatients. *Neuropsychobiology*, 18, 189-194.
- Cohn, J.B. & Wilcox, C.S. (1984). Long-term comparison of alprazolam, lorazepam and placebo in patients with an anxiety disorder. *Pharmacotherapy*, 4(2), 93-98.
- Cragan, M.K. & Deffenbacher, J.L. (1984). Anxiety management training and relaxation as self-control in the treatment of generalized anxiety in medical outpatients. *Journal of Counseling Psychology*, 31(2), 123-131.
- Cutler, N.R., Keppel Hesselink, J.M. & Sramek, J.J. (1994). A phase II multicenter dose-finding, efficacy and safety trial of ipsapirone in outpatients with generalized anxiety disorder. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 18, 447-463.
- Cutler, N.R., Sramek, J.J., Keppel Hesselink, J.M., Krol, A., Roeschen, J., Rickels, K. & Schweizer, E. (1993). A double-blind, placebo-controlled study comparing the efficacy and safety of ipsapirone versus lorazepam in patients with generalized anxiety disorder: A prospective multicenter trial. *Journal of Clinical Psychopharmacology*, 13(6), 429-437
- Davidson, J.R.T., DuPont, R.L., Hedges, D. & Haskins, J.T. (1999). Efficacy, safety, and tolerability of venlafaxine extended release and buspirone in outpatients with generalized anxiety disorder. *Journal of Clinical Psychiatry*, 60(8), 528-535.
- DeMaio, D., Levi Minzi, A., Ceccarelli, G. & Ecari, U. (1989). Treatment of generalized anxiety disorder with a new diazepinic derivative, etizolam. *Current Therapeutic Research*, 46(2), 373-380.
- Diamond, B.I., Nguyen, H., O'Neal, E., Ochs, R., Kaffeman, M. & Borison, R.L.

- (1991). A comparative study of alpidem, a nonbenzodiazepine, and lorazepam in patients with nonpsychotic anxiety. *Psychopharmacology Bulletin*, 27(1), 67-71.
- Enkelmann, R. (1991). Alprazolam versus buspirone in the treatment of outpatients with generalized anxiety disorder. *Psychopharmacology*, 105, 428-432.
- Fontaine, R., Annable, L., Chouinard, G. & Ogilvie, R.I. (1983). Bromazepam and diazepam in generalized anxiety: A placebo-controlled study with measurement of drug plasma concentrations. *Journal of Clinical Psychopharmacology*, 3(2), 80-87.
- Fontaine, R., Beaudry, P., Beauclair, L. & Chouinard, G. (1987). Comparison of withdrawal of buspirone and diazepam: A placebo controlled study. *Progress in Neuropsychopharmacology & Biological Psychiatry*, 11, 189-197.
- Fontaine, R., Beaudry, P., Le Morvan, P., Beauclair, L. & Chouinard, G. (1990). Zopiclone and triazolam in insomnia associated with generalized anxiety disorder: A placebo-controlled evaluation of efficacy and daytime anxiety. *International Clinical Psychopharmacology*, 5, 173-183.
- Fontaine, R., Chouinard, G. & Annable, L. (1984). Bromazepam and diazepam in generalized anxiety: A placebo-controlled study of efficacy and withdrawal. *Psychopharmacology Bulletin*, 20(1), 126-127.
- Fontaine, R., Mercier, P., Beaudry, P., Annable, L. & Chouinard, G. (1986). Bromazepam and lorazepam in generalized anxiety: a placebo-controlled study with measurement of drug plasma concentrations. *Acta psychiatrica Scandinavica*, 74, 451-458.
- Fresquet, A., Sust, M., Lloret, A., Murphy, M.F., Carter, F.J., Campbell, G.M. & Marion-Landais, G. (2000). Efficacy and safety of lesopitron in outpatients with generalized anxiety disorder. *The Annals of Pharmacotherapy*, 34, 147-153.
- Gelenberg, A.J., Lydiard, R.B., Rudolph, R.L., Aguiar, L., Haskins, J.T. & Salinas, E. (2000). Efficacy of venlafaxine extended-release capsules in nondepressed outpatients with generalized anxiety disorder. *JAMA*, 283(23), 3082-3088.
- Goldberg, H.L. & Finnerty, R. (1982). Comparison of buspirone in two separate studies. *Journal of Clinical Psychiatry*, 43(12), 87-91.
- Jannoun; L., Oppenheimer, C. & Gelder, M. (1982). A self-help treatment program for anxiety state patients. *Behavior Therapy*, 13, 103-111.
- Kohli, A., Nehra, V. & Nehra, R. (2000). Comparison of efficacy of psychorelaxation and pharmacotherapy in generalized anxiety disorder. *Journal of Personality and*

---

*Clinical Studies*, 16(1), 43-48.

- Laakmann, G., Schüle, C., Lorkowski, G., Baghai, T., Kuhn, K. & Ehrentraut, S. (1998). Buspirone and lorazepam in the treatment of generalized anxiety disorder in outpatients. *Psychopharmacology*, 136, 357-366.
- Laboratorios Dr. Esteve (1997). Lesopitron in the treatment of outpatients with GAD: A single-center, double-blind, parallel, placebo- and active-controlled comparison with lorazepam. Unpublished Manuscript. SA, Barcelona, Spain.
- Lader, M. & Scotto, J.-C. (1998). A multicentre double-blind comparison of hydroxyzine, buspirone and placebo in patients with generalized anxiety disorder. *Psychopharmacology*, 139, 402-406.
- Ladouceur, R., Dugas, M.J., Freeston, M.H., Léger, E., Gagnon, F. & Thibodeau, N. (2000). Efficacy of cognitive-behavioral treatment of generalized anxiety disorder: Evaluation in a controlled clinical trial. *Journal of Consulting and Clinical Psychology*, 68(6), 957-964.
- Lecrubier, Y., Puech, A. J., Azcona, A., Bailey, P.E. & Lataste, X. (1993). A randomized double-blind placebo-controlled study of tropisetron in the treatment of outpatients with generalized anxiety disorder. *Psychopharmacology*, 112, 129-133.
- Linden, M., Bär, T., Zubrägel, D., Ahrens, B. & Schlattmann, P. (2002). Wirksamkeit der kognitiven Verhaltenstherapie bei generalisierten Angsterkrankungen – Ergebnisse der Berliner KVT-GAD-Studie. *Verhaltenstherapie*, 12, 173-181.
- Lindsay, W.R., Gamsu, C.V., McLaughlin, E., Hood, E.M. & Espie, C.A. (1987). A controlled trial of treatments for generalized anxiety. *British Journal of Clinical Psychology*, 26, 3-15.
- Lydiard, R.B., Ballenger, J.C. & Rickels, K. (1997). A double-blind evaluation of the safety and efficacy of abecarnil, alprazolam, and placebo in outpatients with generalized anxiety disorder. *Journal of Clinical Psychiatry*, 58(11), 11-18.
- McLeod, D.R., Hoehn-Saric, R., Porges, S.W. & Zimmerli, W.D. (1982). Effects of alprazolam and imipramine on parasympathetic cardiac control in patients with generalized anxiety disorder. *Psychopharmacology*, 107, 535-540.
- Mendels, J., Krajewski, T.F., Huffer, V., Taylor, R.J., Secunda, S., Schless, A., Sebastian, J.A., Semchishyn, G., Durr, M.J., Melmed, A.S. & Whyte A. (1986). Effective short-term treatment of generalized anxiety disorder with trifluoperazine. *Journal of Clinical Psychiatry*, 47(4), 170-174.

- 
- Möller, H.-J., Volz, H.-P., Reimann, I.W. & Stoll, K.-D. (2001). Opipramol for the treatment of generalized anxiety disorder: A placebo-controlled trial including an alprazolam-treated group. *Journal of Clinical Psychopharmacology*, 21(1), 59-65.
- Öst, L.-G. & Breitholtz, E. (2000). Applied relaxation vs. cognitive therapy in the treatment of generalized anxiety disorder. *Behaviour Research and Therapy*, 38, 777-790.
- Pecknold, J.C., Matas, M., Howarth, B.G., Ross, C., Swinson, R., Vezeau, C. & Ungar, W. (1989). Evaluation of buspirone as an antianxiety agent: Buspirone and diazepam versus placebo. *Canadian Journal of Psychiatry*, 34(8), 766-770.
- Pollack, M.H., Worthington, J.J., Manfro, G.G., Otto, M.W. & Zucker, B.G. (1997). Abecarnil for the treatment of generalized anxiety disorder: A placebo-controlled comparison of two dosage ranges of abecarnil and buspirone. *Journal of Clinical Psychiatry*, 58(11), 19-23.
- Pollack, M.H., Zaninelli, R., Goddard, A., McCafferty, J.P., Bellew, K.M., Burnham, D.B. & Iyengar, M.K. (2001). Paroxetine in the treatment of generalized anxiety disorder: Results of a placebo-controlled, flexible-dosage trial. *Journal of Clinical Psychiatry*, 62(5), 350-357.
- Pourmotabbed, T., McLeod, D.R., Hoehn-Saric, R., Hipsley, P. & Greenblatt, D.J. (1996). Treatment, discontinuation, and psychomotor effects of diazepam in women with generalized anxiety disorder. *Journal of Clinical Psychopharmacology*, 16(3), 202-207.
- Power, K.G., Jerrom, D.W.A., Simpson, R.J., Mitchell, M.J. & Swanson, V. (1989). A controlled comparison of cognitive-behaviour therapy, diazepam and placebo in the treatment of generalized anxiety. *Behavioural Psychotherapy*, 17, 1-14.
- Power, K.G., Simpson, R.J., Swanson, V. & Wallace, L.A. (1990). A controlled comparison of cognitive-behaviour therapy, diazepam, and placebo, alone and in combination, for the treatment of generalised anxiety disorder. *Journal of Anxiety Disorders*, 4, 267-292.
- Rickels, K., DeMartinis, N. & Aufdembrinke, B. (2000). A double-blind, placebo-controlled trial of abecarnil and diazepam in the treatment of patients with generalized anxiety disorder. *Journal of Clinical Psychopharmacology*, 20(1), 12-18.
- Rickels, K., Downing, R., Schweizer, E. & Hassman, H. (1993). Antidepressants for the treatment of generalized anxiety disorder. *Archives of General Psychiatry*, 50,

884-895.

- Rickels, K., Pollack, M.H., Sheehan, D.V. & Haskins, J.T. (2000). Efficacy of extended-release venlafaxine in nondepressed outpatients with generalized anxiety disorder. *American Journal of Psychiatry*, *157*(6), 968-974.
- Rickels, K., Schweizer, E., DeMartinis, N., Mandos, L. Mercer, C. (1997). Gepirone and diazepam in generalized anxiety disorder: A placebo-controlled trial. *Journal of Clinical Psychopharmacology*, *17*(4), 272-277.
- Rickels, K., Weisman, K., Norstad, N., Singer, M., Stoltz, D., Brown, A. & Danton, J. (1982). Buspirone and diazepam in anxiety: A controlled study. *Journal of Clinical Psychiatry*, *43*(12), 81-86.
- Sarkar, P., Rathee, S.P. & Neera, N. (1999). Comparative efficacy of pharmacotherapy and biofeed back among cases of generalised anxiety disorder. *Journal of Projective Psychology and Mental Health*, *6*, 69-77.
- Scarpini, E., Baron, P.G., Bet, L., Bottini, G., Bresolin, N., Meola, G., Pezzoli, G., Vallar, G., Monza, G.C. & Scarlato, G. (1988). Low doses of ketazolam in anxiety: A double-blind, placebo-controlled study. *Neuropsychobiology*, *20*, 74-75.
- Schweizer, E., Rickels, K., Csanalosi, I., London, J. & Turner, D. (1990). A placebo-controlled study of enciprazine in the treatment of generalized anxiety disorder: A preliminary report. *Psychopharmacology Bulletin*, *26*(2), 215-217.
- Silverstone, P.H. & Salinas, E. (2001). Efficacy of venlafaxine extended release in patients with major depressive disorder and comorbid generalized anxiety disorder. *Journal of Clinical Psychiatry*, *67*(7), 523-529.
- Sramek, J.J., Tansman, M., Suri, A., Hornig-Rohan, M., Amsterdam, J.D., Stahl, S.M., Weisler, R.H. & Cutler, N.R. (1996). Efficacy of buspirone in generalized anxiety disorder with coexisting mild depressive symptoms. *Journal of Clinical Psychiatry*, *57*(7), 287-291.
- Tollefson, G. D., Montague-Clouse, J. & Tollefson, S. L. (1992). Treatment of comorbid generalized anxiety in a recently detoxified alcoholic population with a selective serotonergic drug (buspirone). *Journal of Clinical Psychopharmacology*, *12*(1), 19-26.
- Tollefson, G.D., Lancaster, S.P. & Montague-Clouse, J. (1991). The association of buspirone and its metabolite 1-pyrimidinylpiperazine in the remission of comorbid anxiety with depressive features and alcohol dependency. *Psychopharmacology*

*Bulletin*, 27(2), 163-170.

- White, J., Keenan, M. & Brooks, N. (1992). Stress control: A controlled comparative investigation of large group therapy for generalized anxiety disorder. *Behavioural Psychotherapy*, 20, 97-114.
- Wilcox, C.S., Ryan, P.J., Morrissey, J.L., Cohn, J.B., Defrancisco, D.F., Linden, R.D. & Heiser, J.F. (1994). A fixed-dose study of adinazolam-SR tablets in generalized anxiety disorder. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 18, 979-993.
- Woodward, R. & Jones, R.B. (1980). Cognitive restructuring treatment: A controlled trial with anxious patients. *Behaviour Research and Therapy*, 18, 401-407.
- Zung, W.W.K. (1987). Effect of clorazepate on depressed mood in anxious patients. *Journal of Clinical Psychiatry*, 48(1), 13-14.

## 6.2 Eingeschlossene Studien Panikstörung

Anmerkung: es wurde nur eine Publikation pro Studie aufgenommen.

- Arntz, A. & van den Hout, M. (1996). Psychological treatments of panic disorder without agoraphobia: Cognitive therapy versus applied relaxation. *Behaviour Research and Therapy*, 34(2), 113-121.
- Arntz, A. (2002). Cognitive therapy versus interoceptive exposure as treatment of panic disorder without agoraphobia. *Behaviour Research and Therapy*, 40, 325-341.
- Asnis, G.M., Hammeedi, F.A., Goddard, A.W. , Potkin, S.G., Black, D., Jameel, M., Desagani, K. & Woods, S.W. (2001). Fluvoxamine in the treatment of panic disorder: a multi-center, double-blind, placebo-controlled study in outpatients. *Psychiatry Research*, 103, 1-14.
- Azhar, M.Z. (2000). Comparison of fluvoxamine alone, fluvoxamine and cognitive psychotherapy alone in the treatment of panic disorder in Kelantan – Implications for management by family doctors. *Medical Journal of Malaysia*, 55(4), 402-408.
- Bakish, D., Hooper, C.L., Filteau, M.-J., Fraser, G., West, D.L., Thibaudeau, C. & Raine, D. (1996). A double-blind placebo-controlled trial comparing fluvoxamine and imipramine in the treatment of panic disorder with or without agoraphobia. *Psychopharmacology Bulletin*, 32(1), 135-141.
- Bakker, A., van Dyck, R., Spinhoven, P. & van Balkom, A.J.L.M. (1999). Paroxetine, clomipramine, and cognitive therapy in the treatment of panic disorder. *Journal of Clinical Psychiatry*, 60, 831-838.
- Ballenger, J.C., Burrows, G.D., DuPont, R. L., Lesser, I.M., Noyes, R., Pecknold, J.C., Rifkin, A. & Swinson, R.P. (1988). Alprazolam in panic disorder and agoraphobia: Results from a multicenter trial. *Archives of General Psychiatry*, 45, 413-422.
- Ballenger, J.C., Wheadon, D.E., Steiner, M., Bushnell, W. & Gergel, I.P. (1998). Double-blind, fixed-dose, placebo-controlled study of paroxetine in the treatment of panic disorder. *American Journal of Psychiatry*, 155(1), 36-42.
- Barlow, D.H., Cohen, A.S., Waddell, M.T., Vermilyea, B.B., Klosko, J.S., Blanchard, E.B. & DiNardo, P.A. (1984). Panic and generalized anxiety disorder: Nature and treatment. *Behavior Therapy*, 15, 431-449.
- Barlow, D.H., Craske, M.G., Cerny, J.A. & Klosko, J.S. (1989). Behavioral treatment of panic disorder. *Behavior Therapy*, 20, 261-282.

- 
- Barlow, D.H., Gorman, J.M., Shear, M.K. & Woods, S.W. (2000). Cognitive-behavioral therapy, imipramine, or their combination for panic disorder. A randomized controlled trial. *JAMA*, *283*(19), 2529-2536.
- Beauclair, L., Fontaine, R., Annable, L., Holobow, N. & Chouinard, G. (1994). Clonazepam in the treatment of panic disorder: A double-blind, placebo-controlled trial investigating the correlation between clonazepam concentrations in plasma and clinical response. *Journal of Clinical Psychopharmacology*, *14*(2), 111-118.
- Beck, A.T. (1988). Cognitive approaches to panic disorder: Theory and therapy. In: Rachman, S. & Maser, J.D. (Eds.), *Panic: Psychological Perspectives*. Hillsdale, NJ: Lawrence Erlbaum.
- Beck, A.T., Sokol, L., Clark, D.A., Berchick, R. & Wright, F. (1992). A crossover study of focused cognitive therapy for panic disorder. *American Journal of Psychiatry*, *149*, 778-783.
- Beck, J.G., Stanley, M.A., Baldwin, L.E., Deagle, E.A. & Averill, P.M. (1994). Comparison of cognitive therapy and relaxation training for panic disorder. *Journal of Consulting and Clinical Psychiatry*, *62*(4), 818-826.
- Black, D.W., Wesner, R., Bowers, W. & Gabel, J. (1993). A comparison of fluvoxamine, cognitive therapy, and placebo in the treatment of panic disorder. *Archives of General Psychiatry*, *50*, 44-50.
- Borden, J.W., Clum, G.A. & Salomon, P.G. (1991). Mechanisms of change in the treatment of panic. *Cognitive Therapy and Research*, *15*(4), 257-272.
- Bouchard, S., Gauthier, J., Laberge, B., French, D., Pelletier, M.-H. & Godbout, C. (1996). Exposure versus cognitive restructuring in the treatment of panic disorder with agoraphobia. *Behaviour Research and Therapy*, *34*(3), 213-224.
- Broocks, A., Bandelow, B., Pekrun, G., George, A., Meyer, T., Bartmann, U., Hillmer-Vogel, U. & R  ther, E. (1998). Comparison of aerobic exercise, clomipramine, and placebo in the treatment of panic disorder. *American Journal of Psychiatry*, *155*, 603-609.
- Burke, M., Drummond, L.M. & Johnston, D.W. (1997). Treatment choice for agoraphobic women: Exposure or cognitive-behaviour therapy? *British Journal of Clinical Psychology*, *36*, 409-420.
- Caillard, V., Rouillon, F., Viel, J.F., Markabi, S. and the French University Antidepressant Group. (1999). Comparative effects of low and high doses of clomipramine and placebo in panic disorder: a double-blind controlled study. *Acta*



---

*Psychiatrica Scandinavica*, 99, 51-58.

- Carter, C.S., Fawcett, J., Hertzman, M., Papp, L.A., Jones, W., Patterson, W.M., Swinson, R.P., Weise, C.C., Maddock, R.J., Denahan, A.Q. & Liebowitz, M. (1995). Adinazolam-SR in panic disorder with agoraphobia: Relationship of daily dose to efficacy. *Journal of Clinical Psychiatry*, 56(5), 202-210.
- Chambless, D.L., Foa, E.B., Groves, G.A. & Goldstein, A.J. (1979). Flooding with brevitall in the treatment of agoraphobia: Countereffective? *Behaviour Research and Therapy*, 17, 243-251.
- Clark, D.M., Salkovskis, P.M., Hackmann, A., Middleton, H., Anastasiades, P. & Gelder, M. (1994). A comparison of cognitive therapy, applied relaxation and imipramine in the treatment of panic disorder. *British Journal of Psychiatry*, 164, 759-769.
- Clark, D.M., Salkovskis, P.M., Hackmann, A., Wells, A., Ludgate, J. & Gelder, M. (1999). Brief cognitive therapy for panic disorder: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, 67(4), 583-589.
- Clum, G.A., Watkins, P.L., Borden, J.W., Broyles, S.E. & Hayes, J. (1993). A comparison of guided imaginal coping and imaginal exposure in the treatment of panic disorder. *Journal of Rational-Emotive and Cognitive Behaviour Therapy*, 11, 179-193.
- Cottraux, J., Note, I.-D., Cungi, C., Légeron, P., Heim, F., Chneiweiss, L., Bernard, G. & Bouvard, M. (1995). A controlled study of cognitive behaviour therapy with buspirone or placebo in panic disorder with agoraphobia. *British Journal of Psychiatry*, 167, 635-641.
- Craske, M.G., Maidenberg, E. & Bystritsky, A. (1995). Brief cognitive-behavioral versus non-directive therapy for panic disorder. *Journal of Behavior Therapy and Experimental Psychiatry*, 26(2), 113-120.
- Dager, S.R., Roy-Byrne, P., Hendrickson, H., Cowley, D.S., Avery, D.H., Hall, K.C. & Dunner, D.L. (1992). Long-term outcome of panic states during double-blind treatment and after withdrawal of alprazolam and placebo. *Annals of Clinical Psychiatry*, 4(4), 251-258.
- Davidson, J.R.T., Beitman, B., Greist, J.H., Maddock, R.J., Lewis, C.P., Sheridan, A.Q., Carter, C., Krishnan, R., Liebowitz, M.R. & Haack, D.G. (1994). Adinazolam sustained-release treatment of panic disorder: A double-blind study. *Journal of Clinical Psychopharmacology*, 14(4), 255-263.

- 
- De Ruiter, C., Rijken, H., Garssen, B. & Kraaimaat, F. (1989). Breathing retraining, exposure and a combination of both, in the treatment of panic disorder with agoraphobia. *Behaviour Research and Therapy*, 27(6), 647-655.
- DeBeurs, E., van Balkom, A.J.L.M., Lange, A., Koele, P. & van Dyck, R. (1994). Treatment of panic disorder with agoraphobia: Comparison of fluvoxamine, placebo, and psychological panic management combined with exposure and of exposure in vivo alone. *American Journal of Psychiatry*, 152, 683-691.
- Den Boer, J.A. & Westenberg, H.G.M. (1990). Serotonin function in panic disorder: A double blind placebo controlled study with fluvoxamine and ritanserin. *Psychopharmacology*, 102, 85-94.
- Dunner, D.L., Ishiki, D., Avery, D.H., Wilson, L.G. & Hyde, T.S. (1986). Effect of alprazolam and diazepam on anxiety and panic attacks in panic disorder: A controlled study. *Journal of Clinical Psychiatry*, 47(9), 458-460.
- Echeburúa, E., De Corral, P., Bajos, E.G. & Borda, M. (1993). Interactions between self-exposure and alprazolam in the treatment of agoraphobia without current panic: An exploratory study. *Behavioural and Cognitive Psychotherapy*, 21, 219-238.
- Emmelkamp, P.M.G. & Mersch, P.P. (1982). Cognition and exposure in vivo in the treatment of agoraphobia: Short-term and delayed effects. *Cognitive Therapy and Research*, 6(1), 77-88.
- Emmelkamp, P.M.G. (1974). Self-observation versus flooding in the treatment of agoraphobia. *Behaviour Research and Therapy*, 12, 229-237.
- Emmelkamp, P.M.G., Brillman, E., Kuiper, H. & Mersch, P.P. (1986). The treatment of agoraphobia. A comparison of self-instructional training, rational emotive therapy, and exposure in vivo. *Behavior Modification*, 10(1), 37-53.
- Emmelkamp, P.M.G., Kuipers, A.C.M. & Eggeraat, J.B. (1978). Cognitive modification versus prolonged exposure in vivo: A comparison with agoraphobic subjects. *Behaviour Research and Therapy*, 16, 33-41.
- Evans, L., Holt, C. & Oei, T.P.S. (1991). Long term follow-up of agoraphobics treated by brief intensive group cognitive behavioural therapy. *Australian and New Zealand Journal of Psychiatry*, 25, 343-349.
- Evans, L., Kenardy, J., Schneider, P. & Hoey, H. (1986). Effect of a selective serotonin uptake inhibitor in agoraphobia with panic attacks. *Acta psychiatrica scandinavica*, 73, 49-53.

- 
- Fahy, T.J., O'Rourke, D., Brophy, J., Schazmann, W. & Sciascia, S. (1992). The Galway study of panic disorder I: Clomipramine and lofepramine in DSM III-R panic disorder: A placebo controlled trial. *Journal of Affective Disorders*, 25, 63-76.
- Febbraro, G.A.R., Clum, G.A., Roodman, A.A. & Wright, J.H. (1999). The limits of bibliotherapy: A study of the differential effectiveness of self-administered interventions in individuals with panic attacks. *Behavior Therapy*, 30, 209-222.
- Gentil, V., Lotufo-Neto, F., Andrade, L., Cordás, T., Bernik, M., Ramos, R., Maciel, L., Miyakawa, E. & Gorenstein, C. (1993). Clomipramine, a better reference drug for panic/agoraphobia. *Journal of Psychopharmacology*, 7(4), 316-324.
- Gladsjo, J.A., Rapaport, M.H., McKinney, R., Auerbach, M., Hahn, T., Rabin, A., Oliver, T., Hazen, A. & Judd, L.L. (2001). Absence of neuropsychologic deficits in patients receiving long-term treatment with alprazolam-XR for panic disorder. *Journal of Clinical Psychopharmacology*, 21(2), 131-138.
- Gould, R.A. & Clum, G.A. (1995). Self-help plus minimal therapist contact in the treatment of panic disorder: A replication and extension. *Behavior Therapy*, 26, 533-546.
- Gould, R.A., Clum, G.A. & Shapiro, D. (1993). The use of bibliotherapy in the treatment of panic: A preliminary investigation. *Behavior Therapy*, 24, 241-252.
- Griegel, L.A. (1995). Breathing retraining in panic disorder: Physiological mechanisms or perceived controllability. *Dissertational Abstracts International*, 55(9), 4120-B.
- Hecker, J.E., Fink, C.M., Vogelanz, N.D., Thorpe, G.L. & Sigmon, S.T. (1998). Cognitive restructuring and interoceptive exposure in the treatment of panic disorder: A crossover study. *Behavioural and Cognitive Psychotherapy*, 26, 115-131.
- Hoehn-Saric, R., McLeod, D.R. & Hipsley, P.A. (1993). Effect of fluvoxamine on panic disorder. *Journal of Clinical Psychopharmacology*, 13(5), 321-326.
- Hoffart, A. (1995). A comparison of cognitive and guided mastery therapy of agoraphobia. *Behaviour Research and Therapy*, 33(4), 423-434.
- Johnston, D.G., Troyer, I.E. & Whitsett, S.F. (1988). Clomipramine treatment of agoraphobic women. *Archives of General Psychiatry*, 45, 453-459.
- Keller, M.B., Lavori, P.W., Goldenberg, I.M., Baker, L.A., Pollack, M.H., Sachs, G.S., Rosenbaum, J.F., Deltito, J.A., Leon, A., Shear, K. & Klerman, G.L. (1993). Influence of depression on the treatment of panic disorder with imipramine,

- 
- alprazolam and placebo. *Journal of Affective Disorders*, 28, 27-38.
- Klein, B. & Richards, J.C. (2001). A brief internet-based treatment for panic disorder. *Behavioural and Cognitive Psychotherapy*, 29, 113-117.
- Klerman, G.L. (1988). Overview of the cross-national collaborative panic study. *Archives of General Psychiatry*, 45, 407-412.
- Klosko, J.S., Barlow, D.H., Tassinari, R. & Cerny, J.A. (1990). A comparison of alprazolam and behavior therapy in treatment of panic disorder. *Journal of Consulting and Clinical Psychiatry*, 58(1), 77-84.
- Lecrubier, Y., Bakker, A., Dunbar, G., Judge, R. and the Collaborative Paroxetine Panic Study Investigators. (1997). A comparison of paroxetine, clomipramine and placebo in the treatment of panic disorder. *Acta Psychiatrica Scandinavica*, 95, 145-152.
- Lidren, D.M., Watkins, P.L., Gould, R.A., Clum, G.A., Asterino, M. & Tulloch, H.L. (1994). A comparison of bibliotherapy and group therapy in the treatment of panic disorder. *Journal of Consulting and Clinical Psychiatry*, 62(4), 865-869.
- Lipsedge, M.S., Hajioff, J., Huggins, P., Napier, L., Pearce, J., Pike, D.J. & Rich, M. (1973). The management of severe agoraphobia: A comparison of iproniazid and systematic desensitization. *Psychopharmacologia*, 32, 67-80.
- Loerch, B., Graf-Morgenstern, M., Hautzinger, M., Schlegel, S., Hain, C., Sandmann, J. & Benkert, O. (1999). Randomised placebo-controlled trial of moclobemide, cognitive-behavioural therapy and their combination in panic disorder with agoraphobia. *British Journal of Psychiatry*, 174, 205-212.
- Londborg, P.D., Wolkow, R., Smith, W.T., DuBoff, E., England, D., Ferguson, J., Rosenthal, M. & Weise, C. (1998). Sertraline in the treatment of panic disorder. *British Journal of Psychiatry*, 173, 54-60.
- Lydiard, R.B., Lesser, I.M., Ballenger, J.C., Rubin, R.T., Laraia, M. & DuPont, R. (1992). A fixed-dose study of alprazolam 2 mg, alprazolam 6 mg, and placebo in panic disorder. *Journal of Clinical Psychopharmacology*, 12(2), 96-103.
- Lydiard, R.B., Morton, W.A., Emmanuel, N.P., Zealberg, J.J., Laraia, M.T., Stuart, G.W., O'Neil, P.M. & Ballenger, J.C. (1993). Preliminary report: Placebo-controlled, double-blind study of the clinical and metabolic effects of desipramine in panic disorder. *Psychopharmacology*, 29(2), 183-189.
- Marchione, K.E., Michelson, L., Greenwald, M. & Dancu, C. (1987). Cognitive behavioral treatment of agoraphobia. *Behaviour Research and Therapy*, 25(5),

319-328.

- Margraf, J. & Schneider, S. (1991). *Outcome and active ingredients of cognitive-behavioural treatments for panic disorder*. Paper presented at the 25<sup>th</sup> Conference of the Association for Advancement of Behavior Therapy, New York.
- Marks, I.M., Gray, S., Cohen, D., Hill, R., Mawson, D., Ramm, E. & Stern, R.S. (1983). Imipramine and brief therapist-aided exposure in agoraphobics having self-exposure homework. *Archives of General Psychiatry*, *40*, 153-162.
- Marks, I.M., Swinson, R.P., Basoglu, M., Kuch, K., Noshirvani, H., O'Sullivan, G., Lelliott, P.T. Kirby, M., McNamee, G., Sengun, S. & Wickwire, K. (1993). Alprazolam and exposure alone and combined in panic disorder with agoraphobia. A controlled study in London and Toronto. *British Journal of Psychiatry*, *162*, 776-787.
- Mavissakalian, M. & Michelson, L. (1983). Self-directed in vivo exposure practice in behavioral and pharmacological treatments of agoraphobia. *Behavior Therapy*, *14*, 506-519.
- Mavissakalian, M.R. & Perel, J.M. (1995). Imipramine treatment of panic disorder with agoraphobia: Dose ranging and plasma level-response relationships. *American Journal of Psychiatry*, *152*, 673-682.
- Michelson, D., Lydiard, R.B., Pollack, M.H., Tamura, R.N., Hoog, S.L., Tepner, R., Demitrack, M.A., Tollefson, G.D. and the Fluoxetine Panic Disorder Study Group (1998). Outcome assessment and clinical improvement in panic disorder: Evidence from a randomized controlled trial of fluoxetine and placebo. *American Journal of Psychiatry*, *155*(11), 1570-1577.
- Michelson, L., Mavissakalian, M. & Marchione, K. (1988). Cognitive, behavioral, and psychophysiological treatments of agoraphobia: A comparative outcome investigation. *Behavior Therapy*, *19*, 97-120.
- Michelson, L.K., Marchione, K.E., Greenwald, M., Testa, S. & Marchione, N.J. (1996). A comparative outcome and follow-up investigation of panic disorder with agoraphobia: The relative and combined efficacy of cognitive therapy, relaxation training, and therapist-assisted exposure. *Journal of Anxiety Disorders*, *10*(5), 297-330.
- Modigh, K., Westberg, P. & Eriksson, E. (1992). Superiority of clomipramine over imipramine in the treatment of panic disorder: A placebo-controlled trial. *Journal of Clinical Psychopharmacology*, *12*(4), 251-261.

- 
- Moroz, G. & Rosenbaum, J.F. (1999). Efficacy, safety, and gradual discontinuation of clonazepam in panic disorder: A placebo-controlled, multicenter study using optimized dosages. *Journal of Clinical Psychiatry, 60(9)*, 604-612.
- Munjack, D.J., Crocker, B., Cabe, D., Brown, R., Usigli, R., Zulueta, A., McManus, M., McDowell, D., Palmer, R. & Leonard, M. (1989). Alprazolam, propranolol, and placebo in the treatment of panic disorder and agoraphobia with panic attacks. *Journal of Clinical Psychopharmacology, 9(1)*, 22-27.
- Noyes, R., Burrows, G.D., Reich, J.H., Judd, F.K., Garvey, M.J., Norman, T.R., Cook, B.L. & Marriot, P. (1996). Diazepam versus alprazolam for the treatment of panic disorder. *Journal of Clinical Psychiatry, 57(8)*, 349-355.
- Oehrberg, S., Christiansen, P.E., Behnke, K., Borup, A.L., Severin, B., Soegaard, J., Calberg, H., Judge, R., Ohrstrom, J.K. & Manniche, P.M. (1995). Paroxetine in the treatment of panic disorder. A randomized, double-blind, placebo-controlled study. *British Journal of Psychiatry, 167*, 374-379.
- Oei, T.P.S., Llamas, M. & Evans, L. (1997). Does concurrent drug intake affect the long-term outcome of group cognitive behaviour therapy in panic disorder with or without agoraphobia? *Behaviour Research and Therapy, 35(9)*, 851-857.
- Öst, L.-G. & Westling, B.E. (1995). Applied relaxation vs cognitive behavior therapy in the treatment of panic disorder. *Behaviour Research and Therapy, 33(2)*, 145-158.
- Öst, L.-G., Westling, B.E. & Hellström, K. (1993). Applied relaxation, exposure in vivo and cognitive methods in the treatment of panic disorder with agoraphobia. *Behaviour Research and Therapy, 31(4)*, 383-394.
- Pande, A.C., Pollack, M.H., Crockatt, J., Greiner, M., Chouinard, G., Lydiard, R.B., Taylor, C.B., Dager, S.R. & Shiovitz, T. (2000). Placebo-controlled study of gabapentin treatment of panic disorder. *Journal of Clinical Psychopharmacology, 20(4)*, 467-471.
- Pecknold, J., Luthe, L., Munjack, D. & Alexander, P. (1994). A double-blind, placebo-controlled, multicenter study with alprazolam and extended-release alprazolam in the treatment of panic disorder. *Journal of Clinical Psychopharmacology, 14(5)*, 314-321.
- Petterson, K. & Cesare, S. (1996). Panic disorder: A cognitive-behavioural approach to treatment. *Counselling Psychology Quarterly, 9(2)*, 191-202.
- Pohl, R., Balon, R., Yeragani, V.K. & Gershon, S. (1989). Serotonergic anxiolytics in the treatment of panic disorder: A controlled study with buspirone.

---

*Psychopathology*, 22(11), 60-67.

- Pohl, R.B., Wolkow, R.M. & Clary, C.M. (1998). Sertraline in the treatment of panic disorder: A double-blind multicenter trial. *American Journal of Psychiatry*, 155(9), 1189-1195.
- Pollack, M.H., Otto, M.W., Worthington, J.J., Manfro, G.G. & Wolkow, R. (1998). Sertraline in the treatment of panic disorder. *Archives of General Psychiatry*, 55, 1010-1016.
- Pollack, M.H., Worthington, J.J., Otto, M.W., Maki, K.M., Smoller, J.W., Manfro, G.G., Rudolph, R. & Rosenbaum, J.F. (1996). Venlafaxine for panic disorder: Results from a double-blind, placebo-controlled trial. *Psychopharmacology Bulletin*, 32(4), 667-670.
- Reich, J. & Chaudry, D. (1987). Personality of panic disorder alcohol abusers. *The Journal of Nervous and Mental Disease*, 175(4), 224-228.
- Rosenbaum, J.F., Moroz, G. & Bowden, C.L. (1997). Clonazepam in the treatment of panic disorder with or without agoraphobia: A dose-response study of efficacy, safety, and discontinuance. *Journal of Clinical Psychopharmacology*, 17(5), 390-400.
- Sandford, J.J., Forshall, S., Bell, C., Argyropoulos, S., Rich, A., D'Orlando, K. J., Gammans, R.E. & Nutt, D.J. (2001). Crossover trial of pagoclone and placebo in patients with DSM-IV panic disorder. *Journal of Psychopharmacology*, 15(3), 205-208.
- Sandmann, J., Lörch, B., Bandelow, B., Härtter, S., Winter, P., Hiemke, C. & Benkert, O. (1998). Fluvoxamine or placebo in the treatment of panic disorder and relationship to blood concentrations of fluvoxamine. *Pharmacopsychiatry*, 31, 117-121.
- Schmidt, N.B. & Trakowski, M.H. (1997). Efficacy of a brief psychosocial treatment for panic disorder in an active duty sample: Implications for military readiness. *Military Medicine*, 162, 123-129.
- Schmidt, N.B., Trakowski, J.H. & Staab, J.P. (1997). Extinction of panicogenic effects of a 35% CO<sub>2</sub> challenge in patients with panic disorder. *Journal of Abnormal Psychology*, 106(4), 630-638.
- Schmidt, N.B., Woolaway-Bickel, K., Trakowski, J., Santiago, H., Storey, J., Koselka, M. & Cook, J. (2000). Dismantling cognitive-behavioral treatment for panic disorder: Questioning the utility of breathing retraining. *Journal of Consulting and*

---

*Clinical Psychiatry*, 68(3), 417-424.

- Schweizer, E., Patterson, W., Rickels, K. & Rosenthal, M. (1993). Double-blind, placebo-controlled study of a once-a-day, sustained-release preparation of alprazolam for the treatment of panic disorder. *American Journal of Psychiatry*, 150(8), 1210-1215.
- Schweizer, E., Rickels, K., Weiss, S. & Zavodnick, S. (1993). Maintenance drug treatment of panic disorder. *Archives of General Psychiatry*, 50, 51-60.
- Sharp, D.M., Power, K.G., Simpson, R.J., Swanson, V., Moodie, E., Anstee, J.A. & Ashford, J.J. (1996). Fluvoxamine, placebo, and cognitive behaviour therapy used alone and in combination in the treatment of panic disorder and agoraphobia. *Journal of Anxiety Disorders*, 10(4), 219-242.
- Shear, M.K., Pilkonis, P.A., Cloitre, M. & Leon, A.C. (1994). Cognitive behavioral treatment compared with nonprescriptive treatment of panic disorder. *Archives of General Psychiatry*, 51, 395-401.
- Sheehan, D. V., Raj, A. B., Sheehan, K. H. & Soto, S. (1990). Is buspirone effective for panic disorder? *Journal of Clinical Psychopharmacology*, 10(1), 3-11.
- Sheehan, D.V., Raj, A.B., Harnett-Sheenhan, K., Soto, S. & Knapp, E. (1993). The relative efficacy of high-dose buspirone and alprazolam in the treatment of panic disorder: a double-blind placebo-controlled study. *Acta Psychiatrica Scandinavica*, 88, 1-11.
- Sinnott, A., Jones, R.B., Scott-Fordham, A. & Woodward, R. (1981). Augmentation of in vivo exposure treatment for agoraphobia by the formation of neighbourhood self-help groups. *Behaviour Research and Therapy*, 19, 339-347.
- Stein, M.B., Norton, G.R., Walker, J.R., Chartier, M.J. & Graham, R. (2000). Do selective serotonin re-uptake inhibitors enhance the efficacy of very brief cognitive-behavioral therapy for panic disorder? A pilot study. *Psychiatry Research*, 94, 191-200.
- Swinson, R.P., Fergus, K.D., Cox, B.J. & Wickwire, K. (1995). Efficacy of telephone-administered behavioral therapy for panic disorder with agoraphobia. *Behaviour Research and Therapy*, 33(4), 465-469.
- Taylor, C.B., Hayward, C., King, R., Ehlers, A., Margraf, J., Maddock, R., Clark, D., Roth, W.T. & Agras, S. (1990). Cardiovascular and symptomatic reduction effects of alprazolam and imipramine in patients panic disorder: Results of a double-blind, placebo-controlled trial. *Journal of Clinical Psychopharmacology*, 10(2),



112-118.

- Telch, M.J., Agras, W.S., Taylor, C.B., Roth, W.T. & Gallen, C.C. (1985). Combined pharmacological and behavioral treatment for agoraphobia. *Behaviour Research and Therapy*, 23(3), 325-335.
- Telch, M.J., Lucas, J.A., Schmidt, N.B., Hanna, H.H., Jaimez, T.L. & Lucas, R.A. (1993). Group cognitive-behavioral treatment of panic disorder. *Behaviour Research and Therapy*, 31(3), 279-287.
- Tesar, G.E., Rosenbaum, J.F., Pollack, M.H., Otto, M.W., Sachs, G.S., Herman, J.B., Cohen, L.S. & Spier, S.A. (1991). Double-blind, placebo-controlled comparison of clonazepam and alprazolam for panic disorder. *Journal of Clinical Psychiatry*, 52(2), 69-76.
- Teusch, L., Böhme, H. & Gastpar, M. (1997). The benefit of an insight-oriented and experiential approach on panic and agoraphobia symptoms. *Psychotherapy and Psychosomatics*, 66, 293-301.
- Uhlenhuth, E. H., Matuzas, W., Glass, R. M. & Easton, C. (1989). Response of panic disorder to fixed doses of alprazolam or imipramine. *Journal of Affective Disorders*, 17, 261-270.
- Valenca, A.M., Nardi, A.E., Nascimento, I., Mezzasalma, M.A., Lopes, F.L. & Zin, W. (2000). Double-blind clonazepam vs placebo in panic disorder treatment. *Arquivos Neuropsiquiatrica*, 58(4), 1025-1029.
- Van den Hout, M., Arntz, A. & Hoekstra, R. (1994). Exposure reduced agoraphobia but not panic, and cognitive therapy reduced panic but not agoraphobia. *Behaviour Research and Therapy*, 32(4), 447-451.
- van Megen, H.J.G.M., Westenberg, H.G.M., den Boer, J.A., Slaap, B. & Scheepmakers, A. (1997). Effect of the selective serotonin reuptake inhibitor fluvoxamine on CCK-4 induced panic attacks. *Psychopharmacology*, 129, 357-364.
- van Vliet, I. M., Westenberg, H. G. M. & den Boer, J. A. (1996). Effects of the 5-HT<sub>1A</sub> receptor agonist flesinoxan in panic disorder. *Psychopharmacology*, 127, 174-180.
- van Vliet, I.M., Westenberg, H.G.M. & den Boer, J.A. (1993). MAO inhibitors in panic disorder: Clinical effects of treatment with brofaromine. *Psychopharmacology*, 112, 483-489.
- Versiani, M., Cassano, G., Perugi, G., Benedetti, A., Mastalli, L., Nardi, A. & Savino, M. (2002). Reboxetine, a selective norepinephrine reuptake inhibitor, is an effective and well-tolerated treatment for panic disorder. *Journal of Clinical*

- 
- Psychiatry*, 63(1), 31-37.
- Wade, A.G., Lepola, U., Koponen, H.J., Pedersen, V. & Pedersen, T. (1997). The effect of citalopram in panic disorder. *British Journal of Psychiatry*, 170, 549-553.
- Wardle, J., Hayward, P., Higgitt, A., Stabl, M., Blizard, R. & Gray, J. (1994). Effects of concurrent diazepam treatment on the outcome of exposure therapy in agoraphobia. *Behaviour Research and Therapy*, 32(2), 203-215.
- Williams, S.L. & Falbo, J. (1996). Cognitive and performance-based treatments for panic attacks in people with varying degrees of agoraphobic disability. *Behaviour Research and Therapy*, 34(3), 253-264.
- Williams, S.L. & Rappoport, A. (1983). Cognitive treatment in the natural environment for agoraphobics. *Behavior Therapy*, 14, 299-313.
- Wulsin, L., Liu, T., Storrow, A., Evans, S., Dewan, N. & Hamilton, C. (2002). A randomized, controlled trial of panic disorder treatment initiation in an emergency department chest pain center. *Annals of Emergency Medicine*, 39(2), 139-143.
- Wulsin, L.R., Maddock, R., Beitman, B., Dawaher, R. & Wells, V.E. (1999). Clonazepam treatment of panic disorder in patients with recurrent chest pain and normal coronary arteries. *International Journal of Psychiatry in Medicine*, 29(1), 97-105.
- Zane, G. & Williams, S.L. (1993). Performance-related anxiety in agoraphobia: Treatment procedures and cognitive mechanisms of change. *Behavior Therapy*, 24, 625-643.
- Zitrin, C.M., Klein, D.F. & Woerner, M.G. (1980). Treatment of agoraphobia with group exposure in vivo and imipramine. *Archives of General Psychiatry*, 37, 63-72.
- Zitrin, C.M., Klein, D.F., Woerner, M.G. & Ross, D.C. (1983). Treatment of phobias. I. Comparison of imipramine hydrochloride and placebo. *Archives of General Psychiatry*, 40, 125-138.

### 6.3 Eingeschlossene Studien Posttraumatische Belastungsstörung

Anmerkung: es wurde nur eine Publikation pro Studie aufgenommen.

- Baker, D.G., Diamond, B.I., Gilette, G., Hamner, M., Katzelnick, D., Keller, T., Mellman, T.A., Pontius, E., Rosenthal, M., Tucker, P., van der Kolk, B.A. & Katz, R. (1995). A double-blind, randomized, placebo-controlled, multi-center study of brofaromine in the treatment of post-traumatic stress disorder. *Psychopharmacology*, *122*, 386-389.
- Blanchard, E.B., Hickling, E.J., Devineni, T., Veazey, C. H., Galovski, T.E., Mundy, E., Malta, L. S. & Buckley, T.C. (in press). A controlled evaluation of cognitive behavioral therapy for posttraumatic stress in motor vehicle accident survivors. *Behaviour Research and Therapy*.
- Brady, K., Pearlstein, T., Asnis, G.M., Baker, D., Rothbaum, B., Sikes, C.R. & Farfel, G.M. (2000). Efficacy and safety of sertraline treatment of posttraumatic stress disorder. *JAMA*, *283(14)*, 1837-1844.
- Braun, P., Greenberg, D., Dasberg, H. & Lerer, B. (1990). Core symptoms of posttraumatic stress disorder unimproved by alprazolam treatment. *Journal of Clinical Psychiatry*, *51(6)*, 236-238.
- Connor, K.M., Sutherland, S.M., Tupler, L.A., Malik, M.M. & Davidson, J.R.T. (1999). Fluoxetine in post-traumatic stress disorder. *British Journal of Psychiatry*, *175*, 17-22.
- Davidson, J., Kudler, H., Smith, R., Mahorney, S. L., Lipper, S., Hammett, E., Saunders, W.B. & Cavenar, J. O. (1990). Treatment of posttraumatic stress disorder with amitriptyline and placebo. *Archives of General Psychiatry*, *47*, 259-266.
- Davidson, J.R.T., Rothbaum, B. O., van der Kolk, B. A., Sikes, C.R. & Farfel, G. M. (2001). Multicenter, double-blind comparison of sertraline and placebo in the treatment of posttraumatic stress disorder. *Archives of General Psychiatry*, *58*, 485-492.
- Ehlers, A., Clark, D.M., Hackmann, A., McManus, F., Fennell, M., Herbert, C. & Mayou, R. (subm.). A randomized controlled trial of cognitive therapy, a self-help booklet, and repeated assessments as early interventions for posttraumatic stress disorder.

- Fecteau, G. & Nicki, R. (1999). Cognitive behavioural treatment of post traumatic stress disorder after motor vehicle accident. *Behavioural and Cognitive Psychotherapy*, 27, 201-214.
- Foa, E.B., Dancu, C.V., Hembree, E.A., Jaycox, L.H., Meadows, E.A. & Street, G. P. (1999). A comparison of exposure therapy, stress inoculation training, and their combination for reducing posttraumatic stress disorder in female assault victims. *Journal of Consulting and Clinical Psychology*, 67(2), 194-200.
- Foa, E.B., Rothbaum, B. O., Riggs, D.S. & Murdock, T. B. (1991). Treatment of posttraumatic stress disorder in rape victims: A comparison between cognitive-behavioral procedures and counseling. *Journal of Consulting and Clinical Psychology*, 59(5), 715-723.
- Glynn, S.M., Eth, S., Randolph, E.T., Foy, D.W., Urbaitis, M., Boxer, L., Paz, G.G., Leong, G.B., Firman, G., Salk, J.D., Katzman, J.W. & Crothers, J. (1999). A test of behavioral family therapy to augment exposure for combat-related posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology*, 67(2), 243-251.
- Hertzberg, M. A., Butterfield, M. I., Feldman, M.E., Beckham, J.C., Sutherland, S.M., Connor, K.M. & Davidson, J.R.T. (1999). A preliminary study of lamotrigine for the treatment of posttraumatic stress disorder. *Biological Psychiatry*, 45, 1226-1229.
- Hertzberg, M.A., Feldman, M.E., Beckham, J.C., Kudler, H.S. & Davidson, J.R.T. (2000). Lack of efficacy for fluoxetine in PTSD: A placebo controlled trial in combat veterans. *Annals of Clinical Psychiatry*, 12(2), 101-105.
- Keane, T.M., Fairbank, J.A., Cadell, J.M. & Zimering, R.T. (1989). Implosive (flooding) therapy reduces symptoms of PTSD in vietnam combat veterans. *Behavior Therapy*, 20, 245-260.
- Kosten, T.R., Frank, J.B., Dan, E., McDougale, C. J. & Giller, E.L. (1991). Pharmacotherapy for posttraumatic stress diorder using phenelzine or imipramine. *Journal of Nervous and Mental Disease*, 179, 366-370.
- Marks, I.M., Lovell, K., Noshirvani, H., Livanou, M. & Thrasher, S. (1998). Treatment of posttraumatic stress disorder by exposure and/or cognitive restructuring. A controlled study. *Archives of General Psychiatry*, 55, 317-325.
- Marshall, R.D., Beebe, K.L., Oldham, M. & Zaninelli, R. (2001). Efficacy and safety of paroxetine treatment for chronic PTSD: A fixed-dose, placebo-controlled study.

---

*American Journal of Psychiatry*, 158, 1982-1988.

- Martenyi, F., Brown, E.B., Zhang, H., Prakash, A. & Koke, S.C. (2002). Fluoxetine versus placebo in posttraumatic stress disorder. *Journal of Clinical Psychiatry*, 63(3), 199-206.
- Paunovic, N. & Öst, L.-G. (2001). Cognitive-behavior therapy vs. exposure therapy in the treatment of PTSD in refugees. *Behaviour Research and Therapy*, 39, 1183-1197.
- Resick, P.A., Nishith, P., Weaver, T. L., Astin, M.C. & Feuer, C.A. (2002). A comparison of cognitive processing therapy with prolonged exposure and a waiting condition for the treatment of chronic posttraumatic stress disorder in female rape victims. *Journal of Consulting and Clinical Psychology*, 70(4), 867-879.
- Shestatzky, M., Greenberg, D. & Lerer, B. (1987). A controlled trial of phenelzine in posttraumatic stress disorder. *Psychiatry Research*, 24, 149-155.
- Tarrier, N., Pilgrim, H., Sommerfield, C., Faragher, B., Reynolds, M., Graham, E. & Barrowclough, C. (1999). A randomized controlled trial of cognitive therapy and imaginal exposure in the treatment of chronic posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology*, 67(1), 13-18.
- Tucker, P., Zaninelli, R., Yehuda, R., Ruggiero, L., Dillingham, K. & Pitts, C.D (2001). Paroxetine in the treatment of chronic posttraumatic stress disorder: Results of a placebo-controlled, flexible-dosage trial. *Journal of Clinical Psychiatry*, 62(11), 860-868.
- van der Kolk, B.A., Dreyfuss, D., Michaels, M., Shera, D., Berkowitz, R., Fislser, R. & Saxe, G. (1994). Fluoxetine in posttraumatic stress disorder. *Journal of Clinical Psychiatry*, 55(12), 517-522.

## 6.4 Eingeschlossene Studien Soziale Phobie

Anmerkung: es wurde nur eine Publikation pro Studie aufgenommen.

- Akillas, E. & Efran, J.S. (1995). Symptom prescription and reframing: Should they be combined? *Cognitive Therapy and Research*, 19(3), 263-279.
- Alden, L. (1989). Short-term structured treatment for avoidant personality disorder. *Journal of Consulting and Clinical Psychology*, 57(6), 756-764.
- Allgulander, C. (1999). Paroxetine in social anxiety disorder: A randomized placebo-controlled study. *Acta Psychiatrica Scandinavica*, 100, 193-198.
- Baldwin, D., Bobes, J., Stein, D.J., Scharwächter, I. & Faure, M. (1999). Paroxetine in social phobia/social anxiety disorder. *British Journal of Psychiatry*, 175, 120-126.
- Blomhoff, S., Haug, T.T., Hellström, K., Holme, I., Humble, M., Madsbu, H.P. & Wold, J.E. (2001). Randomized controlled general practice trial of sertraline, exposure therapy and combined treatment in generalized social phobia. *British Journal of Psychiatry*, 179, 23-30.
- Butler, G., Cullington, A., Munby, M., Amies, P. & Gelder, M. (1984). Exposure and anxiety management in the treatment of social phobia. *Journal of Consulting and Clinical Psychology*, 52(4), 642-650.
- Clark, D.B. & Agras, W.S. (1991). The assessment and treatment of performance anxiety in musicians. *American Journal of Psychiatry*, 148, 5,598-605.
- Cottraux, J., Note, I., Albuissou, E., Yao, S.N., Note, B., Mollard, E., Bonasse, F., Jalenques, I., Guerin, J. & Coudert, A.J. (2000). Cognitive behavior therapy versus supportive therapy in social phobia: a randomized controlled trial. *Psychotherapy and Psychosomatics*, 69, 137-146.
- Davidson, J.R.T., Potts, N., Richichi, E., Krishnan, R., Ford, S.M., Smith, R. & Wilson, W.H. (1993). Treatment of social phobia with clonazepam and placebo. *Journal of Clinical Psychopharmacology*, 13(6), 423-428.
- Emmelkamp, P.M.G., Mersch, P.-P., Vissia, E. & Van der Helm, M. (1985). Case histories and shorter communications. Social phobia: a comparative evaluation of cognitive and behavioral interventions. *Behaviour Research and Therapy*, 23(3), 365-369.
- Fahlén, T., Nilsson, H. L., Borg, K., Humble, M. & Pauli, U. (1995). Social phobia: The clinical efficacy and tolerability of the monoamine oxidase-A and serotonin uptake inhibitor brofaromine. *Acta Psychiatrica Scandinavica*, 92, 351-358.

- 
- Fallon, I.R., Lloyd, G.G. & Harpin, R.E. (1981). The treatment of social phobia. Real life rehearsal with nonprofessional therapists. *The Journal of Nervous and Mental Disease*, 169, 3, 180-184.
- Gelernter, C.S., Uhde, T.W. Cimbolic, P., Arnkoff, D.B., Vittone, B.J., Tancer, M.E. & Bartko, J.J. (1991). Cognitive-behavioral and pharmacological treatments of social phobia. A controlled study. *Archives of General Psychiatry*, 48, 938-945.
- Heimberg, R.G., Dodge, C.S., Hope, D.A., Kennedy, C.R. & Zollo, L.J. (1990). Cognitive behavioral group treatment for social phobia: comparison with a credible placebo control. *Cognitive Therapy and Research*, 14(1), 1-23.
- Heimberg, R.G., Liebowitz, M.R., Hope, D.A., Schneier, F.R., Holt, C.S., Welkowitz, L.A., Juster, H.R., Campeas, R., Bruch, M.A., Cloitre, M., Fallon, B. & Klein, D.F. (1998). Cognitive behavioral group therapy vs phenelzine therapy for social phobia. *Archives of General Psychiatry*, 55, 1133-1141.
- Hoffmann, S., Persönliche Kommunikation, 06.02.2002
- Hope, D.A., Heimberg, R.G. & Bruch, M.B. (1995). Dismantling cognitive-behavioral group therapy for social phobia. *Behaviour Research and Therapy*, 33(6), 637-650.
- Jerremalm, A., Jansson, L. & Öst, L.-G. (1986). Cognitive and physiological reactivity and the effects of different behavioral methods in the treatment of social phobia. *Behaviour Research and Therapy*, 24(2), 171-180.
- Katzelnick, D.J., Kobak, K.A., Greist, J.H., Jefferson, J.W., Mantle, J.M. & Serlin, R.C. (1995). Sertraline for social phobia: A double-blind, placebo-controlled crossover study. *American Journal of Psychiatry*, 152(9), 1368-1371.
- Liebowitz, M.R., Schneier, F., Campeas, R., Hollander, E., Hatterer, J., Fyer, A., Gorman, J., Papp, L., Davies, S., Gully, R. & Klein, D.F. (1992). Phenelzine vs. atenolol in social phobia. *Archives of General Psychiatry*, 49, 290-300.
- Lott, M., Greist, J.H., Jefferson, J.W., Kobak, K.A., Katzelnick, D.J., Katz, R.J. & Schaettle, S.C. (1997). Brofaromine for social phobia: A multicenter, placebo-controlled, double-blind study. *Journal of Clinical Psychopharmacology*, 17(4), 255-260.
- Lucas, R.A. (1994). A comparative treatment outcome study of social phobia. *Dissertational Abstracts International*, 54(12), 6465-B.
- Marziller, J.S., Lambert, C. & Kellett, J. (1976). A controlled evaluation of systematic desensitisation and social skills training for socially inadequate psychiatric

- 
- patients. *Behaviour Research and Therapy*, 14, 225-238.
- Mattick, R.P. & Peters, L. (1988). Treatment of severe social phobia: effects of guided exposure with and without cognitive restructuring. *Journal of Consulting and Clinical Psychology*, 56(2), 251-260.
- Mattick, R.P., Peters, L. & Clarke, J.C. (1989). Exposure and cognitive restructuring for social phobia: a controlled study. *Behavior Therapy*, 20, 3-23.
- Mersch, P.P.A. (1995). The treatment of social phobia: the differential effectiveness of exposure in vivo and an integration of exposure in vivo, rational emotive therapy and social skills training. *Behaviour Research and Therapy*, 33(3), 259-269.
- Mersch, P.P.A., Emmelkamp, P.M.G., Bögels, S.M. & van der Sleen, J. (1989). Social phobia: Individual response patterns and the effects of behavioral and cognitive interventions. *Behaviour Research and Therapy*, 27(4), 421-434.
- Montgomery, S.A. (1998). Implications of the severity of social phobia. *Journal of Affective Disorders*, 50, 17-22.
- Munjack, D.J., Baltazar, P.L., Bohn, P.B., Cabe, D.D. & Appleton, A.A. (1990). Clonazepam in the treatment of social phobia: A pilot study. *Journal of Clinical Psychiatry*, 51(5), 35-53.
- Newman, M.G., Hofmann, S.G., Trabert, W., Roth, W.T. & Taylor, C.B. (1994). Does behavioural treatment of social phobia lead to cognitive changes? *Behavior Therapy*, 25, 503-517.
- Oosterbaan, D.B. (2000). *Social Phobia. Cognitive and pharmacological treatment*. Unveröffentlichte Dissertation. Universität Amsterdam.
- Otto, M.W., Pollack, M.H., Gould, R.A., Worthington, J.J., McArdle, E.T. & Rosenbaum, J.F. (2000). A comparison of the efficacy of clonazepam and cognitive-behavioral group therapy for the treatment of social phobia. *Journal of Anxiety Disorders*, 14(4), 345-358.
- Pande, A.C., Davidson, J.R.T., Jefferson, J.W., Janney, C.A., Katzelnick, D.J., Weisler, R.H., Greist, J.H. & Sutherland, S.M. (1999). Treatment of social phobia with gabapentin: A placebo-controlled study. *Journal of Clinical Psychopharmacology*, 19(4), 341-348.
- Salaberria, K. & Echeburua, E. (1998). Long-term outcome of cognitive therapy's contribution to self-exposure in vivo to the treatment of generalized social phobia. *Behavior Modification*, 22(3), 262-284.
- Schneier, F.R., Goetz, D., Campeas, R., Fallon, B., Marshall, R. & Liebowitz, M.R.



- (1998). Placebo-controlled trial of moclobemide in social phobia. *British Journal of Psychiatry*, 172, 70-77.
- Scholing, A. & Emmelkamp, P.M.G. (1993a). Cognitive and behavioural treatments of fear of blushing, sweating or trembling. *Behaviour Research and Therapy*, 31(2), 155-170.
- Scholing, A. & Emmelkamp, P.M.G. (1993b). Exposure with and without cognitive therapy for generalized social phobia: effects of individual and group treatment. *Behaviour Research and Therapy*, 31(7), 667-681
- Stein, M.B., Fyer, A.J., Davidson, J.R.T., Pollack, M.H. & Wiita, B. (1999). Fluvoxamine treatment of social phobia (social anxiety disorder): A double-blind, placebo-controlled study. *American Journal of Psychiatry*, 156(5), 756-760.
- Stein, M.B., Liebowitz, M.R., Lydiard, R.B., Pitts, C.D., Bushnell, W. & Gergel, I. (1998). Paroxetine treatment of generalized social phobia (social anxiety disorder). *JAMA*, 280(8), 708-713.
- Stravynski, A., Arbel, N., Bounader, J., Gaudette, G., Lachance, L., Borgeat, F., Fabian, J., Lamontagne, Y., Sidoun, P. & Todorov, C. (2000). Social phobia treated as a problem in social functioning: a controlled comparison of two behavioral group approaches. *Acta Psychiatrica Scandinavica*, 102, 188-198.
- Stravynski, A., Marks, I. & Yule, W. (1982). Social skill problems in neurotic outpatients. Social skills training with and without cognitive modification. *Archives of General Psychiatry*, 39, 1378-1385.
- Taylor, S., Woody, S., Koch, W.J., McLean, P., Paterson, R.J. & Anderson, K.W. (1997). Cognitive restructuring in the social phobia. Efficacy and mode of action. *Behavior Modification*, 21(4), 487-511.
- The International Multicenter Clinical Trial Group on Moclobemide in Social Phobia (1997). Moclobemide in social phobia. A double-blind, placebo-controlled clinical study. *European Archives of Psychiatry and Clinical Neuroscience*, 247, 71-80.
- Turner, S.M., Beidel, D.C. & Jacob, R.G. (1994). Social phobia: a comparison of behavior therapy and atenolol. *Journal of Consulting and Clinical Psychology*, 62(2), 350-358.
- Van Ameringen, M.A., Lane, R.M., Walker, J.R., Bowen, R.C., Chokka, P.R., Goldner, E.M., Johnston, D.G., Lavalley, Y.-J., Nandy, S., Pecknold, J.C., Hadrava, V. & Swinson, R.P. (2001). Sertraline treatment of generalized social phobia: A 20-week, double-blind, placebo-controlled study. *American Journal of Psychiatry*,

158(2), 275-281.

- van Vliet, I.M., den Boer, J.A. & Westenberg, H.G.M. (1992). Psychopharmacological treatment of social phobia: clinical and biochemical effects of brofaromine, a selective MAO-A inhibitor. *European Neuropsychopharmacology*, 2, 21-29.
- van Vliet, I.M., den Boer, J.A. & Westenberg, H.G.M. (1994). Psychopharmacological treatment of social phobia; a double blind placebo controlled study with fluvoxamine. *Psychopharmacology*, 115, 128-134.
- van Vliet, I.M., den Boer, J.A., Westenberg, H.G.M. & Ho Pian, K.L. (1997). Clinical effects of buspirone in social phobia: A double-blind placebo-controlled study. *Journal of Clinical Psychiatry*, 58(4), 164-168.
- Versiani, M., Nardi, A.E., Figueira, I., Mendlowicz, M. & Marques, C. (1997). Double-blind placebo controlled trial with bromazepam in social phobia. *Journal Brasileiro de Psiquiatria*, 46(3), 167-171.
- Versiani, M., Nardi, A.E., Mundim, F.D., Alves, A.B., Liebowitz, M.R. & Amrein, R. (1992). Pharmacotherapy of social phobia. *British Journal of Psychiatry*, 161, 353-360.

## 6.5 Eingeschlossene Studien Spezifische Phobie

Anmerkung: es wurde nur eine Publikation pro Studie aufgenommen.

- Arntz, A., Merckelbach, H. & de Jong, P. (1993). Opioid antagonist affects behavioral effects of exposure in vivo. *Journal of Consulting and Clinical Psychology, 61*(5), 865-870.
- Baker, B.L., Cohen, D.C. & Saunders, J.T. (1973). Self-directed desensitization for acrophobia. *Behaviour Research and Therapy, 11*, 79-89.
- Benjamin, J., Ben-Zion, I.Z., Karbofsky, E. & Dannon, P. (2000). Double-blind placebo-controlled pilot study of paroxetine for specific phobia. *Psychopharmacology, 149*, 194-196.
- Bernstein, D.A. & Kleinknecht, R.A. (1982). Multiple approaches to the reduction of dental fear. *Journal of Behavior Therapy and Experimental Psychiatry, 13*(4), 287-292.
- Booth, R. & Rachman, S. (1992). The reduction of claustrophobia-I. *Behaviour Research and Therapy, 30*, 207-221.
- Bornas, X., Tortella-Feliu, M., Llabrés, J. & Fullana, M.A. (2001). Computer-assisted exposure treatment for flight phobia: A controlled study. *Psychotherapy Research, 11*(3), 259-273.
- Cohen, D.C. (1977). Comparison of self-report and overt-behavioral procedures for assessing acrophobia. *Behavior Therapy, 8*, 17-23.
- DeJongh, A., Muris, P., ter Horst, G., van Zuuren, F., Schoenmakers, N. & Makkes, P. (1995). One-session cognitive treatment of dental phobia: Preparing dental phobics for treatment by restructuring negative cognitions. *Behaviour Research and Therapy, 33*(8), 947-954.
- Ekeberg, O., Kjeldsen, S.E., Greenwood, D.T. & Enger, E. (1990). Effects of selective beta-adrenoceptor blockade on anxiety associated with flight phobia. *Journal of Psychopharmacology, 4*(1), 35-41.
- Getka, E.J. & Glass, C.R. (1992). Behavioral and cognitive-behavioral approaches to the reduction of dental anxiety. *Behavior Therapy, 23*, 433-448.
- Garcia-Palacios, A., Hoffmann, H., Carlin, A., Furness, T.A. & Botella, C. (2002). Virtual reality in the treatment of spider phobia: A controlled study. *Behaviour Research and Therapy, 40*, 983-993.
- Howard, W.A., Murphy, S.M. & Clarke, J.C. (1983). The nature and treatment of fear

- of flying: A controlled investigation. *Behavior Therapy*, 14, 557-567.
- Jerremalm, A., Öst, L.G. & Johansson, J. (1986). Individual response patterns and the effects of different behavioral methods in the treatment of dental phobia. *Behaviour Research and Therapy*, 24(5), 587-596.
- Ladouceur, R. (1983). Participant modeling with or without cognitive treatment for phobias. *Journal of Consulting and Clinical Psychology*, 51(6), 942-944.
- Moore, R. (1991). Dental fear treatment: Comparison of a video training procedure and clinical rehearsals. *Scandinavian Journal of Dental Research*, 99, 229-235.
- Morris, R.J. & Margrath, K.H. (1979). Contribution of therapist warmth to the contact desensitization treatment of acrophobia. *Journal of Consulting and Clinical Psychology*, 47(4), 786-788.
- Moses, A.N. & Hollandsworth, J.G. (1985). Relative effectiveness of education alone versus stress inoculation training in the treatment of dental phobia. *Behavior Therapy*, 16, 531-537.
- Muris, P. & Merckelbach, H. (1997). Treating spider phobics with eye movement desensitization and reprocessing: A controlled study. *Behavioural and Cognitive Psychotherapy*, 25, 39-50.
- Öst, L.G. (1978). Fading vs systematic desensitization in the treatment of snake and spider phobia. *Behaviour Research and Therapy*, 16, 379-389.
- Öst, L.G., Johansson, J. & Jerremalm, A. (1982). Individual response patterns and the effects of different behavioral methods in the treatment of claustrophobia. *Behaviour Research and Therapy*, 20, 445-460.
- Pendleton, M.G. & Higgins, R.L. (1983). A comparison of negative practice and systematic desensitization in the treatment of acrophobia. *Journal of Behavior Therapy and Experimental Psychiatry*, 14(4), 317-323.
- Rothbaum, B.O., Hodges, L., Smith, S., Lee, J.H. & Price, L. (2000). A controlled study of virtual reality exposure therapy for the fear of flying. *Journal of Consulting and Clinical Psychology*, 68(6), 1020-1026.
- Shaw, D.W. & Thoresen, C.E. (1974). Effects of modeling and desensitization in reducing dentist phobia. *Journal of Counseling Psychology*, 21(5), 415-520.
- Thom, A., Sartory, G. & Jöhren, P. (2000). Comparison between one-session psychological treatment and benzodiazepine in dental phobia. *Journal of Consulting and Clinical Psychology*, 68(3), 378-387.
- Whitehead, W.E., Blackwell, B. & Robinson, A. (1978). Effects of diazepam on phobic

- avoidance behavior and phobic anxiety. *Biological Psychiatry*, 13(1), 59-64.
- Wilhelm, F.H. & Roth, W.T. (1998). Akute und verzögerte Effekte von Alprazolam auf Flugphobiker während Exposition in vivo. *Verhaltenstherapie*, 8, 38-47.
- Williams, S.L., Dooseman, G. & Kleifield, E. (1984). Comparative effectiveness of guided mastery and exposure treatments for intractable phobias. *Journal of Consulting and Clinical Psychology*, 52(4), 505-518.
- Williams, S.L., Turner, S.M. & Peer, D.F. (1985). Guided mastery and performance desensitization treatments for severe acrophobia. *Journal of Consulting and Clinical Psychology*, 53(2), 237-247.
- Wroblewski, P.F., Jacob, T. & Rehm, L.P. (1977). The contribution of relaxation to symbolic modeling in the modification of dental fears. *Behaviour Research and Therapy*, 15, 113-117.

## 6.6 Eingeschlossene Studien Zwangsstörung

Anmerkung: es wurde nur eine Publikation pro Studie aufgenommen.

- Austin, L.S., Lydiard, R.B., Ballenger, J.C., Cohen, B.M., Laraia, M.T., Zealberg, J.J., Fossey, M.D. & Ellinwood, E.H. (1991). Dopamine blocking activity of clomipramine in patients with obsessive-compulsive disorder. *Biological Psychiatry*, 30, 225-232.
- Baxter, L.R., Schwartz, J.M., Bergman, K.S., Szuba, M.P., Guze, B.H., Mazziotta, J.C., Alazraki, A., Selin, C.E., Ferng, H.-K., Munford, P. & Phelps, M.E. (1992). Caudate glucosemetabolic rate changes with both drug and behavior therapy for obsessive-compulsive disorder. *Archives of General Psychiatry*, 49, 681-689.
- Chouinard, G., Goodman, W.K., Greist, J., Jenike, M.A., Rasmussen, S.A., White, K., Hackett, E., Gaffney, M. & Bick, P.A. (1990). Results of a double-blind placebo controlled trial of a new serotonin uptake inhibitor, sertraline, in the treatment of obsessive-compulsive disorder. *Psychopharmacology Bulletin*, 26(3), 279-284.
- Cottraux, J., Mollard, E., Bouvard, M., Marks, I.M., Sluys, M., Nury, A.M., Douge, R. & Ciadella, P. (1990). A controlled study of fluvoxamine and exposure in obsessive-compulsive disorder. *International Clinical Psychopharmacology*, 5, 17-30.
- Cottraux, J., Note, I., Yao, S.N., Lafont, S., Note, B., Mollard, E., Bouvard, M., Sauteraud, A., Bourgeois, M. & Dartigues, J.-F. (2001). A randomized controlled trial of cognitive therapy versus intensive behavior therapy in obsessive compulsive disorder. *Psychotherapy and Psychosomatics*, 70(6), 288-297.
- Den Boer, J.A. & Westenberg, H.G.M. (1992). Oxytocin in obsessive compulsive disorder. *Peptides*, 13, 1083-1085.
- Emmelkamp, P.M.G. & Beens, H. (1991). Cognitive therapy with obsessive-compulsive disorder: A comparative evaluation. *Behaviour Research and Therapy*, 29, 293-300
- Emmelkamp, P.M.G., van der Helm, M., van Zanten, B.L. & Plochg, I. (1980). Treatment of obsessive-compulsive patients: The contribution of self-instructional training to the effectiveness of exposure. *Behaviour Research and Therapy*, 18, 61-66.
- Emmelkamp, P.M.G., Visser, S. & Hoekstra, R.J. (1988). Cognitive therapy vs exposure in vivo in the treatment of obsessive-compulsives. *Cognitive Therapy and Research*, 12(1), 103-114.

- 
- Foa, E.B., Kozak, M.J., Steketee, G.S. & McCarthy, P.R. (1992). Treatment of depressive and obsessive-compulsive symptoms in OCD by imipramine and behaviour therapy. *British Journal of Clinical Psychology*, *31*, 279-292.
- Foa, E.B., Steketee, G.S., Kozak, M.J. & Dugger, D. (1987). Imipramine and placebo in the treatment of obsessive-compulsives: Their effect on depression and on obsessional symptoms. *Psychopharmacology Bulletin*, *23(1)*, 8-12.
- Franklin, M.E., Abramowitz, J.S., Bux Jr., D.A., Zoellner, L.A. & Feeny, N.C. (2002). Cognitive-behavioural therapy with and without medication in the treatment of obsessive-compulsive disorder. *Professional Psychology: Research and Practice*, *33(2)*, 162-168.
- Fux, M., Levine, J., Aviv, A. & Belmaker, R.H. (1996). Inositol treatment of obsessive-compulsive disorder. *American Journal of Psychiatry*, *153*, 1219-1221.
- Goodman, W.K., Kozak, M.J., Liebowitz, M. & White, K.L. (1996). Treatment of obsessive-compulsive disorder with fluvoxamine: A multicentre, double-blind, placebo-controlled trial. *International Clinical Psychopharmacology*, *11*, 21-29.
- Goodman, W.K., Price, L.H., Rasmussen, S.A., Delgado, P.L., Heninger, G.R. & Charney, D.S. (1989). Efficacy of fluvoxamine in obsessive-compulsive disorder. *Archives of General Psychiatry*, *46*, 36-44.
- Greist, J., Chouinard, G., DuBoff, E., Halaris, A., Kim, S.W., Koran, L., Liebowitz, M., Lydiard, R.B., Rasmussen, S., White, K. & Sikes, C. (1995). Double-blind parallel comparison of three dosages of sertraline and placebo in outpatients with obsessive-compulsive disorder. *Archives of General Psychiatry*, *52*, 289-295.
- Hohagen, F., Winkelmann, G., Rasche-Räuchle, H., Hand, I., König, A., Münchau, N., Hiss, H., Geiger-Kabisch, C., Käppler, C., Schramm, P., Rey, E., Aldenhoff, J. & Berger, M. (1998). Combination of behaviour therapy with fluvoxamine in comparison with behaviour therapy and placebo. *British Journal of Psychiatry*, *173 (suppl. 35)*, 71-78.
- Jenike, M.A., Baer, L., Minichiello, W.E., Rauch, S.L. & Buttolph, M.L. (1997). Placebo-controlled trial of fluoxetine and phenelzine for obsessive-compulsive disorder. *American Journal of Psychiatry*, *154*, 1261-1264.
- Jenike, M.A., Hyman, S., Baer, L., Holland, A., Minichiello, W.E., Buttolph, M.L., Summergrad, P., Seymour, R. & Ricciardi, J. (1990). A controlled trial of fluvoxamine in obsessive-compulsive disorder: Implications of serotonergic theory. *American Journal of Psychiatry*, *147*, 1209-1215.

- 
- Krone, K.P., Himle, J.A. & Nesse, R.M. (1991). A standardized behavioural group treatment program for obsessive-compulsive disorder: Preliminary outcomes. *Behaviour Research and Therapy*, *29*(6), 627-631.
- Kronig, M.H., Apter, J., Asnis, G., Bystritsky, A., Curtis, G., Perguson, J., Landbloom, R., Munjack, D., Riesenber, R., Robinson, D., Roy-Byrne, P., Phillips, K. & Du Pont, I.J. (1999). Placebo-controlled multicenter study of sertraline treatment for obsessive-compulsive disorder. *Journal of Clinical Psychopharmacology*, *19*, 172-176.
- Lakatos, A. (1997). *Kognitiv-behaviorale Therapie für Zwangsstörungen. Eine Therapievergleichsstudie*. Regensburg: S.Roderer.
- Lovell, K., Marks, I.M., Noshirvani, H. & O'Sullivan, G. (1994). Should treatment distinguish anxiogenic from anxiolytic obsessive-compulsive ruminations? *Psychotherapy and Psychosomatics*, *61*, 150-155.
- Marks, I.M., Lelliott, P., Basoglu, M., Noshirvani, H., Monteiro, W., Cohen, D. & Kasvikis, Y. (1988). Clomipramine, self-exposure and therapist-aided exposure for obsessive-compulsive rituals. *British Journal of Psychiatry*, *152*, 522-534.
- Marks, I.M., Stern, R.S., Mawson, D., Cobb, J. & McDonald, R. (1980). Clomipramine and exposure for obsessive-compulsive rituals. *British Journal of Psychiatry*, *136*, 1-25.
- Mavissakalian, M., Turner, S.M., Michelson, L. & Jacob, R. (1985). Tricyclic antidepressants in obsessive-compulsive disorder: Antiobsessional or antidepressant agents? *American Journal of Psychiatry*, *142*, 572-576.
- McLean, P.D., Whittal, M.L., Thordarson, D.S., Taylor, S., Söchting, I., Koch, W.J., Paterson, R. & Anderson, K.W. (2001). Cognitive versus behavior therapy in the group treatment of obsessive-compulsive disorder. *Journal of Consulting and Clinical Psychology*, *69*(2), 205-214.
- Montgomery, S.A. (1980). Clomipramine in obsessional neurosis. A placebo-controlled trial. *Pharmaceutical Medicine*, *1*(2), 189-192.
- Montgomery, S.A., Kasper, S., Stein, D.J., Bang Hedegaard, K. & Lemming, O.M. (2001). Citalopram 20 mg, 40 mg and 60 mg are all effective and well tolerated compared with placebo in obsessive-compulsive disorder. *International Clinical Psychopharmacology*, *16*, 75-86.
- Montgomery, S.A., McIntyre, A., Osterheider, M., Sarteschi, P., Zitterl, W., Zohar, J., Birkett, M., Wood, A.J. & The Lilly European OCD Study Group. A double-blind,



- placebo-controlled study of fluoxetine in patients with DSM-III-R obsessive-compulsive disorder. *European Neuropsychopharmacology*, 3, 143-152.
- O'Connor, K., Todorov, C., Robillard, S., Borgeat, F. & Brault, M. (1999). Cognitive-behaviour therapy and medication in the treatment of obsessive-compulsive disorder: A controlled study. *Canadian Journal of Psychiatry*, 44, 64-71.
- Pigott, T.A., L'Heureux, F., Rubenstein, C.S., Bernstein, S.E., Hill, J.L. & Murphy, D.L. (1992). A double-blind, placebo-controlled study of trazodone in patients with obsessive-compulsive disorder. *Journal of Clinical Psychopharmacology*, 12, 156-162.
- Stein, D.J., Hollander, E., Mullen, L.S., DeCaria, C.M. & Liebowitz, M.R. (1992). Comparison of clomipramine, alprazolam and placebo in the treatment of obsessive-compulsive disorder. *Human Psychopharmacology*, 7, 389-395.
- The Clomipramine Collaborative Study Group (1991). Clomipramine in the treatment of patients with obsessive-compulsive disorder. *Archives of General Psychiatry*, 48, 730-738.
- Thoren, P., Asberg, M., Cronholm, B., Jörnstedt, L. & Träskman, L. (1980). Clomipramine treatment of obsessive-compulsive disorder. *Archives of General Psychiatry*, 37, 1281-1285.
- Tollefson, G.D., Rampey, A.H., Potvin, J.H., Jenike, M.A., Rush, A.J., Dominguez, R.A., Koran, L.M., Shear, M.K., Goodman, W. & Genduso, L.A. (1994). A multicenter investigation of fixed-dose fluoxetine in the treatment of obsessive-compulsive disorder. *Archives of General Psychiatry*, 51, 559-567.
- Van Balkom, A.J.L.M., de Haan, E., van Oppen, P., Spinhoven, P., Hoogdun, K. & van Dyck, R. (1998). Cognitive and behavioral therapies alone versus in combination with fluvoxamine in the treatment of obsessive compulsive disorder. *Journal of Nervous and Mental Disease*, 186, 492-499.
- Van Oppen, P., de Haan, E., van Balkom, A.J.L.M., Spinhoven, P., Hoogdun, K. & van Dyck, R. (1995). Cognitive Therapy and exposure in vivo in the treatment of obsessive compulsive disorder. *Behaviour Research and Therapy*, 33(4), 379-390.
- Yaryura-Tobias, J.A. & Neziroglu, F.A. (1996). Venlafaxine in obsessive-compulsive disorder. *Archives of General Psychiatry*, 53, 653-654.
- Zohar, J., Judge, R. & OCD Paroxetine Study Investigators. Paroxetine versus clomipramine in the treatment of obsessive-compulsive disorder. *British Journal of Psychiatry*, 169, 468-474.

## 7 Anhang

## 7.1 Anhang A: Bisherige Meta-Analysen

In der folgenden Tabelle sind bisherige Meta-Analysen zur Effektivität verschiedener Behandlungen bei Angststörungen einschließlich ihrer methodischen Eigenschaften und Ergebnisse im Überblick dargestellt. Ausgeschlossen sind Arbeiten, die älter als 10 Jahre sind (Einschränkung der Ergebnisse, aufgrund einer Vielzahl an jüngeren Primärstudien), Arbeiten, die nur eine spezielle Substanzklasse untersuchten sowie Arbeiten, die keinerlei Angaben über die eingesetzte Methoden enthielten.

Autoren	Primärstudien	Methode (für Posteffektstärken)	Ergebnisse
Abramowitz (1996)	Zwang; PsT N = 24 Datenbanken, Handsuche, Referenzen	Prä-Post Cohens $d$ aus $M$ u. $SD$ , Signifikanzen; Gewichtung mit Stichprobengröße standardstatistische Verfahren FEM; k.A. bzgl. $Q$ Publikationsbias: k.A.	Werte zeigen Effektivität Effekte im Follow up stabil $f \gg s$ sign. pos. Korrelation zu Dauer der Therapie, neg. Korrelation zu Dauer der Symptomatik, therapeutenkontrolliert höhere ES als selbstkontrolliert
Ackerman et al. (2002)	Zwang; PhT N = 25 Datenbanken, Referenzen	Prä-Post Mittelwertsdifferenz für Y-BOCS; Gewichtung mit Stichprobengröße Gewichtete Regressionen FEM; $Q$ berechnet Publikationsbias: k.A.	ES $\gg 0$ höhere Effekte bei schwerer Symptomatik zu Prä, späterer Publikation, längerer Dauer der Zeit vor Randomisierung

Boyer (1995)	Panik; PhT N = 27 Datenbanken, Referenzen	Cohens $d$ zwischen Gruppen; Verhältnis der Patienten, die sich, basierend auf dem BESD (binomial effect size display nach Rosenthal), verbesserten in der aktiven Substanz zu Placebo; Gewichtung: k.A. standardstatistische Verfahren (?) FEM; Homogenität überprüft ( $Q$ ?) Publikationsbias: k.A.	ES >> 0 SSRI >> Alprazolam, Imipramin höhere ES, wenn höhere Dosis bei Alprazolam und Imipramin
Chambless et al. (1993)	Generalisierte Angststörung; PsT N = 7 k.A.	Glass $\Delta$ zwischen Gruppen und Prä zu Post aus $M$ u. $SD$ ; Gewichtung: k.A. standardstatistische Verfahren FEM; k.A. bzgl. $Q$ Publikationsbias: k.A.	Werte zeigen Effektivität Stabilität über die Zeit
Chambless et al. (1993)	Panik; PsT N = 7 k.A.	Glass $\Delta$ zwischen Gruppen und Prä zu Post aus $M$ u. $SD$ ; Gewichtung: k.A. standardstatistische Verfahren FEM; k.A. bzgl. $Q$ Publikationsbias: k.A.	Werte zeigen Effektivität Stabilität über die Zeit
Chambless et al. (1993)	Soziale Phobie; PsT N = 10	Glass $\Delta$ zwischen Gruppen und Prä zu Post aus $M$ u. $SD$ ; Gewichtung: k.A.	Werte zeigen Effektivität Stabilität über die Zeit

	k.A.	standardstatistische Verfahren FEM; k.A. bzgl. $Q$ Publikationsbias: k.A.	
Clum et al. (1993)	Panik; PsT, PhT, PsT+PhT N = 29 Datenbanken, Handsuche, Referenzen	Glass $\Delta$ zwischen Gruppen aus $M$ u. $SD$ , Signifikanzen; Gewichtung: k.A. standardstatistische Verfahren FEM; k.A. bzgl. $Q$ Publikationsbias: k.A.	PsT >> PhT bei Ausschluss von Antidepressiva und bei KG = Placebo Dropouts untersucht Prädiktoren untersucht, aber kein sign. Einfluss
Cox et al. (1993)	Zwang; Exposition, PhT N = 25 Datenbanken	Prä-Post Cohens $d$ aus $M$ u. $SD$ , Signifikanzen; Gewichtung: k.A. standardstatistische Verfahren FEM; k.A. bzgl. $Q$ Publikationsbias: k.A.	ES >> 0, außer für Exposition bei depressiver Symptomatik Dropouts untersucht
Federoff et al. (2001)	Soziale Phobie; PsT, PhT N = 56 Datenbanken, Referenzen, Konferenzen, Kontaktaufnahme mit Wissenschaftlern und Pharmafirmen	Prä-Post Cohens $d$ aus $M$ u. $SD$ ; Gewichtung mit Stichprobengröße standardstatistische Verfahren REM; $Q$ berechnet Publikationsbias: k.A.	ES (Ausnahme WL, Exp) >> 0 s: BZD, SSRI, MAOI >> KG, AE BZD >> MAOI, KT, KVT, SST f: BDZ, SSRI, MAOI, KVT >> WL, PP Effekte stabil über Follow up Prädiktoren untersucht: kein sign. Einfluss

Feske et al. (1995)	Soziale Phobie; PsT N = 15 Datenbanken, Handsuche, Kontaktaufnahme mit Wissenschaftlern	Glass $\Delta$ zwischen Gruppen und Prä zu Post aus <i>M</i> u. <i>SD</i> ; Gewichtung mit Stichprobengröße standardstatistische Verfahren FEM; <i>Q</i> berechnet Publikationsbias: k.A.	Werte zeigen Effektivität Dropouts untersucht VT: höhere ES bei längerer Therapiedauer
Gould et al. (1995)	Panik; PsT, PhT N = 43 Datenbanken, Referenzen	Glass $\Delta$ zwischen Gruppen aus <i>M</i> u. <i>SD</i> , Signifikanzen und Häufigkeiten; Gewichtung: k.A. standardstatistische Verfahren FEM; k.A. bzgl. <i>Q</i> Publikationsbias: k.A.	ES >> 0 Effekte stabil über Follow up, aber höheres Absinken bei PhT PsT >> PhT, aber unterschiedliche KG Dropouts: PsT << PhT Prädiktoren untersucht, aber kein sign. Einfluss
Gould & Otto et al. (1997)	Generalisierte Angststörung; PsT, PhT N = 35 Datenbanken, Konferenzen, Referenzen, Kontaktaufnahme mit Wissenschaftlern	Glass $\Delta$ zwischen Gruppen aus <i>M</i> u. <i>SD</i> , Signifikanzen; Gewichtung: k.A. standardstatistische Verfahren FEM; <i>Q</i> berechnet Publikationsbias: k.A.	ES >> 0 Effekte stabil über Follow up Dropouts untersucht Prädiktoren untersucht, aber kein sign. Einfluss

Gould & Buckminster et al. (1997)	Soziale Phobie; PsT, PhT, PsT + PhT N = 24; Datenbanken, Referenzen; Konferenzen	Glass $\Delta$ zwischen Gruppen aus $M$ u. $SD$ , Signifikanzen; Gewichtung: k.A. standardstatistische Verfahren Ausreißeranalyse FEM; $Q$ berechnet Publikationsbias: k.A.	ES $\gg 0$ Effekte stabil über Follow-up Dropouts untersucht Gruppensetting > Einzelsetting
Kobak et al. (1998)	Zwang; PsT, PhT N = 77 Datenbanken, Konferenzen	Referenzen; mit Korrektur für kleine Stichproben; Gewichtung mit Stichprobengröße Verfahren mit Kontrolle der Verletzung der Varianzhomogenität Kontrolle methodischer Variablen FEM; $Q$ berechnet Publikationsbias: Fail-safe N	Hedges $g$ zwischen Gruppen und Prä zu Post aus $M$ u. $SD$ , Signifikanzen und Häufigkeiten Analysen ES $\gg 0$ VT $\gg$ SSRI Dropouts untersucht geringere ES wenn kontrolliertes Design, randomisiert, unpublizierte Studien oder Journale ohne Reviewer, Fremdrating, spätere Publikation
Piccinelli et al. (1995)	Zwang; PhT N = 36 Datenbanken, Handsuche, Refe-	Prä-Post Hedges $g$ aus $M$ u. $SD$ , Signifikanzen und Häufigkeiten mit Korrektur für kleine Stichproben; Gewichtung mit Stich-	Homogenitätstest sign. in einigen Analysen Werte zeigen Effektivität SSRI, Clomipramine, TCA $\gg$ PP

	renzen, Kontaktaufnahme mit Pharmafirmen	probengröße standardstatistische Verfahren FEM; $Q$ berechnet Publikationsbias: Fail-safe N	Clomipramine >> SSRI, TCA geringere ES bei längerer Dauer der Störung
Oei et al. (1999)	Panik; PsT N = 35 Datenbanken	Standardisierte Mittelwertsdifferenz zur Normpopulation aus $M$ und $SD$ für den FQ; Gewichtung: k.A. standardstatistische Verfahren FEM; k.A. bzgl. $Q$ Publikationsbias: k.A.	Werte zeigen Effektivität Effekte stabil über Follow up
Ruhmland et al. (2001a)	Spezifische Phobie; PsT N = 19 Datenbanken, Referenzen	Prä-Post Hedges $g$ aus $M$ u. $SD$ mit Korrektur für kleine Stichproben; Gewichtung: k.A. standardstatistische Verfahren FEM; k.A. bzgl. $Q$ Publikationsbias: k.A.	ES >> 0, außer Infovermittlung, Stressmanagement Effekte stabil über Follow up Dropouts: Exposition << Systematische Desensibilisierung
Ruhmland et al. (2001a)	Zwang; PsT N = 23 Datenbanken, Referenzen	Prä-Post Hedges $g$ aus $M$ u. $SD$ mit Korrektur für kleine Stichproben; Gewichtung: k.A. standardstatistische Verfahren	ES >> 0 Effekte stabil über Follow-up Dropouts untersucht



		FEM; k.A. bzgl. $Q$ Publikationsbias: k.A.	
Ruhmland et al. (2001b)	Generalisierte Angststörung; PsT N = 14 Datenbanken, Referenzen	Prä-Post Hedges $g$ aus $M$ u. $SD$ mit Korrektur für kleine Stichproben; Gewichtung: k.A. standardstatistische Verfahren FEM; k.A. bzgl. $Q$ Publikationsbias: k.A.	ES >> 0 Effekte stabil über Follow-up Dropouts untersucht
Ruhmland et al. (2001b)	Soziale Phobie; PsT N = 15 Datenbanken, Referenzen	Prä-Post Hedges $g$ aus $M$ u. $SD$ mit Korrektur für kleine Stichproben; Gewichtung: k.A. standardstatistische Verfahren FEM; k.A. bzgl. $Q$ Publikationsbias: k.A.	ES >> 0 Effekte stabil über Follow-up Dropouts untersucht
Ruhmland et al. (2001c)	Panik; PsT N = 51 Datenbanken, Referenzen	Prä-Post Hedges $g$ aus $M$ u. $SD$ mit Korrektur für kleine Stichproben; Gewichtung: k.A. standardstatistische Verfahren FEM; k.A. bzgl. $Q$ Publikationsbias: k.A.	ES >> 0 Exposition in vivo >> KVT Effekte stabil über Follow-up mit Ausnahme der Exposition in vivo Dropouts: KVT << Exposition in vivo
Sherman (1998)	Posttraumatische Belastungsstörung; PsT	Cohens $d$ zwischen Gruppen aus $M$ u. $SD$ und Signifikanzen; Gewichtung: k.A.	ES >> 0 Effekte stabil über Follow up

	N = 17	standardstatistische Verfahren	
	Datenbanken, Referenzen;	Ausreißeranalyse	
	Kontaktaufnahme mit Wissen-	FEM; $Q$ berechnet	
	schaftlern	Publikationsbias: k.A.	
Stein et al. (1995)	Zwang; PhT	Glass $\Delta$ zwischen Gruppen und Prä zu Post	PhT >> PP
	N = 28	aus $M$ u. $SD$ ; Gewichtung: k.A.	SSRI >> andere PhT
	Datenbanken, Referenzen	standardstatistische Verfahren	
		FEM; k.A. bzgl. $Q$	
		Publikationsbias: k.A.	
Stein et al. (2001)	Posttraumatische Belastungs-	Odds ratio; Gewichtung: k.A.	ES >> 0
	störung; PhT	standardstatistische Verfahren	Dropouts untersucht
	N = 9	REM; $Q$ berechnet	
	Datenbanken, Referenzen, Kon-	Publikationsbias: k.A.	
	taktaufnahme mit Wissenschaft-		
	lern und Pharmafirmen		
Taylor (1996)	Soziale Phobie; PsT	Prä-Post Cohens $d$ aus $M$ u. $SD$ ; Gewichtung:	PsT >> WL
	N = 25	k.A.	KVT >> Placebo
	Datenbanken, Konferenzen,	standardstatistische Verfahren	Effekte stabil mit Trend zur Verbesserung

	Referenzen, Kontaktaufnahme mit Wissenschaftlern	Ausreißeranalyse FEM; k.A. bzgl. $Q$ Publikationsbias: Fail-safe N	Dropouts: WL < PsT
van Balkom et al. (1994)	Zwang; PsT, PhT, PsT+PhT N = 86 Datenbanken, Handsuche, Referenzen	Prä-Post Cohens $d$ aus $M$ u. $SD$ und Signifikanzen; Gewichtung: k.A. Verfahren mit Kontrolle der Verletzung der Varianzhomogenität, aber keine Gewichtung Ausreißeranalyse FEM; k.A. bzgl. $Q$ Publikationsbias: Fail-safe N	Werte zeigen Effektivität VT, SA, VT+SA >> PP; KVT, KT > PP; f >> s bei Zwangssymptomatik s: VT >> SA s: VT+SA > SA f: kein sign. Unterschied Effekte im Follow up stabil VT: sign. pos. Korrelation zu Dauer der Therapie, sign. neg. zu Dauer der Störung (Fremdrating)
van Balkom et al. (1997)	Panik; PsT, PhT, PsT+PhT N = 106 Datenbanken, Referenzen	Prä-Post Cohens $d$ aus $M$ u. $SD$ und Signifikanzen; Gewichtung: k.A. Verfahren mit Kontrolle der Verletzung der Varianzhomogenität, aber keine Gewichtung Ausreißeranalyse	ES >> 0 PsT, PhT, Kombination >> KG; PsT+Antidepressiva >> PsT Effekte stabil über Follow up höhere ES bei älteren Publikationen

		FEM; k.A. bzgl. $Q$ Publikationsbias: Fail-safe N	Dropouts untersucht
Van Etten et al. (1998)	Posttraumatische Belastungs- störung; PsT, PhT N = 38 Datenbanken, Konferenzen, Referenzen; Kontaktaufnahme mit Wissenschaftlern	Prä-Post Cohens $d$ aus $M$ u. $SD$ ; Gewichtung mit Stichprobengröße standardstatistische Verfahren FEM; k.A. bzgl. $Q$ Publikationsbias: Fail-safe N	ES $\gg 0$ s: PsT > PhT (VT = SSRI) Dropouts: PsT $\ll$ PhT Prädiktoren untersucht, aber kein sign. Ein- fluss
Westen et al. (2001)	Panik, Generalisierte Angst- störung; PsT N = 22 Datenbanken	Cohens $d$ zwischen Gruppen aus $M$ u. $SD$ ; Gewichtung mit Stichprobengröße standardstatistische Verfahren FEM; k.A. bzgl. $Q$ Publikationsbias: k.A.	Werte zeigen Effektivität Höherer Prozentsatz verbesserter Patienten bei größerem Anteil ausgeschlossener Pati- enten

Anmerkungen: BZD = Benzodiazepine, ES = Effektstärke, f = Fremdrating, FEM = Fixed Effects-Modell, KG = Kontrollgruppe, KT = kognitive Therapie, KVT = kognitiv-behaviorale Therapie, MAOI = Monoaminoxidasehemmer, PhT = Pharmakotherapie, PP = Pillenplacebo, PsT = Psychotherapie, REM = Random Effects Modell, s = Selbstrating, SA = serotonerges Antidepressivum, SSRI = Selektive Serotoninwiederaufnahmehemmer, SST = Training Sozialer Kompetenz, TCA = Tricyklische Antidepressiva, VT = Verhaltenstherapie, WL = Warteliste; signifikante Unterschiede markiert als „ $\gg$ “ bzw. „ $\ll$ “, Tendenzen mit „ $\gt$ “ bzw. „ $\lt$ “

## 7.2 Anhang B: Codierungskatalog

### Codierungsbogen zur Metaanalyse - VT, KVT und PhT bei Angststörungen -

Nummer der Studie:

Autoren (bei mehr als 6 et al.):

Titel der Studie:

Jahr der Publikation:

Quelle:

- Journal
- Buch / -beitrag
- Internet
- Kongressbeitrag
- unveröffentlichte Diplomarbeit, Dissertation oder Habilitation
- unveröffentlichtes Manuskript etc.

Welche Angststörung(en) werden behandelt?

- Panik mit Agoraphobie
- Panik ohne Agoraphobie
- Agoraphobie ohne Panikstörung
- Zwang (welcher):
- GAS
- PTB (welches Trauma):
- Soziale Phobie
  - generalisierter Typus
  - spezifischer Typus
- Spezifische Phobie (welche):

Nach welchem Instrument wurden die Störungen diagnostiziert?

- DSM II
- DSM III
- DSM III-R
- DSM IV
- ICD 8
- ICD 9
- ICD 10
- nicht bestimmbar

Wie wurde die Diagnose gestellt ?

- diagnostisches Interview nach DSM oder ICD
- Screeninginterview
- Cut-off Werte von Fragebögen
- keine Angaben
- sonstiges (wie):

Liegen komorbide Störungen vor (gilt nur, wenn mindestens 25% der Patienten die Störung aufweisen)?

- Affektive Störungen  25-49%  50-74%  75-100%
  - Depressive Störungen (MD, Dysthyme Störung)  25-49%  50-74%  75-100%
  - Bipolare Störungen  25-49%  50-74%  75-100%
  - Andere Affektive Störungen  25-49%  50-74%  75-100%
- Substanzmissbrauch  25-49%  50-74%  75-100%
  - Alkohol  25-49%  50-74%  75-100%
  - Medikamente  25-49%  50-74%  75-100%
  - illegale Drogen  25-49%  50-74%  75-100%
- Somatoforme Störungen  25-49%  50-74%  75-100%
  - Somatisierungsstörung  25-49%  50-74%  75-100%
  - Hypochondrie  25-49%  50-74%  75-100%
  - Andere Somatoforme Störungen  25-49%  50-74%  75-100%
- PSK-Störung  25-49%  50-74%  75-100%
  - Paranoid  25-49%  50-74%  75-100%
  - Schizoid  25-49%  50-74%  75-100%
  - Schizotypisch  25-49%  50-74%  75-100%
  - Antisozial  25-49%  50-74%  75-100%
  - Borderline  25-49%  50-74%  75-100%
  - Histrionisch  25-49%  50-74%  75-100%
  - Narzistisch  25-49%  50-74%  75-100%
  - Vermeidend-Selbstunsicher  25-49%  50-74%  75-100%
  - Dependent  25-49%  50-74%  75-100%
  - Zwanghaft  25-49%  50-74%  75-100%
- andere Angststörung  25-49%  50-74%  75-100%
  - Panik mit/ohne Agoraphobie  25-49%  50-74%  75-100%
  - Zwang  25-49%  50-74%  75-100%
  - GAS  25-49%  50-74%  75-100%
  - PTB  25-49%  50-74%  75-100%
  - Soziale Phobie  25-49%  50-74%  75-100%
  - Spezifische Phobie  25-49%  50-74%  75-100%
  - Sonstiges (welches+Prozent):

Wie wurden die komorbiden Störungen diagnostiziert?

- nach DSM oder ICD mittels diagnostischen Interview

- Screeninginterview
- durch Cut-off-Werte von Fragebögen
- nicht bestimmbar
- sonstiges (wie):

## Sprache der Veröffentlichung

- englisch
- deutsch
- andere (welche):

## Wo wurde die Studie durchgeführt?

- Universität und angeschlossene Institutionen, Forschungszentrum, universitäre Psychiatrische / Medizinische Klinik
- Psychiatrische / Medizinische Klinik (nicht Uni)
- Ambulanz / Beratungsstelle (außerhalb Uni)
- Privatpraxis
- Multicenterstudie
- nicht erkennbar

## Rahmenbedingungen

- stationär (Inpatient)
- teilstationär (Tagesklinik ...)
- ambulant (Outpatient)
- nicht erkennbar

## Art der Therapie (im Zweifel Einzel)

- Einzeltherapie
- Gruppentherapie
- Paartherapie

## Auf welche Weise wurden die Probanden rekrutiert?

- durch Aushänge, Informationen in Presse, Funk oder Fernsehen etc.
- im Rahmen eines regulären Aufenthalts in der Psychiatrie bzw. Behandlung
- nicht erkennbar

## Welche Exklusionskriterien wurden zur Auswahl der Probanden genutzt?

- Vorliegen einer weiteren Achse-I-Störung
- andere Angststörungen

- Affektive Störungen (Major Depression, Bipolare Störung,...)
- Substanzabusus
- Somatoforme Störungen
- Schizophrene Störungen
- Vorliegen einer PSK-Störung
  - Paranoid
  - Schizoid
  - Schizotypisch
  - Antisozial
  - Borderline
  - Histrionisch
  - Narzistisch
  - Vermeidend-Selbstunsicher
  - Dependent
  - Zwanghaft
- akute Suizidalität
- körperliche Erkrankung einschließlich Epilepsie, Demenz etc.
- mentale Retardation
- gleichzeitige Medikation
  - wenn die Substanz erst seit kurzer Zeit eingenommen wird
- Sonstiges (welches):

-----

Wie wurden die Störungen für die Exklusion diagnostiziert?

- nach DSM oder ICD mittels diagnostischen Interview
- Screening-Interview
- durch Cut-off-Werte von Fragebögen
- nicht bestimmbar
- sonstiges (wie):

-----

Handelt es sich bei den Patienten um eine treatmentresistente Stichprobe? (im Zweifel nein)

- ja, bzgl. eines (anderen) Medikamentes
- ja, bzgl. einer (anderen) Psychotherapie
- nein

-----

Betrachtet man die Gesamtanzahl an angesprochenen Betroffenen bzw. Interessenten, wie groß war die Anzahl der Verweigerer, die sich entschlossen, nicht an der Studie teilzunehmen?

%

-----

Aufteilung der Probanden auf die verschiedenen Gruppen

- vollständige Randomisierung



- Matchen / Parallelisieren
- "äußere Gegebenheiten", z.B. weil in einem Ort keine Therapie angeboten wurde
- Selbstselektion
- nicht erkennbar

-----

In welchen bedeutsamen Variablen unterscheiden sich die verschiedenen Gruppen prä?


- in keinen Variablen
- nicht erkennbar

-----

	Stichprobenmerkmale		
	Alter	Female (%)	Male (%)
KG1	<input type="text"/>	<input type="text"/>	<input type="text"/>
KG2	<input type="text"/>	<input type="text"/>	<input type="text"/>
BG1	<input type="text"/>	<input type="text"/>	<input type="text"/>
BG2	<input type="text"/>	<input type="text"/>	<input type="text"/>
BG3	<input type="text"/>	<input type="text"/>	<input type="text"/>
BG4	<input type="text"/>	<input type="text"/>	<input type="text"/>
BG5	<input type="text"/>	<input type="text"/>	<input type="text"/>

-----

Stichprobenmerkmale für die gesamte Stichprobe

<input type="text"/>	Alter
<input type="text"/>	% Frauen
<input type="text"/>	Jahre der Störung

-----

Waren die Probanden gegenüber ihrer Gruppenzugehörigkeit blind?

- ja
- nein
- nicht erkennbar

-----

Im folgenden sollen die einzelnen Treatment-Gruppen beschrieben werden. In das obere Kästchen bitte die Benennung im "study input" eingeben.

**BG1** 

- VT  Systematische Desensibilisierung (SD)  
 in sensu (is)  in vivo (iv)  
 Angewandte Anspannung (AA)
- Exposition (E)  
 in sensu (is)  in vivo (iv)  
 massiert (m)  gestuft (g)
- Reaktionsverhinderung (R)
- Modellernen (M)
- Rollenspiel (RS)
- Gedankenstop (G)
- Angewandte Entspannung (AE)
- Entspannung (PMR)
- Biofeedback (Bfb)
- Fading (F)
- Training Sozialer Kompetenz (TSK)
- unter Einsatz des Internets (I) bzw. computergestützt (C)
- Selbsthilfe (S)
- Bibliotherapie (BT)
- KVT  Training sozialer Kompetenz (TSK)
- Problemlösetraining (PL)
- Modifikation kognitiver Fehler durch kognitive Techniken (KM)
- Verhaltensexperimente (V)
- Selbstinstruktionstraining (SIT)
- unter Einsatz des Internets (I) bzw. computergestützt (C)
- Selbsthilfe
- Bibliotherapie (BT)
- PT  Medikament (welches):
- Dosis:  mg  
 feste Dosis für jeden Patienten  
 Mittelwert, da Anpassung der Dosis an jeweiligen Patienten
- Dauer:  Tage
- Pillenplacebo (PP)
- Psychotherapieplacebo (TP)
- 

**BG2** 

- VT  Systematische Desensibilisierung (SD)  
 in sensu (is)  in vivo (iv)
- Angewandte Anspannung (AA)
- Exposition (E)  
 in sensu (is)  in vivo (iv)  
 massiert (m)  gestuft (g)
- Reaktionsverhinderung (R)
- Modellernen (M)
- Rollenspiel (RS)
- Gedankenstop (G)

- Angewandte Entspannung (AE)
- Entspannung (PMR)
- Biofeedback (Bfb)
- Fading (F)
- Training Sozialer Kompetenz (TSK)
- unter Einsatz des Internets (I) bzw. computergestützt (C)
- Selbsthilfe (S)
- Bibliothherapie (BT)
- KVT  Training sozialer Kompetenz (TSK)
- Problemlösetraining (PL)
- Modifikation kognitiver Fehler durch kognitive Techniken (KM)
- Verhaltensexperimente (V)
- Selbstinstruktionstraining (SIT)
- unter Einsatz des Internets (I) bzw. computergestützt (C)
- Selbsthilfe (S)
- Bibliothherapie (BT)
- PT  Medikament (welches):
- Dosis:  mg
  - feste Dosis für jeden Patienten
  - Mittelwert, da Anpassung der Dosis an jeweiligen Patienten
- Dauer:  Tage
- Pillenplacebo (PP)
- Psychotherapieplacebo (TP)

-----

- BG3**
- VT  Systematische Desensibilisierung (SD)
    - in sensu (is)  in vivo (iv)
  - Angewandte Anspannung (AA)
  - Exposition (E)
    - in sensu (is)  in vivo (iv)
    - massiert (m)  gestuft (g)
  - Reaktionsverhinderung (R)
  - Modellernen (M)
  - Rollenspiel (RS)
  - Gedankenstop (G)
  - Angewandte Entspannung (AE)
  - Entspannung (PMR)
  - Biofeedback (Bfb)
  - Fading (F)
  - Training Sozialer Kompetenz (TSK)
  - unter Einsatz des Internets (I) bzw. computergestützt (C)
  - Selbsthilfe (S)
  - Bibliothherapie (BT)
  - KVT  Training sozialer Kompetenz (TSK)
  - Problemlösetraining (PL)
  - Modifikation kognitiver Fehler durch kognitive Techniken (KM)

- Verhaltensexperimente (V)  
 Selbstinstruktionstraining (SIT)  
 unter Einsatz des Internets (I) bzw. computergestützt (C)  
 Selbsthilfe (S)  
 Bibliothherapie (BT)  
 PT  Medikament (welches):   
     Dosis:  mg  
          feste Dosis für jeden Patienten  
          Mittelwert, da Anpassung der Dosis an jeweiligen Patienten  
     Dauer:  Tage  
 Pillenplacebo (PP)  
 Psychotherapieplacebo (TP)

**Bg4**

- VT  Systematische Desensibilisierung (SD)  
      in sensu (is)                       in vivo (iv)  
 Angewandte Anspannung (AA)  
 Exposition (E)  
      in sensu (is)                       in vivo (iv)  
      massiert (m)                       gestuft (g)  
 Reaktionsverhinderung (R)  
 Modellernen (M)  
 Rollenspiel (RS)  
 Gedankenstop (G)  
 Angewandte Entspannung (AE)  
 Entspannung (PMR)  
 Biofeedback (Bfb)  
 Fading (F)  
 Training Sozialer Kompetenz (TSK)  
 unter Einsatz des Internets (I) bzw. computergestützt (C)  
 Selbsthilfe (S)  
 Bibliothherapie (BT)  
 KVT  Training sozialer Kompetenz (TSK)  
 Problemlösetraining (PL)  
 Modifikation kognitiver Fehler durch kognitive Techniken (KM)  
 Verhaltensexperimente (V)  
 Selbstinstruktionstraining (SIT)  
 unter Einsatz des Internets (I) bzw. computergestützt (C)  
 Selbsthilfe (S)  
 Bibliothherapie (BT)  
 PT  Medikament (welches):   
     Dosis:  mg  
          feste Dosis für jeden Patienten  
          Mittelwert, da Anpassung der Dosis an jeweiligen Patienten  
     Dauer:  Tage  
 Pillenplacebo (PP)  
 Psychotherapieplacebo (TP)

**Bg5**

- VT  Systematische Desensibilisierung (SD)  
 in sensu (is)  in vivo (iv)  
 Angewandte Anspannung (AA)  
 Exposition (E)  
 in sensu (is)  in vivo (iv)  
 massiert (m)  gestuft (g)  
 Reaktionsverhinderung (R)  
 Modellernen (M)  
 Rollenspiel (RS)  
 Gedankenstop (G)  
 Angewandte Entspannung (AE)  
 Entspannung (PMR)  
 Biofeedback (Bfb)  
 Fading (F)  
 Training Sozialer Kompetenz (TSK)  
 unter Einsatz des Internets (I) bzw. computergestützt (C)  
 Selbsthilfe (S)  
 Bibliothherapie (BT)
- KVT  Training sozialer Kompetenz (TSK)  
 Problemlösetraining (PL)  
 Modifikation kognitiver Fehler durch kognitive Techniken (KM)  
 Verhaltensexperimente (V)  
 Selbstinstruktionstraining (SIT)  
 unter Einsatz des Internets (I) bzw. computergestützt (C)  
 Selbsthilfe (S)  
 Bibliothherapie (BT)
- PT  Medikament (welches):   
 Dosis:  mg  
 feste Dosis für jeden Patienten  
 Mittelwert, da Anpassung der Dosis an jeweiligen Patienten  
 Dauer:  Tage  
 Pillenplacebo (PP)  
 Psychotherapieplacebo (TP)

-----  
 Medikamentenstudie: Welches Design wurde gewählt?

- Rekrutierung - Einnahme aktive Substanz bzw. Placebo  
 Rekrutierung - Placebo washout - Einnahme aktive Substanz bzw. Placebo  
 Rekrutierung (=stabiler Medikamentenspiegel) - Einnahme aktive Substanz bzw. Placebo  
 - Taper phase (=Ausschleichen des alten Medikamentes) - Einnahme aktive Substanz bzw. Placebo  
 Rekrutierung (=stabiler Medikamentenspiegel) - Taper phase (=Ausschleichen des alten Medikamentes) - Einnahme aktive Substanz bzw. Placebo  
 Rekrutierung (=stabiler Medikamentenspiegel)- zusätzliche Einnahme aktive Substanz/Placebo  
  
 nicht erkennbar
-

Falls eine Placebotherapie als Kontrollbedingung (bei Psychotherapie) eingesetzt wurde, ist davon auszugehen, dass

- eine positive Erwartungshaltung bei den Probanden geweckt wurde (z.B. über ein gegebenes Rational)?

- ja
- nein
- nicht erkennbar

- eine Problemdiskussion stattfand oder Techniken eingesetzt wurden, die zwar keinen direkten Bezug zur

Störung aufweisen im Sinne einer aktiven Kognitions- oder Verhaltensänderung (z.B. Counselling), aber sinnvoll erscheinen

- ja
- nein
- nicht erkennbar

-----

Falls eine aktive Kontrollgruppe eingesetzt war, welche Intervention wurde durchgeführt? (nur Psychotherapie)

- nondirective Counselling (Rogers)
- supportive Counselling (Problemlösung für allg. Probleme)
- Clinical Management
- Pillenplacebo
- sonstiges (wie):

-----

Erhielten einige der Patienten zusätzlich weitere Therapien (außerhalb der untersuchten - einschließlich Medikamente)?

	post	follow 1	follow 2	follow 3
ja, mit Gleichverteilung zwischen den Gruppen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ja, mit Unterschieden zwischen den Gruppen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
nein	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
nicht erkennbar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

-----

Aufteilung der Therapeuten auf die verschiedenen Gruppen

- Randomisierung oder anders zufälliges Verfahren
- nicht zufälliges Verfahren (z.B. Experten auf Gebiet)
- "äußere Gegebenheiten", z.B. abhängig vom Krankenhaus
- nicht erkennbar

-----

Waren die Therapeuten gegenüber dem eingesetzten Verfahren blind? (betrifft Pharmakotherapie)

- ja
- nein
- nicht erkennbar

-----

Wie war die Grundausbildung der Therapeuten der einzelnen Bedingungen? (betrifft Psychotherapie)

	KG-aktiv	BG1	BG2	BG3	BG4
Psychologen	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Ärzte	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Nurse specialist	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

-----

Anzahl der Therapeuten pro Erfahrungslevel und Bedingung? (betrifft Psychotherapie)

	KG-aktiv	BG1	BG2	BG3	BG4
wenig Erfahrung (Student, Doktorand, ...)	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
mittlere oder viel Erfahrung	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
nicht bestimmbar	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

-----

Erhielten die Therapeuten ein Training in der jeweiligen Methode? (betrifft Psychotherapie)

	KG-aktiv	BG1	BG2	BG3	BG4
ja	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
nein	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
nicht erkennbar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

-----

Anzahl der Sitzungen der verschiedenen Psychotherapien bis post

<input type="text"/>	KG-aktiv
<input type="text"/>	KG-aktiv 2
<input type="text"/>	BG1
<input type="text"/>	BG2
<input type="text"/>	BG3
<input type="text"/>	BG4
<input type="text"/>	BG5

-----

Dauer der einzelnen Behandlungssitzungen im Durchschnitt, d.h. Mittelwert und Einbezug auch längerer Sitzungen (in Minuten)

<input type="text"/>	min in KG-aktiv
<input type="text"/>	min in KG-aktiv 2
<input type="text"/>	min in BG1

- min in BG2  
 min in BG3  
 min in BG4  
 min in BG5

-----

Falls im Rahmen der Psychotherapien zusätzliche Sitzungen zwischen post und follow-up durchgeführt wurden, wie hoch war die Dauer des Kontaktes (Sitzung x Länge)?

- min in KG-aktiv  
 min in KG-aktiv 2  
 min in BG1  
 min in BG2  
 min in BG3  
 min in BG4  
 min in BG5

-----

Wurden in den einzelnen Behandlungsgruppen Hausaufgaben eingesetzt? (betrifft PsT)

	KG-aktiv	BG1	BG2	BG3	BG4
ja	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
nein	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
nicht erkennbar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

-----

Manualisierung der Therapie (betrifft Psychotherapie)

	KG-aktiv	BG1	BG2	BG3	BG4
ja	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
nein	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
nicht erkennbar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

-----

Überwachung der Therapiedurchführung durch Dritte (einschl. Supervision)(betrifft PsT)

	KG-aktiv	BG1	BG2	BG3	BG4
ja	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
nein	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
nicht erkennbar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

-----

Wie erfolgte die Überwachung der Therapiedurchführung?

- per Video/Audio  
 durch Stundenbögen, die die Therapeuten ausfüllten  
 durch persönliche Rücksprache mit dem Therapeuten (einschl. Supervision)



- nicht erkennbar
- sonstiges (wie):

-----

In welchem Umfang erfolgte die Überwachung der Therapiedurchführung?

- komplette Therapie
- einzelne Sitzungen
- Bruchstücke einzelner Sitzung
- nicht bestimmbar

-----

Anzahl der Drop-outs pro Gruppe nach Behandlungsbeginn (in Prozent) bzw. Gesamt

 % in BG1

 % in BG2

 % in BG3

 % in BG4

 % in BG5

 % in KG1

 % in KG2

 % Gesamt

-----

Gab es eine Dropout-Analyse?

- ja  
wenn ja, in welchen Variablen unterscheiden sich Completer sign. von Abbrechern:
- nein

-----

In welcher Weise wurden die Therapieabbrüche in der Auswertung berücksichtigt?

- Abbrecher wurden ohne Ersatz von der Auswertung ausgeschlossen (completer analysis)
- Abbrecher wurden durch Aufnahme neuer Patienten ersetzt
- es wurden Schätzwerte oder die letzten erreichbaren Werte eingesetzt (intent to treat bzw. endpoint analysis)
- 0 drop-out
- Auswahl war abhängig von der Anzahl an Sitzungen
- nicht erkennbar

-----

Gab es eine Analyse für Prädiktoren des Therapieerfolgs?

- ja  
in welchen Variablen unterscheiden sich erfolgreiche Patienten von weniger erfolgreichen?:

nein

-----

Falls es unabhängige Fremdbeurteiler für den Therapieerfolg gab, waren die Beurteiler gegenüber der Gruppenzugehörigkeit der Probanden "offiziell" blind?

- ja
- nein
- nicht erkennbar

-----

Medikamentenstudie: Falls Daten vorliegen zur Ausprägung der Symptomatik für den Zeitpunkt nach Absetzen des Medikamentes, welches Design wurde gewählt?

- Absetzen des Medikamentes ohne weitere Gabe von Kapseln
- Ausschleichen des Medikamentes ohne weitere Gabe von Kapseln
- abruptes Ersetzen der aktiven Substanz durch Placebo
- stufenweises Ersetzen der aktiven Substanz durch Placebo
- nicht erkennbar

-----

Medikamentenstudie: Erfassten die unabhängigen Fremdbeurteiler gleichzeitig auch die Nebenwirkungen der verabreichten Substanzen?

- ja
- nein
- nicht erkennbar

Nicht vergessen:

Falls einer der folgenden Punkte auf die Studie zutrifft, bitte auf einem Extrazettel vermerken.

- wurden im Rahmen der Studie Variablen variiert (z.B. Ausmaß des Therapeutenkontakts, gestufte vs. massierte Exposition)?
- gab es sonstige wichtige Variablen von Therapeuten oder Patienten, die bis jetzt nicht abgefragt wurden?
- waren Reliabilitätskoeffizienten der Fragebögen mitgeteilt?

### 7.3 Anhang C: Beschreibung des Programms

Die Programmierung der nachfolgenden Software erfolgte mit Visual Basic (Microsoft, Version 6.0) mit Excel als ActiveX-Komponente.

Das Programm sollte folgende Funktionen enthalten:

1. die Organisation der eingeschlossenen Studien
2. die Organisation der Daten innerhalb einer Studie
3. die Berechnung der einzelnen Effektstärken
4. die Berechnung der mittleren Effektstärken im FEM und REM
5. die Berechnung von Regressionen

#### 1. Die Organisation der eingeschlossenen Studien

Es erfolgte eine Gliederung der Studien nach untersuchter Störung, die Tabellenblätter entsprachen dabei den unterschiedlichen Störungen.

#### 2. Die Organisation der Daten innerhalb einer Studie

Mit Hilfe eines Moduls, welches speziell für die Eingabe neuer Studien entwickelt wurde, erfolgte eine Standardisierung des Formats einer Studie aufbauend auf den grundlegenden Eigenschaften einer Studie: Gruppen (Typ, Kurzbeschreibung), Messzeitpunkte, Symptom- und Unterkategorie, Fragebogen und Rating. Es erfolgte dann die Eingabe der zur Berechnung der Effektstärken notwendigen Maße. Der Hauptdialog ist auf der folgenden Seite abgebildet.

#### 3. Die Berechnung der einzelnen Effekte

Es erfolgte eine Automatisierung der Berechnung der einzelnen Effektstärken und der Varianz aus Mittelwerten und Standardabweichungen sowie aus Häufigkeiten. Die Berechnung über andere statistische Kennwerte erfolgte indirekt über einen Rechner. Dieser beinhaltet Formeln für Effektstärken aus Mittelwerten und Standardabweichungen,  $t$ - und  $F$ -Werten, Veränderungswerten, sowie  $t$ - und  $F$ -Werten von Veränderungswerten (vgl. Abbildung 19).

Aktuelle Studiennummer  
 Neue Studie hinzufügen  
 Auslesen der manuell geänderten Excel Daten  
 Studie löschen  
 Berechnungen der Studien, Analysen und Regressionen  
 Ausgewählte Gruppe  
 Eingabe der Kennwerte

Assessment	Sk	Lk	Instrumente	Rating	n VT	Ex_R	mean	std
prä	Angst	OCD_G	Y-BOCS	f	8		21,3	9,6
			HZI	s			3,5	2
		OCD_V	f	11,1			5,1	
		OCD_K	f	10,1			5,7	
	gK	A	BES1	s			2,8	0,8
			BES2	s			2,9	0,8
BES3			s	2,5	1,1			
post	Angst	OCD_G	Y-BOCS	f	8		12,8	7,1
			HZI	s			2	1,8
		Umpolen GAS	s	4,5			0,9	
		Umpolen GAS	f	4,2			1,1	
		OCD_V	Y-BOCS	f			7,1	4,7
			OCD_K	f			5,6	4,5
	gK	A	BES1	s			1,6	1,1
			BES2	s			1,5	0,6
			BES3	s			1,7	0,8
ks	Angst	Responder 30%	f			Häufigkeit	5	

Studys: 1    Analysis: 0    Regressions: 0    Typ: Zwang    Counts: 17    27.03.2003    20:35

Anzahl an Studien, Analysen und Regressionen im File    Störung

	1	2
n	16	16
t	2,04	

biased	0,72124892
unbiased	0,70306617
varunbiased	0,13323837

Abbildung 19: Direkte Berechnung der Effektstärken über den integrierten Rechner

#### 4. Die Berechnung der mittleren Effektstärken im FEM und REM

Nach manueller Zusammenfassung der Einzel-Effektstärken zu Studien-Effektstärken ermöglichte das Programm die Berechnung der statistischen Kennwerte für Fixed und Random Effects-Modell. Einen Überblick über ein entsprechendes Excel-File bietet Abbildung 20.

#### 5. Die Berechnung der Regressionen im REM

Nach manueller Zusammenstellung von Studien-Effektstärken und Prädiktoren ermöglichte das Programm die Berechnung der statistischen Kennwerte für die Regressionen.

##	Angst insg.	Substanz	Outcome	g (bzw. T)	v	#	fe_model	re_model	re_model_gew	##	
1	a1	151 Phenelzine	C	0,87117206	0,1366901	g	0,26457398	g	0,26457398	0,26959411	
3		154 Amtriptiline	C	0,08202582	0,45362203	var(g)	0,00247035	var(g)	0,00247035	0,00314232	
4		165 Alprazolam	C	0,29899256	0,08665043	Cl u	0,18715683	Cl u	0,17796185	0,16935309	
5		163 Fluoxetine	C	0,14576862	0,36919266	Cl o	0,36199112	Cl o	0,39118611	0,40963513	
6		153 Lamotrigine	I	0,27689666	0,02213075	Homogeneity	15,757745	re_varianz	-0,02582078	re_varianz gew	0,00905274
7		157 Sertraline	I	1,01180436	0,08679949	Significance	ns				
8		158 Sertraline	I	0,90377713	0,1223568						
9		160 Imipramine	I	0,28341702	0,1014815						
10		161 Imipramine	I	0	0,0352586						
11		162 Fluoxetine	I	-0,25874552	0,35938723						
12		150 Paroxetine	I	0,24878054	0,01315656						
13		152 Paroxetine	I	0,31256989	0,01098966						
14		156 Sertraline	I	0,24011723	0,02001548						
15		164 Fluoxetine	I	0,21113851	0,01785219						
16											
17											
18	a1 ohne direk	Substanz	Outcome	g (bzw. T)	v	#	fe_model	re_model	re_model_gew	##	
19	a2	154 Amtriptiline	C	0,87117206	0,1366901	g	0,46167351	g	0,51969661	g gew	0,52378385
20		165 Alprazolam	C	0,08202582	0,45362203	var(g)	0,01148957	var(g)	0,02289768	var(g) gew	0,02607888
21		157 Sertraline	I	0,27689666	0,02213075	Cl u	0,25158228	Cl u	0,14941709	Cl u gew	0,12661892
22		158 Sertraline	I	1,01180436	0,08679949	Cl o	0,67176473	Cl o	0,88997613	Cl o gew	0,91894879
23		160 Phenelzine	I	0,90377713	0,1223568	Homogeneity	9,92851714	re_varianz	0,04914581	re_varianz gew	0,06612709
24		160 Imipramine	I	0,28341702	0,1014815	Significance	ns				
25		162 Fluoxetine	I	-0,25874552	0,35938723						
26											
27											
28	a2	PTB_K	Substanz	Outcome	g (bzw. T)	v	#	fe_model	re_model	##	
29		154 Amtriptiline	C	0,74705682	0,13318394	g	0,23033388	g	0,23612247		
30		163 Fluoxetine	C	0,27632526	0,08650101	var(g)	0,0027877	var(g)	0,00357825		
31		153 Lamotrigine	I	0,33318542	0,37363324	Cl u	0,12684853	Cl u	0,10081313		
32		157 Sertraline	I	0,1907153	0,02202681	Cl o	0,33381923	Cl o	0,3714318		
33		160 Phenelzine	I	1,07184224	0,12744306	Homogeneity	9,11493884	re_varianz	0,004769		
34		160 Imipramine	I	0,18641154	0,10095953	Significance	ns				
35		150 Paroxetine	I	0,15702486	0,01309504						
36		152 Paroxetine	I	0,28711225	0,01096873						
37		156 Sertraline	I	0,156284	0,0199314						
38		164 Fluoxetine	I	0,14267673	0,01782236						
39											
40	a3	PTB_V	Substanz	Outcome	g (bzw. T)	v	#	fe_model	re_model	##	
41		154 Amtriptiline	C	0,69791556	0,13751601	g	0,26577672	g	0,26577672		
42		163 Fluoxetine	C	0,36094987	0,08711894	var(g)	0,00279137	var(g)	0,00279137		
43		153 Lamotrigine	I	0,24988907	0,37132239	Cl u	0,16222323	Cl u	0,14626754		
44		157 Sertraline	I	0,25148975	0,02210205	Cl o	0,3693302	Cl o	0,38526588		
45		160 Phenelzine	I	0,81429323	0,12000173	Homogeneity	6,64142181	re_varianz	-0,021209		
46		160 Imipramine	I	0,24736681	0,10122037	Significance	ns				
47		150 Paroxetine	I	0,25797514	0,01316426						
48		152 Paroxetine	I	0,27826347	0,01096187						
49		156 Sertraline	I	0,21359314	0,01998503						
50		164 Fluoxetine	I	0,13525676	0,01781888						

Abbildung 20: Berechnung der Modelle

## 7.4 Anhang D: Berechnung der Effektstärken

Berechnung von Hedges  $g$  zwischen Gruppen aus Mittelwerten und Standardabweichungen

$$g = \frac{M_{\text{KG}} - M_{\text{BG}}}{\sqrt{\frac{(n_{\text{KG}} - 1)s_{\text{KG}}^2 + (n_{\text{BG}} - 1)s_{\text{BG}}^2}{n_{\text{KG}} + n_{\text{BG}} - 2}}} \quad (1)$$

mit  $M_{\text{KG}}$  als Mittelwert der Kontrollgruppe,  $M_{\text{BG}}$  als Mittelwert der Behandlungsgruppe,  $s_{\text{KG}}^2$  als Varianz der Kontrollgruppe,  $s_{\text{BG}}^2$  als Varianz der Behandlungsgruppe,  $n_{\text{KG}}$  als Stichprobengröße der Kontrollgruppe und  $n_{\text{BG}}$  als Stichprobengröße der Behandlungsgruppe

Korrektur für kleine Stichprobenumfänge

$$g_{\text{U}} = g * \left( 1 - \frac{3}{4(n_{\text{KG}} + n_{\text{BG}} - 2) - 1} \right) \quad (2)$$

Hedges  $g$  aus  $t$ -Werten

$$g = t \sqrt{\frac{1}{n_{\text{KG}}} + \frac{1}{n_{\text{BG}}}} \quad (3)$$

Hedges  $g$  aus  $F$ -Werten

$$g = \sqrt{F \left( \frac{1}{n_{\text{KG}}} + \frac{1}{n_{\text{BG}}} \right)} \quad (4)$$

Hedges  $g$  aus Veränderungswerten

$$g = \frac{(M_{KG} - M_{BG}) * \sqrt{2(1 - r_{xy})}}{\sqrt{\frac{(n_{KG} - 1)s_{KG}^2 + (n_{BG} - 1)s_{BG}^2}{n_{KG} + n_{BG} - 2}}} \quad (5)$$

mit  $M_{KG}$  als Mittelwert der Veränderungswerte der Kontrollgruppe,  $M_{BG}$  als Mittelwert der Veränderungswerte der Behandlungsgruppe,  $s_{KG}^2$  als Varianz der Veränderungswerte der Kontrollgruppe,  $s_{BG}^2$  als Varianz der Veränderungswerte der Behandlungsgruppe,  $n_{KG}$  als Stichprobengröße der Kontrollgruppe,  $n_{BG}$  als Stichprobengröße der Behandlungsgruppe und  $r_{xy}$  als Prä-Post-Korrelation

Hedges  $g$  aus  $t$ -Werten von Veränderungswerten

$$g = t \sqrt{\frac{1}{n_{KG}} + \frac{1}{n_{BG}}} [2(1 - r_{xy})] \quad (6)$$

Hedges  $g$  aus  $F$ -Werten von Veränderungswerten

$$g = \sqrt{F \left( \frac{1}{n_{KG}} + \frac{1}{n_{BG}} \right) [2(1 - r_{xy})]} \quad (7)$$

Odds Ratio aus Häufigkeiten

$$o = \frac{n_{11}n_{22}}{n_{12}n_{21}} \quad (8)$$

mit  $n_{11}$  als Anzahl an Personen mit Kriterium erfüllt in Gruppe 1,  $n_{21}$  als Anzahl an Personen mit Kriterium erfüllt in Gruppe 2,  $n_{12}$  als Anzahl an Personen mit Kriterium nicht erfüllt in Gruppe 1 und  $n_{22}$  als Anzahl an Personen mit Kriterium nicht erfüllt in Gruppe 2



Hedges  $g$  aus Odds Ratio

$$g = \log o * \frac{\sqrt{3}}{\pi} \quad (9)$$

Berechnung von Hedges  $g$  innerhalb der Gruppen aus Mittelwerten und Standardabweichungen

$$g = \frac{M_{M2} - M_{M1}}{\sqrt{\frac{(n_{M2} - 1)s_{M2}^2 + (n_{M1} - 1)s_{M1}^2}{n_{M2} + n_{M1} - 2}}} \quad (10)$$

mit  $M_{M1}$  als Mittelwert zu Messzeitpunkt 1,  $M_{M2}$  als Mittelwert zu Messzeitpunkt 2,  $s_{M1}^2$  als Varianz zu Messzeitpunkt 1,  $s_{M2}^2$  als Varianz zu Messzeitpunkt 2,  $n_{M1}$  als Stichprobengröße zu Messzeitpunkt 1 und  $n_{M2}$  als Stichprobengröße zu Messzeitpunkt 2

## 7.5 Anhang E: Retest-Reliabilitäten

In der folgenden Tabelle sind die in die Berechnung der mittleren Retest-Reliabilität eingeflossenen Instrumente dargestellt. Es wurden sowohl Instrumente eingeschlossen, die in Studien mehrerer Angststörungen eingesetzt wurden, als auch für die einzelnen Angststörungen spezielle Instrumente.

Instrument	Retest	Autoren
Acrophobia Questionnaire	$r = .84$	Baker, Cohen & Saunders, 1973
Agoraphobic Cognitions Questionnaire	$r = .86$	Chambless, 1984
Anxiety Sensitivity Index	$r = .75$	Peterson & Reis, 1992
Attitudes to Flying	$r = .92$	Howard, Murphy & Clarke, 1983
Beck Anxiety Inventory	$r = .75$	Beck, Epstein, Brown, & Steer, 1988
Beck Depression Inventory	$r = .91$	Beck, Ward, Mendelsohn, Mock & Erbaugh, 1961
Body Sensations Questionnaire	$r = .67$	Chambless, 1984
Clinical Global Impressions	$r = .50$	Guy, 1976
Clinician Administered PTSD Scale	$r = .94$	Weathers et al., 1992
Davidson Trauma Scale	$r = .86$	Davidson et al., 1997
Fear of Negative Evaluation	$r = .78$	Watson & Friend, 1969
Fear Questionnaire	$r = .85$	Marks & Mathews, 1979
Impact of Event Scale	$r = .87$	Horowitz, Wilner, Alvarez, 1979
Maudsley Obsessional-Compulsive Inventory	$r = .76$	Hodgson & Rachman, 1977
Mobility Inventory	$r = .90$	Chambless, Caputo, Gracely, Jasin & Williams, 1985
Padua Inventory	$r = .80$	Sanavio, 1988
Penn State Worry Questionnaire	$r = .92$	Meyer, Miller, Metzger & Borkovec, 1990
Social Avoidance and Distress Scale	$r = .78$	Watson & Friend, 1969
Social Interaction Anxiety Scale	$r = .91$	Heimberg, Müller, Holt, Hope & Liebowitz, 1992
Social Phobia and Anxiety Inventory	$r = .86$	Turner, Beidel, Dancu & Stanley, 1989
Social Phobia Scale	$r = .91$	Heimberg, Müller, Holt, Hope & Liebowitz, 1992
State-Trait Anxiety Inventory	$r = .79$	Spielberger, Gorsuch & Lushene, 1970
Yale-Brown Obsessive Compulsive Scale	$r = .60$	Goodman et al., 1989

## 7.6 Anhang F: Ausreißeranalyse

In der folgenden Tabelle sind für die einzelnen Störungen und Vergleiche die jeweiligen Definitionen der Ausreißer sowie die prozentuale Anzahl an Effektstärken angegeben, die in den Hauptanalysen zu Post ausgeschlossen wurden.

Störung	Vergleich	Definition Ausreißer	% Ausreißer
Generalisierte	PsT : WL	$g < -0.50$ und $g > 2.00$	3.57
Angststörung	PsT : TP/PP	$g < -0.50$ und $g > 1.50$	1.47
	VT : KVT	$g < -1.00$ und $g > 1.00$	0
	PhT : PP	$g < -1.50$ und $g > 2.50$	1.27
	PhT : PsT	$g < -1.50$ und $g > 2.00$	3.45
Panikstörung	PsT : WL	$g < -1.00$ und $g > 2.50$	2.56
	PsT : TP/PP	$g < -1.00$ und $g > 2.00$	1.36
	VT : KVT	$g < -1.00$ und $g > 1.00$	1.01
	PhT : PP	$g < -1.00$ und $g > 2.00$	0.85
	PhT : PsT	$g < -3.00$ und $g > 2.50$	2.94
	PsT : PsT+PhT	$g < -1.00$ und $g > 1.50$	0.68
Posttraumatische	PsT : WL	$g < -0.50$ und $g > 2.50$	1.20
Belastungsstörung	PsT : TP/PP	$g < -0.50$ und $g > 1.50$	0
	VT : KVT	$g < -1.00$ und $g > 1.00$	2.78
	PhT : PP	$g < -0.50$ und $g > 1.50$	1.60
Soziale Phobie	PsT : WL	$g < -0.50$ und $g > 2.00$	2.27
	PsT : TP/PP	$g < -1.00$ und $g > 1.50$	2.81
	VT : KVT	$g < -1.00$ und $g > 1.00$	2.65
	PhT : PP	$g < -1.00$ und $g > 2.00$	5.28
	PhT : PsT	$g < -1.50$ und $g > 1.00$	1.27
Spezifische Phobie	PsT : WL	$g < -1.00$ und $g > 3.00$	2.56
	PsT : TP/PP	$g < -2.00$ und $g > 3.00$	0
	VT : KVT	$g < -1.50$ und $g > 1.00$	0
Zwangsstörung	PsT : WL	$g < -0.50$ und $g > 2.00$	0
	VT : KVT	$g < -1.00$ und $g > 1.00$	1.23
	PhT : PP	$g < -0.50$ und $g > 1.50$	1.01
	PsT : PsT+PhT	$g < -1.00$ und $g > 1.50$	2.86

## 7.7 Anhang G: RE-Schätzer

Die Berechnung des gewichteten Momentmethodenschätzers erfolgte über die zweite Methode nach Shadish et al. (1994) und wurde implementiert in Study Input. Die Berechnungen für den Maximum Likelihood-Schätzer erfolgten mittels HLM 5 mit maximal 10 000 Iterationen (Raudenbush, Bryk & Congdon, 2000).

Über den Momentmethodenschätzer wird die RE-Varianz auf Null gesetzt bei einem negativen Ergebnis der Berechnung, der Maximum Likelihood-Schätzer nähert sich dagegen asymptotisch Null. Es zeigte sich in der überwiegenden Mehrzahl an Analysen eine Übereinstimmung der Ergebnisse über die verschiedenen Schätzer. Wie aufgrund der zugrunde liegenden Formeln zu erwarten, zeigte sich eine interpretative Übereinstimmung der Prüfgröße des Homogenitätstests mit der RE-Varianz für den gewichteten Momentmethodenschätzer und den Maximum Likelihood-Schätzer, im Speziellen auffallend für den Vergleich von PhT mit PP, der der Auslöser der Post hoc-Sensitivitätsanalyse war. Eine Entscheidung für einen anderen Schätzer hätte hier die Durchführung einer Regressionsanalyse erfordert. Generelle Empfehlungen zum Einsatz eines bestimmten Schätzers in Abhängigkeit von Eigenschaften der Meta-Analyse (z.B. Anzahl an Studien, Stichprobengröße etc.) liegen nach Wissen der Autorin nicht vor. In der folgenden Tabelle sind die Ergebnisse der Analysen des Random Effects-Modells mit verschiedenen Schätzern der RE-Varianz dargestellt (MoM ug = Momentmethoden ungewichtet, MoM g = Momentmethoden gewichtet, ML = Maximum Likelihood).

Es sei betont, dass sich keine substantziellen Veränderungen in der Interpretation der Ergebnisse fast aller Analysen ergaben.

Störung	Vergleich	Schätzer	g	SD(g)	95%-CI	$\sigma_{\theta}^2$	$\chi^2$	n
GAS	PsT : WL	MoM ug	0.82	0.09	0.62 – 1.01	0	10.91	19
		MoM g	0.82	0.09	0.62 – 1.01	0	10.91	19
		ML	0.82	0.09	0.62 – 1.01	< 0.001	10.91	19
	PsT : TP	MoM ug	0.57	0.12	0.30 – 0.85	0	7.28	9
		MoM g	0.57	0.12	0.30 – 0.85	0	7.28	9
		ML	0.57	0.12	0.30 – 0.85	< 0.001	7.28	9
	VT : KVT	MoM ug	0.08	0.11	-0.16 – 0.31	0	5.84	11
		MoM g	0.08	0.11	-0.16 – 0.31	0	5.84	11
		ML	0.08	0.11	-0.16 – 0.31	< 0.001	5.84	11
PhT : PP	MoM ug	0.32	0.03	0.25 – 0.38	0.03	74.89	68	
	MoM g	0.31	0.03	0.25 – 0.36	0.01	74.89	68	

		ML	0.30	0.02	0.25 – 0.35	< 0.001	74.89	68
	PsT : PhT	MoM ug	0.33	0.15	-0.02 – 0.67	0	6.80	8
		MoM g	0.33	0.15	-0.02 – 0.67	0	6.80	8
		ML	0.33	0.15	-0.02 – 0.67	0.01	6.81	8
OCD	PsT : WL	MoM ug	0.94	0.15	0.56 – 1.32	0	5.28	6
		MoM g	0.94	0.15	0.54 – 1.33	0.01	5.28	6
		ML	0.94	0.16	0.53 – 1.34	0.02	5.27	6
	VT : KVT	MoM ug	0.09	0.13	-0.22 – 0.41	0	3.04	7
		MoM g	0.09	0.13	-0.22 – 0.41	0	3.04	7
		ML	0.09	0.13	-0.22 – 0.41	< 0.001	3.04	7
	PhT : PP	MoM ug	0.42	0.04	0.33 – 0.51	0	35.63 <sup>+</sup>	22
		MoM g	0.39	0.06	0.26 – 0.52	0.03	35.63 <sup>+</sup>	22
		ML	0.39	0.06	0.26 – 0.52	0.03	36.02 <sup>+</sup>	22
	PsT :	MoM ug	0.28	0.10	0.06 – 0.49	0	11.42	12
	PsT+PhT	MoM g	0.28	0.10	0.06 – 0.49	0	11.42	12
			ML	0.28	0.10	0.06 – 0.49	< 0.001	10.31
PDA	PsT : WL	MoM ug	0.87	0.08	0.71 – 1.03	0.04	41.11	32
		MoM g	0.87	0.08	0.70 – 1.04	0.05	41.11	32
		ML	0.87	0.08	0.69 – 1.04	0.06	41.11	32
	SV : WL	MoM ug	0.80	0.21	0.29 – 1.30	0.16	15.37 <sup>+</sup>	8
		MoM g	0.81	0.23	0.27 – 1.36	0.22	15.37 <sup>+</sup>	8
		ML	0.82	0.23	0.27 – 1.36	0.23	15.99 <sup>+</sup>	8
	PsT : TP	MoM ug	0.51	0.10	0.30 – 0.72	0.01	13.12	13
		MoM g	0.51	0.10	0.30 – 0.72	0.01	13.12	13
		ML	0.51	0.10	0.30 – 0.72	0.01	13.13	13
	VT : KVT	MoM ug	0.09	0.07	-0.07 – 0.24	0	19.61	26
		MoM g	0.09	0.07	-0.07 – 0.24	0	19.61	26
		ML	0.09	0.07	-0.07 – 0.24	< 0.001	19.61	26
	PhT : PP	MoM ug	0.38	0.03	0.31 – 0.45	0.03	104.89 <sup>+</sup>	78
		MoM g	0.37	0.03	0.31 – 0.43	0.02	104.89 <sup>+</sup>	78
		ML	0.36	0.03	0.30 – 0.42	0.01	108.53 <sup>+</sup>	78
	PsT : PhT	MoM ug	0.27	0.16	-0.07 – 0.62	0.21	37.06 <sup>+</sup>	13
		MoM g	0.27	0.15	-0.07 – 0.60	0.20	37.06 <sup>+</sup>	13
		ML	0.27	0.16	-0.08 – 0.63	0.22	39.85 <sup>+</sup>	13
	PsT :	MoM ug	0.23	0.07	0.09 – 0.37	0	11.48	19
	PsT+PhT	MoM g	0.23	0.07	0.09 – 0.37	0	11.48	19
			ML	0.23	0.07	0.09 – 0.37	< 0.001	11.48
PTB	PsT : WL	MoM ug	0.91	0.13	0.62 – 1.20	0.09	17.48 <sup>+</sup>	12
		MoM g	0.91	0.12	0.63 – 1.18	0.06	17.48 <sup>+</sup>	12

		ML	0.91	0.12	0.64 – 1.17	0.06	17.53 <sup>+</sup>	12
	PsT : TP	MoM ug	0.46	0.19	-0.16 – 1.08	0	1.29	4
		MoM g	0.46	0.19	-0.16 – 1.08	0	1.29	4
		ML	0.46	0.19	-0.16 – 1.08	< 0.001	1.29	4
	VT : KVT	MoM ug	0.03	0.10	-0.21 – 0.28	0	8.39	8
		MoM g	0.01	0.12	-0.26 – 0.29	0.02	8.39	8
		ML	0.001	0.13	-0.30 – 0.30	0.03	8.49	8
	PhT : PP	MoM ug	0.28	0.05	0.18 – 0.39	0	15.76	15
		MoM g	0.29	0.06	0.17 – 0.41	0.01	15.76	15
		ML	0.28	0.05	0.18 – 0.39	< 0.001	15.76	15
SoP	PsT : WL	MoM ug	0.75	0.09	0.56 – 0.94	0	9.95	21
		MoM g	0.75	0.09	0.56 – 0.94	0	9.95	21
		ML	0.75	0.09	0.56 – 0.94	< 0.001	9.95	21
	PsT : TP	MoM ug	0.25	0.08	0.07 – 0.43	0	8.64	11
		MoM g	0.25	0.08	0.07 – 0.43	0	8.64	11
		ML	0.26	0.08	0.07 – 0.44	0.003	8.64	11
	VT : KVT	MoM ug	0.04	0.10	-0.17 – 0.25	0	6.11	15
		MoM g	0.04	0.10	-0.17 – 0.25	0	6.11	15
		ML	0.04	0.10	-0.17 – 0.25	< 0.001	6.11	15
	PhT : PP	MoM ug	0.46	0.08	0.29 – 0.63	0.12	65.58 <sup>+</sup>	29
		MoM g	0.45	0.07	0.31 – 0.60	0.07	65.58 <sup>+</sup>	29
		ML	0.46	0.08	0.30 – 0.61	0.09	67.78 <sup>+</sup>	29
	PsT : PhT	MoM ug	-0.06	0.14	-0.39 – 0.27	0.06	10.91	8
		MoM g	-0.06	0.13	-0.38 – 0.25	0.05	10.91	8
		ML	-0.06	0.13	-0.38 – 0.25	0.05	10.91	8
SpP	PsT : WL	MoM ug	1.08	0.10	0.87 – 1.29	0.15	93.78 <sup>+</sup>	33
		MoM g	1.07	0.13	0.80 – 1.33	0.35	93.78 <sup>+</sup>	33
		ML	1.07	0.12	0.82 – 1.31	0.27	94.22 <sup>+</sup>	33
	PsT : TP	MoM ug	0.32	0.20	-0.19 – 0.83	0	1.19	6
		MoM g	0.32	0.20	-0.19 – 0.83	0	1.19	6
		ML	0.32	0.20	-0.19 – 0.83	< 0.001	1.19	6
	VT : KVT	MoM ug	-0.20	0.22	-0.90 – 0.51	0	1.24	4
		MoM g	-0.20	0.22	-0.90 – 0.51	0	1.24	4
		ML	-0.20	0.22	-0.90 – 0.51	< 0.001	1.24	4

<sup>+</sup>p < .10

## **Ehrenwörtliche Erklärung**

Hiermit erkläre ich, dass mir die geltende Promotionsordnung bekannt ist. Ich habe die Dissertation selbst angefertigt, insbesondere die Hilfe eines Promotionsberaters nicht in Anspruch genommen, und alle von mir benutzten Hilfsmittel und Quellen sowie Personen, die mich unterstützten, in der Arbeit angegeben. Darüber hinaus haben keine Dritten weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen. Die Dissertation wurde noch nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht. Weder wesentliche Teile bzw. eine ähnliche oder andere Abhandlung wurden von mir bei einer anderen Hochschule bzw. anderen Fakultät als Dissertation eingereicht.

Hiermit versichere ich, dass ich nach bestem Wissen die reine Wahrheit gesagt und nichts verschwiegen habe.

Jena, 14.04.03

Jena, 14.04.03

## Lebenslauf

Geburtsdatum: 18. März 1977  
Geburtsort: Annaberg-Buchholz (Sachsen)  
Familienstand: Ledig  
Nationalität: deutsch

Ausbildung:

1983-1991	Polytechnische Oberschule Elterlein (Sachsen)
1991-1995	Gymnasium Schlettau (Sachsen)
1995	Reifeprüfung (Abitur); Note sehr gut
1995-2001	Studium an der Friedrich-Schiller-Universität Jena; Diplomstudiengang Psychologie
2001	Diplom; Note sehr gut
2001-	Eingeschrieben als Promotionsstudentin an der Friedrich-Schiller-Universität Jena Gefördert durch ein Landesgraduiertenstipendium

Mitarbeit an verschiedenen Forschungsprojekten

1998-2001	In den Bereichen Informationsverarbeitung (Prof. Dr. E.R. Straube), Entwicklungspsychologie (Dr. E. Schmitt-Rothermund) und Entwicklungspsychopathologie (Dr. R. Steil)
2001-2003	In den Bereichen Psychotherapieforschung (Dr. R. Steil) und Informationsverarbeitung

Sonstiges:

1998	Praktikum Personalentwicklung bei "Carl Zeiss" (6 Wochen)
1998-2001	Zusätzliche Kurse in Soziologie und Informatik
1999	Praktikum Psychiatrie, Klinikum der FSU Jena (Behandlung von Patienten mit emotionalen Störungen; 14 Wochen)