

# Placebos in der Psychotherapieforschung – eine systematische Analyse am Beispiel der systematischen Desensibilisierung

Cosima Locher · Sebastian Hasler · Jens Gaab

Klinische Psychologie und Psychotherapie, Fakultät für Psychologie, Universität Basel, Basel, Schweiz

## Schlüsselwörter

Systematische Desensibilisierung · Placeboeffekte · Attention placebo control · Glaubwürdigkeit

## Zusammenfassung

**Hintergrund:** Die Voraussetzungen der randomisierten, Placebo-kontrollierten Evaluation – die Ununterscheidbarkeit der Therapiebedingungen für die Patienten und die Verblindung der Therapeuten – sind in der Psychotherapieforschung nicht uneingeschränkt gegeben. Das Ziel dieser qualitativen systematischen Übersichtsarbeit ist es, am Beispiel der systematischen Desensibilisierung (SD) die Vorgehensweisen zur Bestimmung der Behandlungsspezifität zu beschreiben und deren theoretische und praktische Implikationen zu diskutieren. **Methodik:** Auf der Basis einer systematischen Literaturrecherche in PsycINFO und PubMed für den Zeitraum von 1976–2015 wurden Psychotherapiestudien gesucht, in denen Patienten mit Angstsymptomen einer SD- oder Placebo-Gruppe zugewiesen und verglichen wurden. Nach Extraktion der gewählten Moderator-Variablen konnten 11 Studien eingeschlossen werden. **Ergebnisse:** Die Ergebnisse zu den spezifischen Effekten der SD waren nicht einheitlich. Eine durchgeführte Moderator-Analyse zeigte, dass der Nachweis spezifischer Effekte von der Glaubwürdigkeit des Placebos und dem dabei erlebten Ausmaß emotionaler Erfahrungen abhängig ist. **Schlussfolgerung:** Ob ein bestimmtes psychotherapeutisches Verfahren als spezifisch einzustufen ist, hängt von der Operationalisierung der Placebo-Kontrolle ab. Insbesondere bei SD scheint die Einschränkung des zu bearbeitenden Themas und in der Folge die Unterbindung der emotionalen Erfahrung des Patienten innerhalb der Kontrollbedingung eine Voraussetzung dafür zu sein, spezifische Effekte der SD zu finden.

© 2016 S. Karger GmbH, Freiburg

Das englische Original ist online abrufbar unter [www.karger.com/doi/10.1159/000443464](http://www.karger.com/doi/10.1159/000443464)

## Keywords

Systematic desensitization · Placebo effects · Attention placebo control · Credibility

## Summary

**Background:** The requirements of the randomized placebo-controlled trial design – indistinguishability of comparators for patients and blinding of allocation for therapists – are not unequivocally warranted in psychotherapy research. Our qualitative review exemplary set out to systematically review studies aiming to establish specificity of a circumscribed psychotherapeutic intervention (systematic desensitization, SD) and to discuss possible theoretical and practical implications. **Methods:** We performed systematic searches in PsycINFO and PubMed on studies from 1976 to 2015 that compared individuals with anxiety symptoms in a SD and psychotherapeutic placebo treatment. After extracting the moderators of interest, 11 studies met all inclusion criteria. **Results:** Regarding treatment efficacy of SD, the included studies led to inconsistent results. Studies with similar credibility in both SD and placebo groups reported no differences between SD and psychotherapeutic placebos; and that emotional experiencing may be the key element for a psychotherapeutic placebo to be effective. **Conclusion:** We conclude that whether a given psychotherapeutic intervention is to be considered a specific treatment or not is influenced by the operationalization of the placebo control condition. Specifically for SD, the restriction of topic and in consequence of emotional experiencing in the control condition seems to be a prerequisite to find specific effects of SD.

## Hintergrund

Vor rund 80 Jahren nahm Robert D. Rudolf, klinischer Professor an der Universität von Toronto und Verfasser eines renommierten medizinischen Lehrbuches, in einem Artikel zum *post hoc ergo propter hoc*-Fehlschluss folgenden Standpunkt ein: «Ich habe beinahe das Gefühl, mich für das Anliegen dieses Artikels entschuldigen zu müssen, aber es scheint mir unumgänglich. Gehen wir nicht alle zu leichtfertig davon aus, dass – weil etwas auf etwas anderes folgt – ersteres automatisch durch letzteres bedingt ist?» [Rudolf, 1938]. Die Annahme, dass Interventionseffekte nicht ausschließlich auf die Intervention selbst zurückzuführen sind, ist Grundlage und Rechtfertigung für das randomisierte Placebo-kontrollierte Studiendesign in der Medizin [Beecher, 1955] sowie in der Psychotherapie [Rosenthal und Frank, 1956]. Für die Psychotherapie folgt daraus, dass sie «mit einem Pillen- oder psychologischen Placebo (...)» verglichen werden muss, um den «Effekt der Behandlung (als) spezifisch (klassifizieren) zu können» [Zitat aus Chambless et al., 1998, p 4, Übersetzung durch die Autoren]. Die dazu notwendigen Voraussetzungen, «um die therapeutische Spreu vom Weizen zu trennen» [Jones und Podolsky, 2015, p 1502, Übersetzung durch die Autoren] – die Ununterscheidbarkeit der Interventionsbedingungen für die Patienten und die Verblindung der Therapeuten – sind in der Psychotherapieforschung nicht uneingeschränkt gegeben [Borkovec und Sibrava, 2005]. Daraus entstehen konzeptionelle sowie praktische Probleme [Kirsch, 2005], die auch die Unterscheidung der Prozesse und Effekte von Placebo sowie Psychotherapie erschweren [Gaab et al., 2015].

Vor diesem Hintergrund ist es für den Nachweis einer möglichen spezifischen Wirkung einer psychotherapeutischen Methode – hier am Beispiel der systematischen Desensibilisierung (SD) – von Interesse, die Auswirkung unterschiedlicher Operationalisierungen des Placebo-Arms in der Psychotherapieforschung zu untersuchen. Obwohl das akademische wie auch das klinische Interesse an der SD seit den 1980er-Jahren wesentlich nachgelassen hat [McGlynn et al., 2004], ist diese psychotherapeutische Methode seit ihren Anfängen Gegenstand wissenschaftlicher Untersuchungen. Die Ergebnisse dieser Forschung wurden vor rund 40 Jahren von Kazdin und Wilcoxon [1976] in einer wegweisenden Übersichtsarbeit zusammengefasst. Auch wenn – im Gegensatz zur genannten frühen Übersichtsarbeit – es nicht das explizite Ziel dieser Arbeit ist, die Spezifität der SD zu untersuchen, sollen hier am Beispiel der SD die methodologischen Schwierigkeiten aufgezeigt werden, die den Nachweis der Spezifität in der Psychotherapie erschweren.

Die von Joseph Wolpe entwickelte SD [Wolpe, 1954] ist eine der ersten therapeutischen Implementierungen der klassischen Konditionierung und verfügt über eine nachgewiesene Wirksamkeit bei der Behandlung von phobischen Ängsten [McGlynn et al., 1981]. Tatsächlich wurde hinsichtlich der SD die Behauptung aufgestellt, dass «zum ersten Mal in der Geschichte der Psychotherapie ein spezifisches therapeutisches Behandlungspaket existiert, das reliabel zu einem messbaren Nutzen für Patienten über ein breites Spektrum an Beschwerden führt» [Paul, 1969, zitiert in McGlynn et al., 1981, p 150, Übersetzung durch die Autoren]. In der Regel

besteht die SD darin, Patienten in Tiefenmuskulentspannung zu trainieren, eine Hierarchie angstauslösender Szenarien zu erstellen und anschließend diese Szenarien wiederholend und in hierarchischer Abfolge mit der Entspannung zu koppeln [Landau et al., 1984; McGlynn et al., 1981].

Pauls enthusiastische Aussage eines «messbaren Nutzen[s] für Patienten» gründete auf seiner eigenen Forschung, in der er die SD mit Placebo-Bedingungen verglich, die den Aspekt der therapeutischen Zuwendung und Aufmerksamkeit kontrollierten, und die weitere Forschungsarbeiten initiierte. Im Placebo-Arm der Originalstudie von Paul – das sogenannte Stresstoleranztraining – mussten Probanden nach Einnahme einer Placebo-Pille eine repetitive Aufgabe absolvieren [Paul und Shannon, 1966]. Interessanterweise wurde die vorgetäuschte Wirksamkeit dieser Placebo-Bedingung zusätzlich dadurch unterstützt, dass es sich bei der «stressenden» Aufgabe tatsächlich um eine stressfreie Aufgabe handelte, bei der nicht unangenehme Geräusche von einer Tonbandaufnahme identifiziert werden mussten. Dieses Stresstoleranztraining von Paul [1969; Paul und Shannon, 1966] diente als Blaupause für nachfolgende Versuche, unspezifische Behandlungseffekte in SD-Studien zu kontrollieren. Dabei konnte wiederholt die Überlegenheit der SD gegenüber den durchgeführten Placebo-Bedingungen repliziert werden [z.B. Lick und Bootzin, 1970; Zeisset, 1968]. Allerdings wurde das Design der Kontrollbedingungen in diesen frühen Studien stark hinterfragt, besonders da die Glaubwürdigkeit des Placebo-Arms nur selten erfasst wurde, auch wenn schon früh auf diese Notwendigkeit hingewiesen worden war [z.B. Nau et al., 1974]. Dieser Forderung wurde auch dadurch Nachdruck verliehen, dass bestimmte Placebo-Bedingungen gegenüber entsprechenden SD-Bedingungen tatsächlich in zahlreichen Studien als weniger glaubwürdig bewertet wurden [Borkovec und Nau, 1972; McReynolds et al., 1973; Nau et al., 1974].

Dieser Umstand führte dazu, dass einerseits glaubwürdige psychotherapeutische Placebo-Bedingungen entworfen wurden (z.B. Dissonanzsteigerungs-Kontrolle [McReynolds et al., 1973]). Andererseits wurde die Erwartung durch experimentelle Variation [Marcia et al., 1969] kontrolliert oder in Bezug auf die therapeutischen Veränderungen und die Glaubwürdigkeit des Behandlungsrationalen [z.B. D’Zurilla et al., 1973; Gelder et al., 1973; McReynolds et al., 1973; McReynolds und Tori, 1972; Steinmark und Borkovec, 1974; Wilson, 1973] erfasst.

Auch wenn diese Vorgehensweise einen ersten Schritt zur Sicherstellung der Ununterscheidbarkeit der Interventionsbedingungen für die Patienten darstellt, wurde die Verblindung nur selten thematisiert. Dies könnte darauf zurückzuführen sein, dass die Therapeuten unweigerlich darüber im Bilde waren, welche Behandlung eingesetzt wurde [O’Leary und Borkovec, 1978]. Eine Ausnahme bildet die Studie von McReynolds und Kollegen, deren «Therapeuten verblindet (waren) bezüglich der Art der Placebo-Kontrolle (...) und lediglich darüber informiert (wurden), dass es sich um eine kognitive Vergleichsbehandlung handle, die die wirklichen kognitiven Komponenten der systematischen Desensibilisierung enthalte» [McReynolds et al., 1973, p. 86, Übersetzung durch die Autoren].

Die Ergebnisse dieser sehr aktiven Forschungsperiode zur Spezifität der SD wurden in der einflussreichen Übersichtsarbeit von Kazdin und Wilcoxon [1976] in der folgenden ernüchternden Schlussfolgerung zusammengefasst: «Der gegenwärtige Stand der Desensibilisierungsforschung erlaubt die konkurrierende Interpretation, dass unspezifische Behandlungseffekte – und nicht spezifische therapeutische Bestandteile – deren Wirksamkeit bedingen. Über diese konkurrierende Interpretation hinaus gibt es keine überzeugende Evidenz dafür, dass die Desensibilisierung eine spezifische Behandlung ist» [Kazdin und Wilcoxon, 1976, p 753, Übersetzung durch die Autoren]. Die Grundlage dieser Schlussfolgerung war die Untersuchung von Studien mit Placebo-Bedingungen, Komponentenstudien sowie von Studien, bei denen die Erwartung der Probanden empirisch erfasst und kontrolliert wurden. Dabei zeigte sich, dass die SD dann mit den Kontrollbedingungen wirkungsgleich und damit «unspezifisch» war, wenn bei letzteren die Erwartungen erfasst und kontrolliert wurden.

Da in der Forschung zur Spezifität der SD bislang eine Vielfalt an verschiedenen psychotherapeutischen Placebo-Bedingungen eingesetzt wurde, eignet sich die SD hervorragend, der Frage nachzugehen, unter welchen Bedingungen psychotherapeutische Placebos wirken und wie dies den Nachweis der Spezifität psychotherapeutischer Behandlungen beeinflusst. Es ist dabei nicht das Ziel dieser Analyse, die Mechanismen oder die Spezifität der SD an sich zu untersuchen, sondern vielmehr Studien, die die Behandlungsspezifität von SD anhand eines Vergleichs mit Placebo-Bedingungen untersucht haben, systematisch zusammenzufassen und die Implikationen dieser Analyse zu diskutieren. Dabei werden Studien berücksichtigt, die im Zeitraum von 1976 bis heute, d.h. nach der Übersichtsarbeit von Kazdin und Wilcoxon [1976] veröffentlicht wurden.

## Methode

Das grundlegende Prinzip Placebo-kontrollierter Untersuchungen besteht darin, verschiedene Interventionen (z.B. Verum und Placebo) mit einem identischen Rational zu versehen – oder anders ausgedrückt, die zufälligen Behandlungsbestandteile zu kontrollieren und die charakteristischen Behandlungsbestandteile zu manipulieren [Grünbaum, 1981]. Spezifische, d.h. Verum-Effekte werden dann als das Ausmaß verstanden, in dem die Effekte der charakteristischen Behandlungsbestandteile die der zufälligen übertreffen. Entscheidend dabei ist, dass eine entsprechende Behandlungstheorie jeweils festlegt, welche Behandlungsbestandteile als charakteristisch und zufällig betrachtet werden.

Für die vorliegende Übersichtsarbeit soll die ursprünglich vorgeschlagene Theorie der reziproken Hemmung – unter der Annahme, dass die SD-Behandlung eine valide Operationalisierung dieser Behandlungstheorie ist – als Grundlage der Behandlung spezifischer Angst- und Phobiesymptome angenommen werden [Wolpe, 1954]. Aufgrund dieser Annahme können Theorie-definierte, charakteristische (d.h. Muskelentspannung, Erstellen einer Hierarchie angstauslösender Szenarien, Koppeln angstauslösender Szenarien mit Entspannung) und zufällige (z.B. Vorstellungen, Erwartungen, Allianz [Colloca und Miller, 2011; Hunter, 2007; Kaptchuk et al., 2008]) Behandlungsbestandteile unterschieden werden. Auf der Grundlage der Hypothese, dass «jede Therapie, die mindestens einen charakteristischen Faktor enthält, der therapeutisch [für eine gegebene Störung] ist, als Verum angesehen werden kann» [Grünbaum, 1981, p 161, Übersetzung durch die Autoren], ist die Debatte, ob die von Wolpe [1968] ursprünglich vorgeschlagene reziproke Hemmung oder doch eher alternative Behandlungstheorien (z.B. Ex-

tinktion, Habituation, Gegenkonditionierung, kognitiv-soziale Verstärkung [McGlynn et al., 1981]) für die vorliegende Störung von Bedeutung sind, nicht Gegenstand dieser Analyse. Deswegen beschränkt sich der Fokus auf Studien mit Kontrollstrategien, die aus der Perspektive der erwähnten SD-Behandlungstheorie aus zufälligen Behandlungsbestandteilen bestehen. Dabei handelt es sich entweder um eine rein konzeptuelle Annahme (d.h. «attention placebo control»-Strategien), oder die Erwartung der Probanden wird in den Kontrollstrategien empirisch überprüft (d.h. empirisch abgeleitete Kontrollstrategien) [Kazdin und Wilcoxon, 1976]. Studien, die den Einsatz bestimmter Behandlungselemente manipulieren (wie z.B. sogenannte Dismantling- oder additive Komponentenstudien), sowie direkte Vergleiche zwischen verschiedenen psychotherapeutischen Methoden (wie z.B. Flooding, rationale-emotive Therapie oder Lernförderung) werden hierbei nicht untersucht, da Vergleiche von Verum-Therapien – gemäß obiger Definition – nicht Gegenstand dieser Arbeit sind.

### Suchstrategie und Einschlusskriterien

Zum Zweck einer qualitativen Übersichtsarbeit wurde für den Suchbegriff «systematische Desensibilisierung» eine systematische Literaturrecherche in PsycINFO und PubMed für den Zeitraum von 1976 bis 2015 durchgeführt. Zusätzlich wurden die Referenzen aller eingeschlossenen Artikel durchsucht. Eingeschlossen wurden randomisierte, klinische, wissenschaftlich begutachtete Studien, die auf Englisch oder Deutsch veröffentlicht wurden und Probanden mit Angstsymptomen untersuchten, die einem SD- oder einem Placebo-Arm zugeteilt wurden. Es wurden nur Studien nach dem Einreichdatum der systematischen Übersichtsarbeit von Kazdin und Wilcoxon [1976], d.h. nach dem 27. Juni 1975, berücksichtigt.

### Studienoutcomes und mögliche Moderatoren

Als Studienoutcomes wurden Selbstbeurteilungsskalen zu Angstsymptomen sowie behaviorale und physiologische Parameter extrahiert. Wenn sowohl Intra-gruppen- als auch Intergruppenvergleiche berichtet wurden, wurden bevorzugt letztere für die Extraktion verwendet.

Da die Validität von Placebo-kontrollierten Studien auf der Ununterscheidbarkeit der Therapiebedingungen für die Patienten und der Verblindung der Therapeuten hinsichtlich der Zuweisung der Patienten beruht, wurden die strukturelle Äquivalenz, die Erwartungen der Probanden bezüglich therapeutischer Veränderung sowie die Glaubwürdigkeit des Behandlungsrationalen und die thematische Einschränkung (siehe unten) als Moderatoren extrahiert. Zusätzlich wurden die Studien hinsichtlich der Verblindung der Therapeuten und der Beurteiler beurteilt.

**Strukturelle Äquivalenz:** Die strukturelle Äquivalenz von SD und psychotherapeutischer Placebo-Behandlung wurde anhand folgender Kriterien beurteilt [siehe dazu Baskin et al., 2003]: a) Anzahl der Sitzungen, b) Sitzungslänge, c) Format (Gruppen- oder Einzelsitzungen), d) Therapeutentraining, e) ob die Interventionen für die Teilnehmer individualisiert waren. In Anlehnung an Baskin et al. [2003] wurden der SD-Arm und der Placebo-Arm dann als strukturell äquivalent angesehen, wenn sie sich hinsichtlich keiner der genannten Dimensionen unterschieden. Entsprechend wurde dann eine strukturelle Inäquivalenz kodiert, wenn Abweichungen in einer oder mehreren Dimensionen auftraten.

**Erwartungen der Probanden bezüglich therapeutischer Veränderung sowie Glaubwürdigkeit des Behandlungsrationalen:** Die Studien wurden anhand der folgenden Operationalisierung von Erwartung und Glaubwürdigkeit beurteilt: a) positives Behandlungsrational und/oder Erwartung eines positiven Behandlungsergebnisses [Bowers und Clum, 1988; Stevens et al., 2000], b) Erfassung der Erwartung (d.h. Selbstbeurteilung, Simulatoren, sowie implizite Maße [Kazdin und Wilcoxon, 1976]) oder der Glaubwürdigkeit der Placebo-Bedingung während der Behandlung [Kazdin und Wilcoxon, 1976].

**Einschränkung des zu bearbeitenden Themas:** Das Kriterium der thematischen Einschränkung wurde gemäß dem Vorgehen von Baskin und Kollegen [2003] gestaltet. Dabei wird erhoben, ob es Patienten erlaubt ist, Themen und Inhalte anzusprechen, die sinngemäß mit dem Behandlungsansatz verbunden sind, z.B. ob phobische Patienten über ihre Ängste sprechen konnten oder die Gespräche auf neutrale Themen beschränkt wurden. Dieses Vorgehen entspricht dem anderer Untersuchungen, die beurteilten, ob die «Behandlung ein Besprechen von Problemen beinhaltet» [Stevens et al., 2000, p 277].

**Tab. 1.** Deskriptive Resultate der qualitativen Übersichtsarbeit (N Studien = 11; N Probanden = 477)

Studie	N	Störung	Sitzungen (SD/Kontrollbedingung)	Beschreibung der Kontrollbedingung	Art der Kontrollbedingung
Barrera und Rosen, 1977	24	Phobie (Schlangen)	(17,25/n.a.)	selbstverabreichtes Bibliotherapie-Placeboprogramm mit Selbstbelohnungskontrahierung	«attention placebo control»
Donovan und Gershman, 1979	30	n.a.	7/7	«Tachistoskop-Therapie (T-scope therapy, TST) – diese Placebobehandlung nutzte simuliertes GSR-Feedback und Exposition zu tachistoskopischen Lichtbildpräsentationen. Jeder Testperson wurde gesagt, dass tachistoskopische Präsentationen subliminal (unterbewusst wahrgenommen) wären und angstauslösende Stimuli darstellten. Es handelte sich in Wirklichkeit um Bilder von nicht vertrauten Personen und Orten, die aufgrund der kurzzeitigen Darstellungszeit (ungefähr 5 Millisekunden) nicht genau wahrgenommen werden konnten.»	empirisch abgeleitete Kontrolle
Holroyd, 1976	60	Prüfungsangst	7/7 (60 min)	Pseudotherapie/Gruppenmeditation: «Die «Meditationsübungen» waren tatsächlich von einer von McCrystals et al. [1973] entwickelten Placeboprozedur abgeleitet. Die Übungen erforderten, dass die Klienten a) sich auf unterschiedliche Empfindungen der Hände, Arme und des Körpers konzentrieren, um ihr «Körperbewusstsein zu steigern»; b) sich alltägliche Situationen vorstellen, z.B. «einen Knopf an ein Paar Hosen annähen» um «mentale Kontrolle» zu entwickeln; und c) sich an zeitgleichen Mentalkontroll- und Körperbewusstseinsübungen für «sachgerechte Meditation» zu beteiligen.»	empirisch abgeleitete Kontrolle
Kelley, 1976	40	Angst vor Dunkelheit	3/3 (30 min innerhalb von 3 Wochen)	Spielplacebo: «Testpersonen in der Placebogruppe beteiligten sich an Sitzungen, in denen das Spielen aus 15 Items zu alltäglichen Aktivitäten in der Vorschule bestand, die jedoch keinen Bezug zu Dunkelheit oder zur BAT-Situation hatten.»	«attention placebo control»
Kirsch und Henry, 1977	33	Angst vor öffentlichem Sprechen	5/5 (60 min)	«Systematische Ventilation wurde den Testpersonen beschrieben als eine Kombination aus «gerichteter Aufmerksamkeit» und «freier Assoziation» zu einer Reihe von Vorfällen in der Kindheit, die gemäß neuester klinischer Forschung Lampenfieber begünstigen. Die Liste von Situationen beinhaltete verschiedenste Items wie erste Schelte von den Eltern, erstes Haustier, erste Erinnerung an einen engen Freund usw. Die Testperson wurde angewiesen, das Ereignis zuerst zu beschreiben und dann frei zu assoziieren, indem dem Therapeuten alles erzählt wurde, was in den Sinn kam, ungeachtet dessen, wie irrelevant es erschien.»	empirisch abgeleitete Kontrolle
Levine und O'Brien, 1980	72	Prüfungsangst	6/6 (wöchentliche Sitzungen, max. 7 h)	psychoanalytisch orientierte Aufmerksamkeits-Placebogruppe: «... die Teilnehmer wurden über die Grundideen Freud'scher Psychotherapie informiert ...»	«attention placebo control»
Peck, 1977	20	Phobie (Ratten, Höhen)	max. 15/15 (3 × 30 min pro Woche)	modifiziertes Stresstoleranztraining: «Eine modifizierte Version des «Stresstraining»-Placeboprinzips von Paul [1966] wurde angewendet. Testpersonen wurden angewiesen, eine «Pille» einzunehmen (ein passives Placebo wurde benutzt), die sie entspannen würde, und anschließend für 5 Minuten im Raum allein gelassen.»	«attention placebo control»
Rosen et al., 1976	43	Phobie (Schlangen)	8/8 Wochen (2 Sitzungen pro Woche)	selbstverabreichte doppelblinde Placebokontrolle: «... selbstverabreichte doppelblinde Placebokontrolle mit der Bezeichnung «systematisches Umlernen» ... jede Person studierte ein Manual, das faktische Information zu Schlangen in 10 Kapitel gliederte.»	empirisch abgeleitete Kontrolle
Russell und Lent, 1982	52	Prüfungsangst	5/5 Sitzungen	unterbewusstes Rekonditionierungsplacebo: «Diese Placebomanipulation, an die von Lent et al. [1981] dargestellte Prozedur anknüpfend, wurde entworfen, um das Ausmaß an Verbesserung aufgrund nichtspezifischer Behandlungsfaktoren festzustellen.»	«attention placebo control»

**Tab. 1.** Fortsetzung nächste Seite

**Tab. 1.** Fortsetzung

Studie	N	Störung	Sitzungen (SD/Kontrollbedingung)	Beschreibung der Kontrollbedingung	Art der Kontrollbedingung
Slutsky und Allen, 1978	67	Angst vor öffentlichem Sprechen	5/5 (50 min)	«Tachistoskop-Therapie» (siehe oben): «Diese Placebo-manipulation replizierte die von Tori und Worell [1973] benutzten Prozeduren. Die Testpersonen erfuhren, dass ihnen tachistoskopische Szenen von Personen, die während des Sprechens gestikulieren, gezeigt würden.»	«attention placebo control»
Trent und Fournet, 1987	36	Mathematikphobie	5/5 (35 min)	Erwartungskontrolle: «... eine Behandlung bestand aus Sitzungen mit aufgenommener Musik.»	«attention placebo control»

SD = Systematische Desensibilisierung; n.a. = nicht angegeben; BAT = Behavioral Avoidance Test.

## Ergebnisse

Die deskriptiven Ergebnisse der qualitativen Analyse der 11 Studien mit insgesamt 477 Probanden sind in Tabelle 1 zusammengefasst. Die Studienteilnehmer waren größtenteils Studenten mit spezifischen Ängsten oder Phobien, d.h. Schlangenphobie, Angst vor Dunkelheit, Prüfungsangst, Angst vor öffentlichem Sprechen und Angst vor dem Lösen mathematischer Aufgaben. Dabei wurden unterschiedliche Kontrollbedingungen und Scheinbehandlungen eingesetzt, die entweder auf früheren Vorgehensweisen beruhten (z.B. Stresstoleranztraining [Peck, 1977], ursprünglich beschrieben von Paul [1969; Paul und Shannon, 1966]; Tachistoskop-Therapie [Donovan und Gershman, 1979; Slutsky und Allen, 1978], ursprünglich beschrieben von Marcia et al. [1969]; und Dissonanzsteigerung [Holroyd, 1976], ursprünglich entwickelt von McReynolds et al. [1973]) oder eigens entworfen wurden (z.B. Bibliotherapie-Placebo [Barrera und Rosen, 1977; Rosen et al., 1976], Spielplacebo [Kelley, 1976], psychoanalytisch orientierte Placebo-Kontrollbedingung [Levine und O'Brien, 1980], unterbewusstes Rekonditionierungsplacebo [Russell und Lent, 1982], Musikplacebo [Trent und Fournet, 1987], und systematische Ventilation [Kirsch und Henry, 1977]). Ausnahmslos alle Placebo-Bedingungen wurden dahingehend konzipiert, hohe Veränderungserwartungen hervorzurufen und eine hohe Glaubwürdigkeit aufzuweisen.

### Wirkungsunterschiede zwischen SD und Placebo

Insgesamt führten die eingeschlossenen Studien, die die SD jeweils mit einem Placebo-Arm verglichen, zu uneinheitlichen Ergebnissen (Tab. 2). In 5 Fällen war die SD, zumindest in Bezug auf bestimmte Parameter, signifikant wirksamer als die Placebo-Bedingung [Donovan und Gershman, 1979; Levine und O'Brien, 1980; Rosen et al., 1976; Russell und Lent, 1982; Slutsky und Allen, 1978]. Im Gegensatz dazu berichteten 6 Studien keine Unterschiede zwischen der SD und der jeweiligen Placebo-Bedingung [Barrera und Rosen, 1977; Holroyd, 1976; Kelley, 1976; Kirsch und Henry, 1977; Peck, 1977; Trent und Fournet, 1987]. Um diese uneinheitlichen Ergebnisse besser zu verstehen, wurden verschiedene Moderatorvariablen hinsichtlich der Ununterscheidbarkeitsvo-

raussetzung einbezogen, d.h. in Bezug auf strukturelle Äquivalenz, Erwartung und Glaubwürdigkeit, Einschränkung des Themas und Verblindung.

### Moderatorvariablen

**Strukturelle Äquivalenz:** In insgesamt 10 Studien wurde der Placebo-Arm als strukturell äquivalent zum SD-Arm eingeschätzt, d.h. in Bezug auf die Anzahl und Länge der Sitzungen, das Therapieformat, und die Individualisierung des Vorgehens. Bei einer Studie wurde keine Sitzungsanzahl für den Placebo-Arm angegeben, weshalb die Einstufung der strukturellen Äquivalenz unklar blieb [Barrera und Rosen, 1977; Tab. 3]. In einigen Studien wurde nicht auf die Ausbildung der Therapeuten und deren Verteilung über die Studien-Arme eingegangen [Barrera und Rosen, 1977; Kelley, 1976; Peck, 1977], wohingegen in weiteren Studien entweder Studenten in fortgeschrittenen Semestern, Berater oder erfahrene Therapeuten in beiden Versuchsbedingungen eingesetzt wurden [Donovan und Gershman, 1979; Holroyd, 1976; Kirsch und Henry, 1977; Levine und O'Brien, 1980; Russell und Lent, 1982; Slutsky und Allen, 1978]. Eine Studie setzte «Therapeuten, die mindestens ein Semester supervisierter Arbeit im Rahmen eines klinischen Praktikums aufwiesen» [Rosen et al., 1976, p. 210, Übersetzung durch die Autoren] in der SD ein, wobei die Probanden im Placebo-Arm die Behandlung eigenständig durchführten. Um eine mögliche Verzerrung der Ergebnisse durch diese Inäquivalenz zu vermeiden, wurde die zweite SD-Interventionsgruppe berücksichtigt, in der ebenfalls kein therapeutischer Kontakt stattfand (d.h. selbstverabreichte Desensibilisierung). Die Ergebnisse zwischen diesen beiden Bedingungen unterschieden sich allerdings nicht. Um die Voraussetzung der strukturellen Äquivalenz zu erfüllen, extrahierten wir dennoch die Ergebnisse der selbstverabreichten Desensibilisierung in Tabelle 2.

**Erwartungen der Probanden bezüglich therapeutischer Veränderung sowie Glaubwürdigkeit des Behandlungsrationalen:** Von den 11 eingeschlossenen Studien wiesen 7 ein positives Rational für die psychotherapeutische Placebo-Bedingung auf (Tab. 3). Dabei wurden verschiedene Begründungen verwendet, d.h. den Patienten wurde mitgeteilt, dass sie einer «Therapiegruppe» beitreten [Levine und O'Brien, 1980], dass die Behandlung bei der Überwindung

**Tab. 2.** Outcomes

Studie	Selbstbeurteilungsskala	Ergebnis(se)	Behaviorale oder physiologische Daten	Ergebnis(se)
Barrera und Rosen, 1977	FT; SNAQ; FSS; Discomfort Ratings; General Fear Change Rating	SD = Kontrolle (signifikante Intergruppen-Differenzen von Posttest- oder Follow-up-Werten wurden in keinem Studienoutcome gefunden)	BAT	SD = Kontrolle (signifikante Intergruppen-Differenzen von Posttest- oder Follow-up-Werten wurden in keinem Studienoutcome gefunden)
Donovan und Geshman, 1979	Self-Report Anxiety	SD = Kontrolle (signifikante Gruppenunterschiede wurden gefunden, $F(2,26) = 4,76$ , $p = 0,017$ ; der Vergleich SD versus TST war nicht signifikant)	GSR	SD > Kontrolle (für GSR wurden signifikante Gruppenunterschiede gefunden, $F(2,26) = 6,03$ , $p = 0,007$ ; SD-Gruppe signifikant tiefer auf GSR verglichen mit der TST-Gruppe ( $p < 0,01$ ) und der PPC-Gruppe ( $p < 0,05$ ))
			RC	SD = Kontrolle (für RC wurden keine Unterschiede zwischen den Gruppen gefunden, F-Werte < 1)
			FP	SD = Kontrolle (für FP wurden keine Unterschiede zwischen den Gruppen gefunden, F-Werte < 1)
Holroyd, 1976	Achievement Anxiety Test	SD = Kontrolle (die drei Behandlungsprozeduren und die Pseudotherapie-Kontrolle bewirkten vergleichbare Reduktionen hemmender Angst; zudem zeigte keine der Gruppen wesentliche Veränderungen bei der unterstützenden Angst)		**
	Analogue Testing Situation (Digit Symbol)	SD = Kontrolle (systematische Desensibilisierung, Pseudotherapie und Warteliste-Kontrollgruppe unterschieden sich nicht signifikant voneinander)		
	Analogue Testing Situation (State Anxiety)	SD = Kontrolle (die Posttest-«State Anxiety»-Werte der SD-Gruppe, der kombinierten Gruppe und der Pseudotherapie-Gruppe unterschieden sich nicht signifikant voneinander)		
	Analogue Testing Situation (Anxiety Differential)	SD = Kontrolle (die kombinierte Gruppe, die SD-Gruppe und die Pseudotherapie-Gruppe unterschieden sich von der Warteliste-Kontrollgruppe auf einem Level, das sich der Signifikanz annäherte ( $p < 0,10$ ))		
Kelley, 1976	FT	SD = Kontrolle (die Behandlungsgruppen unterschieden sich nicht signifikant hinsichtlich der Totalanzahl von Furchtsignalen ( $F(2,21) = 0,6711$ , $p < 0,05$ ))	BAT	SD = Kontrolle (Vergleiche aller fünf Behandlungsbedingungen waren nicht signifikant)
Kirsch und Henry, 1977	PRCS	SD = Kontrolle (SD, operante Desensibilisierung, systematische Ventilation: $F(2,88) = 2,82$ )	BC	SD = Kontrolle (SD, operante Desensibilisierung, systematische Ventilation: $F(2,88) = 0,35$ )
	AD	SD = Kontrolle (SD, operante Desensibilisierung, systematische Ventilation: $F(2,88) = 0,07$ )		
Levine und O'Brien, 1980	Test Anxiety Questionnaire	SD > Kontrolle (die fünf Behandlungen unterschieden sich signifikant ( $F(4,61) = 4,42$ , $p < 0,01$ ); nur die SD-Gruppe erklärte einen signifikanten Varianzanteil (9%) der Postbehandlungswerte)		**

**Tab. 2.** Fortsetzung nächste Seite

**Tab. 2.** Fortsetzung

Studie	Selbstbeurteilungsskala	Ergebnis(se)	Behaviorale oder physiologische Daten	Ergebnis(se)
Peck, 1977	FT	SD = Kontrolle (Gruppen unterschieden sich nicht, SD berichtete die meiste Angst)	BAT BC	SD = Kontrolle (Gruppen unterschieden sich nicht) SD = Kontrolle (Gruppen unterschieden sich nicht)
Rosen et al., 1976	SNAQ FSS	SD = Kontrolle (SD: $t(7) = 3,90$ , $p < 0,001$ ; Kontrolle: $t(9) = 3,41$ , $p < 0,01$ )* SD = Kontrolle (SD: $t(7) = 3,05$ , $p < 0,05$ ; Kontrolle: $t(9) = 2,71$ , $p < 0,01$ )*	BAT	SD > Kontrolle (SD: $t(7) = 3,17$ , $p < 0,05$ ; Kontrolle: nicht signifikant)*
Russell und Lent, 1982	TAS S-R Exam (R-R Inventory of Anxiousness, Exam Form) TMAS all Outcomes	SD = Kontrolle (SD: $t = 3,86$ , $p < 0,01$ ; Kontrolle: $t = 1,89$ , $p < 0,05$ )* SD = Kontrolle (SD: $t = 3,93$ , $p < 0,01$ ; Kontrolle: $t = 2,93$ , $p < 0,01$ )* SD > Kontrolle (SD: $t = 2,75$ , $p < 0,01$ ; Kontrolle: $t = 1,46$ , nicht signifikant)* SD = Kontrolle (keine signifikanten F-Werte wurden festgestellt, was darauf hinweist, dass keine Experimentalbedingung effektiver war als eine andere im Bewirken von Veränderungen)	-	-
Slutsky und Allen, 1978	AD Trait Anxiety Questionnaire Form PRCS	SD > Kontrolle (SD: $t = 3,38$ , $p < 0,01$ ; Kontrolle: $t = 2,18$ , nicht signifikant)* SD > Kontrolle (SD: $t = 3,47$ , $p < 0,01$ ; Kontrolle: $t = -0,17$ , nicht signifikant)* SD = Kontrolle (SD: $t = 3,33$ , $p < 0,01$ ; Kontrolle: $t = 3,84$ , $p < 0,01$ ) SD = Kontrolle (SD: $t = 6,62$ , $p < 0,001$ ; Kontrolle: $t = 2,49$ , $p < 0,01$ )	PBC PR PS	SD = Kontrolle (SD: $t = 4,48$ , $p < 0,01$ ; Kontrolle: $t = 3,23$ , $p < 0,01$ )* SD > Kontrolle (SD: $t = 2,10$ , $p < 0,01$ ; Kontrolle: $t = -0,68$ , nicht signifikant)* SD = Kontrolle (SD: $t = 0,87$ , nicht signifikant; Kontrolle: $t = -0,63$ , nicht signifikant)*
Trent und Fournet, 1987	MARS RMAS Mathematics Performance	SD = Kontrolle ( $F(2,30) = 0,86$ , $p > 0,05$ ) SD = Kontrolle ( $F(2,30) = 0,58$ , $p > 0,05$ ) SD = Kontrolle	-	-

SD = Systematische Desensibilisierung; n.a. = nicht angegeben; AD = Anxiety Differential; BAT = Behavioral Avoidance Test; BC = Behavioral Checklist; FP = Finger Pulse; FSS = Fear Survey Schedule; FT = Fear Thermometer; GSR = Galvanic Skin Response; MARS = Mathematics Anxiety Rating Scale; PBC = Paul's Behavioral Checklist; PPC = Pretest-Posttest Control; PR = Pulse Rate; PRCS = Personal Report of Confidence as a Speaker; PS = Palmar Sweat; RC = Respiration Cycle; RMAS = Revised Mathematics Attitude Scale; SNAQ = Snake Attitude Questionnaire; TAS = Test Anxiety Scale; TMAS = Taylor Manifest Anxiety Scale; TST = T-Scope Therapy.

\*Durchschnittliche Veränderung von Baseline zu Endpunkt innerhalb jeder Gruppe (Intragruppenvergleiche, da keine Intergruppenvergleiche berichtet wurden);

\*\*Notendurchschnitt (d.h. akademische Leistung) berichtet, aber nicht als behaviorales Maß aufgelistet, da Variable von mehreren Faktoren abhängig ist.

von Ängsten helfen würde [Kirsch und Henry, 1977; Peck, 1977; Rosen et al., 1976; Slutsky und Allen, 1978], dass «sie durch das Erlernen der Stressbewältigung in der einen Situation lernen würden, in der Gegenwart des phobischen Stimulus nicht mehr ängstlich zu sein» [Peck, 1977, p. 139, Übersetzung der Autoren], oder den Teilnehmern wurde eine Erklärung für den Ursprung ihrer Ängste und die Art und Weise, wie die Behandlung diese verringern würde, gegeben ([Russell und Lent, 1982; Trent und Fournet, 1987], zum Beispiel: «Zu Beginn der ersten Sitzung wurde den Studenten gesagt, dass der Ursprung der Angst vor dem Lösen mathematischer Aufgaben im Unterbewusstsein liegt. Daher könnten unbewusste Muster der Mathematikphobie durch korrigierende und subliminale Instruktionen über Musiktonbänder verändert werden» [Trent und Fournet, 1987, p. 140, Übersetzung durch die

Autoren]). Sieben Studien verwendeten eine «attention placebo control»-Strategie. In 4 der 11 Studien wurden die Erwartung eines positiven Behandlungsergebnisses und/oder die Glaubwürdigkeit der Interventionsbedingung erfasst, d.h. empirisch abgeleitete Kontrollgruppen lagen vor (Tab. 1). In 3 dieser 4 Studien wurden keine signifikanten Unterschiede bezüglich der Glaubwürdigkeit zwischen den Gruppen beobachtet [Donovan und Gershman, 1979; Holroyd, 1976; Kirsch und Henry, 1977; Holroyd, 1976], und in der Studie, in der nur die initialen Erwartungen einer positiven therapeutischen Veränderung erfasst wurden, wurden entsprechend keine signifikanten Gruppenunterschiede beobachtet [Rosen et al., 1976]. Es zeigte sich, dass Studien, die nur eine Placebo-Bedingung ohne Erfassung der Veränderungserwartung und/oder Glaubwürdigkeit beinhalteten, uneinheitliche Ergebnisse auf-

**Tab. 3.** Moderatoren (Tabelle gruppiert nach spezifischen Effekten des Outcomes)

Studie	Spezifische Effekte des Outcomes <sup>a,b</sup>	Strukturelle Äquivalenz	Erwartung		Begründung	Bewertung	Glaubwürdigkeitsbewertung	Thematische Einschränkung		Verblindung
			Positive Begründung	Begründung				Themen	Themenbeschreibung	
Barreira und Rosen, 1977	nein	unklar	nein	n.a.	/	/	/	neutral	faktische Information	n.a.
Holroyd, 1976	nein	ja	nein	n.a.	nein	ja	ja	neutral	Imaginieren alltäglicher Situationen	verblindete Untersucher
Kelley, 1976	nein	ja	nein	n.a.	/	/	/	neutral	kein Bezug zum phobischen Stimulus (Dunkelheit)	verblindete Untersucher
Kirsch und Henry, 1977	nein	ja	ja	«(...) der Therapeut las eine Behandlungsbeschreibung und -begründung vor und beantwortete alle Fragen, die die Testperson bezüglich der Behandlung, die er oder sie erhalten würde, hatte.»	nein	ja	ja	der Behandlunglung sinngemäß	Reihe von Vorfällen in der Kindheit mit Bezug zu Lampenfieber; hoher Grad an Teilnehmerbeteiligung	verblindete Untersucher
Peck, 1977	nein	ja	ja	«Den Testpersonen wurde gesagt, dass sie durch das Erlernen der Stressbewältigung in der einen Situation lernen würden, in der Gegenwart des phobischen Stimulus nicht mehr ängstlich zu sein.»	/	/	/	neutral	Zahlen ablesen	n.a.
Trent und Fournet, 1987	nein	ja	ja	«Zu Beginn der ersten Sitzung wurde den Studenten gesagt, dass der Ursprung der Angst vor dem Lösen mathematischer Aufgaben im Unterbewusstsein liegt. Daher könnten unbewusste Muster der Mathematikphobie durch korrigierende und subliminale Instruktionen über Musiktonbänder verändert werden.»	/	/	/	neutral	keine subliminalen Instruktionen	n.a.
Donovan und Gershman, 1979	gemischt	ja	nein	n.v.	nein	ja	ja	neutral	Bilder nicht vertrauter Personen und Orte	n.a.
Rosen et al., 1976	gemischt	ja	ja	«Die wesentliche Begründung des Behandlungsprogramms war, dass Personen ihre Ängste erheblich reduzieren können, indem sie fehlerhafte Auffassungen zum gefürchteten Objekt durch korrekte Information ersetzen.»	ja	nein	nein	neutral	faktische Information	verblindete Untersucher

**Tab. 2.** Fortsetzung nächste Seite



Tab. 3. Fortsetzung

Studie	Spezifische Effekte des Outcome <sup>a,b</sup>	Strukturelle Äquivalenz	Erwartung		Begründung	Bewertung	Glaubwürdigkeitsbewertung	Thematische Einschränkung		Verblindung
			Positive Begründung					Themen	Themenbeschreibung	
Russell und Lent, 1982	gemischt*	ja	ja	/	«Die Begründung erklärte, dass, da Prüfungsgänge von unterbewussten Gedanken und Gefühlen herrühren, die Behandlung auf eine Rekonditionierung des Unterbewusstseins abzielt, indem Anti-Angst-Nachrichten auf subliminaler Ebenen präsentiert werden. Tatsächlich bestanden diese Nachrichten aus unsinnigen Silben, die wie Worte erschienen, wenn sie mit hoher Geschwindigkeit in einem Tachistoskop aufleuchteten.»	/	/	neutral	unsinnige Silben	n.a.

N.a. = nicht angegeben.  
<sup>a</sup> Für Details, siehe Tabelle 2.  
<sup>b</sup> Kein signifikanter Gruppenunterschied in selbstberichteten und verhaltens- bzw. physiologischen Angaben; gemischt = einige signifikante Gruppenunterschiede in entweder selbstberichteten oder physiologischen Angaben; ja = signifikante Gruppenunterschiede in allen berichteten Bereichen.  
<sup>\*</sup> Signifikante Veränderung von der Baseline zum Endpunkt einiger Parameter in der Versuchsgruppe, jedoch nicht in der Kontrollgruppe. Gruppe × Zeit-Interaktionen wurden nicht berichtet.

wiesen, wohingegen Studien mit Erfassung dieser Variablen entweder nur Unterschiede auf behavioralen und physiologischen Skalen [Donovan und Gershman, 1979; Rosen et al., 1976] oder überhaupt keine Gruppenunterschiede berichteten [Holroyd, 1976; Kirsch und Henry, 1977]. Entsprechend finden sich auf den Selbstbeurteilungsskalen dieser Studien keine Gruppendifferenzen zwischen der SD und dem jeweiligen Placebo-Arm (Tab. 2).

*Einschränkung des zu bearbeitenden Themas:* Hinsichtlich des Kriteriums der thematischen Einschränkung wurde das Behandlungsrational und die tatsächliche Aufgabe innerhalb des Placebo-Arms eingehend analysiert (Tab. 3). Dabei konnten verschiedene Placebo-Bedingungen unterschieden werden:

- Täuschung der Patienten, d.h. den Patienten wurde mitgeteilt, dass ihnen gewisse subliminale, stressende und angstauslösende Lärmgeräusche oder Bilder bei der Überwindung ihrer Ängste helfen würden, obwohl tatsächlich keine bzw. nur unerhebliche subliminale Signale verwendet wurden [Donovan und Gershman, 1979; Peck, 1977; Russell und Lent, 1982; Slutsky und Allen, 1978; Trent und Fournet, 1987].

- Aufgaben ohne direkten Bezug zu Ängsten und ohne die Möglichkeit, angstrelevante Themen zu besprechen, oder angstrelevante emotionale Erfahrungen zu machen, wie beispielsweise Spielplacebo [Kelley, 1976], Meditationsübungen [Holroyd, 1976] oder das Sammeln faktischer Informationen zum angstauslösenden Objekt [Barrera und Rosen, 1977; Rosen et al., 1976].

- Thematische Kontrolle: In der Evaluation einer psychoanalytisch orientierten Behandlung wurde eine Placebo-Bedingung verwendet, bei der Konfrontationen zwischen Teilnehmern verhindert, intensive emotionale Reaktionen sowie emotionale oder persönliche Äußerungen vermieden und persönliche Interpretationen unterlassen wurden [Levine und O'Brien, 1980, p. 825]. Damit waren der Gesprächsgegenstand sowie die emotionale Reaktion der Patienten nicht nur eingeschränkt, sondern diese wurden sogar kontrolliert. In diesem Fall zeigten sich im selbstbeurteilten Prüfungsangst-Fragebogen signifikante Unterschiede zwischen den Gruppen.

- Emotionale Aktivierung: In einer weiteren Studie wurde eine entgegengesetzte Strategie angewendet: In der Placebo-Bedingung «Systematische Ventilation» hatte das Thema starken Angstbezug, entsprechend wurde die Prozedur «den Versuchspersonen als «gerichtete Aufmerksamkeit» und «freie Assoziation» zu einer Reihe von Vorfällen in der Kindheit beschrieben, die «gemäß neuester klinischer Forschung Lampenfieber begünstigen» [Kirsch und Henry, 1977, p. 1054, Übersetzung durch die Autoren]. Zudem waren die Teilnehmer emotional engagiert und die «Therapeuten berichteten einen hohen Grad an Teilnehmerbeteiligung in diesem Verfahren, zeitweise sogar begleitet von beträchtlichen Affektäußerungen» [Kirsch und Henry, 1977, p. 1054, Übersetzung durch die Autoren]. Die Ergebnisse dieser Studie zeigten keine Unterschiede zwischen der Placebo-Bedingung und der SD in den eingesetzten Selbstbeurteilungsverfahren und behavioralen Parametern.

*Verblindung:* Die Maßnahmen der Verblindung innerhalb der Studien und gegenüber den Therapeuten und Beurteilern sind in Tabelle 3 dargestellt. In keiner der eingeschlossenen Studien wurde

eine Verblindung der Therapeuten angestrebt oder umgesetzt. In 4 Studien wurden die Beurteiler hinsichtlich der Gruppenzuweisung der Teilnehmer verblindet [Holroyd, 1976; Kirsch und Henry, 1977; Rosen et al., 1976; Kelley, 1976].

## Diskussion

Das Ziel dieser qualitativen Übersichtsarbeit war die systematische Evaluierung von Studien, die die Spezifität der SD durch den Vergleich mit unterschiedlichen Placebo-Bedingungen nachweisen wollten und die nach der einflussreichen Übersichtsarbeit von Kazdin und Wilcoxon [1976] veröffentlicht wurden. In Übereinstimmung mit Kazdin und Wilcoxon zeigten sich zunächst uneinheitliche Ergebnisse, d.h. einige Studien wiesen eine Überlegenheit der SD über die psychotherapeutische Placebo-Behandlung nach, wohingegen andere Studien dies nicht berichteten. Diese widersprüchlichen Ergebnisse wurden in der Folge unter Berücksichtigung der Voraussetzungen der Ununterscheidbarkeit und der Verblindung analysiert.

Dabei fällt erstens auf, dass bei der Mehrheit der Studien die Kriterien der strukturellen Äquivalenz, d.h. die Anzahl und Länge der Sitzungen, das Therapieformat, die Individualisierung der Behandlung und die Ausbildung der Therapeuten, für die SD und die psychotherapeutische Placebo-Behandlung – mit Ausnahme einer Studie [Barrera und Rosen, 1977] – identisch waren. Somit kann die Anforderung der strukturellen Äquivalenz in den eingeschlossenen Studien als erfüllt betrachtet werden. Es muss jedoch berücksichtigt werden, dass diese Anforderung im Fall der SD leicht erfüllbar ist, weil es sich hierbei um eine stark strukturierte Intervention handelt [Wolpe, 1968]. Da dieses Kriterium nicht zur Ausdifferenzierung der uneinheitlichen Ergebnisse beitrug, kann die strukturelle Äquivalenz als notwendiges, aber nicht hinreichendes Kriterium für die Ununterscheidbarkeit angesehen werden.

Zweitens zeigten sich zwischen Studien, die nur eine Placebo-Bedingung durchführten (d.h. «attention placebo control»-Strategien), und solchen, in denen zusätzlich die Erwartungen der Probanden bezüglich therapeutischer Veränderungen oder die Glaubwürdigkeit des Behandlungsrationalen erfasst wurden (d.h. empirisch abgeleitete Kontrollstrategien), Unterschiede in den Ergebnissen. Während Studien ohne Berücksichtigung der Erwartung oder Glaubwürdigkeit der Placebo-Bedingung weiterhin uneinheitliche Resultate aufwiesen, konnten in den Studien, die die Glaubwürdigkeit des Behandlungsrationalen erfassten und über die untersuchten Behandlungsbedingungen hinweg analog hielten, keine Unterschiede zwischen der SD und den psychotherapeutischen Placebos auf den Selbstbeurteilungsskalen nachgewiesen werden. Zur Begründung kann aufgeführt werden, dass entweder das Behandlungsrational dann tatsächlich glaubhafter war oder aber dass die Bewertungen reaktiv sind, d.h. dass der Behandlungsfortschritt an sich die Bewertungen beeinflusst [Bernstein, 1973; Bernstein und Paul, 1971; O'Leary und Borkovec, 1978]. Dementsprechend «führt die Anwendung von Glaubwürdigkeits- und Erwartungsskalen vor Behandlungsbeginn zu Antworten, die einen

vernachlässigbaren Informationsgehalt aufweisen und darum unzuverlässig sind» [Baskin et al., 2003, p 974, Übersetzung durch die Autoren]. Obwohl die Glaubwürdigkeit des Behandlungsrationalen notwendig und wichtig ist, reicht sie folglich nicht aus, um die Ununterscheidbarkeit in psychotherapeutischen Vergleichsstudien sicherzustellen. Entsprechend kann die Glaubwürdigkeit als einflussreicher, jedoch lediglich proximaler Indikator angesehen werden. Diese Ergebnisse stimmen dabei mehrheitlich mit denen von Kazdin und Wilcoxon [1976] überein. Jedoch ist zu berücksichtigen, dass in der Arbeit von Kazdin und Wilcoxon auch sogenannte Dismantling-Studien und Studien von Therapievergleichen einbezogen wurden, wohingegen diese Studiendesigns in der hier vorliegenden Analyse ausgeschlossen wurden.

Ein weiteres Ergebnis ist, dass die thematische Einschränkung im Placebo-Arm starke Auswirkungen auf die Ergebnisse der jeweiligen Studie hat. In der Mehrheit der Studien durften die Patienten lediglich neutrale Themen einbringen, und entsprechend waren die zu erfüllenden Aufgaben bedeutungslos oder hatten keinen Bezug zum eigentlichen Behandlungsgegenstand, wie z.B. Angst vor der Dunkelheit. Das Kriterium der thematischen Einschränkung gemäß Baskin et al. [2003] wurde dabei durch die Berücksichtigung der emotionalen Erfahrung erweitert. Entsprechend könnte das Sammeln faktischer Informationen zu einem gefürchteten Objekt thematisch uneingeschränkt sein, jedoch zugleich emotionale Erfahrungen unterbinden (z.B. Bibliothherapie-Placebo oder systematisches Umlernen [Barrera und Rosen, 1977; Rosen et al., 1976]). Während eine der eingeschlossenen Studien sogar aktiv der emotionalen Erfahrung entgegenwirkte [Levine und O'Brien, 1980], ermöglichte das Vorgehen in einer weiteren Untersuchung dies ausdrücklich, was dann auch mit bedeutsamen emotionalen Reaktionen einherging [Kirsch und Henry, 1977]. In dieser Studie bewerteten die Probanden nicht nur das Placebo-Verfahren als genauso glaubwürdig wie die SD, sondern zeigten auch Verbesserungen bezüglich selbstbeurteilter sowie behavioraler Maße. Die Bedeutung der emotionalen Erfahrung für die Glaubwürdigkeit einer psychotherapeutischen Placebo-Behandlung zeigt sich auch in einer Studie, in der das Stresstoleranztraining von Paul angewendet wurde [Paul 1969; Paul und Shannon, 1966]. Dort wurde vermerkt, dass ein Teilnehmer «lachte, als ihm mitgeteilt wurde, dass er mithilfe dieser Behandlung seine Höhenangst überwinden würde. Er sagte, dass es doch nicht die Zahlen seien, die ihn ängstigen würden» [Peck, 1977, p 143, Übersetzung durch die Autoren]. Vergleichbare Schilderungen finden sich auch in einer der letzten Untersuchungen zur SD, die ein Dismantling-Verfahren einsetzte und in der ein «Modified Avoidance Response Placebo» (MARP) mit einer Pseudotherapie verglichen wurde. Diese beiden Kontrollstrategien unterschieden sich nur bezüglich der verwendeten Bilder; während in der MARP-Bedingung vier Bilder aus der Hierarchie der Zahnbehandlungsphobie eingesetzt wurden, erhielten die Teilnehmer der Pseudotherapie die Anweisung, sich «vier bedeutungslose und affektiv neutrale Bilder» vorzustellen [Landau et al., 1984, p 118, Übersetzung durch die Autoren]. Es überrascht kaum, dass die Pseudotherapie gegenüber dem MARP als signifikant weniger glaubwürdig eingestuft wurde und

zusätzlich auch signifikant weniger Verbesserung als die Vergleichspsychotherapie bewirkte.

Schlussfolgernd kann die Hypothese aufgestellt werden, dass die emotionale Erfahrung des Patienten während einer psychotherapeutischen Placebo-Behandlung die Grundvoraussetzung für die Glaubwürdigkeit eben dieser Behandlung ist. Die objektiven strukturellen Kriterien hinsichtlich der Ununterscheidbarkeit werden entsprechend durch die emotionale Erfahrung [Castonguay et al., 1996] aufseiten des Patienten vervollständigt. Die emotionale Erfahrung könnte also das Schlüsselement für die erfolgreiche Wirksamkeit eines psychotherapeutischen – oder sogar jeglicher [LeBlanc, 2014] – Placebos sein.

Wenig überraschend ist das vollkommene Fehlen der Verblindung von Therapeuten in den untersuchten Studien. Es gilt dabei zu bedenken, dass die Verblindung in Psychotherapiestudien nahezu unmöglich ist [Kirsch, 2005]. Des Weiteren berichteten nur 4 Studien, dass die Beurteiler verblindet wurden, wobei diese Bewertungen auf behavioralen Skalen abgeben mussten. Das vergleichsweise bescheidene Resultat steht in Einklang mit der generellen Kritik, dass «Forscher und Zeitschriften es versäumen, eine erfolgreiche Verblindung zu dokumentieren» [Fergusson et al., 2004, p 2, Übersetzung durch die Autoren].

Bevor die Hauptschlussfolgerungen dargestellt werden, soll auf eine nennenswerte Limitation der gewählten Methode einer qualitativen Übersichtsarbeit hingewiesen werden. Wie die meisten interpretativen Verfahren besteht die Gefahr eines möglichen Mangels an Objektivität, da keine systematischen und quantitativen Arbeitsschritte vorliegen [Dixon-Woods et al., 2005]. Es wurde versucht, diese Fehlerquelle zu umgehen, indem sowohl eine systematische Literatursuche durchgeführt wurde als auch strenge und präzise Einschlusskriterien verwendet wurden. Zudem scheint der methodische Zugang hier gerechtfertigt, da uneinheitliche Studienoutcomes untersucht wurden und der Fokus eindeutig auf der qualitativen Beschaffenheit der Placebo-Kontrollbedingungen lag.

Die wichtigste Schlussfolgerung aus den oben aufgeführten Ergebnissen ist, dass die Beurteilung der Spezifität eines bestimmten psychotherapeutischen Verfahrens – hier exemplarisch die SD – wesentlich von der Operationalisierung der Placebo-Bedingung abhängt. Hierbei scheint die Einschränkung des zu bearbeitenden Themas und in der Folge die Unterbindung der emotionalen Erfahrung des Patienten innerhalb der Kontrollbedingung eine Voraussetzung dafür zu sein, spezifische Effekte der SD zu finden. Je nachdem, welche Bestandteile der SD als zufällig oder charakteristisch definiert werden, liegen unterschiedliche Implikationen unserer Schlussfolgerung vor [Grünbaum, 1981, 1986].

Einerseits kann die emotionale Erfahrung des Patienten während der SD als zufälliger Behandlungsbestandteil angesehen werden. Die Bedeutung der emotionalen Erfahrung wird in verschiedenen Modellen hervorgehoben. Gemäß Frank und Frank [1993] stellen die Emotionen in der Therapie eines der sechs Elemente dar, die allen therapeutischen Verfahren gemein sind. Gleichermaßen betrachten Goldfried [1980] und auch Castonguay [1993; Castonguay et al., 1996] die emotionale Aktivierung bzw. die emotionale Erfahrung des Patienten als Grundvoraussetzung für eine po-

sitive therapeutische Veränderung. Wird also die emotionale Aktivierung als zufälliger Behandlungsbestandteil betrachtet, dann sollte in Psychotherapiestudien die emotionale Erfahrung des Patienten im Placebo-Arm gewährleistet sein. Entsprechend sollte die Placebo-Bedingung so angepasst werden, dass diese nicht nur über ein glaubwürdiges Behandlungsrational verfügt, sondern auch die affektive Erfahrung der Patienten ermöglicht. Entsprechend könnte «(...) das Placebo von heute (...) die Therapie von morgen sein» [Kirsch und Henry, 1977, p 1053]. Auch von einer konkreten Umsetzung der «Placebo-Therapie» [Fish, 1973] zu sprechen, liegt hier nicht fern.

Andererseits kann die emotionale Erfahrung auch als charakteristischer Behandlungsbestandteil beschrieben werden, wobei dann die Behandlungstheorie der SD revidiert werden sollte. Die Integration der emotionalen Erfahrung in die Theorie der SD wäre jedoch mit erheblichen Hindernissen verbunden, da die Berücksichtigung der emotionalen Erfahrung des Patienten mit der ursprünglich vorgeschlagenen Behandlungstheorie der SD nicht unbedingt vereinbar ist. Wolpe [1954, p 210, Übersetzung durch die Autoren] erklärte hierzu: «Es ist immer vorzuziehen, eher zu langsam als zu schnell vorzugehen. Während früher Experimente mit dieser Methode habe ich bei zwei Patienten ernsthafte Rückschläge durch das voreilige Präsentieren von Stimuli mit hohem angstausslösenden Potenzial erzeugt.» Oder, in anderen Worten: «Laut Wolpe [1958] können intensivere Erfahrungen im Pawlow'schen Konstrukt der *transmarginalen Hemmung* oder in einem Zustand der völligen Stilllegung des Organismus resultieren» [Barlow, 2010, p 13, Übersetzung durch die Autoren]. Die aufgeführten Überlegungen dienen jedoch hauptsächlich der logischen Vollständigkeit, da die postulierten Mechanismen der SD seit den 1970er-Jahren reformuliert wurden. Zweifellos kann es als Stärke einer empirisch orientierten Therapie betrachtet werden, dass die SD als Methode der Wahl in der Behandlung spezifischer Phobien abgelöst und durch neue Konzepte wie Exposition [Fiegenbaum, 1988] und Erwartungsverletzung [Craske, 2015; Craske et al., 2014] ersetzt wurde.

In Anbetracht der vorangehenden Diskussion sollen die konzeptuellen und theoretischen Implikationen psychotherapeutischer Placebos besprochen werden. Erstens ist die Wirkung und Wirksamkeit einer bestimmten Behandlung kein direkter Nachweis für die Validität der zugrunde liegenden Definition der charakteristischen Behandlungsbestandteile. Zweitens ist es unter Berücksichtigung von Grünbaum [1981, 1986] entscheidend festzulegen, für welche Faktoren eine psychotherapeutische Placebo-Behandlung kontrollieren soll, und zu definieren, welches die theoriebasierten charakteristischen Behandlungsbestandteile der untersuchten Verum-Intervention sind. So wie auch von der vorliegenden Analyse ausgeschlossene Studiendesigns, z.B. Komponentenstudien, mit methodischen Problemen einhergehen [Bell et al., 2013; Kazdin und Whitley, 2003], birgt auch die randomisierte Placebo-kontrollierte Studie – als Goldstandard der klinischen Forschung – ihre eigenen Herausforderungen [Wager und Atlas, 2015]. Dabei ist die Ununterscheidbarkeit der untersuchten Behandlungsbedingungen kein fixer Wert, sondern vielmehr eine Funktion von struktureller Äquivalenz, Erwartung und Glaubwürdigkeit. Die Validität rando-

misiertes Placebo-kontrollierter Psychotherapiestudien ist dabei abhängig von der Gewährleistung der Ununterscheidbarkeit und Verblindung beider Studien-Arme, wobei der Placebo-Arm dann idealerweise auch keine charakteristischen Behandlungsbestandteile enthält [Grünbaum, 1981, 1986]. Entsprechend der Maßgabe, dass nicht nur «ein einziger Mechanismus des Placebo-Effekts und nicht ein einziger Placebo-Effekt (existiert) – sondern viele» [Benedetti, 2009, p 38, Übersetzung durch die Autoren], kann es auch nicht eine einzige, sondern es muss – je nach Behandlungstheorie – unterschiedliche psychotherapeutische Placebo-Bedingungen geben. Dementsprechend sind die vielen verschiedenen Operationalisierungen der Placebo-Bedingung in der medizinischen Forschung (z.B. Zuckerpillen, Kochsalzinjektionen und Scheinoperationen) hinsichtlich Aussehen, Geschmack, Klang und Empfindung voneinander verschieden, aber zur jeweiligen korrespondierenden Verum-Intervention identisch gehalten. Analog kann in der Psychotherapieforschung von einem fließenden Übergang von Placebo-Kontrollen zu Dismantling-Designs ausgegangen werden, oder anders formuliert: Da es keine qualitativen Unterschiede zwischen Placebos und Psychotherapie gibt (d.h. beide basieren auf psychologischen Prozessen), können Placebo-kontrollierte Studien als Extremvarianten von Dismantling-Studien betrachtet werden. Die unterschiedliche Gewichtung einzelner Komponenten in der vorliegenden qualitativen Übersichtsarbeit gleicht demnach der ge-

stufen Vorgehensweise elaborierter Placebo-Dismantling-Studien [Kaptchuk et al., 2008].

Unter welchen Bedingungen also können psychotherapeutische Placebo-Bedingungen als angemessene Kontrollbedingungen in der Untersuchung der Spezifität psychotherapeutischer Methoden erachtet werden? Aus der vorangehend dargestellten Perspektive sind alle Bestrebungen, die Ununterscheidbarkeit der zu vergleichenden Behandlungsbedingungen zu fördern, ein Indikator für die Validität der jeweiligen Untersuchung. Darüber hinaus sollten zukünftige Untersuchungen in der Psychotherapieforschung durchdachte und innovative Experimentaldesigns aus der Placebo-Forschung übernehmen, wie z.B. sogenannte «blinded balanced» [Lund et al., 2014] und «open/hidden» Studiendesigns [Colloca et al., 2004].

## Disclosure Statement

Die Autoren erklären, dass keine Interessenskonflikte in Bezug auf das bestehende Manuskript vorliegen.

---

*Übersetzt von Cosima Locher,  
Sebastian Hasler und  
Prof. Dr. Jens Gaab*

## Literatur

- Barlow DH: Negative effects from psychological treatments: a perspective. *Am Psychol* 2010;65:13–20.
- Barrera M, Rosen GM: Detrimental effects of a self-reward contracting program on subjects' involvement in self-administered desensitization. *J Consult Clin Psychol* 1977;45:1180–1181.
- Baskin TW, Tierney SC, Minami T, Wampold BE: Establishing specificity in psychotherapy: a meta-analysis of structural equivalence of placebo controls. *J Consult Clin Psychol* 2003;71:973–979.
- Beecher HK: The powerful placebo. *J Am Med Assoc* 1955; 159:1602–1606.
- Bell EC, Marcus DK, Goodlad JK: Are the parts as good as the whole? A meta-analysis of component treatment studies. *J Consult Clin Psychol* 2013;81:722–736.
- Benedetti F: *Placebo Effects: Understanding the Mechanisms in Health and Disease*. New York, Oxford University Press, 2009.
- Bernstein DA: Behavioral fear assessment: anxiety or artifact? in Adams H, Unikel P (eds): *Issues and Trends in Behavior Therapy*. Springfield, Charles C Thomas, 1973, pp 225–268.
- Bernstein DA, Paul GL: Some comments on therapy analogue research with small animal 'phobias'. *J Behav Ther Exp Psy* 1971;2:225–237.
- Borkovec TD, Nau SD: Credibility of analogue therapy rationales. *J Behav Ther Exp Psy* 1972;3:257–260.
- Borkovec TD, Sibrava NJ: Problems with the use of placebo conditions in psychotherapy research, suggested alternatives, and some strategies for the pursuit of the placebo phenomenon. *J Clin Psychol* 2005;61:805–818.
- Bowers TG, Clum GA: Relative contribution of specific and nonspecific treatment effects: meta-analysis of placebo-controlled behavior therapy research. *Psychol Bull* 1988;103:315–323.
- Castonguay LG: 'Common factors' and 'nonspecific variables': clarification of the two concepts and recommendations for research. *J Psychother Integr* 1993;3:267–286.
- Castonguay LG, Goldfried MR, Wisner S, Raue PJ, Hayes AM: Predicting the effect of cognitive therapy for depression: a study of unique and common factors. *J Consult Clin Psychol* 1996;64:497–504.
- Chambless DL, Baker MJ, Baucom DH, et al.: Update on empirically validated therapies, II. *Clin Psychol* 1998; 51:3–16.
- Colloca L, Lopiano L, Lanotte M, Benedetti F: Overt versus covert treatment for pain, anxiety, and Parkinson's disease. *Lancet Neurol* 2004;3:679–684.
- Colloca L, Miller FG: Role of expectations in health. *Curr Opin Psychiatry* 2011;24:149–155.
- Craske MG: Optimizing exposure therapy for anxiety disorders: an inhibitory learning and inhibitory regulation approach. *Verhaltenstherapie* 2015;25:134–143.
- Craske MG, Treanor M, Conway CC, Zbozinek T, Vervliet B: Maximizing exposure therapy: an inhibitory learning approach. *Behav Res Ther* 2014;58:10–23.
- D'Zurilla TJ, Wilson GT, Nelson R: A preliminary study of the effectiveness of graduated prolonged exposure in the treatment of irrational fear. *Behav Ther* 1973;4: 672–685.
- Dixon-Woods M, Agarwal S, Jones D, Young B, Sutton A: Synthesising qualitative and quantitative evidence: a review of possible methods. *J Health Serv Res Policy* 2005;10:45–53.
- Donovan TR, Gershman L: Experimental anxiety reduction: systematic-desensitization versus a false-feedback expectancy manipulation. *J Behav Ther Exp Psy* 1979; 10:173–179.
- Fergusson D, Glass KC, Waring D, Shapiro S: Turning a blind eye: the success of blinding reported in a random sample of randomised, placebo controlled trials. *BMJ* 2004;328:432.
- Fiegenbaum W: Long-term efficacy of ungraded versus graded massed exposure in agoraphobics; in Hand I, Wittchen H-U (eds): *Panic and Phobias 2: Treatments and Variables Affecting Course and Outcome*. Berlin, Springer, 1988, pp 83–88.
- Fish JM: *Placebo Therapy: A Practical Guide to Social Influence in Psychotherapy*. San Francisco, Jossey-Bass, 1973.
- Frank JD, Frank JB: *Persuasion and Healing: A Comparative Study of Psychotherapy*. Baltimore, Johns Hopkins University Press, 1993.