

Stellungnahme

zum „Gutachten zur Wissenschaftlichen Anerkennung der Gestalttherapie“ des Wissenschaftlichen Beirats Psychotherapie nach §11 PsychThG*

Harald Walach (1,2)

- 1 Change Health Science Institut, Berlin
- 2 Next Society Institute, Kazimieras Simonavicius University, Vilnius, Litauen

Kontakt

Prof. Dr.Dr. Harald Walach

CHS-Institut

Schönwalder Str. 17

13347 Berlin

030 467 97 436

hw@chs-institute.org

<https://harald-walach.de>

* PsychThG, W. B. P.. (2018). Gutachten zur wissenschaftlichen Anerkennung der Gestalttherapie. *Deutsches Ärzteblatt PP*, 17(7), 335.

Der Wissenschaftliche Beirat Psychotherapie (WBP), ein Beratungsorgan der Bundesärzte- und Bundespsychotherapeutenkammer, hat in seinem Gutachten zur wissenschaftlichen Anerkennung der Gestalttherapie festgestellt, dass lediglich bei Erwachsenen für den Bereich der Affektiven Störungen ausreichend viel Erkenntnismaterial vorliegt, das eine Wirksamkeit belegt. [1] Daher wird die Gestalttherapie weder als wirksames Verfahren zur Ausbildung für Psychotherapeuten für Erwachsene noch für Kinder empfohlen. Der WBP legt dazu eine Liste von Studien vor, die der Bewertung zugrunde lagen und hat eine ganze Reihe von Studien gar nicht in die Betrachtung einbezogen. Diese negative Ergänzungsliste wurden der Antragstellerin zugestellt, aber nicht publiziert.

Ich habe die Studien gesichtet, die als Belege nicht anerkannt wurden (Tabelle 2 in [1]), sowie einige der vorgelegten Studien, die nicht berücksichtigt wurden und bin der Frage nachgegangen, ob der WBP seine eigenen Kriterien sauber angewandt hat. Diese Betrachtung lässt die Frage unberührt, ob diese Kriterien sachlich gerechtfertigt sind. Die Einsichten sind in Tabelle 1 festgehalten.

Die Zusammenschau ergibt, dass die Auswahl der Studien und die Bewertung an manchen Stellen den Verdacht nahelegt, dass die Lektüre der Studien bzw. vermutlich die Vorauswahl der nicht in Frage kommenden Studien an Hilfspersonal delegiert wurde, das zu wenig Sachverstand hatte. Manche Studien werden hier beispielsweise aufgeführt, die klarerweise gar nicht ins Portfolio gehören, während andere gar nicht erscheinen. Außerdem legt sich der Verdacht nahe, dass Studien, die negativ ausgefallen sind zur Bewertung herangezogen wurden, obwohl sie dem eigenen Kriterienkatalog widersprechen (z.B. [2]) während positive Studien mit Verweis auf die zu kleine Patientenzahl ausgeschieden wurden. Dies wiederum lässt vermuten, dass die Auswertung des Erkenntnismaterials von unterschiedlichen Personen und mit erkennbarem Bias gegen die zu bewertende Therapie vorgenommen wurde. So wurde beispielsweise eine Studie [3] als negativ bewertet, obwohl sie in einem Zielparаметer klar signifikant positiv ist und im anderen eine Tendenz mit klinisch relevanter Effektstärke aufweist. An sich wäre die Studie zu klein. Wäre sie auch aufgrund der Größe, wie manche andere positiven Studien ausgeschieden worden, wenn man bei der Lektüre bemerkt hätte, dass es an sich eine positive Studie ist?

Die Tabellen mit den ausgeschiedenen Studien, die gar nicht im Gutachten auftauchen, habe ich nur stichprobenartig geprüft und das Ergebnis im zweiten Teil der Tabelle 1 festgehalten. Überblickt man diese Studien, dann erkennt man meiner Meinung nach folgendes Muster: Offenkundig wurde weniger sachkundiges Personal – Hilfskräfte, Sekretärinnen – mit der Aufgabe betraut, vermeintlich einfach zu klärende Kriterien zu sichten, um das Material zu reduzieren. Zu diesen Kriterien gehörten offensichtlich die Frage der Diagnosesicherung durch formale Maßnahmen (strukturelle Interviews, Diagnosefragebögen, Angaben von ICD-Codes), deren Vorhandensein man relativ leicht aus Publikationen extrahieren kann. Dass dies oft nicht sachkundig vorgenommen wurde, ergibt sich aus den wenigen Beispielen, die ich anführe. Das legt den Verdacht nahe, der durch genauere Interviews zu erhärten wäre, dass das Material einem Sichtungsalgorithmus unterworfen wurde, bei dem erst einfach zu klärende Kriterien (u.a. Diagnosestellung, Studiengröße, Randomisation, etc.) abgefragt wurden, was manchmal auch von weniger gut ausgebildetem Personal geleistet werden kann. Der Rest dürfte dann sorgfältiger gesichtet worden sein. Dadurch schlichen sich aber Fehler ein. Ein solches Vorgehen ist außerdem dem Ernst der Lage und der Tragweite der Entscheidung nicht angemessen.

Wie man außerdem der genaueren Sichtung der sechs Studien im zweiten Teil von Tabelle 1 entnehmen kann, sind einige Fehler zu verzeichnen. Einige Ablehnungen sind nicht nachvollziehbar und entweder schlecht oder falsch begründet. Drei von den sechs Studien wurden aus meiner Sicht zu Unrecht aussortiert. Bei einer Studie wäre zu fragen ob Fremdsprachigkeit (in diesem Fall Farsi) bei einer aus dem englischen Abstract erkennbar guten Studie ein Grund ist, sie gar nicht erst zur Kenntnis zu nehmen. Bei einer anderen Studie kann man erkennen, dass die Personen, die das Screening übernahmen zu wenig Sachkenntnis hatten und bei der dritten Studie wäre zu diskutieren, ob tatsächlich kein Krankheitswert erkennbar ist. Außerdem sieht man, dass große Outcomestudien, die durchaus von Bedeutung sein können, apriori nicht zur Kenntnis genommen werden. Dies ist eine Grundsatzentscheidung, die dem Stand der Wissenschaft nicht entspricht. Bei den anderen zwei Studien sind die Ablehnungskriterien aus meiner Sicht nicht erfüllt bzw. falsch oder unklar.

Da diese sechs Studien nur einen kleinen Bruchteil des Materials ausmachen, das nicht in die Bewertung einging und dort bereits gravierende Mängel feststellbar sind, stellt sich aus meiner Sicht die Frage, wie sorgfältig die Zuarbeit zu den Entscheidungen des WBP ausgefallen sind. Ich fürchte, die Qualität ließe zu wünschen übrig, wenn man alle Studien sichten würde.

Von daher wäre meine Anregung:

- a) Das Methodenpapier, das die Grundlage für die Entscheidung darstellt, sollte gründlich überarbeitet werden entlang der Kriterien bzw. Punkte, die ich in meinem Gutachten genannt habe und die auch bereits anderweitig kritisiert wurden [4].
- b) Anschließend sollte das gesamte Material neu gesichtet werden entsprechend den neuen Kriterien inclusive des Erkenntnismaterials, das seither angefallen ist.

Tabelle 1 - Bewertung der Studien, die der WBP als mangelnden Wirksamkeitsbeleg qualifiziert (Tabelle 2 in [1]) oder nicht berücksichtigt hat (Negativtabelle an die Antragstellerin geschickt; diese ist nicht vollständig gesichtet)

Studie	Urteil WBP	Ergebnis Studie	Kommentar	Fazit
Zweitsichtung der Studien aus Tabelle 2 des Gutachtens, die als negativ oder unbrauchbar bewertet werden				
Azar et al [5]	Abgelehnt, da kein separater Effekt der Gestalttherapie erkennbar	Die Studie ist klein (n=7 pro Gruppe) und dreiarig (CBT, Gestalt, Wartegruppe). Beide Behandlungsgruppen zeigen eine deutliche, signifikante Überlegenheit. Die Effektstärken sind sehr groß für Gestalttherapie gegenüber Wartegruppe und grösser oder gleich denen von CBT	Die Studie ist in der Tat nicht für einen Vergleich CBT/Gestalt gepowert, das war aber nie das Ziel. Vielmehr legt die Gleichwertigkeit der Interventionen eine Ebenbürtigkeit CBT-Gestalt nahe. Dies wird auch durch die durchwegs gleichen bis höheren Effektstärken der Gestalttherapie im Vergleich zu CBT (d = 2.3 bis 0.4, siehe Anhangstabelle) nahegelegt.	Die Entscheidung ist formal korrekt, lässt aber Anwendung von Sachverstand bei der Bewertung der Studie vermissen. Bei wohlwollender Betrachtung wäre die Studie ein Beleg für die Wirksamkeit
Baron et al [6]	Abgelehnt, da EFT CBT unterlegen ist	Die Studie war nie als Studie der Wirksamkeit von EFT angelegt, sondern EFT war hier als „control for non-specific factors“ in einer Telefon-counselling Studie angelegt. Es zeigen sich keinerlei Gruppenunterschiede (also auch CBT ist nicht „wirksam“), sondern nur Verbesserungen über die Zeit, Effektstärkeberechnung nicht möglich, weil keine Daten präsentiert sind	Die Studie ist für den Vergleich unbrauchbar und erfüllt nicht die Kriterien der „ausreichenden methodischen Qualität“, weil hier EFT klar als schwächere Kontrollprozedur zu CBT konzeptualisiert war	Diese Studie gehört nichts ins Portfolio
Mohr et al [7]	Abgelehnt, da EFT CBT unterlegen ist und keine weitere Kontrolle vorhanden war	Gut konzipierte und berichtete Studie mit 63 Patienten, die eine Überlegenheit von CBT gegenüber EFT vergleichbare Effekte von CBT gegenüber SSRI zeigt	Die Studie ist korrekt verwendet, gut berichtet und zeigt eine klare Überlegenheit von CBT	Richtig bewertet
Beckner et al [8]	Abgelehnt, da EFT CBT unterlegen ist und keine weitere Kontrolle vorhanden war	Dies ist eine sekundäre Analyse von Mohr et al 2005 [9], die die Hypothese nicht bestärken konnte, dass Personen mit niedriger sozialer Unterstützung mehr von EFT profitieren.	Diese Studie ist eine Sekundäranalyse und keine Wirksamkeitsstudie	Diese Studie gehört nicht ins Portfolio und es ist unklar, warum sie überhaupt zitiert wird
Mohr et al [9]	Abgelehnt, da EFT CBT unterlegen ist und keine weitere Kontrolle vorhanden war	Dies ist die originale Studie. Sie erklärt explizit, dass Telefon-EFT als Kontrolle konzipiert war und einige wichtige Gestalt-Elemente nicht implementiert hatte. Die Studie konnte keine Überlegenheit von CBT	Diese Studie ist keine Wirksamkeitsstudie von EFT, sondern verwendet EFT als starke Kontrolle	Diese Studie gehört nicht ins Portfolio

		zeigen; die Effektstärken zwischen CBT und T-EFT sind sehr klein		
Beutler et al [2]	Abgelehnt, keine Überlegenheit über Kontrollbedingung	Die Studie konzipierte in der Tat Gestalttherapie als einer Psychoedukation überlegen, fand aber keinen Unterschied bei Patienten mit chronischen Schmerzen und Depression	Die Studie ist sehr klein (n = 9 pro Gruppe), berichtet keine Daten, sondern nur summarisch das Ergebnis der Testung. An sich ist die Studie lt. Kriterienkatalog WBP zu klein; die Studie wurde zweimal bewertet, einmal für Affektive Störungen, einmal für somatoforme Störungen und beide Male abgelehnt	Richtig bewertet, aber schlecht berichtet und klein; wenn allerdings diese Studie herangezogen wird, so ist nicht klar warum andere kleine Studien ausgeschlossen werden
Denton et al [3]	Abgelehnt, da kein Unterschied zwischen EFT+ Medikation und Medikation allein	Die Studie kombinierte Medikamentenmanagement bei depressiven Frauen mit emotionsfokussierter Familientherapie auch bei den Partnern. Die Studie war mit n=12 pro Gruppe als Pilotstudie angelehnt, aber methodisch sauber, randomisiert mit 12 Wochen Beobachtung, allerdings hohen Ausfällen. Sie zeigt dennoch eine klare Signifikanz für die Überlegenheit der EFT im Quality of Marriage Index und eine Tendenz in der Depression. Die Effektstärke für Depression nach 12 Monaten ist d = 0.5 zugunsten EFT, für Quality of Marriage nach 6 Monaten d = 1.2 und nach 12 Monaten d = 0.3	Die Studie zeigt in einem der Zielparame-ter Überlegenheit, im anderen eine deutliche Tendenz mit klinisch relevanten Effektstärken; sie hätte bei fachgerechter Bewertung durchaus auch positiv eingeordnet werden können.	Die Einordnung als negativer Beleg spricht aus meiner Sicht dafür, dass sie nicht mit ausreichender Sorgfalt und/oder Sachkenntnis bewertet wurde.
Shear et al [10]	Abgelehnt, da keine Überlegenheit über Placebo und CBT EFT überlegen ist	Das ist eine multizentrische, vierarmige Studie, die CBT, Imipramin und Placebo bei Panikstörung verglich und später noch einen Gestalt-Arm dazu nahm. Die Gestalttherapie scheint gut implementiert zu sein. Sie erweist sich dem Placebo gleichwertig und den beiden anderen aktiven Armen deutlich unterlegen	Die Studie wurde richtig interpretiert und eingeordnet	Kein Beleg für Wirksamkeit von EFT bei Panikstörung
Ventegodt et al [11]	Abgelehnt, da kein Effekt der Gestalttherapie nachweisbar ist	In dieser Studie erhielten Patienten mit Schleudertrauma (median 3 Jahre Krankheitsdauer) einen zweitägigen und nach einem Monat einen eintägigen Kurs mit Vorlesung und 6-10 Einzelsitzungen, die individuell angepasst waren und aus Gestalttherapie, Cranio-Sacraltherapie und Rosen-Body Work bestanden. Es zeigte sich nach 3 Monaten kein Effekt, die Drop-out Rate war hoch, die ursprüngliche Intervention war verändert worden	Es ist sehr fraglich, ob diese Studie überhaupt nach den Kriterien des WBP herangezogen werden sollte, weil Gestalt-Therapie darin nur eines von vielen und nicht unbedingt das wichtigste Element war. Außerdem war das Format verändert worden und es ging eher um die Untersuchung eines eigenen Konzeptes.	Studie gehört eigentlich nicht ins Portfolio
Stichprobenartige Prüfung von Studien aus der Studienbewertung des WBP, die entweder ausgeschieden wurden oder als Beleg der Unwirksamkeit angesehen wurde; dies ist keine systematische Auseinandersetzung, sondern eine relativ unsystematische Sichtung, die einzig nach Kriterien der Erreichbarkeit von Studien und rudimentärer Vorkenntnis durchgeführt wurde				

Yousefi et al 2009	Taucht nirgendwo auf, weder in den Tabellen der Publikation noch der erweiterten Stellungnahme	Wurde eingereicht; es liegt nur ein englisches Abstract vor, der Text ist auf Farsi (Universität Teheran). Die Studie belegt die Überlegenheit von Gestalttherapie und Logotherapie bei Studierenden mit klinisch relevanter Depression, Aggression und Angst mit einer 6-Monatskatamnese bei je 30 Personen	Es wäre aus meiner Sicht die Pflicht des WBP sich durch Konsultation mit einem Sprachkundigen vom genauen Inhalt zu überzeugen; die Studie einfach zu ignorieren ist schlechte Praxis	Spricht nicht für sorgfältige Arbeit
Rostami et al [12]	Abgelehnt, da Patienten ohne Diagnose, ohne Krankheitswert, kein reliables Messverfahren, Surrogatparameter	Dies ist eine Studie an 20 Paaren (40 Personen), die das Familienzentrum in Teheran wegen Eheproblemen aufgesucht haben. Die Studie ist randomisiert und verlief über 8 Wochen. Gemessen wurde der Marital Satisfaction Questionnaire von Olson. Hier wurde eine angepasste Version verwendet die ein Cronbach alpha = 0.9 hat, also reliabel ist. Die Ehekonflikte werden signifikant reduziert	Es ist nicht nachzuvollziehen, warum der WBP ausgerechnet die Reliabilität des Fragebogens, die berichtet wird und hoch ist, ankreidet, und warum die Messung der ehelichen Zufriedenheit und der innerfamiliären Konflikte mittels eines Fragebogens ein Surrogatparameter sein soll. Man kann sich darüber streiten, ob das Aufsuchen eines Familienberatungszentrums in einer konservativen Gesellschaft wie dem Iran Krankheitswert hat. Aber die rigide Anwendung dieses Diagnosekriteriums ist fraglich.	Die Ausscheidung dieser Studie ist aus meiner Sicht fragwürdig
Hagl et al [13]	Abgelehnt, da keine adäquate Diagnosestellung und Patienten ohne Störung mit Krankheitswert	Diese Studie behandelte schwer kriegstraumatisierte Frauen im Bosnienkrieg, die während des Krieges ihren Mann entweder verloren oder vermisst haben. Die Autoren bauen ein klares Argument auf, dass diese Kategorie von Traumatisierung, die keine eigene diagnostische Kategorie hat, als schwerer Stressor anzusehen ist. Die Baselinemessungen verdeutlichen, dass die Personen im Rahmen klinisch relevanter Belastung verortet waren. Die Studie ist mit 112 Frauen, die in einer dialogischen Exposition oder in einer Standard-Unterstützungstherapie waren groß. Die Effekte sind klinisch relevant und in verschiedenen Parametern statistisch signifikant.	Es ist formal korrekt, aber sachlich falsch das Kriterium „mangelnder Krankheitswert“ und „fehlende Diagnose“ anzuwenden. Denn jeder klinisch Sachverständige wird den Autoren zustimmen, dass diese Frauen unter schweren Kriegstraumata litten. Dafür gibt es keine relevante Diagnose und es würde sich auch die Frage stellen, ob bei solchen Personen eine formale Diagnosestellung nicht einer Retraumatisierung gleichkäme.	Die Ausscheidung dieser Studie ist aus meiner Sicht unbegründet und dürfte einem „Ausscheidungsalgorithmus“ geschuldet (siehe Text oben)
Stevens et al [14]	Nicht berücksichtigt	Dies ist ein großes Outcome Forschungsprojekt des englischen NHS. In ihm werden unterschiedliche Therapieformen und Therapeuten von Patienten mit dem gleichen Fragebogen, der Clinical Outcome Routine Evaluation (CORE) bewertet. Solche Daten haben sich in der Psychotherapieforschung als nützlich erwiesen [15]. Insgesamt 180 Gestalttherapien werden	Diese Outcome-Studie belegt, dass Gestalttherapeuten in der englischen niedergelassenen Praxis etwa so gut arbeiten wie ihre Kollegen. Die Daten sind einem robusten Evaluationssystem entnommen. Diese Studie ist relevant und aufgrund ihrer Praxisnähe wichtig, kann	Die Tatsache, dass solche praxisnahen Outcomestudien nicht in die Erwägung gezogen werden können, zeigt, wie unbrauchbar der

		mit ca. 12.000 Therapien gemischter Provenienz, 7.000 Therapien mit CBT, Gesprächstherapie und tiefenpsychologischer Therapie sowie mit 118 Therapien durch minimal trainierte Therapeuten verglichen. Die Prä-Post-Effektstärke liegt bei $d = 1.12$ (Vergleich: 1.36 bis 1.42; minimal ausgebildet: $d = 0.7$), zuverlässige und klinisch relevante Besserung bei 56% der Patienten (Vergleich 30 -61%)	aber in keines der Kriterien des WBP eingeordnet werden und wurde vermutlich deshalb nicht einmal erwähnt. Das gleiche gilt für [16]	Kriterienkatalog des WBP ist.
Johnson & Smith [17]	Abgelehnt aufgrund methodischer Mängel	Dies ist eine kleine Studie mit $n = 23$ Schlangenphobikern, die entweder mit Gestalttherapie, mit systematischer Desensibilisierung oder gar nicht behandelt wurden; beide Therapiearme waren signifikant der Wartegruppe überlegen	Es ist nicht erkennbar, welches die methodischen Mängel sein sollten, außer dass die Studie eher klein ist (was aber nicht als Argument angeführt wurde)	Die Entscheidung ist unklar
Greenberg et al [18]	Abgelehnt, da keine Diagnose, kein Krankheitswert, keine randomisierte Zuweisung und keine humanistische Therapie	Die Studie untersuchte 20 Paare, die gestörte Paarbeziehung aufwiesen und die ihre eigene Kontrollgruppe mit einer verlängerten Baseline von ca. 3 Monaten darstellten. Die Therapie war „emotion focused therapy for couples“, also Gestalttherapie, und die Ergebnisse sind signifikant gegenüber Warteperiode und nach 3 Monaten stabil.	Man kann darüber diskutieren, ob dysfunktionale Paarbeziehung ohne formale Diagnose keinen Krankheitswert haben. Die Ausgangswerte der verletzten Partner liegen im GSI der SCL90 nahe an den Werten der deutschen Eichstichprobe mit psychosomatischen Störungen [19]. Von daher ist die Behauptung, die Patienten hätten keine Störung mit Krankheitswert fraglich. Definitiv falsch ist, dass es sich nicht um eine humanistische Therapie handeln würde. Richtig, aber für die Bewertung nur teilweise bedeutsam ist, dass es keine randomisierte Wartegruppe gab. Dies wird durch eine ausführliche intra-individuelle Kontrolle kompensiert.	Die Ablehnungskriterien sind teilweise richtig, teilweise falsch und müssten sicher diskutiert werden.

Literatur

1. PsychThG WBPn. Gutachten zur wissenschaftlichen Anerkennung der Gestalttherapie. Deutsches Ärzteblatt PP. 2018;17(7):335.
2. Beutler LE, Daldrup R, Engle D, Guest P, Corbishley A, Meredith KE. Family dynamics and emotional expression among patients with chronic pain and depression. *Pain*. 1988;32(1):65-72. doi: [https://doi.org/10.1016/0304-3959\(88\)90024-3](https://doi.org/10.1016/0304-3959(88)90024-3).
3. Denton WH, Wittenborn AK, Golden RN. Augmenting Antidepressant Medication Treatment of Depressed Women With Emotionally Focused Therapy for Couples: A Randomized Pilot Study. *Journal of Marital and Family Therapy*. 2012;38(s1):23-38. doi: <https://doi.org/10.1111/j.1752-0606.2012.00291.x>.
4. Kriz J. Perspektiven der "Wissenschaftlichkeit" in der Psychotherapie. In: Hermer M, editor. *Psychotherapeutische Perspektiven am Beginn des 21. Jahrhunderts*. Tübingen: DGVT-Verlag; 2000. p. 43-66.
5. Azar FS, Asadnia S. Efficacy of cognitive behavior therapy and Gestalt therapy on poor sleep quality among college female students with headache. *Journal of Education, Psychology and Social Sciences*. 2013;1(2):15-24.
6. Baron KG, Corden M, Jin L, Mohr DC. Impact of psychotherapy on insomnia symptoms in patients with depression and multiple sclerosis. *Journal of Behavioral Medicine*. 2011;34(2):92-101. doi: 10.1007/s10865-010-9288-2.
7. Mohr DC, Boudewyn AC, Goodkin DE, Bostrom A, Epstein L. Comparative outcomes for individual cognitive-behavior therapy, supportive-expressive group psychotherapy, and sertraline for the treatment of depression in multiple sclerosis. *Journal of Consulting and Clinical Psychology*. 2001;69(6):942-9. doi: <https://doi.org/10.1037/0022-006X.69.6.942>.
8. Beckner V, Howard I, Vella L, Mohr DC. Telephone-administered psychotherapy for depression in MS patients: moderating role of social support. *Journal of Behavioral Medicine*. 2010;33(1):47-59. doi: 10.1007/s10865-009-9235-2.
9. Mohr DC, Hart SL, Julian L, Catledge C, Honos-Webb L, Vella L, et al. Telephone-Administered Psychotherapy for Depression. *Archives of General Psychiatry*. 2005;62(9):1007-14. doi: 10.1001/archpsyc.62.9.1007.
10. Shear MK, Houck P, Greeno C, Masters S. Emotion-Focused Psychotherapy for Patients With Panic Disorder. *American Journal of Psychiatry*. 2001;158(12):1993-8. doi: 10.1176/appi.ajp.158.12.1993.
11. Ventegodt S, Merrick J, Andersen NJ, Bendix T. A Combination of Gestalt Therapy, Rosen Body Work, and Cranio Sacral Therapy did not help in Chronic Whiplash-Associated Disorders (WAD) - Results of a Randomized Clinical Trial. *The Scientific World Journal*. 2004;4:284951. doi: 10.1100/tsw.2004.132.
12. Rostami M, Taheri A, Abdi M, Kermani N. The Effectiveness of Instructing Emotion-focused Approach in Improving the Marital Satisfaction in Couples. *Procedia - Social and Behavioral Sciences*. 2014;114:693-8. doi: <https://doi.org/10.1016/j.sbspro.2013.12.769>.
13. Hagl M, Powell S, Rosner R, Butollo W. Dialogical Exposure with Traumatically Bereaved Bosnian Women: Findings from a Controlled Trial. *Clinical Psychology & Psychotherapy*. 2015;22(6):604-18. doi: <https://doi.org/10.1002/cpp.1921>.
14. Stevens C, Stringfellow J, Waelin K, Waring J. The UK Gestalt psychotherapy CORE research project: The findings. *British Gestalt Journal*. 2011;20(2):22.
15. Seligman M. The effectiveness of psychotherapy. *American Psychologist*. 1995;50:965-74.

16. Tschuschke V, Cramer A, Koemeda M, Schulthess P, von Wyl A, Weber R. Psychotherapieforschung – Grundlegende Überlegungen und erste Ergebnisse der naturalistischen Psychotherapie-Studie ambulanter Behandlungen in der Schweiz (PAPS). Psychotherapie Forum. 2009;17(4):160. doi: 10.1007/s00729-009-0300-7.
17. Johnson WR, Smith EWL. Gestalt empty-chair dialogue versus systematic desensitization in the treatment of a phobia. Gestalt Review. 1997;1(2):150-62. doi: <https://doi.org/10.2307/44394035>.
18. Greenberg L, Warwar S, Malcolm W. Emotion-Focused Couples Therapy and the Facilitation of Forgiveness. Journal of Marital and Family Therapy. 2010;36(1):28-42. doi: <https://doi.org/10.1111/j.1752-0606.2009.00185.x>.
19. Schmitz N, Hartkamp N, Kruse J, Franke GH, Reister G, Tress W. The Symptom Check-List-90-R (SCL-90-R): A German validation study. Quality of Life Research. 2000;9:185-93.

Anhang – Detailtabelle der Studie von Azar et al [5]

Studie	Variable	Effektstärke Gestalt d	Effektstärke CBT d
Azar et al	Schlafqualität	2.3	2.15
	Schlaflatenz	2.0	1.3
	Schlafdauer	0.5	0.48
	Schlafunterbrechung	1.1	1.1
	Schlafmedikation	0	-0.4 (1)
	Funktion untertags	0.4	0.8

(1) Verschlimmerung