**Supplementary Online Content**

**eAppendix 1.** Search Terms

**eAppendix 2.** Addition to Methods

**eTable 1**. Relevant Characteristics of all 44 RCTs Included in the Network Meta-analysis

**eTable 2.** Summary of Study Characteristics per Treatment Group

**eTable 3.** Risk of Bias for Each Included Study

**eFigure 1.** Risk of Bias Assessment

**eTable 4.** Indirectness for Each Included Study

**eFigure 2.** Indirectness

**eTable 5:** Confidence in Network Meta-analysis (CINEMA) Rating

**eAppendix 3.** Baseline Differences

**eAppendix 4.** Additional Results from Network Meta-analyses: Main Analyses

**eAppendix 5.** Additionals Results from Network Meta-analyses: Sensitivity Analyses

**eAppendix 6.** Additional Results from Pairwise Meta-analyses on Symptom Seveirty

**eAppendix 7.** Additional Results from Network Meta-analyses: Comparative Acceptability

**eAppendix 8.** Additional Results from Pairwise Meta-analyses on Comparative Acceptability

**eAppendix 1.** Search Terms

**MEDLINE (Ovid)**

|  |  |
| --- | --- |
| Diagnosis | Exp Anxiety Disorders/ |
| ((trauma\* adj3 stress) or (stress adj3 disorder\*) or PTSD).tw. |
| Design | randomized controlled trial.pt. OR controlled clinical trial.pt. OR randomized.tw. OR placebo.tw. OR clinical trials as topic.sh. OR randomly.tw. OR trial.tw. OR groups.tw. |
| Intervention | self disclosure/ |
| (((express\* or emotion\* or guid\* or experiment\* or self or reflective\*) adj3 (writ\* or disclos\* or express\*)) or Pennebaker or ((psychotherap\* or therap\* or intervention\* or disclos\*) adj3 writ\*) or interapy or ((computer or internet or web or journal\*) adj5 (treatment\* or therap\* or intervention\* or disclos\*))).mp. |

**EMBASE (Ovid)**

|  |  |
| --- | --- |
| Diagnosis | posttraumatic stress disorder/ |
| ((trauma\* adj3 stress) or (stress adj3 disorder\*) or PTSD).tw. |
| Design | random\*.mp. OR clinical trial\*.mp. OR exp treatment outcome/ OR exp controlled clinical trial/ |
| Intervention | self disclosure/ |
| (((express\* or emotion\* or guid\* or experiment\* or self or reflective\*) adj3 (writ\* or disclos\* or express\*)) or Pennebaker or ((psychotherap\* or therap\* or intervention\* or disclos\*) adj3 writ\*) or interapy or ((computer or internet or web or journal\*) adj5 (treatment\* or therap\* or intervention\* or disclos\*))).mp. |

**PsycINFO (Ovid)**

|  |  |
| --- | --- |
| Diagnosis | posttraumatic stress disorder/ OR emotional trauma/ OR stress reactions/ OR traumatic neurosis/ or child abuse/ |
| ((trauma\* adj3 stress) or (stress adj3 disorder\*) or PTSD).tw. |
| Design | exp Treatment Effectiveness Evaluation/ OR exp Treatment Outcomes/ OR \*placebo/ OR exp followup studies/ OR placebo\*.mp. OR random\*.mp. OR "comparative stud\*".mp. OR (clinical adj3 trial).mp. OR (evaluat\* adj3 stud\*).mp. OR (prospectiv\* adj3 stud\*).mp. OR ((singl\* or doubl\* or trebl\* or tripl\*) adj3 (blind\* or mask\*)).mp. OR (research adj3 design).mp. |
| Intervention | \*self disclosure/ |
| (((express\* or emotion\* or guid\* or experiment\* or self or reflective\*) adj3 (writ\* or disclos\* or express\*)) or Pennebaker or ((psychotherap\* or therap\* or intervention\* or disclos\*) adj3 writ\*) or interapy or ((computer or internet or web or journal\*) adj5 (treatment\* or therap\* or intervention\* or disclos\*))).mp. |

**CENTRAL**

|  |  |
| --- | --- |
| Diagnosis | MeSH descriptor: [Stress, Psychological] this term only OR MeSH descriptor: [Stress Disorders, Traumatic] 1 tree(s) exploded |
| (trauma\* NEAR3 stress) OR (stress NEAR3 disorder\*) OR PTSD |
| Design | - |
| Intervention | MeSH descriptor: [Self Disclosure] this term only |
| (((express\* or emotion\* or guid\* or experiment\* or self or reflective\*) NEAR3 (writ\* or disclos\* or express\*)) or Pennebaker or ((psychotherap\* or therap\* or intervention\* or disclos\*) NEAR3 writ\*) or interapy or ((computer or internet or web or journal\*) NEAR5 (treatment or therap\* or intervention\* or disclos\*))) |
| Additional Filter | Trials |

**eAppendix 2.** Addition to Methods

**A. Definition of Population**

We included studies only if all study participants had experienced a traumatic event (as defined in DSM IV).

A study was considered to have included participants with *full PTSD* if at least 80% of all study participants fulfilled diagnostic criteria for PTSD (e.g., DSM III, DSM IV, or DSM 5) or if the authors described the sample as “participants with PTSD”. We considered a study to have included participants with *subclinical / subthreshold / partial PTSD* if authors gave a clear definition of subclinical of subthreshold PTSD. For instance, it was necessary to report symptoms above a pre-defined cutoff score for PTSD for inclusion in a particular study (e.g., scores above 33 in IES-R, above 45 on CAPS, above 44 on PCL, above 2.5 on HTQ), or study participants had to report experiencing at least one, but not all, of the symptom clusters of reexperiencing, avoidance or hyperarousal. For the *presence of PTSD symptoms,* we required the presence of elevated PTSD symptoms on a validated PTSD scale, however it was not required that all participants reported scores higher than a cut-off for PTSD. We included studies also if study authors described the sample as “participants with relevant PTSD symptoms”, for instance.

**B. Definition of Trauma-focused Treatments**

***Definition of established psychotherapeutic PTSD treatments***

Psychotherapeutic PTSD treatments had to be implemented at the level of individual patients, rather than in group, family, or couple therapy; they had to include face-to-face contact between the patient and the therapist, as opposed to telephone or internet-based interactions between patient and therapist; they had to be standardized (similar dose of treatment for all patients and treatment based on the same rational for all patients in one study); they had to consist primarily of verbal communication; and they had to directly address the trauma or subsequent PTSD symptoms. PTSD treatments were not considered for the analyses if they included experimental manipulations of an established treatment approach (e.g., dismantling of individual treatment components).

***Definition of expressive writing treatments***

We defined trauma-focused writing as a writing treatment that targeted the traumatic event the participant had experienced. We classified expressive writing treatments as 1st expressive writing (original; EW) and 2nd enhanced writing (EW+). We allowed any delivery method (e.g., paper and pencil, computerized, or internet-based intervention), as long as it was a purely written intervention and not mixed with any other intervention like verbal cognitive behavioral therapy.

***EW:*** Authors either explicitly referred to the original paradigm by Pennebaker & Beale (1986),1 or writing treatments were similarly structured as the original writing paradigm (e.g., 3 to 4 sessions of 15 to 30 minutes duration). Importantly, to be considered EW no therapist involvement was allowed. Also, no individualized instructions for each writing session were allowed.

***EW+***: The treatment description 1st did not explicitly refer to the original Pennebaker writing paradigm AND 2nd writing treatments included additional elements assumed to increase their efficacy: the treatments included either the presence of a therapist during writing sessions, or any therapist feedback. In many cases experimental manipulation of the writing content was used (e.g., more directive writing instructions which changed for each writing session). Enhanced writing treatments also typically used more or longer writing sessions compared with the original paradigm. However, the use of longer sessions alone was not sufficient to classify a writing treatment as enhanced.

***Exclusion criteria***

Writing treatments that were administered in addition to one of the writing treatments (EW or EW+), and which didn’t fulfill the criteria for neither EW nor EW+ (ie, not writing about one’s own trauma, not being allowed to read the written account, not being allowed to write about the same event every session) were excluded from the analyses.

Studies that used only experimental manipulations of formal aspects of the writing task (e.g., writing in the first person vs writing in the third person, e.g., Andersson & Conley, 2013; Kenardy & Tan, 2016)2,3 but which had no additional comparator were not included in the analyses.

**C. Risk of Bias in the Included Studies**

To evaluate the quality of studies and potential risk of bias (RoB), we rated to the predefined criteria in the “Cochrane Handbook for System­atic Reviews of Interventions”.4 For the application of the RoB criteria in the context of psychotherapy research, we adhered to the recommendations by Munder & Barth.5

For each included study the risk for five potential bias categories was assessed: 1st *selection bias* (*sequence generation,* and *allocation sequence concealment*), 2nd *performance bias (use of only passive control, weiting intervention described as control,* and *non-equivalent duration of active interventions)*, 3rd *detection bias (non-blind assessors in case of observer rated outcome assessment)*, 4th *attrition bias (incomplete outcome dara,* and *inadequate analysis strategy, i.e. no intention-to-treat),* and 5th *reporting bias (no specification of one primary outcome,* and *results for some outcome measures not reported completely)*.

1stwe rated “low” risk of *selection bias*, if both relevant categories (*sequence generation* and *concealment of allocation* were considered as “low”. Risk for *selection bias* was considered “high”, if both categories were considered “high” or one was considered “high” and one “unclear”. All other cases of mixed codings for the two categories were considered “unclear”. *Sequence generation* was considered adequate (i.e. low risk of bias), if participants were randomly assigned to treatment conditions stating a randomization procedure that ensured that similarity of groups at baseline was warranted (e.g., computerized random sequence generation). *Concealment of allocation* was considered adequate if the procedures described ensured that the investigators responsible for patient selection did not suspect which treatment was next before allocation (e.g., if allocation to treatments was conducted by an external third party).

2nd risk of *performance bias* was considered “low” if two equally credible treatments were compared, even if participants and treatment providers were not blinded. Risk of *performance bias* was rated “high”, if participants as well as assessors were not blinded and knew which therapy the participant received and if the treatments differed with respect to their credibility. The credibility rating was based on three sources of information: 1st the use of an active comparator, 2nd whether the writing intervention itself was described as a control treatment within the study, and 3rd whether the amount of treatment was comparable across the different treatment groups. If all three items were considered as “low” risk of bias, risk for *performance bias* was considered “low”, if all at least one item was considered high and none of the items was considered “low” risk of *performance bias* was considered “high”. In all other cases of mixed ratings risk of *performance bias* was considered “unclear”.

3rd risk of detection bias was rated “high” if outcome assessors knew which therapy a participant was assigned to (i.e., non-blind assessment of observer-rated outcomes). Risk of *detection bias* was considered “low” if only self-rated outcome measures were used5, or if observer-rated outcomes were assessed by blind assessors.

4th risk of *attrition bias* was considered “high” if missing outcome data varied largely across conditions and analyses were not conducted according to the intention to treat (ITT) principle. Risk of *attrition bias* was considered “low” if all participants were analysed as randomized. Incase of mixed or unclear evidence risk of *attrition bias* was considered “unclear”.

5th risk of *reporting bias* was considered “low” if a primary outcome was specified and results for effect size generation was reported for all mentioned outcome measures. Risk of *reporting bias* was considered “high”, if either data for effect size generation was not provided for all outcome measures or no orimary outcome was prespecified and nonw of the two relevant items was considered as “low” risk of bias. In cases of mixed evidence we considered risk of bias as “unclear”.

If relevant information on any quality criterion were not reported, or if the reported information was insufficient for a clear “high” or “low” rating we coded the respective criterion as “unclear”.

We rated a study as “high” regarding overall RoB, if one or more of the five dimensions were rated with “high” RoB. We rated a study as “low” regarding overall RoB, if all dimensions were rated “low”. In any other case we considered the study to have “some concerns” regarding RoB.

**C. Indirectness**

We rated the indirectness of the available evidence as recommended by Guyatt et al. 2011.6 We assessed whether 1st a study differed from the studies of interest with respect to 1st the relevant study population, 2nd the applied intervention, 3rd the evaluated outcomes, and 4th whether a study provided direct evidence for at least one of the comparisons of interest.

Overall indirectness was considered “low,” if at least 3 items were rated as “low” and maximum one item was rated “unclear”. If at least one item was rated “high” the overall rating could not be “low”. Overall indirectness was considered “high” if at least 2items were rated as “high” or 2 items were rated “high”. All other combinations were rated “moderate”.

It is important to note the difference between the rating of indirectness, ie the assessment whether a study reflects a typicaly study intended to be included in the meta-analysis, and indirect evidence, which refers to evidence which has not directliy been observed in a study, but can only be inferred from available evidence via indirect paths. This may for instance be the case if two treatments have never been compared in one study, but both have been compared with a third treatment. The comparisons with the third treatment can then be used to infer about the comparison between the two initial treatments.

**D. Confidence in Network Meta-analysis (CINEMA)**

We assessed the quality of the entire network using the CINEMA framework.7 This includes evaluations of within study bias, across study bias, indirectness, imprecision, heterogeneity and incoherence. The RoB rating was used for evaluating within study bias. For across study bias we assumed that the likelihood of unpublished data was small because we included dissertation theses and conducted our searches in 5 databases and screened the reference lists of previous relevant meta-analyses and systematic reviews. For the evaluation of indirectness, we used the rating as described above. For the evaluation of imprecision, heterogeneity and incoherence we defined the clinically important effect size as 0.6.8 For the overall rating of confidence we used the recommendations suggested by Guyatt et al., 2011.9

**E. Hierarchy of PTSD Outcome Scales**

1=IES,10 2= IES-R,11 3=PDS,12 4=PCL-S/M/C,13 5=CAPS,14 6=other (see eTable 1)

**F. Deviations from the published protocol**

Deviations from the protocol10 include the following:

1. We defined the analysis including the longest available follow-up as primary outcome in the manuscript, while initially planned the analyses using end-of-treatment data as primary outcome, because we expected all studies would report results at this time-point. Unexpectedly, a number of studies reported only data which were assessed more than one month after treatment termination, which we defined as long-term outcomes. In order to be able to include all studies in our main analysis, we defined the analysis using the longest available follow-up data as primary outcome data in our manuscript. Nevertheless, we report all results in Table 2 in our manuscript: 1. results relying on end-of-treatment data only, 2. results relying on longer-term data only and 3. results using the longest available follow-up (i.e. including all studies = our main analysis).
2. We did not conduct the sensitivity analysis excluding studies that reported adjusted means, because the available information regarding adjustment of means was insufficient in the primary studies in order to allow for valid interpretations.
3. Due to the very high levels of heterogeneity and inconsistency we conducted additional sensitivity analyses in order to explore sources of variation between treatment effects; we especially conducted an analysis excluding the two-arm comparisons between enhanced writing and waiting-list, which we describe in more detail as exploratory analyses in the manuscript, because this was the analysis in which the largest amount of heterogeneity and inconsistency were explained.
4. We report risk ratios instead of odds ratios, as described in the protocol, because risk ratios were considered easier to interpret.

1. Pennebaker JW, Beall SK. Confronting a traumatic event: toward an understanding of inhibition and disease. *Journal of Abnormal Psychology.* 1986;95(3):274.

2. Andersson MA, Conley CS. Optimizing the perceived benefits and health outcomes of writing about traumatic life events. *Stress and Health: Journal of the International Society for the Investigation of Stress.* 2013;29(1):40-49.

3. Kenardy J, Tan L-J. The Role of Avoidance Coping in the Disclosure of Trauma. *Behav Change.* 2006;23(1):42-54.

4. Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions. In: Version 5.1.0 ed.: The Cochrane Collaboration; updated March 2011: [www.cochrane-handbook.org](file:///Users/hgerger/Dropbox/MA-EW/Remake/09_NMA/06Manuscript/00submit/Psychological%20Medicine/02Resubmission/www.cochrane-handbook.org).

5. Munder T, Barth J. Cochrane’s risk of bias tool in the context of psychotherapy outcome research. *Psychotherapy Research.* 2018;28(3):347-355.

6. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 8. Rating the quality of evidence—indirectness. *J Clin Epidemiol.* 2011;64(12):1303-1310.

7. Salanti G, Del Giovane C, Chaimani A, Caldwell DM, Higgins JPT. Evaluating the quality of evidence from a network meta-analysis. *PLoS One.* 2014;9(7):e99682.

8. Stefanovics EA, Rosenheck RA, Jones KM, Huang G, Krystal JH. Minimal clinically important differences (MCID) in assessing outcomes of post-traumatic stress disorder. *Psychiatr Q.* 2018;89(1):141-155.

9. Guyatt GH, Oxman AD, Vist G, et al. GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias). *J Clin Epidemiol.* 2011;64(4):407-415.

10. Horowitz M, Wilner N, Alvarez W. Impact of Event Scale: A measure of subjective stress. Psychosomatic Medicine. 1979;41(3):209-218.

11. Weiss D, Marmar C. The Impact of Event Scale-Revised. Wilson JP, Keane TM, editors. Assessing psychological trauma and PT-SD. New York Guliford Press; 1997.

12. Foa EB, Cashman L, Jaycox L, Perry K. The validation of a self-report measure of posttraumatic stress disorder: The Posttraumatic Diagnostic Scale. Psychol Assess. 1997;9(4):445-451.

13. Weathers FW, Huska JA, Keane TM. PCL-C for DSM-IV. Boston: National Center for PTSD-Behavioral Science Division. 1991.

14. Blake DD, Weathers FW, Nagy LM, et al. The development of a clinician-administered PTSD scale. J Trauma Stress. 1995;8(1):75-90.

15. Gerger H, Gaab J, Werner C. A network meta-analysis on the efficacy and acceptability of writing interventions for traumatised adults. PROSPERO International Prospective Register of Systematic Reviews [Internet]. 2018.

**eTable 1**. **Relevant Characteristics of all 44 RCTs Included in the Network Meta-analysis**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| First author  (publication year) | Interventions  (number of patients at baseline) | Baseline M | PTSD severity measure | Treatment duration:  weeks / number of sessions | Reference to original paradigm / Therapist contact | Treatment drop-outs | Trauma type | Age  M | Sex female % | Time to last available follow-up |
| Gidron  (1996) | expressive writing (8),  neutral writing control (6) | M=40.0 (SD=5.5),  M=39.7 (SD=18.3) | IES | 3 days /  3 sessions | Explicit / no | NR | Mixed trauma | NR | 57 | 35 days |
| Greenberga (1996) | real-trauma writing (34),  neutral writing control (28) | M=16.6 (SD=10.0),  M=12.5 (SD=10.0) | IES intrusion subscale | 1 day /  1 session | Similar but no reference / no | 6 | NR | 19 | 100 | 28 days |
| Barry\*  (2001) | expressive writing (15),  waitlist control (15) | M=1.8 (SD=1.2),  M=1.3 (SD=1.0) b | IES-R | 4 days /  4 sessions | Similar but no reference / no | 10 | Newborn at NICU | 33 | 100 | 28 days |
| Batten\*  (2001) | expressive writing (30),  neutral writing control (56) | M=14.1 (SD=8.0),  M=11.1 (SD=7.6) | TSI avoidance subscale | 1 week /  4 sessions | Similar but no reference / no | 6 | Sexual abuse / assault | 35 | 100 | 84 days |
| Lange  (2001) | Expressive writing (13),  waitlist control (12) | M=17.5 (SD=6.5),  M=13.6 (SD=7.0) | IES intrusion subscale | 5 weeks /  10 sessions | Enhanced writing / therapist feedback | 5 | Mixed trauma | 22 | 53 | 42 days |
| Yanko\*a  (2001) | own trauma realistic writing (24),  neutral writing control (22) | NR | IES | 3 days /  3 sessions | Similar bat no reference / no | 7 | Mixed trauma | 19 | 100 | 28 days |
| Largo-Marsh\*\* (1996 / 2002) | expressive writing (12),  EMDR (12) | M=1.7 (SD=0.7),  M=2.1 (SD=0.6) b | IES intrusion subscale | NA /  up to 3 sessions | Enhanced writing / session with therapist | NR | Mixed trauma | 34 | 71 | 30 days |
| Schoutrop (2002) | expressive writing (26),  waitlist control (22) | M=15.6 (SD=5.1),  M=16.8 (SD=3.0) | IES intrusion subscale | 2 weeks /  5 sessions | Similar but no reference / no | NR | NR | 22 | 75 | 42 days |
| Lange  (2003) | expressive writing (69),  waitlist control (32) | M=20.2 (SD=7.5),  M=19.9 (SD=8.2) | IES intrusion subscale | 5 weeks /  10 sessions | Enhanced writing / therapist feedback | 44 | Mixed trauma | 39 | NR | 42 days |
| Deters  (2003) | expressive writing (30),  neutral writing control (27) | M=1.7 (SD=0.9),  M=1.6 (SD=0.9) b | IES-R | 2 weeks /  4 sessions | Similar but no reference / no | 0 | Mixed trauma | 23 | 73 | 42 days |
| Sloan  (2004) | expressive writing (26),  neutral writing control (23) | M=17.6 (SD=6.8),  M=16.6 (SD=5.3) | PDS | 3 days /  3 sessions | Explicit / no | 0 | Mixed trauma | 19 | 100 | 28 days |
| Koopman (2005) | expressive writing (25),  neutral writing control (22) | M=46.5 (SD=15.6),  M=44.3 (SD=14.0) | PCL-S | 4 week /  4 sessions | Similar but no reference / no | NR | Intimate partner violence | 37 | 100 | 122 days |
| Nguyen\*  (2005) | expressive writing (30),  neutral writing control/waiting list (33) | M=2.0 (SD=0.6),  M=2.0 (SD=0.7) b | HTQ | 4 days /  4 sessions | Explicit / no | 0 | Ex-political detainees | 61 | 0 | 30 days |
| Sloana  (2005) | same trauma writing (28),  neutral writing control (25) | M=20.0 (SD=6.1),  M=19.8 (SD=9.5) | PDS | 3 days /  3 sessions | Explicit / no | 0 | NR | 19 | 73 | 56 days |
| Freyd  (2005) | expressive writing (49),  neutral writing control (49) | M=79.3 (SD=20.3),  M=79.3 (SD=20.3) | TSC | 3 weeks /  3 sessions | Similar but no reference / no | 22 | Mixed trauma | 42 | 56 | 182 days |
| Wagner  (2006) | interapy (26),  waitlist control (25) | M=24.3 (SD=6.8),  M=26.6 (SD=4.9) | IES intrusion subscale | 5 weeks /  10 sessions | Enhanced writing / therapist feedback | 4 | Complicated grief | 37 | 93 | 91 days |
| Sloan  (2007) | emotional expression writing (28)c,  insight and cognitive writing (27)c,  neutral writing control (27) | M=20.7 (SD=9.1),  M=17.5 (SD=9.1),  M=17.3 (SD=5.7) | PDS | 3 days /  3 sessions | Similar but no reference / no | 1 | Mixed / trauma | 19 | 80 | 28 days |
| Knaevelsrud (2007) | interapy (49),  waitlist control (46) | M=23.0 (SD=6.4),  M=23.3 (SD=7.9) | IES-R | 5 weeks / 1  0 sessions | Enhanced writing / therapist feedback | 9 | Mixed trauma | 35 | 90 | 91 days |
| Possemato \*\* (2008 / 2010) | expressive writing (25),  neutral writing control (23) | M=37.0 (SD=12.0),  M=37.0 (SD=11.0) | PCL-C | 3 days /  3 sessions | Similar but no reference / no | 4 | Kidney transplant | 46 | 54 | 91 days |
| Resicka  (2008) | expressive writing (48),  CPT (53), | M=29.4 (SD=9.7),  M=29.2 (SD=9.5), | PDS | 6 weeks /  12 sessions | Enhanced writing / sessions with therapist | 40 | Mixed trauma | 35 | 100 | 182 days |
| Smyth  (2008) | expressive writing (14),  neutral writing control (10) | NR | PSS-I | 1 day /  3 sessions | Similar but no reference / no | 1 | Mixed trauma | NR | 56 | 91 days |
| van Emmerik (2008) | structured writing (44),  CBT (41),  waitlist control (41) | M=47.9 (SD=13.8),  M=46.4 (SD=12.3),  M=49.1 (SD=14.7) | IES | Acute PTSD:  5 weeks / 5 sessions  Chronic PTSD:  5 weeks /  10 sessions | Enhanced writing / sessions with therapist | NA | Mixed trauma | 40 | 67 | 500 days |
| Bugg  (2009) | expressive writing (31),  information control (36) | M=21.5 (SD=12.0),  M=22.0 (SD=9.8) | PDS | 3 days /  3 sessions | Explicit / no | 15 | Mixed trauma | 37 | 32 | 152 days |
| Kearns  (2010) | expressive writing,  neutral writing control (total 73) | M=34.6 (SD=15.8),  M=36.4 (SD=19.1) | TSC | 4 week /  4 sessions | Similar but no reference / no | 1 | Sexual abuse / assault | 19 | 100 | 30 days |
| Lichtenthal (2010) | expressive writing (16),c  benefit-finding writing (17)c,  sense-making writing (19)c,  neutral writing control (16) | M=29.1 (SD=9.1),  M=27.5 (SD=8.6),  M=28.8 (SD=8.7),  M=31.6 (SD=10.7) | PCL-C | 2 weeks /  3 sessions | Explicit / no | NA | Significant interpersonal loss | 20 | 26 | 91 days |
| Beyer\*  (2011) | instant message expressive writing (41)c,  expressive writing with feedback (41)c,  standard expressive writing (41),  neutral writing control (40) | M=1.9 (SD=0.9),  M=1.6 (SD=0.9),  M=2.2 (SD=1.1),  M=2.0 (SD=0.9) b | IES-R | 7-10 days /  3 sessions | Explicit / no | 16 | Mixed trauma | 22 | 83 | 42 days |
| Kersting  (2011) | expressive writing (45),  waitlist control (33) | M=33.1 (SD=13.2),  M=34.6 (SD=11.4) | IES | 5 weeks /  10 sessions | Enhanced writing / therapist feedback | 19 | Loss of child during pregnancy | 34 | 100 | 91 days |
| Slavin-Spennya (2011) | active fascilitator writing (31)  expressive writing (36),  neutral writing control (34) | M=2.0 (SD=1.2),  M=2.0 (SD=1.1) b | IES intrusion subscale | 1 day /  1 session | Similar but no reference / no | 9 | NR | 22 | 82 | 42 days |
| Sloan  (2011) | expressive writing (21),  neutral writing control (21) | M=24.8 (SD=5.5),  M=25.2 (SD=5.2) | PDS | 3 days /  3 sessions | Explicit / no | 0 | Mixed / traumatic events | 19 | NA | 30 days |
| Zakowski (2011) | expressive writing (43),  neutral writing control (45) | M=8.3 (SD=8.0),  M=7.6 (SD=6.7) | IES avoidance subscale | 3 days /  3 sessions | Explicit / no | 10 | Gynaecological cancer | 58 | 100 | 7 days |
| Ironson  (2013) | augmented-trauma writing (120),  neutral writing control (122) | M=27.7 (SD=23.8),  M=28.7 (SD=23.5) | Davidson PTSD scale | 2-4 weeks /  4 sessions | Enhanced writing / no | 26 | NR | 43 | 39 | 365 days |
| Sloan  (2012) | expressive writing (22),  waitlist control (24) | M=61.4 (SD=15.0),  M=70.6 (SD=18.6) | CAPS | 5 weeks /  5 sessions | Enhanced writing / sessions with therapist | 2 | Mixed trauma | 41 | 65 | 91 days |
| Jensen-Johansen (2013) | expressive writing (243),  neutral writing control (243)  natural course (2912) | M=17.6 (SD=14.5),  M=20.2 (SD=15.9) | IES | 3 weeks /  3 sessions | Explicit / no | 18 | Treatment for breast cancer | 54 | 100 | 273 days |
| Kersting  (2013) | expressive writing (115),  waitlist control (113) | M=30.5 (SD=12.0),  M=31.7 (SD=11.6) | IES-R | 5 weeks /  10 sessions | Enhanced writing / therapist feedback | 29 | Loss of child during pregnancy | 34 | 92 | 365 days |
| Milbury  (2014) | expressive writing (138),  neutral writing control (139) | M=17.8 (SD=15.1),  M=19.6 (SD=15.0) | IES | 10 days /  4 sessions | Explicit / no | 66 | Renal cell carcinoma | 61 | 41 | 304 days |
| Southern\* (2014) | narrative writing (10)c,  repeated expressive writing (8)c,  neutral writing control (10) | M=2.3 (SD=1.0),  M=2.1 (SD=0.9),  M=2.7 (SD=0.8) b | IES-R | 4 days /  4 sessions | Similar but no reference other group/ no | 50 | Domestic violence | 37 | 100 | 30 days |
| Truijensa  (2014) | written imaginal exposure (20),  neutral writing control (19) | M=49.1 (SD=12.1),  M=46.9 (SD=12.3) | IES | 1 day /  1 session | Similar but no reference / no | 0 | Mixed trauma | 24 | 78 | 7 days |
| Stockton (2014) | expressive writing (14),  neutral writing control (10) | M=19.9 (SD=7.1),  M=13.3 (SD=10.2) | IES intrusion subscale | 9 days /  3 sessions | Explicit / no | 19 | Mixed trauma | 33 | 22 | 42 days |
| Sayer  (2015) | expressive writing (508),  neutral writing control (507),  waitlist control (277) | Med=36.0 (IQR=25.5-50.0),  Med=35.0 (IQR=25.0-51.0),  Med=39.0 (IQR=28.0-51.0) | PCL-M | 10 days /  4 sessions | Explicit / no | 220 | Veterans of Afghanistan and Iraque wars | 37 | 39 | 182 days |
| Knaevelsrud (2015) | interapy (79),  waitlist control (80) | M=30.4 (SD=8.2),  M=30.7 (SD=8.1) | PDS | 5 weeks /  10 sessions | Enhanced writing / therapist feedback | 65 | War | 28 | 72 | 91 days |
| Horsch  (2016) | expressive writing (33),  treatment as usual (32) | M=4.1 (SD=2.9),  M=4.1 (SD=3.2) b | PPQ | 3 days /  3 sessions | Explicit / no | 4 | Very preterm birth | 32 | 100 | 91 days |
| Alessandri\* (2017) | expressive writing (44),  directive protocol (42) | M=41.0 (SD=20.6),  M=36.4 (SD=)19.6 | MPSSR-SR | 3 days /  3 sessions | Similar but no reference / no | 1 | Mixed traumat | 20 | 77 | 30 days |
| Knaevelsrud (2017) | integrative testimonial therapy (47),  waitlist control (47) | M=22.8 (SD=8.8),  M=21.9 (SD=8.8) | PDS | 6 weeks /  11 sessions | Enhanced writing / therapist feedback | 9 | War (WW II) | 71 | 65 | 365 days |
| Sloan\*\*\*  (2018) | written exposure therapy (63),  cognitive processing therapy (63) | M=36.1 (SD=8.9),  M=37.1 (SD=10.0) | CAPS | 12 weeks / 12 sessions | Enhanced writing / sessions with therapist | 29 | Mixed trauma | 44 | 48 | 420 days |

CAPS, Clinician Administered PTSD Scale;1 CBT, cognitive behavioral therapy; EMDR, eye movement desensitization and reprocessing; DTS, Davidson PTSD scale,2 HTQ, Harvard Trauma Questionnaire;3 IES, Impact of Event Scale;4 IES-R, Impact of Event Scale-Revised;5 M, mean; MPSS-SR, Modified PTSD Symptoms Scale-Self Report;6 NR, nor teported; PCL, PTSD Checklist;7 PDS, Posttraumatic Diagnostic Scale;8 PPQ, Perinatal PTSD Questionnaire;9 PTSD, posttraumatic stress disorder; PSS-I, PTSD Symptom Scale Interview;10 SD, standard deviation; TSI, Trauma Symptom Inventory;11 TSC, Trauma Symptom Checklist;12

a In these studies additional groups were included, which did not match with the definition of either defined treatment group. They were omitted from the analyses.

b In these studies means and standard deviations were calculated as the average of each item (not as the sum scores of the questionnaire/scale)

c If two treatments within one category were considered very similar, they were combined.

\* Dissertation thesis

\*\* Published article and dissertation thesis identified

\*\*\* The latest follow-up data for this study was extracted from an additional published article (Thompson-Hollands J, Marx BP, Lee DJ, Resick PA,Sloan DM. Long-term treatment gains of a brief exposure -based treatment for PTSD. Depression Anxiety. 2018; 35; 985-991.).

1 Blake DD, Weathers FW, Nagy LM, et al. The development of a clinician-administered PTSD scale. J Trauma Stress. 1995;8(1):75-90.

2 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., Davison, R. M., Katz, R., & Feldman, M. (1997). Assessment of a new self-rating scale for post-traumatic stress disorder. Psychological Medicine, 27, 153-160.

3 Mollica RF, Caspi-Yavin Y, Bollini P, Truong T, Tor S, Lavelle J. The Harvard Trauma Questionnaire: validating a cross-cultural instrument for measuring torture, trauma, and posttraumatic stress disorder in Indochinese refugees. J Nerv Ment Dis. 1992.

4 Horowitz M, Wilner N, Alvarez W. Impact of Event Scale: A measure of subjective stress. Psychosomatic Medicine. 1979;41(3):209-218.

5 Weiss D, Marmar C. The Impact of Event Scale-Revised. Wilson JP, Keane TM, editors. Assessing psychological trauma and PT-SD. New York Guliford Press; 1997.

6 Falsetti SA, Resnick HS, Resick PA, Kilpatrick DG. The modified PTSD symptom scale: a brief self-report measure of posttraumatic stress disorder. The Behavior Therapist. 1993.

7 Weathers FW, Huska JA, Keane TM. PCL-C for DSM-IV. Boston: National Center for PTSD-Behavioral Science Division. 1991.

8 Foa EB, Cashman L, Jaycox L, Perry K. The validation of a self-report measure of posttraumatic stress disorder: The Posttraumatic Diagnostic Scale. Psychol Assess. 1997;9(4):445-451.

9 DeMier RL, Hynan MT, Harris HB, Manniello RL. Perinatal stressors as predictors of symptoms of posttraumatic stress in mothers of infants at high risk. J Perinatol. 1996;16(4):276–280.

10 Foa EB, Riggs DS, Dancu CV. Reliability and validity of a brief instrument for assessing posttraumatic stress disorder. Journal of Traumatic Stress. 1993; 6; 459-473.

11 Briere J, Elliott DM, Harris K, Cotman A. Trauma symptom inventory. Journal of Interpersonal Violence. 1995;10(4): 387-401.

12 Briere J, Runtz M. The Trauma Symptom Checklist (TSC-33) early data on a new scale. Journal of Interpersonal Violence. 1989;4(2):151-163.

**eTable 2.** Summary of Study Characteristics per Treatment Group

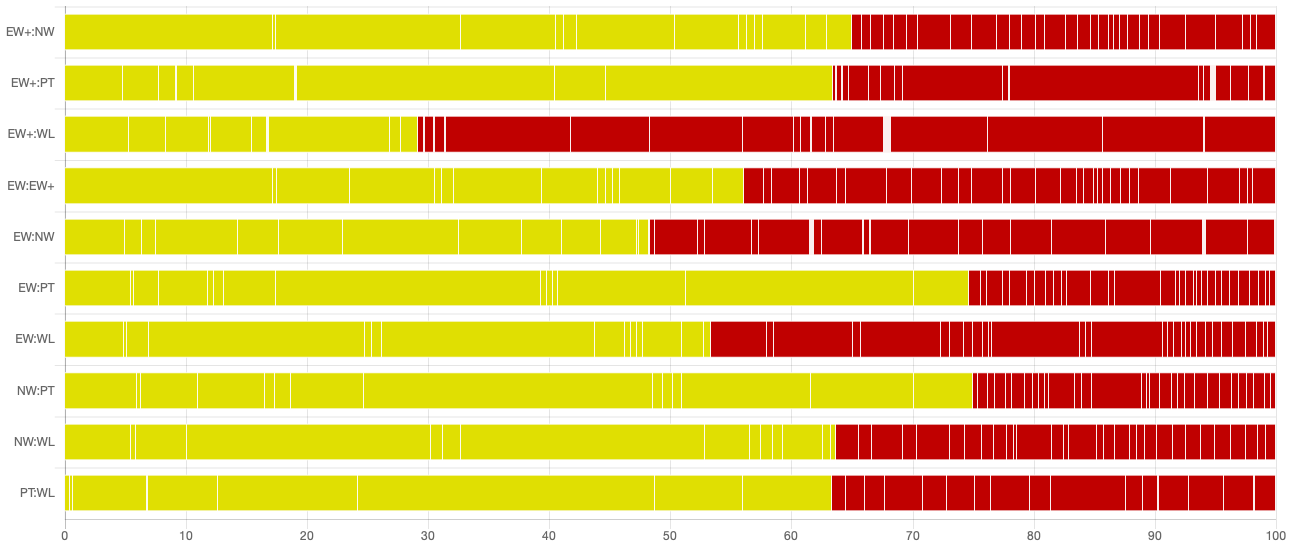
|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Treatment group** | **Number of studies** | **Days of treatment** | **Treatment sessions** | **Age** | **Gender** | **PTSD Diagnosis** | **Time to follow up** | **Sample size** | **Mixed trauma** | **Treatment described as control** | **Treatment location** |
|  |  | Mean of study means | Mean of study means | Mean of study means | Mean of study means | Percent of studies with full or subthreshholdPTSD diagnosis | Median of study means | Median of study means | Percent of studies with no restricted trauma type | Number of studies (%) | Percent of studies with home-based treatment |
| EW | 30 | 7.7 | 3.17 | 31.75 | 72.55 | 33.33 | 42 | 29 | 56.67 | 1 (3.33) | 36.67 |
| EW+ | 15 | 35 | 8.47 | 37.27 | 74.14 | 32 | 91 | 47 | 66.67 | 1 (6.67) | 60 |
| NW | 25 | 8.92 | 3.08 | 31.61 | 73.29 | 30.43 | 42 | 27 | 60 | 25 (100) | 40 |
| PT | 6 | 33 | 6.3 | 32.5 | 74.17 | 83.33 | 112 | 42 | 100 | 1 (16.67) | 0 |
| WL | 23 | - | - | 38.65 | 71.44 | 66.67 | 91 | 33 | 30.43 | 23 (100) | not applicable |
| - not assessed; EW, expressive writing; EW+, enhanced writing; NW, neutral writing; PT, psychotherapy; WL, waiting list | | | | | | | | | | | |

**eTable 3.** Risk of Bias for Each Included Study (Sorted by Publication Year)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Author** | **Publication year** | **Selection bias** | **Performance bias** | **Detection bias** | **Attrition bias** | **Reporting bias** | **Total RoB** |
| Gidron | 1996 | unclear | low | low | unclear | unclear | some concerns |
| Greenberg | 1996 | unclear | low | low | high | unclear | high |
| Barry\* | 2001 | high | high | low | unclear | high | high |
| Batten\* | 2001 | low | low | low | high | high | high |
| Lange | 2001 | low | high | low | high | unclear | high |
| Yanko\* | 2001 | unclear | unclear | low | unclear | high | high |
| Largo-Marsh\*\* | 2002 | high | unclear | low | low | unclear | high |
| Schoutrop | 2002 | unclear | high | low | low | unclear | high |
| Lange | 2003 | low | high | low | high | unclear | high |
| Deters | 2003 | unclear | low | low | high | unclear | high |
| Sloan | 2004 | unclear | low | low | unclear | unclear | some concerns |
| Koopman | 2005 | unclear | low | low | unclear | unclear | some concerns |
| Nguyen\* | 2005 | low | high | low | unclear | low | high |
| Sloan | 2005 | unclear | low | low | high | unclear | high |
| Freyd | 2005 | unclear | low | low | high | high | high |
| Wagner | 2006 | low | high | low | high | unclear | high |
| Sloan | 2007 | unclear | low | low | high | unclear | high |
| Knaevelsrud | 2007 | unclear | high | low | unclear | low | high |
| Possemato\*\* | 2008 | low | low | low | high | unclear | high |
| Resick | 2008 | unclear | low | low | high | unclear | high |
| Smyth | 2008 | unclear | low | unclear | high | high | high |
| van Emmerik | 2008 | low | unclear | low | unclear | unclear | some concerns |
| Bugg | 2009 | unclear | high | low | high | low | high |
| Kearns | 2010 | unclear | low | low | high | low | high |
| Lichtental | 2010 | high | low | low | high | unclear | high |
| Beyer | 2011 | low | low | low | unclear | low | some concerns |
| Kersting | 2011 | unclear | high | low | high | unclear | high |
| Slavin-Spenny | 2011 | low | low | low | low | unclear | some concerns |
| Sloan | 2011 | unclear | low | low | unclear | unclear | some concerns |
| Zakowski | 2011 | unclear | low | low | high | high | high |
| Ironson | 2012 | unclear | low | low | low | unclear | some concerns |
| Sloan | 2012 | low | high | low | low | low | high |
| Jensen-Johansen | 2013 | low | low | low | unclear | unclear | some concerns |
| Kersting | 2013 | unclear | high | low | unclear | low | high |
| Milbury | 2014 | low | low | low | low | unclear | some concerns |
| Southern\* | 2014 | unclear | low | low | high | low | high |
| Truijens | 2014 | unclear | low | low | unclear | low | some concerns |
| Stockton\* | 2014 | high | low | low | high | low | high |
| Sayer | 2015 | low | low | low | low | unclear | some concerns |
| Knaevelsrud | 2015 | low | high | low | low | low | high |
| Horsch | 2016 | low | high | low | unclear | low | high |
| Alessandri\* | 2017 | low | unclear | low | unclear | low | some concerns |
| Knaevelsrud | 2017 | low | high | low | low | low | high |
| Sloan | 2018 | low | unclear | low | low | low | some concerns |

\* Dissertation; \*\* Published article and dissertation thesis identified

**eFigure 2.** Risk of Bias Contributions



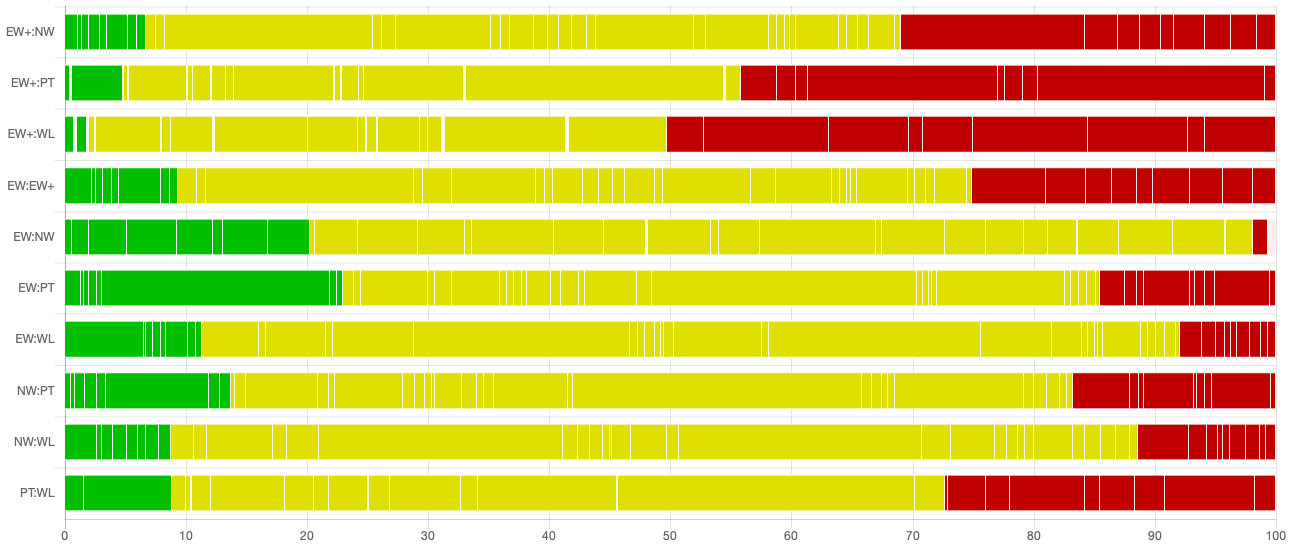
Proportion of evidence that is considered at low (green), moderate (yellow), and high risk of bias (red); EW, expressive writing; EW+, enhanced writing; NW, neutral writing; PT, psychotherapy; WL, waiting list

**eTable 4**. Indirectness for Each Included Study (Sorted by Publication Year)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Author** | **Publication year** | **Population** | **Intervention** | **Outcomes** | **Comparisons** | **Indirectness total** |
| Gidron | 1996 | low | low | low | low | low |
| Greenberg | 1996 | high | unclear | low | low | moderate |
| Barry\* | 2001 | high | low | low | low | moderate |
| Batten\* | 2001 | high | low | low | low | moderate |
| Lange | 2001 | unclear | high | low | low | moderate |
| Yanko\* | 2001 | high | low | low | low | moderate |
| Largo-Marsh\*\* | 2002 | low | high | low | low | moderate |
| Schoutrop | 2002 | unclear | unclear | low | low | moderate |
| Lange | 2003 | low | high | low | low | moderate |
| Deters | 2003 | unclear | low | low | low | low |
| Sloan | 2004 | high | low | low | low | moderate |
| Koopman | 2005 | high | low | low | low | moderate |
| Nguyen\* | 2005 | high | low | low | low | moderate |
| Sloan | 2005 | unclear | low | low | low | low |
| Freyd | 2005 | unclear | unclear | low | low | moderate |
| Wagner | 2006 | high | high | low | low | high |
| Sloan | 2007 | unclear | low | low | low | low |
| Knaevelsrud | 2007 | low | high | low | low | moderate |
| Possemato\*\* | 2008 | high | low | low | low | moderate |
| Resick | 2008 | high | high | low | low | high |
| Smyth | 2008 | low | unclear | high | low | moderate |
| van Emmerik | 2008 | low | unclear | low | low | moderate |
| Bugg | 2009 | unclear | low | low | low | low |
| Kearns | 2010 | high | low | low | low | moderate |
| Lichtental | 2010 | unclear | low | low | low | low |
| Beyer | 2011 | high | low | low | low | moderate |
| Kersting | 2011 | high | high | low | low | high |
| Slavin-Spenny | 2011 | unclear | unclear | low | low | moderate |
| Sloan | 2011 | unclear | low | unclear | low | moderate |
| Zakowski | 2011 | high | low | low | low | moderate |
| Ironson | 2012 | high | high | low | low | high |
| Sloan | 2012 | unclear | high | high | low | high |
| Jensen-Johansen | 2013 | high | low | low | low | moderate |
| Kersting | 2013 | high | high | low | low | high |
| Milbury | 2014 | high | low | low | low | moderate |
| Southern\* | 2014 | high | low | low | low | moderate |
| Truijens | 2014 | low | unclear | low | low | low |
| Stockton\* | 2014 | high | low | low | low | moderate |
| Sayer | 2015 | high | low | low | low | moderate |
| Knaevelsrud | 2015 | high | high | low | low | high |
| Horsch | 2016 | high | low | low | low | moderate |
| Alessandri\* | 2017 | unclear | low | low | low | low |
| Knaevelsrud | 2017 | high | high | low | low | high |
| Sloan | 2018 | low | high | high | low | high |

\* Dissertations; \*\* Published article and dissertation thesis identified

**eFigure 3.** Indirectness Contributions



Proportion of evidence that is contributed by studies which are considered low (green), moderate (yellow), or high red) regarding indirectness; EW, expressive writing; EW+, enhanced writing; NW, neutral writing; PT, psychotherapy; WL, waiting list

**eTable 5:** Confidence in Network Meta-analysis (CINEMA) Rating

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Comparison** | **Number of studies** | **Within-study biasa** | **Across-studies biasb** | **Indirectnessc** | **Imprecisiond** | **Heterogeneityd** | **Incoherenced** | **Confidence rating** |
| EW:EW+ | 1 | Some concerns | Undetected | Some concerns | No concerns | Some concerns | No concerns | Moderate |
| EW:NW | 24 | Major concerns | Undetected | Some concerns | No concerns | Some concerns | No concerns | Low |
| EW:PT | 2 | Some concerns | Undetected | Some concerns | No concerns | Some concerns | No concerns | Moderate |
| EW:WL | 7 | Some concerns | Undetected | Some concerns | No concerns | Some concerns | No concerns | Moderate |
| EW+:NW | 2 | Some concerns | Undetected | Some concerns | No concerns | Some concerns | Some concerns | Moderate |
| EW+:PT | 4 | Some concerns | Undetected | Some concerns | No concerns | Major concerns | No concerns | Low |
| EW+:WL | 10 | Major concerns | Undetected | Major concerns | No concerns | No concerns | Some concerns | Low |
| NW:PT | 1 | Some concerns | Undetected | Some concerns | No concerns | Some concerns | No concerns | Moderate |
| NW:WL | 2 | Some concerns | Undetected | Some concerns | No concerns | Some concerns | Some concerns | Moderate |
| PT:WL | 1 | Some concerns | Undetected | Some concerns | No concerns | No concerns | No concerns | Moderate |

EW, expressive writing; EW+, enhanced writing; NW, neutral writing; PT, psychotherapy; WL, waiting list

a Risk of bias assessment (see eAppendix 2 for a more detailed description)

b We assumed that the likelihood of unpublished data was small because we included dissertation theses and conducted our searches in 5 databases and screened the reference lists of previous relevant meta-analyses and systematic reviews.

c Indirectness rating (see eAppendix 2 for a more detailed description)

d For the evaluation of imprecision, heterogeneity and incoherence we defined the clinically important effect size as 0.6.

**eAppendix 3.** Baseline Differences

##########################################################################

> ## ------------------------------------------------------BASELINE--------

**Original data (with adjusted standard errors for multi-arm studies):**

treat1 treat2 TE seTE seTE.adj narms multiarm

ew004 EW WL 0.3726 0.3688 0.3688 2

ew005 EW NW 0.3748 0.2705 0.2705 2

ew006 EW NW 0.1658 0.2301 0.3206 3 \*

ew006 EW+ NW -0.3328 0.2040 0.2385 3 \*

ew006 EW EW+ 0.4993 0.1986 0.2289 3 \*

ew009 EW WL -0.0436 0.2451 0.2451 2

ew017 EW NW 0.0224 0.5401 0.5401 2

ew018 EW NW 0.4105 0.2580 0.2580 2

ew023 EW WL -0.0162 0.2481 0.2481 2

ew024 EW+ NW -0.0405 0.1286 0.1286 2

ew025 EW NW -0.1642 0.0909 0.2593 3 \*

ew025 EW WL -0.1580 0.0665 0.0686 3 \*

ew025 NW WL 0.0063 0.0665 0.0686 3 \*

ew026 EW NW -0.1021 0.2342 0.2342 2

ew028 EW+ WL -0.0971 0.1325 0.1325 2

ew029 EW+ WL -0.1190 0.2294 0.2294 2

ew030 EW NW 0.1454 0.2927 0.2927 2

ew032 EW+ WL 0.0398 0.2139 0.2139 2

ew034 EW+ WL 0.5592 0.4095 0.4095 2

ew036 EW+ PT -0.6936 0.4227 0.4227 2

ew044 EW NW -0.1193 0.1203 0.1203 2

ew048 EW WL 0.0048 0.2523 0.2523 2

ew054 EW NW 0.0000 0.2889 0.2889 2

ew055 EW+ PT 0.0897 0.2064 0.2064 2

ew060 EW NW 0.0063 0.0628 0.0708 3 \*

ew060 EW WL -0.0806 0.0747 0.0959 3 \*

ew060 NW WL -0.0870 0.0747 0.0960 3 \*

ew061 EW WL -0.2763 0.2912 0.2912 2

ew064 EW PT -0.2067 0.2457 0.3020 3 \*

ew064 NW PT -0.1797 0.2489 0.3103 3 \*

ew064 EW NW -0.0270 0.2392 0.2869 3 \*

ew065 EW NW 0.1602 0.2867 0.2867 2

ew066 EW+ WL -0.5442 0.3011 0.3011 2

ew067 EW NW 0.0250 0.2752 0.2752 2

ew068 EW NW 0.2211 0.2357 0.2357 2

ew069 EW NW -0.0733 0.3087 0.3087 2

ew071 EW NW 0.0000 0.4140 0.4140 2

ew073 EW NW -0.5146 0.4013 0.4013 2

ew075 EW NW -0.1285 0.3207 0.3207 2

ew078 EW+ PT 0.1068 0.2172 0.2643 3 \*

ew078 PT WL -0.1990 0.2214 0.2749 3 \*

ew078 EW+ WL -0.0923 0.2172 0.2642 3 \*

ew082 EW NW 0.0000 0.2952 0.2952 2

ew083 EW NW 0.0892 0.2134 0.2134 2

ew084 EW PT 0.2323 0.2165 0.2165 2

ew086 EW NW 0.0987 0.2654 0.2654 2

ew087 EW NW 0.0000 0.2020 0.2020 2

ew088 EW+ WL -0.0416 0.2053 0.2053 2

ew089 EW+ WL -0.0367 0.1586 0.1586 2

ew090 EW+ WL 0.0997 0.2064 0.2064 2

ew092 EW NW -0.3379 0.2874 0.2874 2

ew093 EW+ PT -0.1045 0.1783 0.1783 2

ew094 EW NW 0.7469 0.4305 0.4305 2

ew095 EW+ WL -0.3811 0.2829 0.2829 2

Number of studies: k = 44

Number of treatments: n = 5

Number of pairwise comparisons: m = 54

Number of designs: d = 10

Random effects model

Treatment estimate (sm = 'SMD'):

EW EW+ NW PT WL

EW . 0.0628 -0.0053 0.0288 -0.0601

EW+ -0.0628 . -0.0680 -0.0339 -0.1229

NW 0.0053 0.0680 . 0.0341 -0.0549

PT -0.0288 0.0339 -0.0341 . -0.0890

WL 0.0601 0.1229 0.0549 0.0890 .

Lower 95%-confidence limit:

EW EW+ NW PT WL

EW . -0.0572 -0.0766 -0.1583 -0.1375

EW+ -0.1827 . -0.1859 -0.2116 -0.2293

NW -0.0661 -0.0498 . -0.1546 -0.1338

PT -0.2159 -0.1437 -0.2228 . -0.2739

WL -0.0172 0.0165 -0.0241 -0.0960 .

Upper 95%-confidence limit:

EW EW+ NW PT WL

EW . 0.1827 0.0661 0.2159 0.0172

EW+ 0.0572 . 0.0498 0.1437 -0.0165

NW 0.0766 0.1859 . 0.2228 0.0241

PT 0.1583 0.2116 0.1546 . 0.0960

WL 0.1375 0.2293 0.1338 0.2739 .

Quantifying heterogeneity / inconsistency:

tau^2 = 0; I^2 = 0%

Tests of heterogeneity (within designs) and inconsistency (between designs):

Q d.f. p-value

Total 37.53 45 0.7778

Within designs 27.73 35 0.8041

Between designs 9.80 10 0.4585

**eAppendix4:** Aditional Results from Network Meta-analyses: Main Analyses

##########################################################################

> ## ---------------------------------------Longest available FU----------

**Original data (with adjusted standard errors for multi-arm studies):**

treat1 treat2 TE seTE seTE.adj narms multiarm

ew004 EW WL -0.7098 0.3782 0.3782 2

ew005 EW NW 0.6342 0.2751 0.2751 2

ew006 EW NW -0.0393 0.2298 0.3234 3 \*

ew006 EW+ NW -0.1480 0.2030 0.2369 3 \*

ew006 EW EW+ 0.1086 0.1959 0.2246 3 \*

ew009 EW WL -0.3373 0.2827 0.2827 2

ew017 EW NW 0.6849 0.5612 0.5612 2

ew018 EW NW 0.1985 0.2559 0.2559 2

ew023 EW WL -0.4167 0.2755 0.2755 2

ew024 EW+ NW 0.0088 0.1523 0.1523 2

ew025 EW NW -0.0329 0.0965 0.2693 3 \*

ew025 EW WL -0.0396 0.0719 0.0745 3 \*

ew025 NW WL -0.0066 0.0695 0.0717 3 \*

ew026 EW NW 0.1367 0.2344 0.2344 2

ew028 EW+ WL -0.8815 0.1388 0.1388 2

ew029 EW+ WL -0.8398 0.2742 0.2742 2

ew030 EW NW 0.1027 0.2925 0.2925 2

ew032 EW+ WL -1.1878 0.2302 0.2302 2

ew034 EW+ WL -0.4949 0.4075 0.4075 2

ew036 EW+ PT -0.4755 0.4151 0.4151 2

ew044 EW NW -0.0755 0.1645 0.1645 2

ew048 EW WL 0.0048 0.2523 0.2523 2

ew054 EW NW -0.4202 0.2924 0.2924 2

ew055 EW+ PT 0.1974 0.2317 0.2317 2

ew060 EW NW -0.1076 0.0628 0.0709 3 \*

ew060 EW WL -0.1794 0.0748 0.0961 3 \*

ew060 NW WL -0.0717 0.0747 0.0958 3 \*

ew061 EW WL -1.1720 0.3154 0.3154 2

ew064 EW PT -0.0000 0.2450 0.3007 3 \*

ew064 NW PT -0.0821 0.2484 0.3098 3 \*

ew064 EW NW 0.0822 0.2393 0.2874 3 \*

ew065 EW NW -0.7802 0.2978 0.2978 2

ew066 EW+ WL -2.6373 0.4121 0.4121 2

ew067 EW NW -1.4261 0.3106 0.3106 2

ew068 EW NW -0.5475 0.2390 0.2390 2

ew069 EW NW 0.1272 0.3090 0.3090 2

ew071 EW NW 0.0000 0.4494 0.4494 2

ew073 EW NW -0.4957 0.4008 0.4008 2

ew075 EW NW -0.5631 0.3273 0.3273 2

ew078 EW+ PT -0.0000 0.2171 0.2604 3 \*

ew078 PT WL -0.6197 0.2264 0.2830 3 \*

ew078 EW+ WL -0.6200 0.2225 0.2727 3 \*

ew082 EW NW 0.1494 0.2956 0.2956 2

ew083 EW NW 0.2860 0.2144 0.2144 2

ew084 EW PT 0.7125 0.2228 0.2228 2

ew086 EW NW 0.2364 0.2663 0.2663 2

ew087 EW NW 0.0000 0.2250 0.2250 2

ew088 EW+ WL -0.9314 0.2166 0.2166 2

ew089 EW+ WL -0.9165 0.1669 0.1669 2

ew090 EW+ WL -0.4163 0.2086 0.2086 2

ew092 EW NW -0.1451 0.2862 0.2862 2

ew093 EW+ PT 0.1462 0.1784 0.1784 2

ew094 EW NW 0.2618 0.4161 0.4161 2

ew095 EW+ WL -1.2042 0.3064 0.3064 2

Number of studies: k = 44

Number of treatments: n = 5

Number of pairwise comparisons: m = 54

Number of designs: d = 10

Random effects model

Treatment estimate (sm = 'SMD'):

EW EW+ NW PT WL

EW . 0.3831 -0.0584 0.3460 -0.4314

EW+ -0.3831 . -0.4414 -0.0370 -0.8145

NW 0.0584 0.4414 . 0.4044 -0.3731

PT -0.3460 0.0370 -0.4044 . -0.7775

WL 0.4314 0.8145 0.3731 0.7775 .

Lower 95%-confidence limit:

EW EW+ NW PT WL

EW . 0.1305 -0.2092 0.0233 -0.6504

EW+ -0.6356 . -0.7015 -0.3373 -1.0191

NW -0.0925 0.1814 . 0.0717 -0.6080

PT -0.6687 -0.2633 -0.7371 . -1.0984

WL 0.2124 0.6099 0.1382 0.4565 .

Upper 95%-confidence limit:

EW EW+ NW PT WL

EW . 0.6356 0.0925 0.6687 -0.2124

EW+ -0.1305 . -0.1814 0.2633 -0.6099

NW 0.2092 0.7015 . 0.7346 -0.1382

PT -0.0233 0.3373 -0.0717 . -0.4565

WL 0.6504 1.0191 0.6080 1.0984 .

Quantifying heterogeneity / inconsistency:

tau^2 = 0.0792; I^2 = 67.6%

Tests of heterogeneity (within designs) and inconsistency (between designs):

Q d.f. p-value

Total 139.05 45 < 0.0001

Within designs 89.82 35 < 0.0001

Between designs 49.23 10 < 0.0001

**> # Inconsistency**

> net4$d

[1] 10

>

> designs4 = as.character(decomp.design(net3)$Q.het.design$design)

> designs4

[1] "EW:NW" "EW:PT" "EW:WL" "EW+:NW" "EW+:PT" "EW+:WL" "EW:EW+:NW" "EW:NW:PT" "EW:NW:WL" "EW+:PT:WL"

>

> split4 = netsplit(net4)

> print(split4, showall = FALSE, digits = 2)

Back-calculation method to split direct and indirect evidence

Random effects model:

comparison k prop nma direct indir. Diff z p-value

EW:EW+ 1 0.14 0.38 0.11 0.43 -0.32 -0.86 0.3878

EW:NW 24 0.94 -0.06 -0.08 0.23 -0.31 -0.92 0.3594

EW:PT 2 0.40 0.35 0.37 0.33 0.04 0.12 0.9041

EW:WL 7 0.67 -0.43 -0.32 -0.66 0.34 1.41 0.1576

EW+:NW 2 0.32 -0.44 -0.06 -0.62 0.55 1.95 0.0516

EW+:PT 4 0.67 -0.04 0.03 -0.18 0.21 0.64 0.5233

EW+:WL 10 0.78 -0.81 -0.94 -0.37 -0.57 -2.27 0.0232

NW:PT 1 0.20 0.40 -0.08 0.53 -0.61 -1.45 0.1463

NW:WL 2 0.34 -0.37 -0.04 -0.55 0.51 2.00 0.0453

PT:WL 1 0.21 -0.78 -0.62 -0.82 0.20 0.49 0.6242

Legend:

comparison - Treatment comparison

k - Number of studies providing direct evidence

prop - Direct evidence proportion

nma - Estimated treatment effect (SMD) in network meta-analysis

direct - Estimated treatment effect (SMD) derived from direct evidence

indir. - Estimated treatment effect (SMD) derived from indirect evidence

Diff - Difference between direct and indirect treatment estimates

z - z-value of test for disagreement (direct versus indirect)

p-value - p-value of test for disagreement (direct versus indirect)

Warnmeldung:

In print.netsplit(split4, showall = FALSE, digits = 2) :

Deprecated argument 'showall' has been replaced by argument 'show'.

>

> decomp.design(net4)

Q statistics to assess homogeneity / consistency

Q df p-value

Total 139.05 45 < 0.0001

Within designs 89.82 35 < 0.0001

Between designs 49.23 10 < 0.0001

Design-specific decomposition of within-designs Q statistic

Design Q df p-value

EW:NW 50.03 19 0.0001

EW:WL 9.13 4 0.0578

EW+:PT 2.17 2 0.3458

EW+:WL 26.66 8 0.0008

EW:NW:WL 1.83 2 0.3997

Between-designs Q statistic after detaching of single designs

Detached design Q df p-value

EW:NW 49.21 9 < 0.0001

EW:PT 47.66 9 < 0.0001

EW:WL 44.99 9 < 0.0001

EW+:NW 32.90 9 0.0001

EW+:PT 47.70 9 < 0.0001

EW+:WL 18.28 9 0.0321

EW:EW+:NW 42.64 8 < 0.0001

EW:NW:PT 41.14 8 < 0.0001

EW:NW:WL 14.87 8 0.0618

EW+:PT:WL 49.15 8 < 0.0001

Q statistic to assess consistency under the assumption of

a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within

Between designs 13.17 10 0.2144 0.2556 0.0653

**eAppendix5:** Additional Results from Network Meta-analyses: Sensitivity Analyses

**## Only end of treatment data**

Number of studies: k = 34

Number of treatments: n = 5

Number of pairwise comparisons: m = 36

Number of designs: d = 7

Random effects model

Treatment estimate (sm = 'SMD'):

EW EW+ NW PT WL

EW . 0.5075 -0.0350 0.6275 -0.3989

EW+ -0.5075 . -0.5425 0.1200 -0.9063

NW 0.0350 0.5425 . 0.6625 -0.3639

PT -0.6275 -0.1200 -0.6625 . -1.0264

WL 0.3989 0.9063 0.3639 1.0264 .

Lower 95%-confidence limit:

EW EW+ NW PT WL

EW . 0.0675 -0.2797 0.0986 -0.8067

EW+ -0.9474 . -1.0135 -0.3011 -1.1887

NW -0.2097 0.0714 . 0.1008 -0.8148

PT -1.1565 -0.5411 -1.2242 . -1.4902

WL -0.0090 0.6240 -0.0871 0.5626 .

Upper 95%-confidence limit:

EW EW+ NW PT WL

EW . 0.9474 0.2097 1.1565 0.0090

EW+ -0.0675 . -0.0714 0.5411 -0.6240

NW 0.2797 1.0135 . 1.2242 0.0871

PT -0.0986 0.3011 -0.1008 . -0.5626

WL 0.8067 1.1887 0.8148 1.4902 .

Quantifying heterogeneity / inconsistency:

tau^2 = 0.1682; I^2 = 73.8%

Tests of heterogeneity (within designs) and inconsistency (between designs):

Q d.f. p-value

Total 118.14 31 < 0.0001

Within designs 103.35 27 < 0.0001

Between designs 14.78 4 0.0052

**## Only follow-up data**

Number of studies: k = 26

Number of treatments: n = 5

Number of pairwise comparisons: m = 36

Number of designs: d = 10

Random effects model

Treatment estimate (sm = 'SMD'):

EW EW+ NW PT WL

EW . 0.3111 -0.0562 0.3081 -0.4233

EW+ -0.3111 . -0.3674 -0.0031 -0.7344

NW 0.0562 0.3674 . 0.3643 -0.3671

PT -0.3081 0.0031 -0.3643 . -0.7314

WL 0.4233 0.7344 0.3671 0.7314 .

Lower 95%-confidence limit:

EW EW+ NW PT WL

EW . -0.0137 -0.2399 -0.0341 -0.6853

EW+ -0.6360 . -0.6901 -0.3203 -1.0889

NW -0.1275 0.0446 . 0.0137 -0.6473

PT -0.6503 -0.3142 -0.7149 . -1.1095

WL 0.1613 0.3800 0.0869 0.3533 .

Upper 95%-confidence limit:

EW EW+ NW PT WL

EW . 0.6360 0.1275 0.6503 -0.1613

EW+ 0.0137 . -0.0446 0.3142 -0.3800

NW 0.2399 0.6901 . 0.7149 -0.0869

PT 0.0341 0.3203 -0.0137 . -0.3533

WL 0.6853 1.0889 0.6473 1.1095 .

Quantifying heterogeneity / inconsistency:

tau^2 = 0.0806; I^2 = 71.3%

Tests of heterogeneity (within designs) and inconsistency (between designs):

Q d.f. p-value

Total 94.14 27 < 0.0001

Within designs 42.20 17 0.0006

Between designs 51.94 10 < 0.0001

**## Direct comparisons between EW+ and WL excluded**

Number of studies: k = 35

Number of treatments: n = 5

Number of pairwise comparisons: m = 45

Number of designs: d = 9

Random effects model

Treatment estimate (sm = 'SMD'):

EW EW+ NW PT WL

EW . 0.1239 -0.0712 0.2294 -0.2862

EW+ -0.1239 . -0.1952 0.1055 -0.4102

NW 0.0712 0.1952 . 0.3007 -0.2150

PT -0.2294 -0.1055 -0.3007 . -0.5157

WL 0.2862 0.4102 0.2150 0.5157 .

Lower 95%-confidence limit:

EW EW+ NW PT WL

EW . -0.1553 -0.2021 -0.0629 -0.4859

EW+ -0.4032 . -0.4719 -0.1685 -0.7215

NW -0.0596 -0.0816 . 0.0037 -0.4297

PT -0.5217 -0.3795 -0.5976 . -0.8391

WL 0.0866 0.0988 0.0003 0.1922 .

Upper 95%-confidence limit:

EW EW+ NW PT WL

EW . 0.4032 0.0596 0.5217 -0.0866

EW+ 0.1553 . 0.0816 0.3795 -0.0988

NW 0.2021 0.4719 . 0.5976 -0.0003

PT 0.0629 0.1685 -0.0037 . -0.1922

WL 0.4859 0.7215 0.4297 0.8391 .

Quantifying heterogeneity / inconsistency:

tau^2 = 0.0460; I^2 = 55.8%

Tests of heterogeneity (within designs) and inconsistency (between designs):

Q d.f. p-value

Total 81.44 36 < 0.0001

Within designs 63.17 27 < 0.0001

Between designs 18.28 9 0.0321

**## Studies with imputed SDs excluded**

Number of studies: k = 40

Number of treatments: n = 5

Number of pairwise comparisons: m = 48

Number of designs: d = 10

Random effects model

Treatment estimate (sm = 'SMD'):

EW EW+ NW PT WL

EW . 0.3341 -0.0763 0.3019 -0.5143

EW+ -0.3341 . -0.4104 -0.0321 -0.8484

NW 0.0763 0.4104 . 0.3782 -0.4380

PT -0.3019 0.0321 -0.3782 . -0.8162

WL 0.5143 0.8484 0.4380 0.8162 .

Lower 95%-confidence limit:

EW EW+ NW PT WL

EW . 0.0336 -0.2607 -0.0694 -0.7894

EW+ -0.6345 . -0.7246 -0.3744 -1.0843

NW -0.1081 0.0962 . -0.0083 -0.7379

PT -0.6732 -0.3101 -0.7648 . -1.1858

WL 0.2392 0.6125 0.1381 0.4467 .

Upper 95%-confidence limit:

EW EW+ NW PT WL

EW . 0.6345 0.1081 0.6732 -0.2392

EW+ -0.0336 . -0.0962 0.3101 -0.6125

NW 0.2607 0.7246 . 0.7648 -0.1381

PT 0.0694 0.3744 0.0083 . -0.4467

WL 0.7894 1.0843 0.7379 1.1858 .

Quantifying heterogeneity / inconsistency:

tau^2 = 0.1179; I^2 = 70.5%

Tests of heterogeneity (within designs) and inconsistency (between designs):

Q d.f. p-value

Total 135.78 40 < 0.0001

Within designs 85.45 30 < 0.0001

Between designs 50.33 10 < 0.0001

**## Studies with high indirectness excluded**

Number of studies: k = 35

Number of treatments: n = 5

Number of pairwise comparisons: m = 45

Number of designs: d = 9

Random effects model

Treatment estimate (sm = 'SMD'):

EW EW+ NW PT WL

EW . 0.4091 -0.0770 0.2653 -0.3296

EW+ -0.4091 . -0.4861 -0.1438 -0.7387

NW 0.0770 0.4861 . 0.3423 -0.2526

PT -0.2653 0.1438 -0.3423 . -0.5949

WL 0.3296 0.7387 0.2526 0.5949 .

Lower 95%-confidence limit:

EW EW+ NW PT WL

EW . 0.1074 -0.2159 -0.0725 -0.5357

EW+ -0.7108 . -0.7974 -0.5203 -1.0220

NW -0.0618 0.1748 . -0.0070 -0.4767

PT -0.6031 -0.2327 -0.6916 . -0.9495

WL 0.1236 0.4555 0.0285 0.2403 .

Upper 95%-confidence limit:

EW EW+ NW PT WL

EW . 0.7108 0.0618 0.6031 -0.1236

EW+ -0.1074 . -0.1748 0.2327 -0.4555

NW 0.2159 0.7974 . 0.6916 -0.0285

PT 0.0725 0.5203 0.0070 . -0.2403

WL 0.5357 1.0220 0.4767 0.9495 .

Quantifying heterogeneity / inconsistency:

tau^2 = 0.0561; I^2 = 59.8%

Tests of heterogeneity (within designs) and inconsistency (between designs):

Q d.f. p-value

Total 89.53 36 < 0.0001

Within designs 63.28 27 < 0.0001

Between designs 26.25 9 0.0019

**## Studies with observer-rated outcomes excluded**

Number of studies: k = 41

Number of treatments: n = 5

Number of pairwise comparisons: m = 51

Number of designs: d = 10

Random effects model

Treatment estimate (sm = 'SMD'):

EW EW+ NW PT WL

EW . 0.3416 -0.0611 0.2944 -0.3917

EW+ -0.3416 . -0.4027 -0.0472 -0.7333

NW 0.0611 0.4027 . 0.3555 -0.3306

PT -0.2944 0.0472 -0.3555 . -0.6861

WL 0.3917 0.7333 0.3306 0.6861 .

Lower 95%-confidence limit:

EW EW+ NW PT WL

EW . 0.1008 -0.2045 -0.0273 -0.5964

EW+ -0.5823 . -0.6493 -0.3660 -0.9304

NW -0.0822 0.1561 . 0.0244 -0.5496

PT -0.6160 -0.2717 -0.6867 . -1.0126

WL 0.1870 0.5361 0.1115 0.3596 .

Upper 95%-confidence limit:

EW EW+ NW PT WL

EW . 0.5823 0.0822 0.6160 -0.1870

EW+ -0.1008 . -0.1561 0.2717 -0.5361

NW 0.2045 0.6493 . 0.6867 -0.1115

PT 0.0273 0.3660 -0.0244 . -0.3596

WL 0.5964 0.9304 0.5496 1.0126 .

Quantifying heterogeneity / inconsistency:

tau^2 = 0.0637; tau = 0.2523; I^2 = 63.5% [49.4%; 73.6%]

Tests of heterogeneity (within designs) and inconsistency (between designs):

Q d.f. p-value

Total 114.98 42 < 0.0001

Within designs 71.83 32 < 0.0001

Between designs 43.16 10 < 0.0001

**## Studies with experimental psychotherapeutic PTSD treatment excluded**

Number of studies: k = 42

Number of treatments: n = 5

Number of pairwise comparisons: m = 50

Number of designs: d = 8

Random effects model

Treatment estimate (sm = 'SMD'):

EW EW+ NW PT WL

EW . 0.3944 -0.0740 0.3914 -0.4256

EW+ -0.3944 . -0.4684 -0.0030 -0.8200

NW 0.0740 0.4684 . 0.4654 -0.3516

PT -0.3914 0.0030 -0.4654 . -0.8170

WL 0.4256 0.8200 0.3516 0.8170 .

Lower 95%-confidence limit:

EW EW+ NW PT WL

EW . 0.1214 -0.2286 -0.0433 -0.6538

EW+ -0.6674 . -0.7466 -0.3567 -1.0281

NW -0.0807 0.1901 . 0.0267 -0.5943

PT -0.8260 -0.3507 -0.9041 . -1.2098

WL 0.1974 0.6119 0.1089 0.4242 .

Upper 95%-confidence limit:

EW EW+ NW PT WL

EW . 0.6674 0.0807 0.8260 -0.1974

EW+ -0.1214 . -0.1901 0.3507 -0.6119

NW 0.2286 0.7466 . 0.9041 -0.1089

PT 0.0433 0.3567 -0.0267 . -0.4242

WL 0.6538 1.0281 0.5943 1.2098 .

Quantifying heterogeneity / inconsistency:

tau^2 = 0.0788; tau = 0.2808; I^2 = 67.8% [55.9%; 76.5%]

Tests of heterogeneity (within designs) and inconsistency (between designs):

Q d.f. p-value

Total 130.60 42 < 0.0001

Within designs 89.82 35 < 0.0001

Between designs 40.77 7 < 0.0001

**## Studies reporting only enhanced PTSD symptoms excluded**

Number of studies: k = 19

Number of treatments: n = 5

Number of pairwise comparisons: m = 23

Number of designs: d = 7

Random effects model

Treatment estimate (sm = 'SMD'):

EW EW+ NW PT WL

EW . 0.3482 -0.3446 0.3877 -0.6034

EW+ -0.3482 . -0.6928 0.0395 -0.9516

NW 0.3446 0.6928 . 0.7323 -0.2588

PT -0.3877 -0.0395 -0.7323 . -0.9910

WL 0.6034 0.9516 0.2588 0.9910 .

Lower 95%-confidence limit:

EW EW+ NW PT WL

EW . -0.1830 -0.7004 -0.1921 -1.1725

EW+ -0.8794 . -1.2831 -0.3808 -1.3038

NW -0.0111 0.1026 . 0.0877 -0.8904

PT -0.9675 -0.4597 -1.3769 . -1.4932

WL 0.0343 0.5994 -0.3729 0.4889 .

Upper 95%-confidence limit:

EW EW+ NW PT WL

EW . 0.8794 0.0111 0.9675 -0.0343

EW+ 0.1830 . -0.1026 0.4597 -0.5994

NW 0.7004 1.2831 . 1.3769 0.3729

PT 0.1921 0.3808 -0.0877 . -0.4889

WL 1.1725 1.3038 0.8904 1.4932 .

Quantifying heterogeneity / inconsistency:

tau^2 = 0.1622; tau = 0.4027; I^2 = 71.7% [54.5%; 82.4%]

Tests of heterogeneity (within designs) and inconsistency (between designs):

Q d.f. p-value

Total 60.00 17 < 0.0001

Within designs 46.23 12 < 0.0001

Between designs 13.77 5 0.0171

**eAppendix 6.** Additional Results from Pairwise Meta-analyses on Symptom Severity

**End of treatment**

Psychotherapy vs enhanced writing

Number of studies combined: k = 4

SMD 95%-CI z p-value

Random effects model -0.2852 [-0.5493; -0.0211] -2.12 0.0343

Quantifying heterogeneity:

tau^2 = 0.0172; H = 1.14 [1.00; 2.92]; I^2 = 23.3% [0.0%; 88.3%]

Test of heterogeneity:

Q d.f. p-value

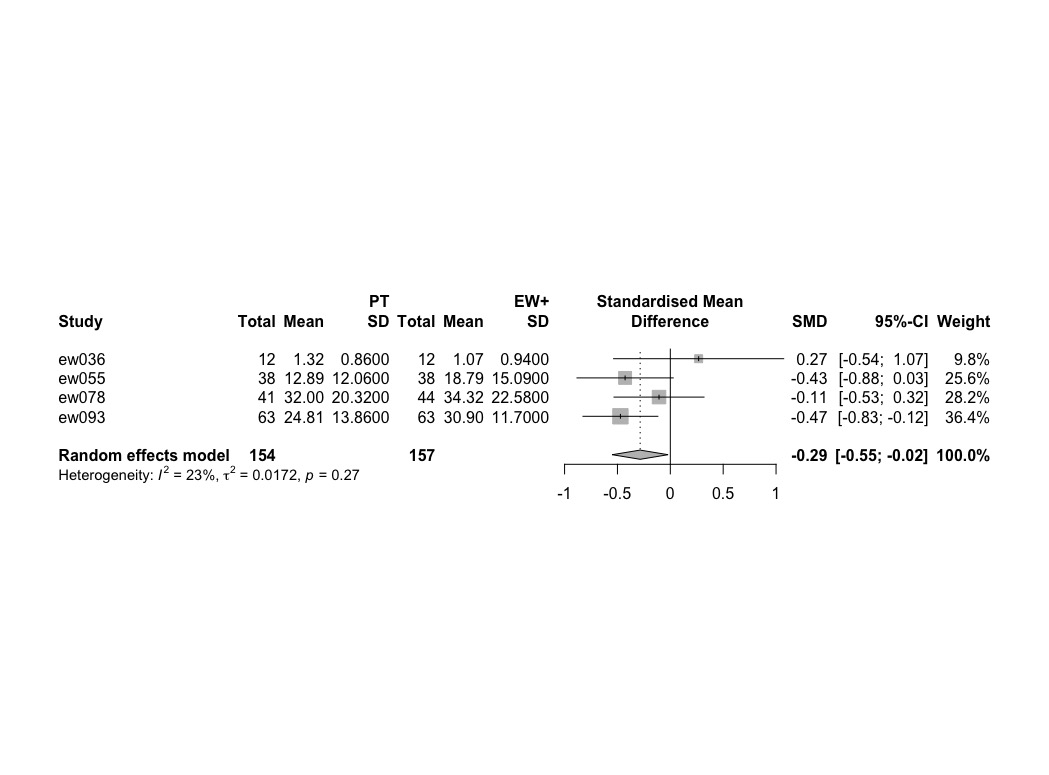
3.91 3 0.2711

Details on meta-analytical method:

- Inverse variance method

- DerSimonian-Laird estimator for tau^2

- Hedges' g (bias corrected standardised mean difference)



Psychotherapy vs expressive writing

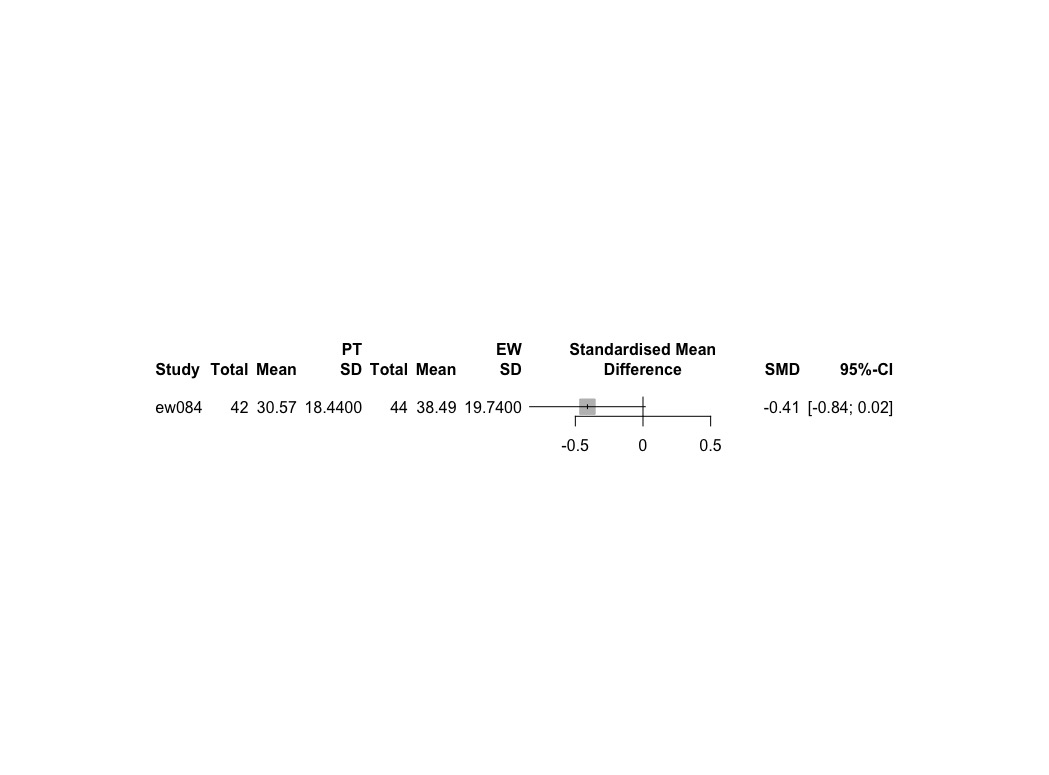
SMD 95%-CI z p-value

-0.4106 [-0.8380; 0.0169] -1.88 0.0597

Details:

- Inverse variance method

- Hedges' g (bias corrected standardised mean difference)



Psychotherapy vs neutral writing

*No studies*

Psychotherapy vs waiting list

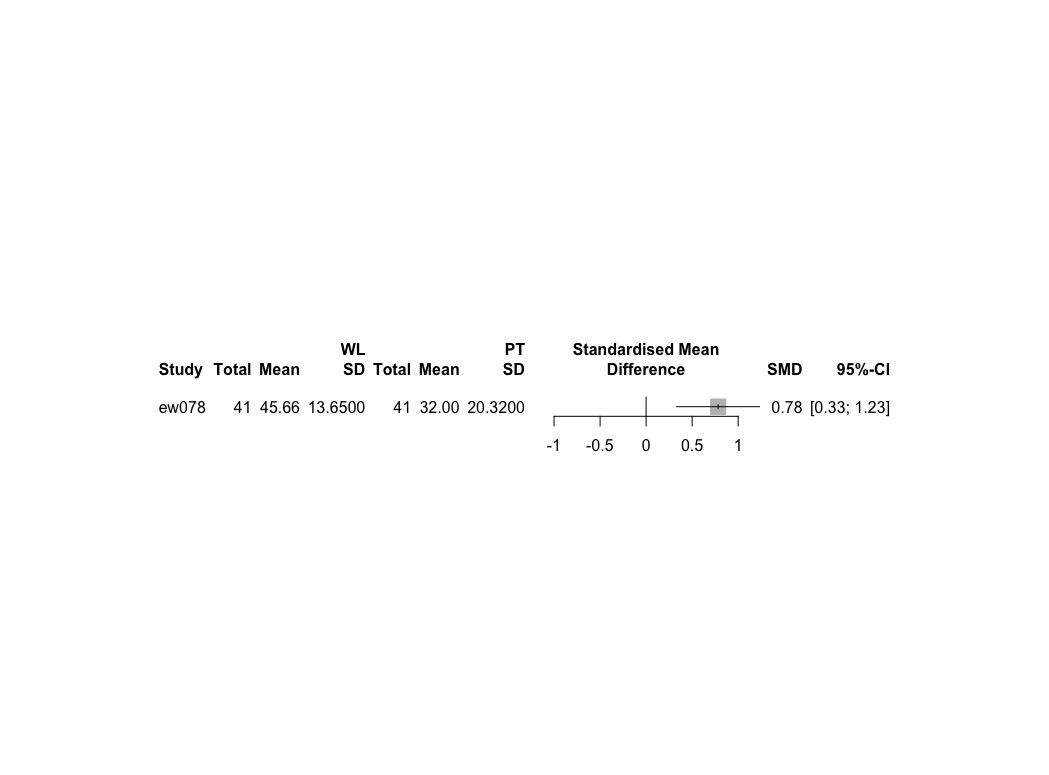
SMD 95%-CI z p-value

0.7817 [0.3318; 1.2317] 3.41 0.0007

Details:

- Inverse variance method

- Hedges' g (bias corrected standardised mean difference)



Enhanced writing vs expressive writing

*No studies*

Enhanced writing vs neutral writing

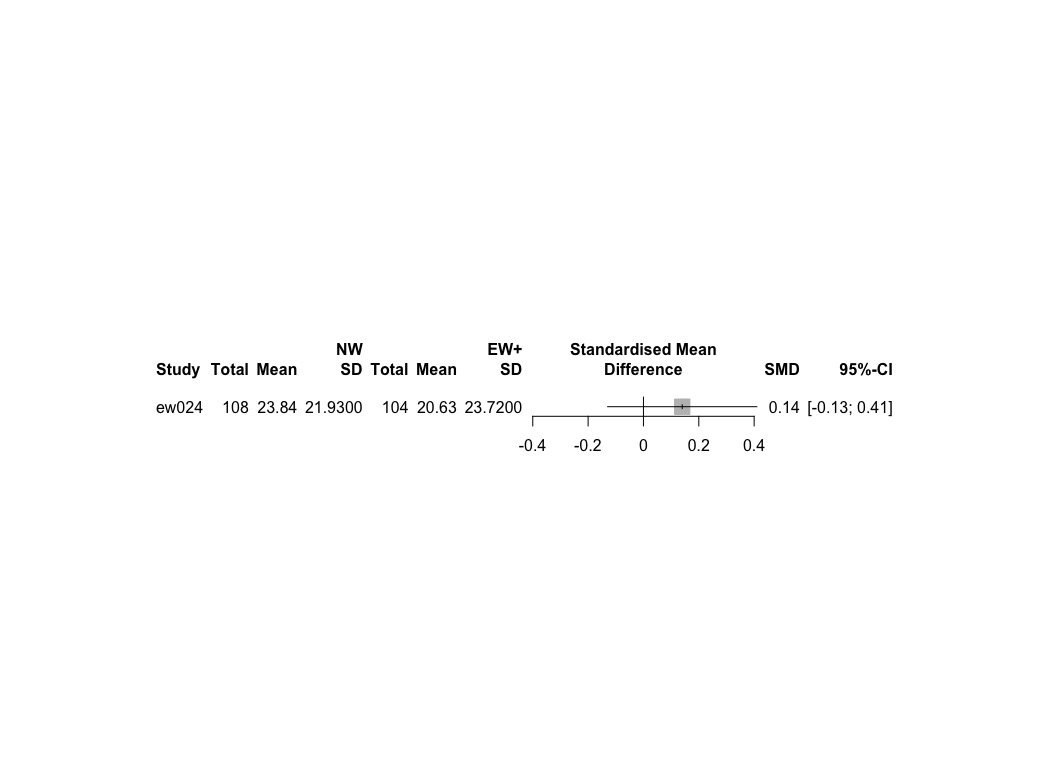
SMD 95%-CI z p-value

0.1401 [-0.1295; 0.4097] 1.02 0.3083

Details:

- Inverse variance method

- Hedges' g (bias corrected standardised mean difference)



Enhanced writing vs waiting list

Number of studies combined: k = 10

SMD 95%-CI z p-value

Random effects model 0.9966 [0.6891; 1.3042] 6.35 < 0.0001

Quantifying heterogeneity:

tau^2 = 0.1777; H = 2.11 [1.56; 2.86]; I^2 = 77.6% [59.0%; 87.8%]

Test of heterogeneity:

Q d.f. p-value

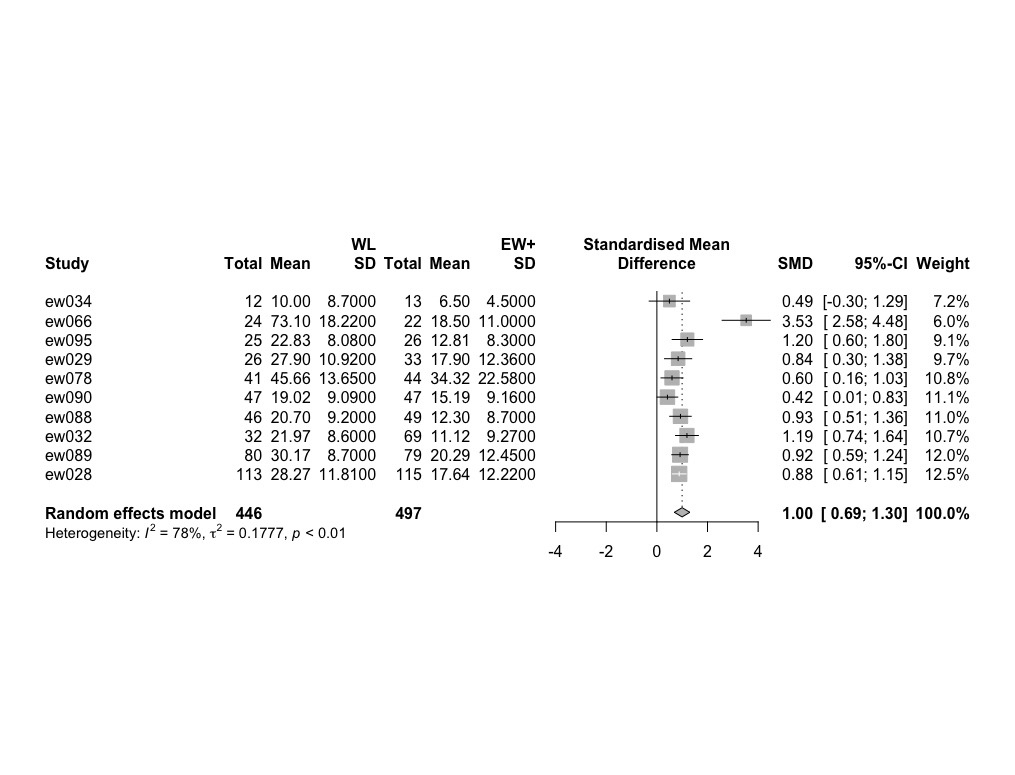
40.25 9 < 0.0001

Details on meta-analytical method:

- Inverse variance method

- DerSimonian-Laird estimator for tau^2

- Hedges' g (bias corrected standardised mean difference)



Expressive writing vs neutral writing

Number of studies combined: k = 15

SMD 95%-CI z p-value

Random effects model 0.0702 [-0.2079; 0.3483] 0.49 0.6206

Quantifying heterogeneity:

tau^2 = 0.2202; H = 2.03 [1.58; 2.61]; I^2 = 75.8% [60.1%; 85.3%]

Test of heterogeneity:

Q d.f. p-value

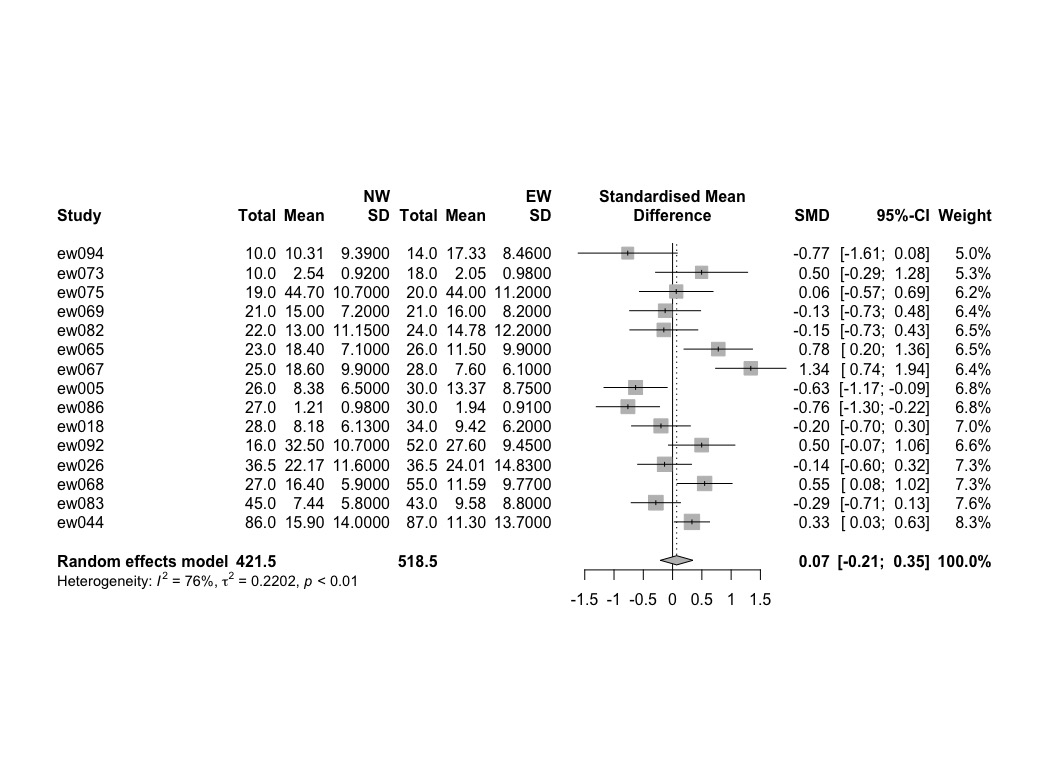
57.89 14 < 0.0001

Details on meta-analytical method:

- Inverse variance method

- DerSimonian-Laird estimator for tau^2

- Hedges' g (bias corrected standardised mean difference)



Expressive writing vs waiting list

Number of studies combined: k = 4

SMD 95%-CI z p-value

Random effects model 0.1798 [-0.1643; 0.5239] 1.02 0.3058

Quantifying heterogeneity:

tau^2 = 0.0403; H = 1.22 [1.00; 2.05]; I^2 = 32.7% [0.0%; 76.1%]

Test of heterogeneity:

Q d.f. p-value

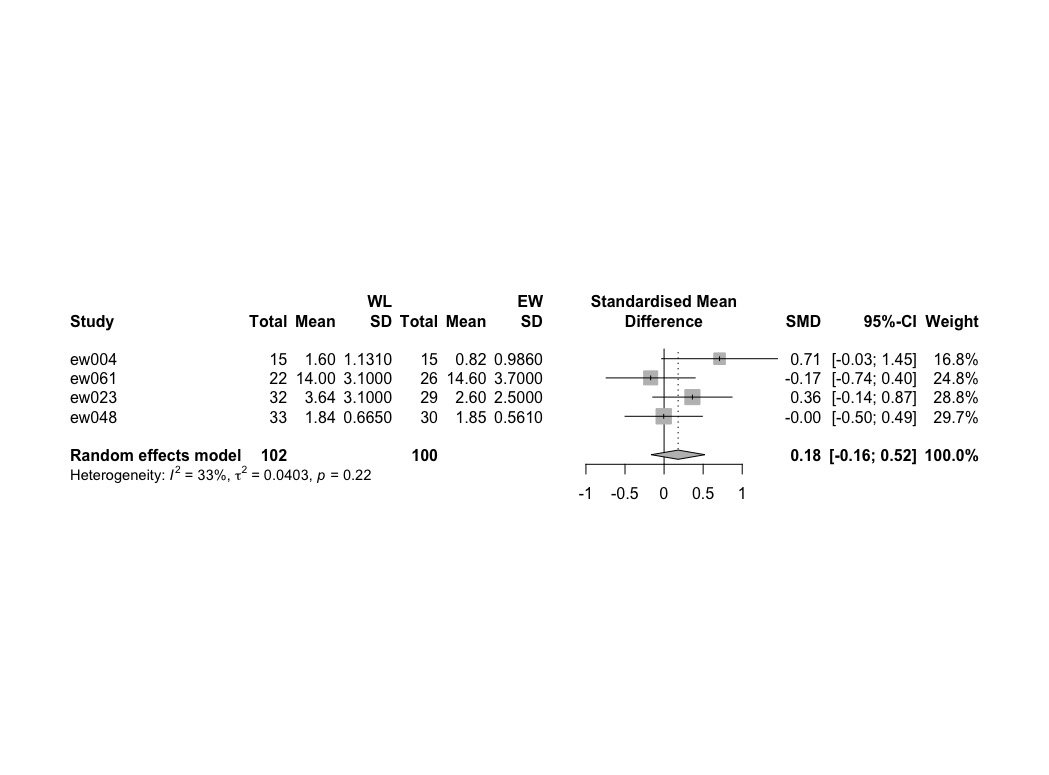
4.46 3 0.2161

Details on meta-analytical method:

- Inverse variance method

- DerSimonian-Laird estimator for tau^2

- Hedges' g (bias corrected standardised mean difference)



Neutral writing vs waiting list

*No studies*

**Longest available follow-up**

Psychotherapy vs enhanced writing

Number of studies combined: k = 4

SMD 95%-CI z p-value

Random effects model -0.0715 [-0.2948; 0.1518] -0.63 0.5303

Quantifying heterogeneity:

tau^2 = 0; H = 1.00 [1.00; 2.25]; I^2 = 0.0% [0.0%; 80.2%]

Test of heterogeneity:

Q d.f. p-value

2.32 3 0.5095

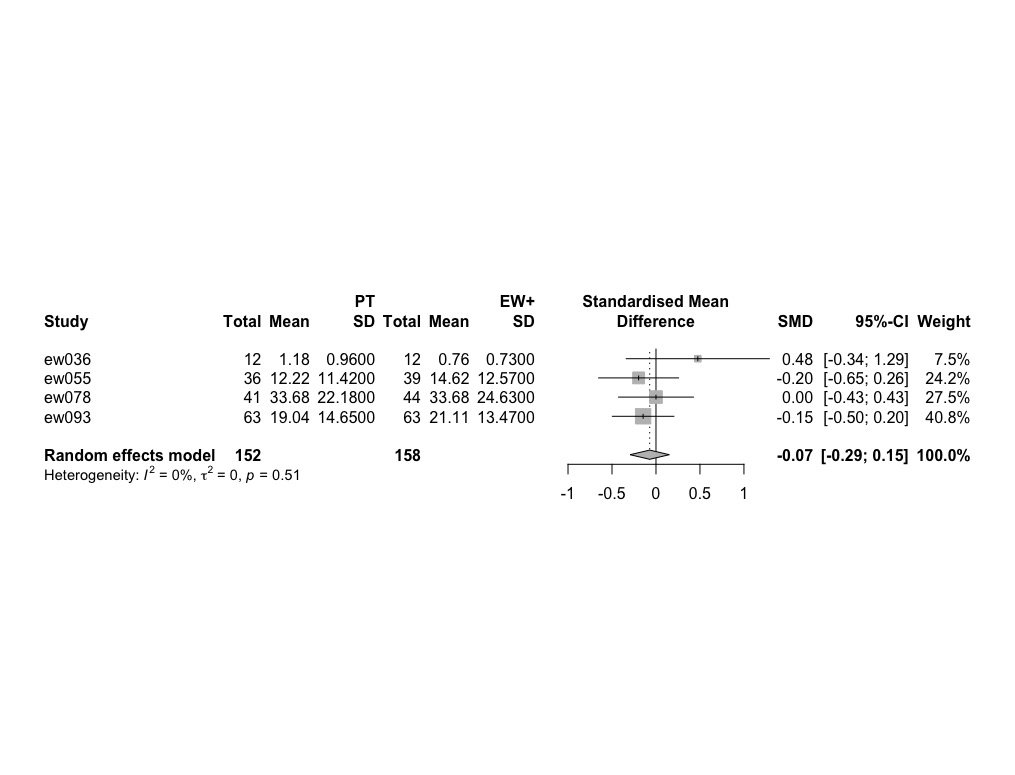
Details on meta-analytical method:

- Inverse variance method

- DerSimonian-Laird estimator for tau^2

- Hedges' g (bias corrected standardised mean difference)

> summary(pooledSMD92)



Psychotherapy vs expressive writing

Number of studies combined: k = 2

SMD 95%-CI z p-value

Random effects model -0.3636 [-1.0616; 0.3345] -1.02 0.3074

Quantifying heterogeneity:

tau^2 = 0.1990; H = 2.15; I^2 = 78.4% [6.0%; 95.0%]

Test of heterogeneity:

Q d.f. p-value

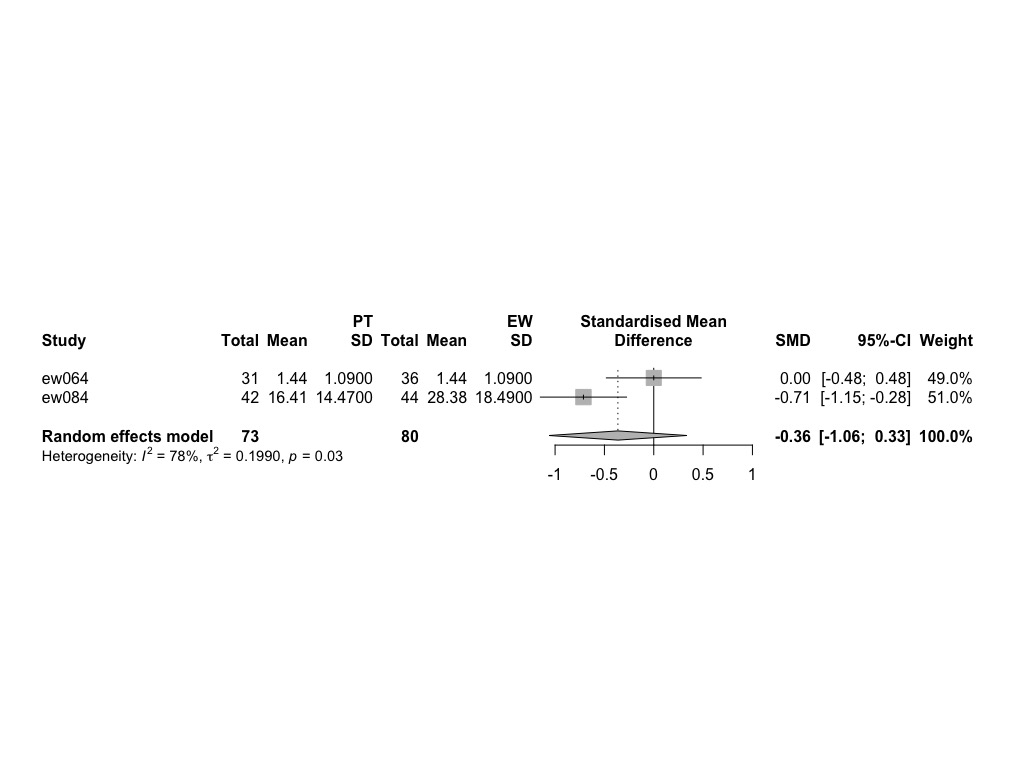
4.63 1 0.0314

Details on meta-analytical method:

- Inverse variance method

- DerSimonian-Laird estimator for tau^2

- Hedges' g (bias corrected standardised mean difference)



Psychotherapy vs neutral writing

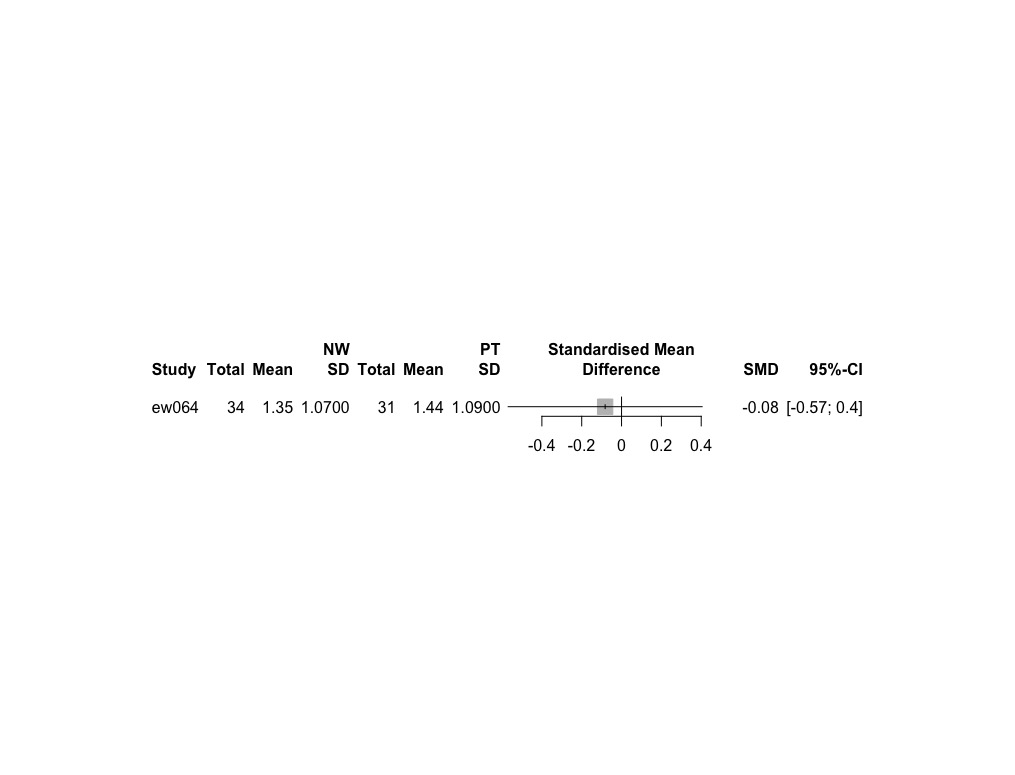
SMD 95%-CI z p-value

-0.0824 [-0.5693; 0.4046] -0.33 0.7402

Details:

- Inverse variance method

- Hedges' g (bias corrected standardised mean difference)



Psychotherapy vs waiting list

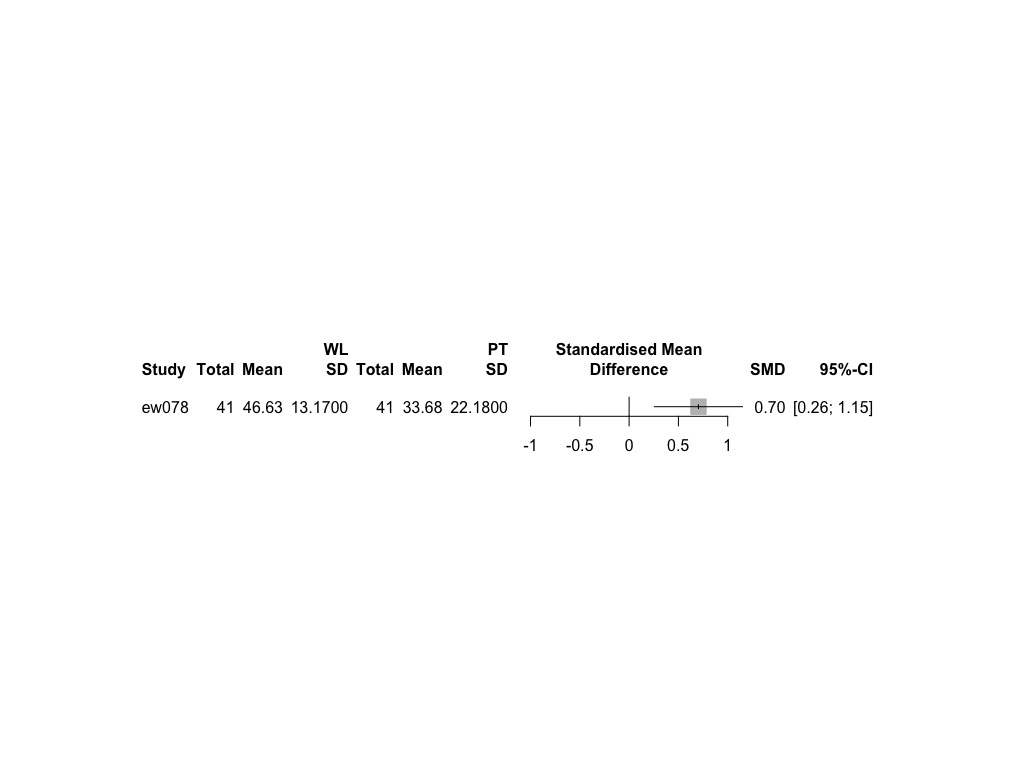
SMD 95%-CI z p-value

0.7033 [0.2566; 1.1500] 3.09 0.0020

Details:

- Inverse variance method

- Hedges' g (bias corrected standardised mean difference)



Enhanced writing vs expressive writing

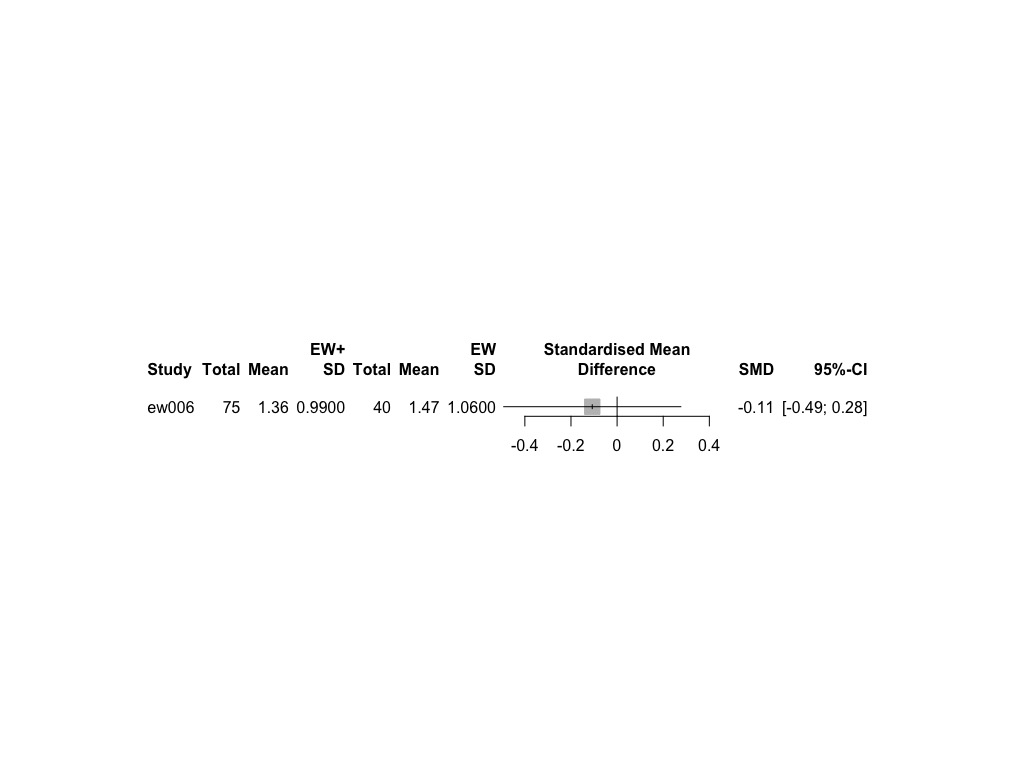
SMD 95%-CI z p-value

-0.1077 [-0.4917; 0.2763] -0.55 0.5826

Details:

- Inverse variance method

- Hedges' g (bias corrected standardised mean difference)



Enhanced writing vs neutral writing

Number of studies combined: k = 2

SMD 95%-CI z p-value

Random effects model 0.0487 [-0.1900; 0.2875] 0.40 0.6892

Quantifying heterogeneity:

tau^2 = 0; H = 1.00; I^2 = 0.0%

Test of heterogeneity:

Q d.f. p-value

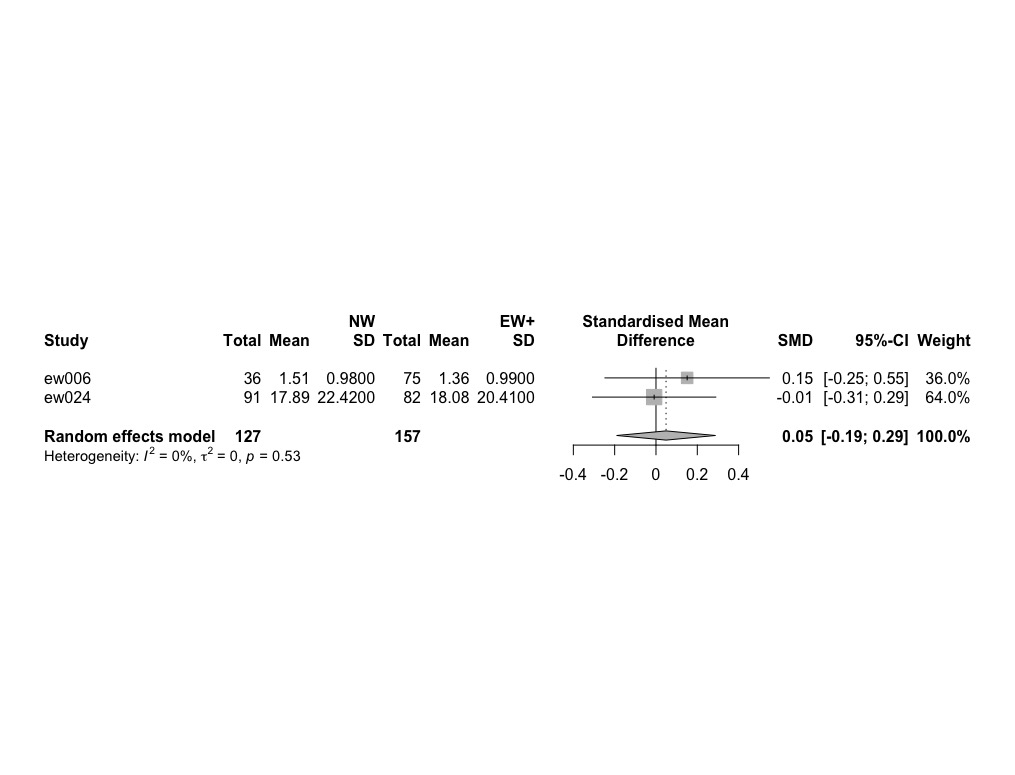
0.40 1 0.5290

Details on meta-analytical method:

- Inverse variance method

- DerSimonian-Laird estimator for tau^2

- Hedges' g (bias corrected standardised mean difference)



Enhanced writing vs waiting list

Number of studies combined: k = 10

SMD 95%-CI z p-value

Random effects model 0.9529 [0.6969; 1.2089] 7.30 < 0.0001

Quantifying heterogeneity:

tau^2 = 0.1076; H = 1.77 [1.27; 2.46]; I^2 = 68.0% [38.0%; 83.5%]

Test of heterogeneity:

Q d.f. p-value

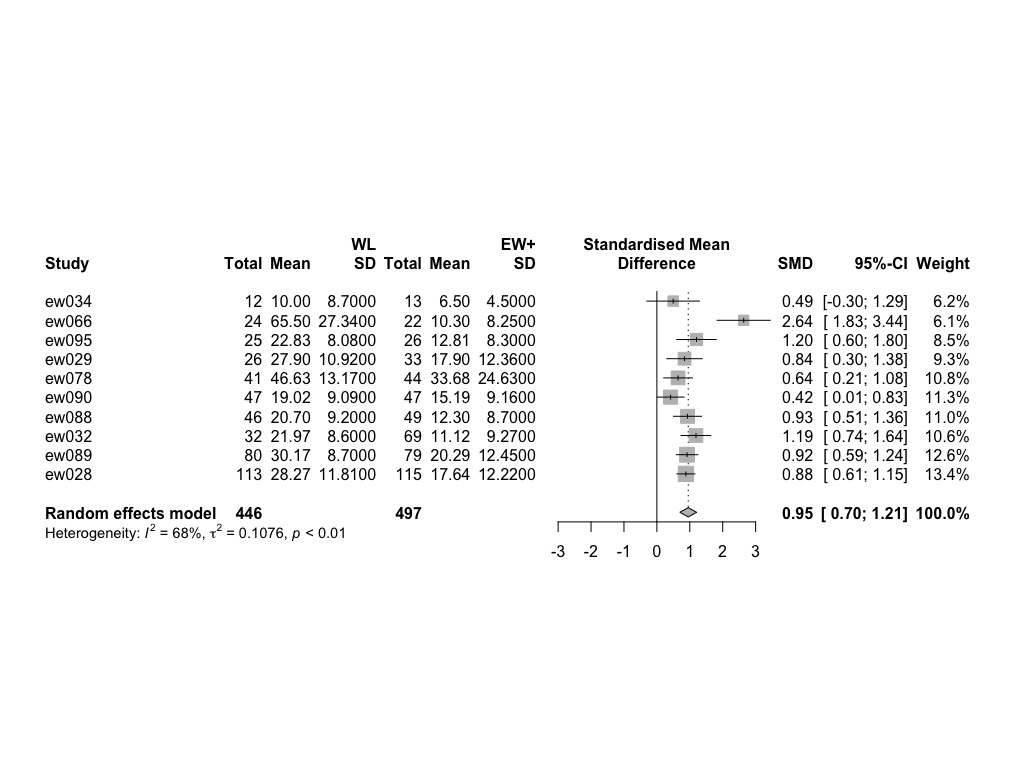
28.11 9 0.0009

Details on meta-analytical method:

- Inverse variance method

- DerSimonian-Laird estimator for tau^2

- Hedges' g (bias corrected standardised mean difference)



Expressive writing vs neutral writing

Number of studies combined: k = 24

SMD 95%-CI z p-value

Random effects model 0.0734 [-0.0648; 0.2116] 1.04 0.2980

Quantifying heterogeneity:

tau^2 = 0.0516; H = 1.49 [1.18; 1.87]; I^2 = 54.8% [28.4%; 71.5%]

Test of heterogeneity:

Q d.f. p-value

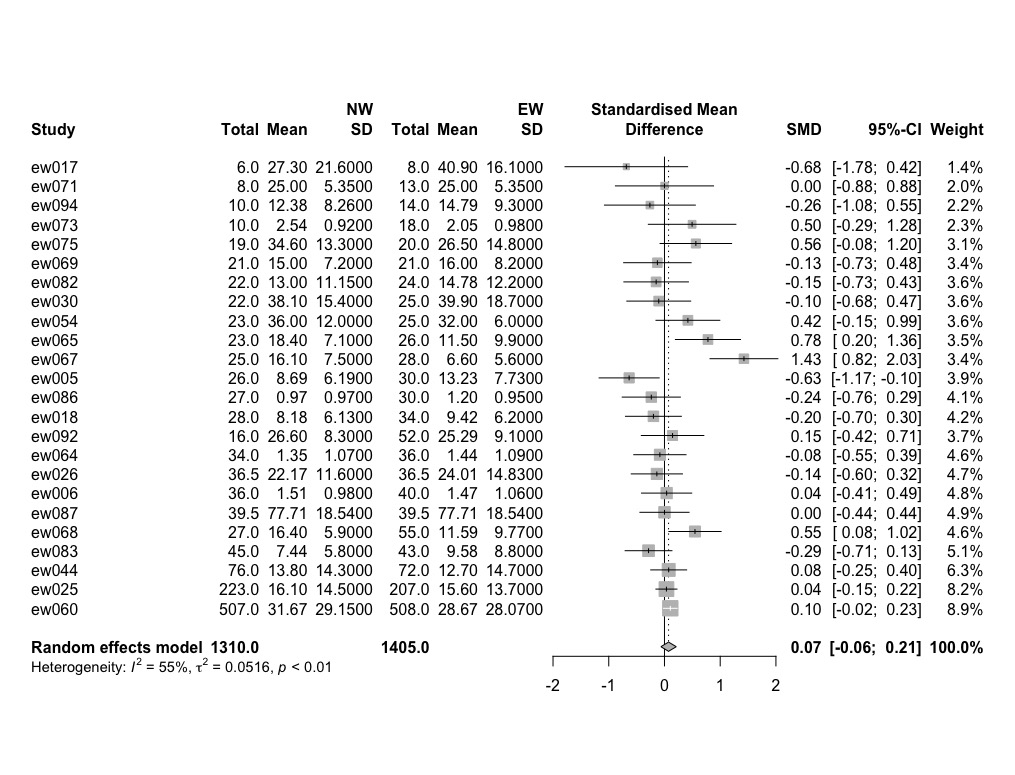
50.88 23 0.0007

Details on meta-analytical method:

- Inverse variance method

- DerSimonian-Laird estimator for tau^2

- Hedges' g (bias corrected standardised mean difference)



Expressive writing vs waiting list

Number of studies combined: k = 7

SMD 95%-CI z p-value

Random effects model 0.2846 [0.0680; 0.5012] 2.58 0.0100

Quantifying heterogeneity:

tau^2 = 0.0411; H = 1.68 [1.12; 2.53]; I^2 = 64.7% [20.6%; 84.3%]

Test of heterogeneity:

Q d.f. p-value

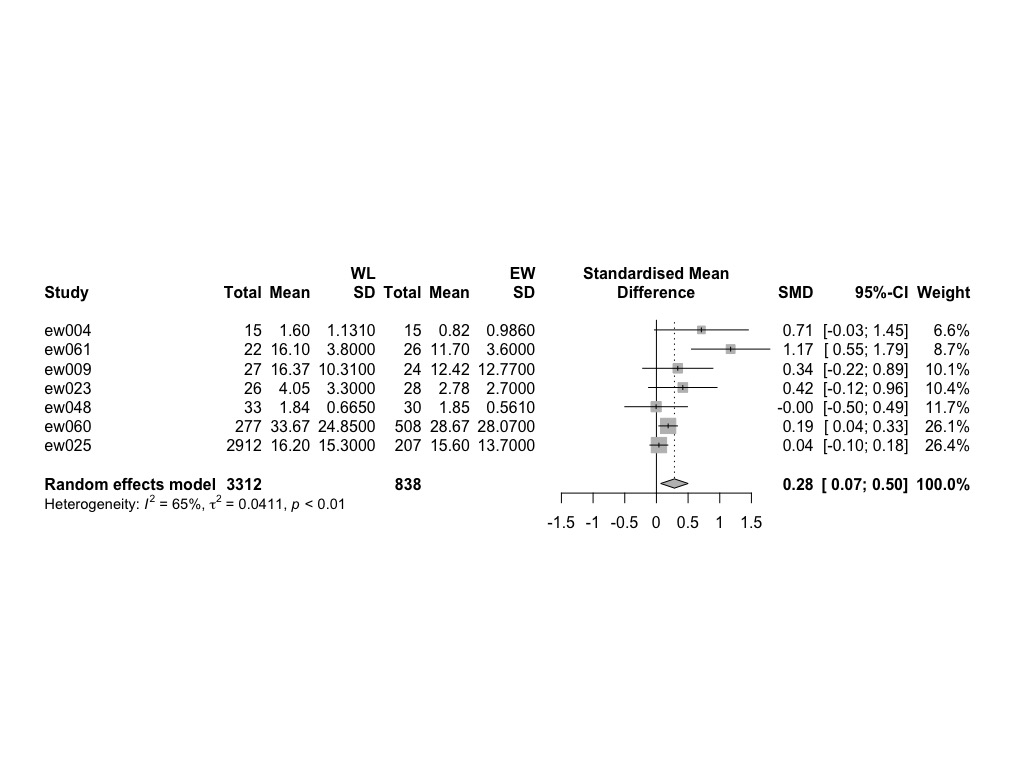
17.01 6 0.0093

Details on meta-analytical method:

- Inverse variance method

- DerSimonian-Laird estimator for tau^2

- Hedges' g (bias corrected standardised mean difference)



Neutral writing vs waiting list

Number of studies combined: k = 2

SMD 95%-CI z p-value

Random effects model 0.0369 [-0.0628; 0.1367] 0.73 0.4678

Quantifying heterogeneity:

tau^2 = 0; H = 1.00; I^2 = 0.0%

Test of heterogeneity:

Q d.f. p-value

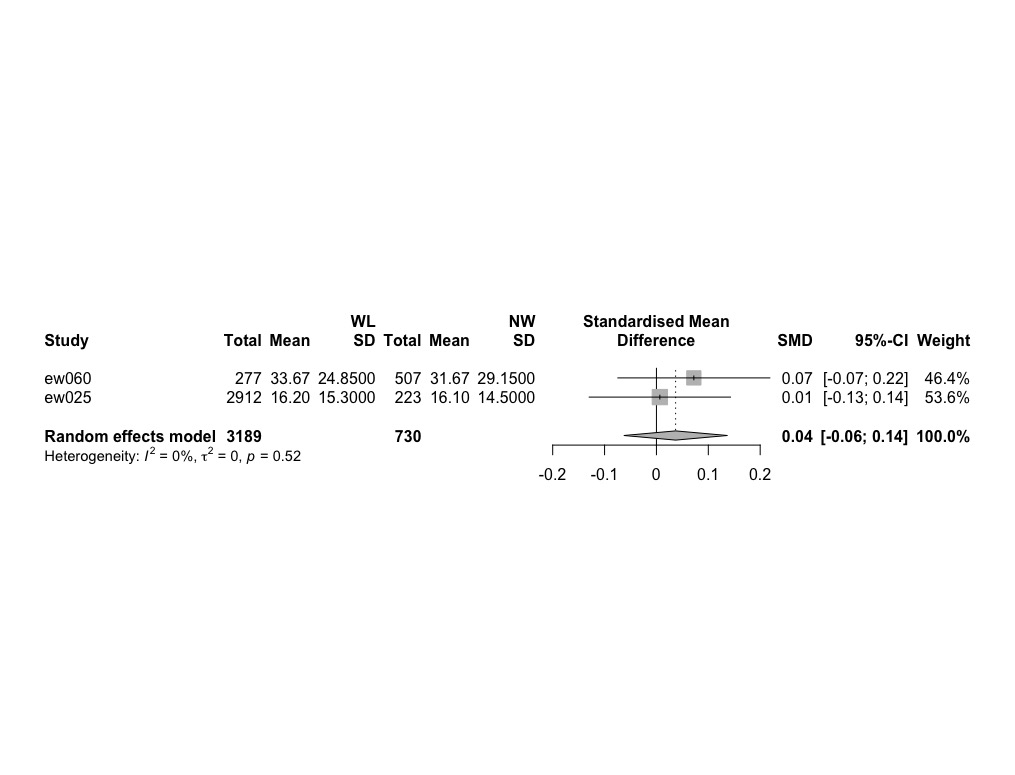
0.41 1 0.5206

Details on meta-analytical method:

- Inverse variance method

- DerSimonian-Laird estimator for tau^2

- Hedges' g (bias corrected standardised mean difference)



**eAppendix 7:** Additional Results from Netowrk Meta-analysis:Comparative Acceptability

##########################################################################

> ## ------------------------------------------Drop-outs--------------

**Original data (with adjusted standard errors for multi-arm studies):**

**treat1 treat2 TE seTE seTE.adj narms multiarm**

**ew004 EW WL 0.0000 0.5164 0.5164 2**

**ew005 EW NW 0.5500 0.8235 0.8235 2**

**ew006 EW NW -1.4917 1.0942 1.6941 3 \***

**ew006 EW+ NW -0.1744 0.5931 0.6175 3 \***

**ew006 EW EW+ -1.3173 1.0510 1.2934 3 \***

**ew018 EW NW -0.8873 1.1979 1.1979 2**

**ew023 EW WL 2.0487 1.4793 1.4793 2**

**ew024 EW+ NW 0.0165 0.3706 0.3706 2**

**ew025 EW NW 1.6094 0.6259 0.6259 2**

**ew028 EW+ WL 0.1901 0.3491 0.3491 2**

**ew029 EW+ WL 0.2288 0.4167 0.4167 2**

**ew034 EW+ WL -0.4855 0.8204 0.8204 2**

**ew044 EW NW -0.1141 0.2154 0.2154 2**

**ew048 EW WL 0.0953 1.9840 1.9840 2**

**ew054 EW NW -1.1820 1.1180 1.1180 2**

**ew055 EW+ PT 0.2198 0.3665 0.3665 2**

**ew060 EW NW 0.2725 0.1208 0.1208 2**

**ew064 EW PT -0.8427 0.8306 1.0140 3 \***

**ew064 NW PT -0.3801 0.7223 0.8135 3 \***

**ew064 EW NW -0.4626 0.8810 1.1980 3 \***

**ew065 EW NW -0.1226 1.9794 1.9794 2**

**ew066 EW+ WL 1.4733 1.5533 1.5533 2**

**ew067 EW NW -0.1133 1.9810 1.9810 2**

**ew069 EW NW 0.0000 1.9760 1.9760 2**

**ew071 EW NW 0.3567 1.6818 1.6818 2**

**ew073 EW NW 0.0606 0.1850 0.1850 2**

**ew075 EW NW -0.0513 1.9742 1.9742 2**

**ew084 EW PT 0.6466 1.7186 1.7186 2**

**ew086 EW NW -0.1054 1.9823 1.9823 2**

**ew088 EW+ WL 2.0163 1.0406 1.0406 2**

**ew089 EW+ WL -0.0182 0.1908 0.1908 2**

**ew090 EW+ WL 0.6931 0.6763 0.6763 2**

**ew093 EW+ PT -1.8326 0.5082 0.5082 2**

**ew094 EW NW 1.0090 0.5583 0.5583 2**

**ew095 EW+ WL 1.0594 1.1202 1.1202 2**

Number of studies: k = 31

Number of treatments: n = 5

Number of pairwise comparisons: m = 35

Number of designs: d = 8

Random effects model

Treatment estimate (sm = 'RR'):

EW EW+ NW PT WL

EW . 1.1019 1.1701 0.6572 1.3472

EW+ 0.9075 . 1.0619 0.5964 1.2226

NW 0.8546 0.9417 . 0.5616 1.1513

PT 1.5217 1.6768 1.7806 . 2.0500

WL 0.7423 0.8179 0.8686 0.4878 .

Lower 95%-confidence limit:

EW EW+ NW PT WL

EW . 0.6320 0.8950 0.3175 0.7402

EW+ 0.5205 . 0.6205 0.3299 0.8585

NW 0.6537 0.5502 . 0.2744 0.6374

PT 0.7352 0.9275 0.8701 . 1.0406

WL 0.4079 0.5744 0.4809 0.2476 .

Upper 95%-confidence limit:

EW EW+ NW PT WL

EW . 1.9211 1.5298 1.3602 2.4517

EW+ 1.5822 . 1.8176 1.0782 1.7410

NW 1.1173 1.6117 . 1.1493 2.0795

PT 3.1497 3.0313 3.6440 . 4.0383

WL 1.3509 1.1648 1.5690 0.9609 .

Quantifying heterogeneity / inconsistency:

tau^2 = 0.0463; I^2 = 17.7%

Tests of heterogeneity (within designs) and inconsistency (between designs):

Q d.f. p-value

Total 35.24 29 0.1969

Within designs 31.96 23 0.1009

Between designs 3.28 6 0.7736

**> # Inconsistency**

Back-calculation method to split direct and indirect evidence

Random effects model:

comparison k prop nma direct indir. RoR z p-value

EW:EW+ 1 0.07 1.10 0.27 1.23 0.22 -1.37 0.1716

EW:NW 16 0.96 1.17 1.17 1.19 0.98 -0.03 0.9788

EW:PT 2 0.23 0.66 0.58 0.68 0.84 -0.19 0.8477

EW:WL 3 0.36 1.35 1.27 1.39 0.92 -0.14 0.8890

EW+:NW 2 0.60 1.06 0.96 1.24 0.77 -0.46 0.6440

EW+:PT 2 0.80 0.60 0.58 0.67 0.87 -0.18 0.8534

EW+:WL 8 0.91 1.22 1.23 1.13 1.09 0.14 0.8890

NW:PT 1 0.24 0.56 0.68 0.53 1.29 0.30 0.7651

Legend:

comparison - Treatment comparison

k - Number of studies providing direct evidence

prop - Direct evidence proportion

nma - Estimated treatment effect (RR) in network meta-analysis

direct - Estimated treatment effect (RR) derived from direct evidence

indir. - Estimated treatment effect (RR) derived from indirect evidence

RoR - Ratio of Ratios (direct versus indirect)

z - z-value of test for disagreement (direct versus indirect)

p-value - p-value of test for disagreement (direct versus indirect)>

Q statistics to assess homogeneity / consistency

Q df p-value

Total 35.24 29 0.1969

Within designs 31.96 23 0.1009

Between designs 3.28 6 0.7736

Design-specific decomposition of within-designs Q statistic

Design Q df p-value

EW:NW 12.85 13 0.4591

EW:WL 1.71 2 0.4246

EW+:PT 10.73 1 0.0011

EW+:WL 6.66 7 0.4646

Between-designs Q statistic after detaching of single designs

Detached design Q df p-value

EW:NW 2.30 5 0.8066

EW:PT 2.92 5 0.7124

EW:WL 3.27 5 0.6588

EW+:NW 3.25 5 0.6611

EW+:PT 3.27 5 0.6584

EW+:WL 3.27 5 0.6588

EW:EW+:NW 0.92 4 0.9212

EW:NW:PT 2.72 4 0.6062

Q statistic to assess consistency under the assumption of

a full design-by-treatment interaction random effects model

Q df p-value tau.within tau2.within

Between designs 3.16 6 0.7887 0.2751 0.0757

**eAppendix 8.** Additional Results from Pairwise Meta-analyses on Comparative Acceptability (Drop-out rates between beginning and end of treatment)

Psychotherapy vs enhanced writing

RR 95%-CI z p-value

0.8027 [0.3913; 1.6464] -0.60 0.5487

Details:

- Inverse variance method

Psychotherapy vs expressive writing

Number of studies combined: k = 2

RR 95%-CI z p-value

Random effects model 1.7518 [0.4045; 7.5870] 0.75 0.4535

Quantifying heterogeneity:

tau^2 = 0; H = 1.00; I^2 = 0.0%

Test of heterogeneity:

Q d.f. p-value

0.61 1 0.4347

Details on meta-analytical method:

- Mantel-Haenszel method

- DerSimonian-Laird estimator for tau^2

Psychotherapy vs neutral writing

RR 95%-CI z p-value

1.4624 [0.3550; 6.0235] 0.53 0.5987

Details:

- Inverse variance method

Psychotherapy vs waiting list

*No studies*

Enhanced writing vs expressive writing

RR 95%-CI z p-value

3.7333 [0.4759; 29.2872] 1.25 0.2101

Details:

- Inverse variance method

Enhanced writing vs neutral writing

Number of studies combined: k = 2

RR 95%-CI z p-value

Random effects model 0.9636 [0.5205; 1.7840] -0.12 0.9061

Quantifying heterogeneity:

tau^2 = 0; H = 1.00; I^2 = 0.0%

Test of heterogeneity:

Q d.f. p-value

0.07 1 0.7849

Details on meta-analytical method:

- Mantel-Haenszel method

- DerSimonian-Laird estimator for tau^2

Enhanced writing vs waiting list

Number of studies combined: k = 8

RR 95%-CI z p-value

Random effects model 1.1542 [0.8618; 1.5457] 0.96 0.3360

Quantifying heterogeneity:

tau^2 = 0.0027; H = 1.01 [1.00; 1.77]; I^2 = 1.1% [0.0%; 67.9%]

Test of heterogeneity:

Q d.f. p-value

7.08 7 0.4206

Details on meta-analytical method:

- Mantel-Haenszel method

- DerSimonian-Laird estimator for tau^2

Expressive writing vs neutral writing

Number of studies combined: k = 16

RR 95%-CI z p-value

Random effects model 1.1747 [0.9589; 1.4390] 1.55 0.1200

Quantifying heterogeneity:

tau^2 = 0.0100; H = 1.03 [1.00; 1.49]; I^2 = 5.4% [0.0%; 54.9%]

Test of heterogeneity:

Q d.f. p-value

15.86 15 0.3913

Expressive writing vs waiting list

Number of studies combined: k = 3

RR 95%-CI z p-value

Random effects model 1.2402 [0.4903; 3.1369] 0.45 0.6493

Quantifying heterogeneity:

tau^2 = 0; H = 1.00 [1.00; 3.08]; I^2 = 0.0% [0.0%; 89.5%]

Test of heterogeneity:

Q d.f. p-value

1.98 2 0.3722

Neutral writing vs waiting list

*No studies*