TECHNICAL COMMENTARY

NeuRA
Discover. Conquer. Cure.
POST-TRAUMATIC STRESS DISORDER LIBRARY

Outcome assessment tools

Introduction

Standardised assessment tools are vital for assessing a range of variables including symptoms, functioning, and quality of life. They are often used within a controlled research environment, but high-quality assessment tools are also useful in practice for both clinical management and outcome prediction.

Method

We have included only systematic reviews (systematic literature search, detailed methodology with inclusion/exclusion criteria) published in full text, in English, from the year 2010 that report results separately for people with a diagnosis of PTSD. Reviews were identified by searching the databases MEDLINE, EMBASE, and PsycINFO. Hand searching reference lists of identified reviews was also conducted. Reviews with pooled data are prioritised for inclusion.

Review reporting assessment was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist that describes a preferred way to present a meta-analysis[1]. Reviews with less than 50% of items checked have been excluded from the library. The PRISMA flow diagram is a suggested way of providing information about studies included and excluded with reasons for exclusion. Where no flow diagram has been presented by individual reviews, but identified studies have been described in the text, reviews have been checked for this item. Note that early reviews may have been guided by less stringent reporting checklists than the PRISMA, and that some reviews may have been limited by journal guidelines.

Evidence was graded using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group approach where high quality evidence such as that gained from randomised controlled trials (RCTs) may be downgraded to moderate, low or very low if review and study quality is limited, if there is inconsistency in results, indirect comparisons, imprecise or sparse data and high probability of reporting bias. It may also be downgraded if risks associated with the intervention or other matter under review are high. Conversely, low quality evidence such as that gained from observational studies may be upgraded if effect sizes are large, there is a dose dependent response or if results are reasonably consistent, precise and direct with low associated risks (see end of table for an explanation of these terms)[2]. The resulting table represents an objective summary of the available evidence, although the conclusions are solely the opinion of staff of NeuRA (Neuroscience Research Australia).

Results

We found five systematic reviews that met our inclusion criteria[3-7].

- Moderate quality evidence finds a model comprising 4-factors of intrusions, avoidance, hyperarousal, and dysphoria/numbing yielded the best fit for clustering PTSD symptoms. Results were not moderated by PTSD measure or sample type.
- Assessment measures for this model included the Clinician-Administered PTSD Scale, Harvard Trauma Questionnaire, Modified PTSD Symptom Scale, PTSD Checklist, PTSD Diagnostic Scale, PTSD Symptom Scale, SCID = Structured Clinical Interview for DSM-IV Diagnosis, Screen for Posttraumatic Stress Symptoms.
- Intrusion symptoms include intrusive thoughts of trauma, recurrent dreams of trauma, flashbacks, emotional reactivity to trauma cues, and physiological reactivity to trauma cues.

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

NeuRA Discover. Conquer. Cure.

POST-TRAUMATIC STRESS DISORDER LIBRARY

Outcome assessment tools

- Avoidance symptoms include avoiding thoughts of trauma and avoiding reminders of trauma.
- Hyperarousal symptoms include hypervigilance, exaggerated startle response, sleep disturbance, irritability, and difficulty concentrating.
- Dysphoria/numbing symptoms include inability to recall aspects of the trauma, loss of interest, detachment, restricted affect, sense of foreshortened future. Sleep disturbance, irritability, and difficulty concentrating may also be classed as dysphoria symptoms.
- Moderate to low quality evidence finds small to medium-sized associations between increased PTSD symptoms and decreased mindfulness scores on the Five Facet Mindfulness Questionnaire.
- Moderate to high quality evidence finds weak to moderate correlations between increased Centrality of Event Scale scores (having a negative event central to one's identity and life story) and increased; PTSD symptoms (all, avoidance, arousal, re-experiencing), post-traumatic growth, grief, trauma cognitions, memory vividness, emotional intensity, shame, physical reaction, depression, anxiety, negative trauma emotions, dissociation, neuroticism, life danger and injury traumas, female sex, and openness. There were no to weak correlations between increased Centrality of Event Scale scores and decreased satisfaction with life, social support, extraversion, conscientiousness, and agreeableness.
- Moderate to high quality evidence finds similar scores on clinician-administered and self-report PTSD symptom rating scales in clinical trials. Subgroup analysis found a trend for more conservative scores on clinician-administered scales in trials of children and adolescents, but not in trials of adults.
- Moderate to low quality evidence is unable to recommend the use of any particular scale for assessing outcomes in youth exposed to traumatic events. Scales assessed were; the Child Behaviour Checklist-PTSD, University of Los Angeles–Post-Traumatic Stress Disorder Scale – reaction index, Child PTSD Symptom Scale, Child Dissociative Checklist, Adolescent Dissociative Experiences Scale, Solution Focused Recovery Scale, Child and Youth Resilience Measure-28, Child and Youth Resilience Measure-12, Minnesota Multiphasic Personality Inventory-Adolescent, Beck Self-Concept Inventory for Youth, Adult Attachment Interview, Global Assessment of Functioning Scale, Children's Global Assessment of Functioning Scale, Adolescent Clinical Sexual Behaviour Inventory, Child Sexual Behaviour Inventory, Vineland Adaptive Behaviour Scale-II, Trauma Symptom Checklist for Young Children, Trauma Symptom Checklist for Young Children–Short Form, Assessment Checklist for Children, Brief Assessment Checklist for Children, Trauma Assessment for Young Children, Child Paediatric Emotional Distress Scale, Trauma Play Scale, Story Stem Assessment Profile, Dominic Interactive Assessment, Assessment Checklist for Adolescents, and the Brief Assessment Checklist for Adolescents.

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Carpenter JK, Conroy K, Gomez AF, Curren LC, Hofmann SG

The relationship between trait mindfulness and affective symptoms: A metaanalysis of the Five Facet Mindfulness Questionnaire (FFMQ)

Clinical Psychology Review 2019; 74: 101785

View review abstract online

Comparison	**Correlations between PTSD symptom scores and scores on the Five Facet Mindfulness Questionnaire.**
Summary of evidence	**Moderate to low quality evidence (direct, unclear sample size, unable to assess consistency or precision) finds small to medium-sized relationships between increased PTSD symptoms and decreased FFMQ scores.**
Five Facet Mindfulness Questionnaire (FFMQ)	
Small to medium-sized correlations between increased PTSD symptoms and decreased FFMQ; Total scores: 12 studies, r = -0.42 Describe internal experiences subscale: 13 studies, r = -0.30, p < 0.01 Act with awareness of the present subscale: 13 studies, r = -0.51, p < 0.01 Non-judgmental stance towards one's inner experiences subscale: 13 studies, r = -0.48, p < 0.01 Non-reacting to one's thoughts and feelings subscale: 14 studies, r = -0.25, p < 0.01 *No significant associations with;* Observe internal and external experiences subscale: 13 studies, r = 0.02, p > 0.05	
Consistency in results‡	Unable to assess; no measure of consistency is reported.
Precision in results§	Unable to assess; no measure of precision is reported.
Directness of results‖	Direct

Denton R, Frogley C, Jackson S, John M, Querstret D

The assessment of developmental trauma in children and adolescents: A systematic review

Outcome assessment tools

Clinical Child Psychology and Psychiatry 2017; 22: 260-87 View review abstract online	
Comparison	**Outcome scales for children and adolescents exposed to trauma.**
Summary of evidence	**Moderate to low quality evidence (direct, unable to assess consistency or precision) is unable to recommend the use of any particular scale for assessing PTSD, functional or other mental health outcomes in children or adolescents exposed to traumatic events.**
Scales for assessing PTSD outcomes	

Child Behaviour Checklist-PTSD

A global measure of emotional and behavioural difficulties in children and young people.

3-point Likert Scale, caregiver rated

One study assessed the CBCL-PTSD 20 item scale in foster children (N = 36) exposed to sexual abuse. This study found reasonable internal consistency (α = 0.73) but poor convergent validity with the Clinician Administered Post-Traumatic Stress Disorder Scale–Child and Adolescent (*r* = 0.21). There was moderate convergent validity with the DSM-IV PTSD (AUC = 0.53), and good convergent validity with the ICD-10 (*r* = 0.75).

One study assessed the CBCL-PTSD 15 item scale in children (N = 61) exposed to multiple-event trauma. This study found 26.2% met PTSD diagnostic criteria (*r* = 0.66). There was good sensitivity (75%) and specificity (84%) for PTSD diagnosis. There were moderate correlations with the parent interview (*r* = 0.66) and internalising (*r* = 0 .57) and externalising (*r* = 0.42) subscales. Factor analysis accounted for 46% of the variance.

Another study assessed the CBCL-PTSD 15 item scale in children (N = 51) exposed to mostly interpersonal, small number single event traumas. 6% met DSM-IV criteria for PTSD. There was moderate sensitivity (60%) and good specificity (80%) with DSM-IV. There was good internal consistency (α = 0.79).

One study assessed three subscales from the CBCL-PTSD; PTSD Scale (7 items), PTSD/Dissociation Scale (16 items), and the Dissociation Scale (3 items) in children (N = 1,293) exposed to sexual abuse compared to controls (N = 419). Factor analysis found adequate fit for a 3-factor model. Inter-scale correlations were moderate to good (*r* = 0.51 to 0.90). Internal consistency was good (*r* = 0.70 to 0.85).

Another study of children (N = 239) exposed to physical neglect found moderate correlation with the Trauma Symptom Checklist for Young Children (*r* = 0.65) and the Child Dissociative Checklist (*r* = 0.58). There was weak to moderate internal consistency (α = 0.39 to 0.55).

Another study of foster children (N = 36) found poor correlation with the Trauma Symptom Checklist for Young Children (*r* = 0.12), and moderate correlation with the DSM-IV (*r* = 0.51) and ICD-10 (*r* = 0.37). Internal consistency was moderate (α = 0.63).

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Outcome assessment tools

University of Los Angeles–Post-Traumatic Stress Disorder Scale – reaction index

Assesses exposure to traumatic events and PTSD symptoms according to DSM-IV criteria.

48-item, semi-structured, clinician rated.

One study of children (N = 6,291) exposed to mixed multiple and single-incident traumas found good correlation with the Trauma Symptom Checklist for Children – post-traumatic symptoms (r = 0.75), and moderate correlation with other subscales (0.54 to 0.67). There was good internal consistency (r = 0.86 to 0.91). A 5-factor model was the best fit. Inter-correlations between some factors were high (r = 0.86 to 0.89).

Child PTSD Symptom Scale

Assesses PTSD symptom severity in children and young people.

24 items, 17 items correspond to DSM-IV symptoms, caregiver rated.

One study of female children with a PTSD diagnosis (N = 91) that were exposed to sexual abuse found good correlation with the PTSD module of the Kiddie Schedule for Schizophrenia and Affective Disorders (r = 74.5 to 76.5). There was good to excellent test–retest reliability after one week on self-report measure (r = 0.86).

Child Dissociative Checklist

Assesses dissociative symptoms.

20-item, self-report.

One study of children (N = 232) exposed to physical and sexual abuse found good convergent validity. A 3-factor model accounted for 46% of the variance. Internal consistency was moderate to good (α = 0.69 to 0.83).

Adolescent Dissociative Experiences Scale

Assesses symptoms of pathological dissociation

30 items, self-report.

One Study of female adolescents (N = 65) exposed to sexual abuse or no sexual abuse found a better predictor of clinical group membership (87%) than non-clinical membership (68%). There was good internal consistency (α = .94).

Scales for assessing general functioning and mental health outcomes

Solution Focused Recovery Scale

Assesses positive coping skills specific to childhood sexual abuse.

36 items, Likert-style, self-report.

One study assessed female adolescents (N = 99) exposed to childhood sexual abuse and found three factors explained 38% of the variance. Internal consistency was good for the total scale (α = 0.89), inter-correlations of individual items with total score (r = 0.06 to 0.67). Authors report an expected lack of correlation with the Child Behaviour Checklist.

Child and Youth Resilience Measure-28

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Assesses resilience across cultural contexts.

28 items, Likert Scale, self-report.

One study of adolescents (total N = 907) found a confirmatory factor analysis, with good fit of three latent variables explaining 40.4% of the variance. Internal consistency was moderate to good (*r* = 0.65 to 0.91).

Child and Youth Resilience Measure-12

Assesses resilience across cultural contexts.

12 items, self-report.

One study of adolescents (total N = 1,516) found factor analyses identified 12 items with one latent structure. Internal consistency was good (α = 0.84).

Minnesota Multiphasic Personality Inventory-Adolescent

Assesses psychopathology among youth.

478 items, self-report.

One study of adolescent juvenile offenders (N = 186) found the model correctly classified 65.5% in the trauma-nominated group, and 84.3% in the no trauma group. Receiver operating characteristic (ROC) curve analysis = 0.771, suggesting reasonable predictor of trauma group membership. Inter-rater reliability was good at 100%.

Beck Self-Concept Inventory for Youth

Assesses competence and self-worth.

20 items, Likert-style, self-report.

One study of adolescents (N = 100) with a sexual abuse history found moderate, negative correlations with the Child Behavioural Checklist internalising subscale (*r* = -0.35 to -0.41). There were moderate negative correlations with the Beck Youth Inventory, anxiety (*r* = -0.45) and anger (*r* = -0.27) scales. Internal consistency was good (total α = 0.94, subscales α = 0.80).

Adult Attachment Interview

Assesses adult representation of attachment.

Interviews recorded, transcribed, and coded. Classified into attachment representation categories: autonomous, dismissing, preoccupied, unresolved or cannot classify, semi-structured interview.

One study of adolescents (N = 62 females) with physical and/or sexual abuse occurring before age 12 found the Adult Attachment Interview elicited more reports on physical abuse (91%) than the Childhood Trauma Interview (55%). Fewer participants reported sexual abuse (62%).

Another study of adolescents (N = 55 females) with emotional abuse, physical abuse, sexual abuse and/or caregiver neglect found sexual abuse history uniquely predicted unresolved status. There was significant agreement between the Adult Attachment Interview and the Childhood Trauma Questionnaire on maltreatment experiences. Inter-rater reliability was good at 80%.

Global Assessment of Functioning Scale

Assesses level of psychological, social, and occupational functioning among adults.

Ten sections with corresponding descriptive characteristics for each 10-point increment. Lower

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

scores were suggestive of low level of functioning, clinician rating.

One study of clinicians (N = 22) rated 15 vignettes of childhood physical abuse, neglect, and sexual abuse. Inter-rater scores were poor to moderate (no trauma history *r* = 0.68 to 0.73, trauma history *r* = 0.33).

Children's Global Assessment of Functioning Scale

Assesses global functioning among children and adolescents.

Comparable scoring to the Global Assessment of Functioning, clinician rating.

One study of clinicians (N = 22) rated 15 vignettes of childhood physical abuse, neglect, and sexual abuse. Inter-rater scores were poor to moderate (no trauma history *r* = 0.55 to 0.60, trauma history *r* = 0.38).

Adolescent Clinical Sexual Behaviour Inventory

Assesses a range of sexual behaviours among adolescents.

45 items, 3-point scale, self-report, and parent-report.

One study of adolescents (N = 174) with a sexual abuse history found a five-factor analysis explained 37.6% of the variance. There was good convergent validity between the Trauma Symptom Checklist for Children and the self-report version of this scale (*r* = 0.54 to 0.74), but there was low correlation with the parent-report version of this scale *r* = 0.36 to 0.44). Total score showed a moderate correlation with the Child Behaviour Checklist total (0.66). There was good internal validity (self-report α = 0.86, parent-report α = .84), and test-retest validity (1 week *r* = 0.74). There was a moderate correlation between the self-report and parent-report versions (*r* = 0.55).

Another study of adolescents (N = 141) referred for sexual abuse evaluations found weak correlations between three factors and Child Behaviour Checklist scales (*r* = 0.23 to 0.33), and weak to moderate correlations with the Symptom Checklist scales (*r* = 0.26 to 0.50). Internal consistency was moderate (*r* = 0.61 to 0.75).

Child Sexual Behaviour Inventory

Assesses the frequency of sexual behaviours in children aged 2-10 years.

38 items, parent report.

One study of children (N = 97) in a residential care home or fostered and a control group found small to high correlations with the Child Behaviour Checklist (*r* = 0.13 to 0.7). There was a weak, non-significant correlation with the number of trauma events (*r* = 0 .33).

Vineland Adaptive Behaviour Scale-II

Assesses personal and social functioning in communication, motor skills, socialisation domain, emotional and behavioural difficulties interfering with functioning.

Semi-structured interview, parent/carer, teacher, or extended interview.

One study with a mixed sample (N = 57) confirmed a diagnosis of reactive attachment disorder and met criteria for complex trauma.

Scales for assessing trauma-related outcomes

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Outcome assessment tools

Trauma Symptom Checklist for Young Children

Assesses trauma-related symptoms in children.

90 items. 8 subscales: anxiety, depression, anger and abnormal sexual behaviour, caregiver rated.

1 study of children (N = 388) exposed to sexual abuse or non-sexual abuse found reasonable sensitivity (76%) and good specificity (87%). The scale showed good internal consistency (α = 0.81 to 0.93).

Another study of children (N = 34) with substantiated sexual abuse found 32% met criteria for PTSD. There was significant correlation with PTSD scales for those with PTSD ($p < 0.001$), and good internal consistency (α = 0.73 to 0.91).

Another study of children (N = 172) exposed to sexual abuse found reasonable convergent validity with the Child Behavioural Checklist ($r = 0.54$ to 0.84), but less so with the UCLA-PTSD ($r = 0.34$ to 0.59), the Trauma Symptom Checklist for Children ($r = 0.47$ to 0.29) and the Child Sexual Behaviour Inventory ($r = 0.44$ to 0.72).

Trauma Symptom Checklist for Young Children–Short Form

Assesses trauma-related symptoms in children.

32 items, caregiver rated.

I study of children (N = 295) exposed to sexual abuse found reasonable convergent validity for scales of anger, sexual concerns, anxiousness, and depression ($r = 0.53$ to 0.83). Moderate correlations were found between the Trauma Symptom Checklist for Young Children and the Child Sexual Behaviour Inventory ($r = 0.42$ to 0.60), the Trauma Symptom Checklist for Children and the UCLA-PTSD-revised ($r = 0.47$ to 0.75). There was good internal consistency (α = 0.77 to 0.91).

Assessment Checklist for Children

Assesses behaviours, emotional states, traits and relating to others, as manifested among children in care.

120 items. 10 clinical scales and 2 self-esteem scales, caregiver rated.

I study of children (N = 412) in long-term foster and kinship care found good convergent correlation to the Child Behavioural Checklist ($r = 0.89$ for boys, $r = 0.90$ for girls). There was good internal consistency (α = 0.70 to 0.96).

Brief Assessment Checklist for Children

Screens/monitors emotional and behavioural difficulties experienced by children in out-of-home care.

20 items developed from the 120-item Assessment Checklist, caregiver rated.

1 study of children (N = 347) showed moderate to strong correlations with the Assessment Checklist for Children subscales ($r = 0.32$ to 0.96), the Child Behavioural Checklist subscales ($r = 0.41$ to 0.82) and DSM-oriented scale scores ($r = 0.34$ to 0.64). Internal consistency was good (α = 0.89).

Trauma Assessment for Young Children

Assesses symptoms of trauma in young children.

10 items, self-report and caregiver rated.

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

1 study of children (N = 47) with or without interpersonal traumas found weak correlations between child and parent reports (r = 0.01 to 0.46), and moderate correlations with Trauma Symptom Checklist for Young Children PTSD subscales (r = 0.59 for trauma, r = 0.41 for total sample). There was a weak correlation in the trauma group with the Child Behavioural Checklist externalising subscale (r = 0.24). Internal consistency was moderate (trauma group α = 0.48, non-trauma group α = 0.56). Test–retest was good at 2 weeks (r = 0.79), based on the non-clinical sample.

Child Paediatric Emotional Distress Scale

Assesses symptomatology in children following stressful and/or traumatic event.

21 items, caregiver rated.

1 study of children (N = 383) who witnessed domestic violence found weak to good correlations with the Revised Behaviour Problem Checklist (r = 0.21 to 0.73). The 3-factor model did not fit (CFI = 0.84) and there were weak to good correlations between the factors (r = 0.38 to 1.0). Factor analysis showed a 2-factor model (CFI = 0.97; r = 0.17). Internal consistency was good for the 2-factor model (α = 0.80 to 0.82).

Trauma Play Scale

Observation-based measure of play behaviours in children exposed to trauma.

Five subscales: intense play, repetitive play, play disruption, avoidant play, behaviour and expression of negative affect. Clinician rated.

1 study of children (N = 12) with and without a trauma history found good inter- and intra-rater reliability (r = 0.85 to 0.98). There was good discriminant validity between groups for average and subscale scores (p < 0.001), and good internal reliability (r = 0.74), and inter-rater reliability (86%).

Story Stem Assessment Profile

Assessment profile captures effects of abuse in young children. Asked to continue a narrative; interview analysed against specific themes.

Clinician rated. 1 study of children (N = 206) who were maltreated or not maltreated with a clinical diagnosis vs controls found the maltreated group were more defensive, avoidant, insecure, and Disorganized. Internal consistency was moderate (α = 0.52).

Dominic Interactive Assessment

Assesses psychiatric symptoms related to common mental health disorders based on seven DSM-III-R and DSM-IV diagnoses.

91 items.

1 study of children (N = 55) exposed to interpersonal violence found moderate correlation with the Child Behaviour Checklist total (r = 0.42). There was weak to moderate correlation with the Diagnostic Interview for Children and Adolescents on number of symptoms (r = 0.05 to 0.34). Sensitivity was moderate (52.6%), and specificity was good (81.5%) for the Child Behaviour Checklist clinical range score.

Assessment Checklist for Adolescents

Assesses behaviours, emotional states, traits, and manners of relating to others.

120 items. Includes 10 clinical scales and 2 self-esteem scales, caregiver rated.

1 study of adolescents (N = 372) in long-term foster and kinship care found a 7-factor model

accounted for 51% score variance. There was strong correlation with Child Behaviour Checklist total scores (boys *r* = 0.90, girls *r* = 0.88). Internal consistency was good (total clinical score *r* = 0.95, clinical scales *r* = 0.73 to 0.89). Brief Assessment Checklist for Adolescents Brief screening tool for use by social or health professionals without child mental health qualification. 20 items. Caregiver rated. 1 study of adolescents (N = 230) found strong correlation with the Brief Assessment Checklist for Adolescents scale (total *r* = 0.94, problem scores *r* = 0.88). There was good internal consistency (α = 0.87), and good correlation with the Child Behaviour Checklist clinical range scores (*r* = 0.93 to 0.94).	
Consistency in results	Unable to assess; no measure of consistency is reported.
Precision in results	Unable to assess; no measure of precision is reported.
Directness of results	Direct

Gehrt TB, Berntsen D, Hoyle RH, Rubin DC

Psychological and clinical correlates of the Centrality of Event Scale: A systematic review

Clinical Psychology Review 2018; 65: 57-80

View review abstract online

Comparison	**Correlations between PTSD symptoms and other outcomes and the Centrality of Event Scale (having a negative event central to one's identity and life story).**
Summary of evidence	**Moderate to high quality evidence (large samples, direct, precise, direct, unable to assess consistency) finds weak to moderate correlations between increased Centrality of Event Scale scores and increased; PTSD symptoms (all, avoidance, arousal, re-experiencing), post-traumatic growth, grief, trauma cognitions, memory vividness, emotional intensity, shame, physical reaction, depression, anxiety, negative trauma emotions, dissociation, neuroticism, life danger and injury traumas, female sex, and openness. There were no to weak correlations between increased Centrality of Event Scale scores and decreased satisfaction with life, social support, extraversion, conscientiousness, and agreeableness.**

Centrality of Event Scale (CES)	
Weak-moderate correlations with increased CES and increased (in descending order of strength); Post-traumatic growth: N = 11,578, *r* = 0.549, 95%CI 0.534 to 0.564 Grief: N = 1,260, *r* = 0.539, 95%CI 0.492 to 0.585 All PTSD symptoms: N = 37,626, *r* = 0.511, 95%CI 0.503 to 0.520 Re-experiencing: N = 4,421, *r* = 0.497, 95%CI 0.472 to 0.523 Arousal: N = 4,096, *r* = 0.456, 95%CI 0.428 to 0.483 Trauma cognitions: N = 2,082, *r* = 0.438, 95%CI 0.399 to 0.477 Avoidance: N = 4,421, *r* = 0.408, 95%CI 0.381 to 0.435 Memory vividness: N = 2,161, *r* = 0.398, 95%CI 0.360 to 0.437 Emotional intensity: N = 2,176, *r* = 0.384, 95%CI 0.346 to 0.423 Shame: N = 3,858, *r* = 0.383, 95%CI 0.354 to 0.412 Physical reaction: N = 2,082, *r* = 0.305, 95%CI 0.264 to 0.346 Depression: N = 21,953, *r* = 0.281, 95%CI 0.268 to 0.293 Anxiety: N = 6,201, *r* = 0.271, 95%CI 0.247 to 0.295 Negative trauma emotions: N = 10,737, *r* = 0.264, 95%CI 0.246 to 0.282 Dissociation: N = 2,211, *r* = 0.246, 95%CI 0.205 to 0.286 Neuroticism: N = 18,987, *r* = 0.203, 95%CI 0.189 to 0.217 Life danger and injury trauma: N = 10,767, *r* = 0.129, 95%CI 0.110 to 0.148 Female (vs male) sex: N = 6,714, *r* = 0.113, 95%CI 0.089 to 0.137 Openness: N = 5,923, *r* = 0.079, 95%CI 0.054 to 0.104 *No-weak correlations with increased CES scores and decreased (in descending order of strength);* Satisfaction with life: N = 3,509, *r* = -0.175, 95%CI -0.207 to -0.142 Social support: N = 7,374, *r* = -0.083, 95%CI -0.106 to -0.061 Extraversion: N = 5,923, *r* = -0.065, 95%CI -0.091 to -0.040 Conscientiousness: N = 5,923, *r* = -0.061, 95%CI -0.087 to -0.036 Agreeableness: N = 5,923, *r* = -0.021, 95%CI -0.046 to 0.005	
Consistency in results	Unable to assess; no measure of consistency is reported.
Precision in results	Precise
Directness of results	Direct

Lenz AS, Luo Y

Differential estimation of treatment effect between clinician-administered and self-reported PTSD assessments

Journal of Counseling and Development 2019; 97: 3-14

View review abstract online

Comparison	**Correlation between clinician-administered and self-report measures of PTSD symptoms in clinical trials.**
Summary of evidence	**Moderate to high quality evidence (large samples, direct, consistent, unable to assess precision) finds similar scores on clinician-administered and self-report PTSD rating scales. There was a trend for more conservative scores on clinician-administered scales in trials of children and adolescents.**

Clinician-administered = Clinician-administered PTSD Scale

Self-report = Child PTSD Symptom Scale, Davidson Trauma Scale, Impact of Events Scale−Revised, PTSD Checklist for DSM-5, Posttraumatic Diagnostic Scale for DSM-5, PTSD Symptoms Scale, Traumatic Stress Symptom Checklist, University of California at Los Angeles PTSD Reaction Index, and the Traumatic Stress Institute Beliefs Scale

There were no significant differences between clinician-administered and self-report assessments;

17 studies, N = 1,405, Q(1,46) = 0.24, $p = 0.61$

Clinician-administered: $g = -0.75$, 95%CI -0.92 to -0.57, $p < 0.05$

Self-report: $g = -0.82$, 95%CI -1.01 to -0.64, $p < 0.05$

Subgroup analyses found a trend effect for clinicians to make more conservative estimates of improvement in symptoms than child self-reports. In adults, treatment effect estimates were similar.

Consistency in results	Within-groups consistency measures were not reported.
Precision in results	Precise
Directness of results	Direct

Yufik T, Simms LJ

A meta-analytic investigation of the structure of posttraumatic stress disorder symptoms

Journal of Abnormal Psychology 2010; 119: 764-76 View review abstract online	
Comparison	**Best fit model for PTSD symptom clusters.**
Summary of evidence	**Moderate quality evidence (large sample, direct, unable to assess consistency or precision) finds a model comprising 4-factors of intrusions, avoidance, hyperarousal, and dysphoria/numbing yielded the best fit. Results were not moderated by PTSD measure or sample type.**
PTSD symptom clusters **Measured using the Clinician-Administered PTSD Scale, Harvard Trauma Questionnaire, Modified PTSD Symptom Scale, PTSD Checklist, PTSD Diagnostic Scale, PTSD Symptom Scale, SCID = Structured Clinical Interview for DSM-IV Diagnosis, Screen for Posttraumatic Stress Symptoms**	
35 studies, N = 14,827 *A model comprising 4-factors of intrusions, avoidance, hyperarousal, and dysphoria/numbing yielded the best fit;* Intrusions Intrusive thoughts of trauma Recurrent dreams of trauma Flashbacks Emotional reactivity to trauma cues Physiological reactivity to trauma cues Avoidance Avoiding thoughts of trauma Avoiding reminders of trauma Hyperarousal Hypervigilance Exaggerated startle response Sleep disturbance Irritability Difficulty concentrating Dysphoria/numbing Inability to recall aspects of trauma Loss of interest	

Outcome assessment tools

Detachment Restricted affect Sense of foreshortened future Sleep disturbance Irritability Difficulty concentrating Authors report these results were not moderated by measure or sample type.	
Consistency in results	Within-groups consistency measures were not reported.
Precision in results	No measure of precision is reported.
Directness of results	Direct

Explanation of acronyms

α = Cronbach's alpha test for internal consistency, CI = confidence interval, *g* = Hedges' g, standardised mean difference, N = number of participants, *p* = statistical probability of obtaining that result, Q = test for heterogeneity between groups of studies, *r* = correlation coefficient, vs. = versus

Outcome assessment tools

Explanation of technical terms

* Bias has the potential to affect reviews of both RCT and observational studies. Forms of bias include; reporting bias – selective reporting of results; publication bias - trials that are not formally published tend to show less effect than published trials, further if there are statistically significant differences between groups in a trial, these trial results tend to get published before those of trials without significant differences; language bias – only including English language reports; funding bias - source of funding for the primary research with selective reporting of results within primary studies; outcome variable selection bias; database bias - including reports from some databases and not others; citation bias - preferential citation of authors. Trials can also be subject to bias when evaluators are not blind to treatment condition and selection bias of participants if trial samples are small[8].

---

† Different effect measures are reported by different reviews.

Prevalence refers to how many existing cases there are at a particular point in time. Incidence refers to how many new cases there are per population in a specified time period. Incidence is usually reported as the number of new cases per 100,000 people per year. Alternatively some studies present the number of new cases that have accumulated over several years against a person-years denominator. This denominator is the sum of individual units of time that the persons in the population are at risk of becoming a case. It takes into account the size of the underlying population sample and its age structure over the duration of observation.

Reliability and validity refers to how accurate the instrument is. Sensitivity is the proportion of actual positives that are correctly identified (100% sensitivity = correct identification of all actual positives) and specificity is the proportion of negatives that are correctly identified (100% specificity = not identifying anyone as positive if they are truly not). A receiver operating characteristic (ROC) curve represents sensitivity/specificity pairs corresponding to different cut-off values. A guide for interpreting the area under the curve (AUC) statistic is; 0.90 to 1.00 = excellent, 0.80 to 0.90 = good, 0.70 to 0.80 = fair, 0.60 to 0.70 = poor, and 0.50 to 0.60 = fail.

Weighted mean difference scores refer to mean differences between treatment and comparison groups after treatment (or occasionally pre to post treatment) and in a randomized trial there is an assumption that both groups are comparable on this measure prior to treatment. Standardized mean differences are divided by the pooled standard deviation (or the standard deviation of one group when groups are homogenous) that allows results from different scales to be combined and compared. Each study's mean difference is then given a weighting depending on the size of the sample and the variability in the data. 0.2 represents a small effect, 0.5 a moderate effect, and 0.8 and over represents a large effect[8].

Odds ratio (OR) or relative risk (RR) refers to the probability of a reduction (< 1) or an increase (> 1) in a particular outcome in a treatment group, or a group exposed to a risk factor, relative to the comparison group. For example, a RR of 0.75 translates to a reduction in risk of an outcome of 25% relative to those not receiving the treatment or not exposed to the risk factor. Conversely, a RR of 1.25 translates to an increased risk of 25% relative to those not receiving treatment or not having been exposed to a risk factor. A RR or OR of 1.00 means there is no difference between groups. A medium effect is considered if RR > 2 or < 0.5 and a large

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

effect if RR > 5 or < 0.2[9]. lnOR stands for logarithmic OR where a lnOR of 0 shows no difference between groups. Hazard ratios measure the effect of an explanatory variable on the hazard or risk of an event.

Correlation coefficients (eg, r) indicate the strength of association or relationship between variables. They can provide an indirect indication of prediction, but do not confirm causality due to possible and often unforseen confounding variables. An r of 0.10 represents a weak association, 0.25 a medium association and 0.40 and over represents a strong association. Unstandardized (*b*) regression coefficients indicate the average change in the dependent variable associated with a 1 unit change in the independent variable, statistically controlling for the other independent variables. Standardized regression coefficients represent the change being in units of standard deviations to allow comparison across different scales.

‡ Inconsistency refers to differing estimates of effect across studies (i.e. heterogeneity or variability in results) that is not explained by subgroup analyses and therefore reduces confidence in the effect estimate. I^2 is the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance) - 0% to 40%: heterogeneity might not be important, 30% to 60%: may represent moderate heterogeneity, 50% to 90%: may represent considerable heterogeneity and over this is considerable heterogeneity. I^2 can be calculated from Q (chi-square) for the test of heterogeneity with the following formula[8];

$$I^2 = \left(\frac{Q - df}{Q}\right) \times 100\%$$

§ Imprecision refers to wide confidence intervals indicating a lack of confidence in the effect estimate. Based on GRADE recommendations, a result for continuous data (standardised mean differences, not weighted mean differences) is considered imprecise if the upper or lower confidence limit crosses an effect size of 0.5 in either direction, and for binary and correlation data, an effect size of 0.25. GRADE also recommends downgrading the evidence when sample size is smaller than 300 (for binary data) and 400 (for continuous data), although for some topics, these criteria should be relaxed[10].

‖ Indirectness of comparison occurs when a comparison of intervention A versus B is not available but A was compared with C and B was compared with C that allows indirect comparisons of the magnitude of effect of A versus B. Indirectness of population, comparator and/or outcome can also occur when the available evidence regarding a particular population, intervention, comparator, or outcome is not available and is therefore inferred from available evidence. These inferred treatment effect sizes are of lower quality than those gained from head-to-head comparisons of A and B.

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

References

1. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMAGroup (2009): Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *British Medical Journal* 151: 264-9.
2. GRADEWorkingGroup (2004): Grading quality of evidence and strength of recommendations. *British Medical Journal* 328: 1490.
3. Carpenter JK, Conroy K, Gomez AF, Curren LC, Hofmann SG (2019): The relationship between trait mindfulness and affective symptoms: A metaanalysis of the Five Facet Mindfulness Questionnaire (FFMQ). *Clinical Psychology Review* 74: 101785.
4. Denton R, Frogley C, Jackson S, John M, Querstret D (2017): The assessment of developmental trauma in children and adolescents: A systematic review. *Clinical Child Psychology and Psychiatry* 22: 260-87.
5. Gehrt TB, Berntsen D, Hoyle RH, Rubin DC (2018): Psychological and clinical correlates of the Centrality of Event Scale: A systematic review. *Clinical Psychology Review* 65: 57-80.
6. Lenz AS, Luo Y (2019): Differential estimation of treatment effect between clinician-administered and self-reported PTSD assessments. *Journal of Counseling and Development* 97: 3-14.
7. Yufik T, Simms LJ (2010): A meta-analytic investigation of the structure of posttraumatic stress disorder symptoms. *Journal of Abnormal Psychology* 119: 764-76.
8. CochraneCollaboration (2008): Cochrane Handbook for Systematic Reviews of Interventions. Accessed 24/06/2011.
9. Rosenthal JA (1996): Qualitative Descriptors of Strength of Association and Effect Size. *Journal of Social Service Research* 21: 37-59.
10. GRADEpro (2008): [Computer program]. Jan Brozek, Andrew Oxman, Holger Schünemann. *Version 32 for Windows*