Available online at www.sciencedirect.com

ScienceDirect

Comprehensive Psychiatry 55 (2014) 11–24

COMPREHENSIVE PSYCHIATRY

www.elsevier.com/locate/comppsych

# Research methodology used in studies of child disaster mental health interventions for posttraumatic stress

Betty Pfefferbaum[a,*], Elana Newman[b,c], Summer D. Nelson[b,c], Brandi D. Liles[b,c], Robert P. Tett[c], Vandana Varma[a], Pascal Nitiéma[a,d]

[a]*Department of Psychiatry and Behavioral Sciences, College of Medicine, and Terrorism and Disaster Center, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA*

[b]*Institute of Trauma, Abuse and Neglect, University of Tulsa, Tulsa, OK, USA*

[c]*Department of Psychology, University of Tulsa, Tulsa, OK, USA*

[d]*Department of Biostatistics and Epidemiology, College of Public Health, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA*

**Abstract**

**Objective:** In the last decade, the development of community-based and clinical interventions to assist children and adolescents after a disaster has become an international priority. Clinicians and researchers have begun to scientifically evaluate these interventions despite challenging conditions. The objective of this study was to conduct a systematic review of the research methodology used in studies of child disaster mental health interventions for posttraumatic stress.

**Method:** This scientifically rigorous analysis used standards for methodological rigor of psychosocial treatments for posttraumatic stress disorder (PTSD) to examine 29 intervention studies.

**Results:** This analysis revealed that further refinement of methodology is needed to determine if certain intervention approaches are superior to other approaches and if they provide benefit beyond natural recovery. Most studies (93.1%) clearly described the interventions being tested or used manuals to guide application and most (89.7%) used standardized instruments to measure outcomes, and many used random assignment (69.0%) and provided assessor training (65.5%). Fewer studies used blinded assessment (44.8%) or measured treatment adherence (48.3%), and sample size in most studies (82.8%) was not adequate to detect small effects generally expected when comparing two active interventions. Moreover, it is unclear what constitutes meaningful change in relation to treatment especially for the numerous interventions administered to children in the general population.

**Conclusions:** Overall, the results are inconclusive about which children, what settings, and what approaches are most likely to be beneficial.


## 1. Introduction

While child disaster mental health research has generated a rich literature documenting the reactions of children to an array of natural, accidental, and human-caused events, the development and evaluation of disaster interventions is relatively recent and understudied. Evaluating interventions is necessary to assure that well-intentioned services include appropriate cost-effective interventions that meet the needs of children, families, and communities affected by disasters. Several reports published since the September 11, 2001, terrorist attacks have reviewed child trauma [1–4], disaster [5], single-incident [6], and school-based [7,8] interventions, but most of these reviews have included relatively few of the disaster interventions advanced and studied in recent years. The meta-analysis of school-based trauma interventions, conducted by Rolfsnes and Idsoe [8], is an exception with 16 studies specifically related to disasters, political conflict, or war and 3 related to community violence. That report did not address interventions delivered in settings other than schools, however. Still lacking is a scientifically rigorous analysis of the child disaster mental health intervention research.

This manuscript includes three tables and no figures. The manuscript has not been published previously and is not currently under consideration for publication elsewhere.

* Corresponding author. Department of Psychiatry and Behavioral Sciences, College of Medicine, University of Oklahoma Health Sciences Center, P.O. Box 26901, WP 3470, Oklahoma City, OK 73126-0901, USA. Tel.: +1 405 271 5121; fax: +1 405 271 8775.

*E-mail address:* betty-pfefferbaum@ouhsc.edu (B. Pfefferbaum).

0010-440X/$ – see front matter © 2014 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.comppsych.2013.08.014

Various criteria advanced to develop clinical guidelines can be used to analyze the quality of intervention efficacy research studies [9–11]. For example, in preparing the practice guidelines for acute stress disorder (ASD) and posttraumatic stress disorder (PTSD), the American Psychiatric Association designated studies according to methodological strength with the highest priority given to randomized, placebo-controlled trials and lesser value given to randomized trials that are not double blind and to open trials, literature reviews, meta-analyses, and other studies [11]. Expanding on their earlier work for the International Society for Traumatic Stress Studies PTSD Treatment Guidelines Task Force, Foa and colleagues [10,12] identified randomized, well-controlled clinical trials, followed by well-designed clinical studies without randomization or placebo comparison, as the highest levels of evidence. The American Academy of Child and Adolescent Psychiatry Practice Parameter for PTSD generated a similar order of preference for evidence—randomized controlled trials in which participants are randomly assigned to two or more treatment conditions, controlled trials in which participants are non-randomly assigned to two or more treatment conditions, uncontrolled trials in which participants are assigned to one treatment condition, and case series or a case report [9].

Guided by recommendations for preparing clinical guidelines, this review identified controlled trials examining PTSD and/or PTSD symptoms or reactions to examine the efficacy of interventions for children used in the context of disasters and related traumatic experiences. The framework for the analysis was shaped by seven gold standards for methodological rigor of PTSD studies enumerated by Foa and Meadows [13] in their critical review of psychosocial treatments for PTSD. The seven parameters are as follows: (1) clearly defined target symptoms; (2) reliable and valid measures; (3) use of blinded evaluators; (4) assessor training; (5) manualized, replicable, specific treatment programs; (6) unbiased assignment to treatment; and (7) treatment adherence. These stringent standards were chosen because despite the challenges in post-disaster environments, rigorous and controlled trials which respect community needs and cultural norms are possible and provide the strictest criteria against which to make critical decisions about mental health resource allocations.

## 2. Methods

Research studies on child disaster mental health interventions were identified by searching the Ovid, Medline, PsycINFO, EBM Reviews, EMBASE, ERIC, and PubMed databases. The search was conducted in the winter of 2012. Keyword search terms addressed interventions and treatment for children and adolescents exposed to disasters and terrorism. The search was confined to materials on children and adolescents, aged 0 to 18 years, and to English language sources. Sixty unduplicated references were generated by the search. Abstracts were reviewed to identify publications describing scientific studies of child interventions for disasters and terrorist incidents, but interventions addressing threat of terrorism, political conflict, war, and various single-incident traumas such as accidents were also included. Throughout the search process, the reference sections of publications, review articles, and studies known to the authors that were not generated in the search were used to locate additional materials. Studies using experimental control conditions were selected for analysis. In all, 28 research papers, which assessed PTSD or PTSD symptoms or reactions as outcomes, were judged appropriate for inclusion in this review. One article described two separate investigations using the same intervention delivered in individual and group format [14]; these studies were analyzed separately. Thus, the final database of independent samples numbered 29.

### 2.1. Coding procedures

Information and categories to be coded and analyzed were determined after a preliminary review of the available studies. Every study was coded to denote the treatment approach and the quality of the description of the intervention; characteristics of the sample, methodology, and instruments; assessor training; treatment adherence; and efficacy of the intervention. After piloting the procedures on publications from the adult trauma and child maltreatment literatures, two investigators coded all studies with respect to various categories. Toward ensuring adequate inter-rater agreement, piloting continued until coders reached an agreement of 80% on the coding variables for five practice articles. The same two coders then coded each of the studies. Cohen's kappa coefficients for the final sample ranged from .29 to .91 (median = .71) and percent agreement ranged from 52% to 94% (median = 85%). Of 19 coding variables, 7 yielded kappa at or below .70 and 12 yielded kappa above .70. There were no statistically significant differences in kappa for the two raters on the first 15 articles and the final sample. The research team met to discuss discrepancies in coding and to reach a reasoned consensus.

The research team sorted interventions into treatment types based on the techniques used in each study. The treatment types used were as follows: strict Cognitive Behavioral Therapy (CBT), eclectic with no cognitive behavioral components, eclectic with some cognitive behavioral components, Eye Movement Desensitization and Reprocessing (EMDR), exposure, massage, relaxation, and psychological debriefing/crisis intervention. Other novel approaches were coded as described in the intervention studies as spiritual hypnosis [15] and as Huggy Puppy [14].

The analysis of each intervention study identified the treatment approach, the use of the diagnosis of PTSD and/or PTSD symptoms or reactions as target symptoms or outcome, and the use of inclusion and exclusion criteria;

the assessment instruments used and the examination of their psychometrics if applied to the study described; the methodological rigor of the studies including the use of blinded assessment and random assignment as well as the type of control; and the use of assessor training. The quality of the description of the intervention, including the use of manuals; the use of methods to assess treatment adherence; and the adequacy of the power to detect small, medium, and large treatment effects were also examined. Finally, the efficacy of the interventions with respect to PTSD and/or posttraumatic stress symptoms or reactions, was determined.

### 2.2. Target symptoms

In operationalizing "clearly defined target symptoms," Foa and Meadows [13] require the study of "significant" trauma-related psychopathology such as PTSD or a related condition to assess improvement associated with the intervention (p. 453). They describe the importance of using diagnostic status and a threshold of symptom severity as well as inclusion criteria for beginning treatment and exclusion criteria to exclude participants with, for example, comorbid conditions. Because disaster interventions are commonly used in nonclinical populations of children in the general community population who suffer a range of exposures to a disaster, the use of diagnosis as an enrollment criterion or an outcome measure is not always appropriate. Thus, in this analysis, which focused on posttraumatic stress outcomes, posttraumatic stress symptoms or reactions were identified as the primary outcome if the study did not address the diagnosis of PTSD. The use of inclusion and/or exclusion criteria was determined as yes or no for each.

### 2.3. Reliable and valid measures

The instruments used in the studies were documented as was psychometric information if it was assessed for the sample studied.

### 2.4. Blinded assessment

Use of blinded assessment was recorded as yes or no.

### 2.5. Assessor training

Assessor training was evaluated to determine if assessors were trained (coded as yes or no), to indicate if the assessor was or was not a mental health professional (clinician, psychologist, psychiatrist, school counselor, graduate psychology student), and to ascertain if inter-rater reliability was calculated and if procedures to prevent evaluator drift were used during the study (each coded as yes or no).

### 2.6. Manualized treatment and intervention description

Studies were coded as having used a manual, well-described treatment, or neither (yes, well described, and no).

### 2.7. Random assignment and type of control group

Random assignment was coded as yes or no. The type of control group utilized in studies was coded with potential options and definitions including the following: (1) "established treatment" referring to treatments with previously determined effectiveness or efficacy; (2) "not well-established" which included treatment as usual and/or approaches being used but not yet established by previous research, (3) "placebo" controls utilizing techniques presumed to be inert (e.g., engaging participants in a neutral discussion unrelated to the trauma); (4) "waitlist" control group including those using assessment alone for participants who received the active intervention at a later time; and (5) "non-treatment" controls including participants who were assessed but never received an active treatment.

### 2.8. Treatment adherence

Fidelity of treatment was coded as yes or no depending on whether studies utilized methods to ensure that adherence to specific treatment protocols was maintained (e.g., independent coders rated clinicians' adherence during sessions).

### 2.9. Statistical approach and analysis

Studies were coded according to whether the authors performed an intent-to-treat analysis (whether all participants assigned to a treatment arm, including those who did not complete the trial, were included in the analysis) [16,17]. Sample size and the type of statistical technique used to assess treatment effects were documented.

The adequacy of the power of 27 studies to detect small, medium, and large treatment effects was evaluated using the GPower 3.1.2 "post hoc" routine [18]. In two studies where a hierarchical linear model assessed the effectiveness of the intervention [19,20], the Optimal Design software version 3.01 was used for the power calculation [21]. Power analysis for repeated-measures ANOVA, ANCOVA and MANOVA designs relied on effect size *f*, with large, medium, and small effects identified as .40, .25, and .10, respectively [22]. The same values of effect sizes were used for the power analysis of the linear mixed-effects models. Since none of the studies reported the correlation between the pretest and posttest measurements, an estimated rho of .70 was used, in keeping with meta-analytic mean test–retest reliability estimates reported for measures similar to PTSD outcomes [23–25]. Estimating power for independent- and dependent-samples *t*-tests utilized *dz* as the effect size index, with large, medium, and small effects identified as .80, .50, and .20, respectively [22]. Wilcoxon signed rank tests and Wilcoxon Mann–Whitney *U* tests utilized *r* as the effect size index, with "strong," "moderate," and "weak" effects identified as .50, .30, and .10, respectively [22]. Finally, for independent chi-square tests, large, medium, and small effect sizes were identified as $w$ = .50, .30, and .10, respectively [22]. The

generally accepted benchmark of power = .80 was used to assess power adequacy in all cases.

### 2.10. Efficacy of interventions

Efficacy was rated with respect to control group(s). Categories identified were "superior," "not significantly different," or "inferior" to established treatment, not well-established treatment, placebo control, waitlist control, or non-treatment control.

## 3. Results

Most studies reviewed in this analysis examined discrete interventions in relation to a control group. One paper described two distinct studies using the same novel intervention ("Huggy Puppy"), delivered individually in one study and to groups of children in the second [14]. Two studies compared group and individual applications of the same intervention—one studied an eclectic intervention with no cognitive behavioral components in children after Hurricane Iniki [26] and one studied an eclectic cognitive behavioral intervention in children after Hurricane Katrina [27]. In their 2001 study, Layne and colleagues [28] compared children who received full and partial treatment with an eclectic cognitive behavioral intervention, and in their later study published in 2008 [29], they compared two tiers of a three-tiered mental health program for war-exposed adolescents with traumatic bereavement and post-war adversities.

### 3.1. Target symptoms

By design, all studies reviewed for this analysis used PTSD and/or PTSD symptoms or reactions as the target symptoms or outcome. All studies reviewed measured posttraumatic stress symptoms and/or reactions and 9 (31.0%) also examined the diagnosis of PTSD. Use of inclusion and exclusion criteria, as well as the treatment approach and other aspects of each study's research design, is presented in Table 1. More studies ($n$ = 21, 72.4%) used inclusion than exclusion ($n$ = 10, 34.5%) criteria.

### 3.2. Reliable and valid measures

The majority (89.7%) of studies used previously developed assessment tools, although two created new instruments [14,15]. Among the 26 studies that used a standardized PTSD assessment, 2 (7.7%) modified the standardized PTSD measures in some way to address specific issues [30,31]. With respect to the reliability of the study measures, 12 studies (41.4%) reported a reliability statistic for the instrument in their study sample, including internal consistency only (8 studies), internal consistency and test–retest reliability (2 studies), and inter-rater reliability (2 studies). The reported statistics ranged from 0.65 to 0.89 with only one below 0.70 (test–retest reliability = 0.65). See Table 1 for information on the instruments used.

### 3.3. Blinded assessment

In 16 (55.2%) studies, the assessment was not blind. Thirteen (44.8%) studies used single-blinded assessment and no studies used double-blinded assessment. See Table 1.

### 3.4. Assessor training

Table 2 details findings on assessor training. Fourteen studies (48.3%) used mental health professionals to conduct assessments, 5 (17.2%) used non-mental health professionals, 1 (3.4%) used both mental health and non-mental health professionals, 2 (6.9%) used child self-report, and 7 (24.1%) did not specify the qualification of the assessor. Almost two-thirds of the studies ($n$ = 19, 65.5%) indicated that assessors were trained, 4 (13.8%) demonstrated inter-rater reliability among assessors, and 4 (13.8%) reported implementing procedures to prevent drift in outcome assessment.

### 3.5. Manualized treatment and intervention description

The majority of studies used a manual to guide the treatment ($n$ = 20, 69.0%) or delivered a well-described intervention ($n$ = 7, 24.1%). See Table 1.

### 3.6. Random assignment and type of control

Fifteen (51.7%) studies used random assignment of individual participants, 5 (17.2%) used random assignment by group (cluster randomization), and 9 (31.0%) did not use random assignment. Of the five cluster randomized studies, two were randomized by classroom [14,32] and three were randomized by school [19,20,33]. See Table 1.

One study used two controlled conditions—a non-treatment control and a not well-established treatment (teacher psychoeducation)—in studying a debriefing intervention [34]. Wait-list comparison was the most commonly used control ($n$ = 12, 41.4%); 8 (27.6%) studies compared an intervention to a non-treatment control; 8 (27.6%) compared an intervention to a not well-established treatment; 1 (3.4%) used a placebo control; and 1 (3.4%) compared an intervention to established treatment.

### 3.7. Treatment adherence

Almost one-half ($n$ = 14, 48.3%) of the studies measured treatment adherence. See Table 1.

### 3.8. Statistical approach and analysis

Table 3 presents information on the statistical analyses used in the studies. Fifteen (75.0%) of the randomized trials performed an intent-to-treat analysis, that is, all the participants randomly assigned to an intervention group were included in the statistical analysis. All 15 studies assessed effects on posttraumatic stress symptoms or

Table 1
Research methods and designs.

| Study | Treatment[a] approach | PTS/ PTSD[b] | Inclusion/ exclusion | Instrument[c] PTS/PTSD | Reliability[d] | Blinded | Manual[e] | Random[f] | Control[g] | Adherence | Outcome[h] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Berger et al., 2007 [32] | Eclectic with CBT | Y/Y | Y/N | UPID/Six questions determining criteria A1 and A2 for PTSD diagnosis | – | Y | Y | C | WC | Y | Superior |
| Catani et al., 2009 [35] | Exposure | Y/Y | Y/Y | UPID/UPID + 5 items on functioning | – | Y | Y | Y | NWE | Y | NSD[i] |
| Chemtob et al., 2002 [43] | EMDR | Y/Y | Y/N | CPTSD-RI/CPTSD-RI (score 12 or above) | α = .87 | Y | Y | Y | WC | Y | NSD[i] |
| Chemtob et al., 2002 [26] | Eclectic, no CBT | Y/N | Y/N | KRI; CPTDS-RI | – | Y | Y | Y | NWE | Y | NSD |
| Ehntholt et al., 2005 [54] | Eclectic with CBT | Y/N | N/Y | CRIES | – | N | Y | N | WC | N | Superior |
| Gelkopf and Berger, 2009 [56] | Eclectic with CBT | Y/Y | N/N | UPID/Exposure to terrorist attack questionnaire determining criteria A1 and A2 for PTSD diagnosis | α = .89<br>α = .82 for subjective exposure questionnaire | Y | Y | Y | WC | Y | Superior |
| Gilboa-Schechtman et al., 2010 [36] | Exposure | Y/Y | Y/Y | CPSS/Diagnostic interview | – | Y | Y | Y | NWE | Y | Superior |
| Goenjian et al., 1997 [46] | Eclectic with CBT | Y/N | N/N | CPTDS-RI | – | N | WD | N | NTC | N | Superior |
| Goenjian et al., 2005 [57] | Eclectic with CBT | Y/N | N/N | CPTDS-RI | – | N | WD | N | NTC | N | Superior |
| Gordon et al., 2008 [41] | Relaxation | Y/N | Y/N | HTQ | α = .84 | N | Y | Y | WC | Y | Superior |
| Jaycox et al., 2010 [30] | Strict CBT | Y/N | Y/N | CPSS | – | N | Y | Y | ET | N | NSD[i] |
| Jordans et al., 2010 [19] | Eclectic with CBT | Y/N | Y/Y | CPSS | α = .81<br>TR = 0.85 | N | Y | C | WC | N | NSD[i] |
| Kemp et al., 2010 [42] | EMDR | Y/N | Y/Y | UPID | – | N | Y | Y | WC | Y | Superior |
| Layne et al., 2001 [28] | Eclectic with CBT | Y/N | Y/N | UPID | – | N | Y | N | NWE | Y | NSD[i] |
| Layne et al., 2008 [29] | Eclectic with CBT | Y/N | Y/Y | UPID | – | N | Y | Y | NWE | Y | Superior |
| Lesmana et al., 2009 [15] | Spiritual Hypnosis | Y/N | Y/N | Authors' own instrument | α = .70 (T1)<br>α = .82 (T2) | Y | WD | Y | NTC | N | Superior |
| Ronan and Johnston, 1999 [33] | Eclectic with CBT | Y/N | N/N | CPTDS-RI | α = .81 | N | WD | C | NWE | N | NSD |
| Ruf et al., 2010 [40] | Exposure | Y/N | Y/Y | UPID | – | Y | Y | Y | WC | N | Superior |

(*continued on next page*)

Table 1 (*continued*)

| Study | Treatment[a] approach | PTS/ PTSD[b] | Inclusion/ exclusion | Instrument[c] PTS/PTSD | Reliability[d] | Blinded | Manual[e] | Random[f] | Control[g] | Adherence | Outcome[h] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Sadeh et al., 2008 [14] (Study I) | Huggy Puppy | Y/N | N/N | SRCL | α = .79 for stress reaction scale | N | N | N | NTC | N | Superior |
| Sadeh et al., 2008 [14] (Study II) | Huggy Puppy | Y/N | N/N | SRCL | α = .76 for stress reaction scale | Y | WD | C | NTC | N | Superior |
| Salloum and Overstreet, 2008 [27] | Eclectic with CBT | Y/N | Y/N | UPID | α = .86 | Y | Y | Y | NWE | Y | NSD[i] |
| Scheeringa et al., 2011 [58] | Strict CBT | Y/Y | Y/Y | PAPA/PAPA (DSM-IV and alternative algorithms for PTSD) | – | N | Y | Y | WC | Y | Superior |
| Shooshtary et al., 2008 [49] | Eclectic with CBT | Y/N | Y/N | IES-R | – | N | Y | N | WC | N | Superior |
| Smith et al., 2007 [37] | Strict CBT | Y/Y | Y/Y | CAPS-CA, CPSS, CRIES/CAPS-CA | kappa = 0.82 (CAPS-CA) | Y | Y | Y | WC | Y | Superior |
| Stallard et al., 2006 [38] | Debriefing | Y/Y | Y/N | CRIES/CAPS-CA | kappa = 0.84 (CAPS-CA) | Y | Y | Y | PC | N | NSD[i] |
| Thabet et al., 2005 [34] | Debriefing | Y/Y | Y/N | CPTSD-RI/CPTSD-RI (Score 25 or above) | – | N | WD | N | NTC and NWE | N | NSD[i] |
| Tol et al., 2008 [20] | Eclectic with CBT | Y/N | Y/Y | CPSS | α = .85 TR = 0.65 | N | Y | C | WC | Y | Superior |
| Vijayakumar et al., 2006 [31] | Eclectic with CBT | Y/N | Y/N | CPTSD-RI | – | N | WD | N | NTC | N | NSD[i] |
| Wolmer et al., 2005 [39] | Eclectic with CBT | Y/N | N/N | CPTSD-RI | – | Y | N | N | NTC | N | NSD[i] |

– denotes data not reported. Y = yes; N = no.

[a] CBT = cognitive behavioral therapy; EMDR = eye movement desensitization and reprocessing.

[b] PTS = posttraumatic stress disorder symptoms or reactions; PTSD = diagnosis of posttraumatic stress disorder.

[c] CAPS-CA = Clinician-Administered PTSD Scale for Children and Adolescents; CPSS = Child PTSD Symptom Scale; CPTSD-RI = Child PTSD Reaction Index; CRIES = Children's Revised Impact of Event Scale; HTQ = Harvard Trauma Questionnaire; IES-R = Impact of Event Scale-Revised; KRI = Kauai Recovery Inventory; PAPA = Preschool Age Psychiatric Assessment; SRCL = Stress Reaction Checklist (instrument devised by the authors); UPID = UCLA PTSD Index for DSM-IV (Child version).

[d] α = internal consistency; kappa = inter-rater reliability; TR = test–retest reliability.

[e] WD = well described (no treatment manual).

[f] C = cluster randomized.

[g] ET = established treatment; NWE = not well-established treatment; NTC = non-treatment control; PC = placebo control; WC = waitlist control.

[h] NSD = not significantly different.

[i] Insufficient statistical power to detect a small difference.

Table 2
Assessor training and procedures.

| Study | Assessor qualification[a] | Training of assessors | Demonstrated inter-rater reliability among assessors | Procedures to prevent drift in outcome assessment |
|---|---|---|---|---|
| Berger et al., 2007 [32] | MHP | Y | N | NS |
| Catani et al., 2009 [35] | Non-MHP | Y | N | Y |
| Chemtob et al., 2002 [43] | Non-MHP | Y | N | NS |
| Chemtob et al., 2002 [26] | Non-MHP (KRI)<br>MHP (CPTSD-RI) | NS | N | NS |
| Ehntholt et al., 2005 [54] | NS | NS | N | NS |
| Gelkopf and Berger, 2009 [56] | Self-administered | NA | NA | NA |
| Gilboa-Schechtman et al., 2010 [36] | MHP | Y | N | Y |
| Goenjian et al., 1997 [46] | MHP | Y | N | NS |
| Goenjian et al., 2005 [57] | Self-administered | NA | NA | NA |
| Gordon et al., 2008 [41] | Non-MHP | NS | N | NS |
| Jaycox et al., 2010 [30] | NS | Y | N | NS |
| Jordans et al., 2010 [19] | Non-MHP | Y | Y | NS |
| Kemp et al., 2010 [42] | MHP | NS | N | NS |
| Layne et al., 2001 [28] | MHP | Y | N | Y |
| Layne et al. 2008 [29] | MHP | Y | N | NS |
| Lesmana et al., 2009 [15] | MHP | Y | N | NS |
| Ronan and Johnston, 1999 [33] | MHP | Y | NA (1 assessor) | NS |
| Ruf et al., 2010 [40] | MHP | Y | N | NS |
| Sadeh et al., 2008 [14] (Study I) | NS | NS | N | NS |
| Sadeh et al., 2008 [14] (Study II) | NS | NS | N | NS |
| Salloum and Overstreet, 2008 [27] | MHP | Y | N | NS |
| Scheeringa et al., 2011 [58] | NS | Y | N | Y |
| Shooshtary et al., 2008 [49] | MHP | Y | N | NS |
| Smith et al., 2007 [37] | MHP | Y | Y | NS |
| Stallard et al., 2006 [38] | NS | Y | Y | NS |
| Thabet et al., 2005 [34] | MHP | NS | N | NS |
| Tol et al., 2008 [20] | MHP | Y | Y | NS |
| Vijayakumar et al., 2006 [31] | NS | NS | N | NS |
| Wolmer et al., 2005 [39] | Non-MHP | Y | N | NS |

NA = not applicable; NS = not specified; Y = yes; N = no.

[a] MHP = Mental health professional includes psychologists, psychiatrists, school counselors, psychology graduate students. Some authors mentioned "Mental Health Professional" [49] or clinicians [26,34]; KRI = Kauai Reaction Index; CPTSD-RI = Child Posttraumatic Stress Disorder Reaction Index.

reactions. Statistical techniques included an independent-samples *t*-test ($n$ = 1, 3.4%), Wilcoxon non-parametric test ($n$ = 2, 6.9%), Fisher's exact ($n$ = 1, 3.4%), ANOVA/ANCOVA/MANOVA ($n$ = 23, 79.3%), and hierarchical linear models ($n$ = 2, 6.9%). The majority ($n$ = 20, 87.0%) of the studies utilizing ANOVA/ANCOVA/MANOVA tested both within-subject (i.e., pre/post) and between-subject variables (i.e., intervention versus control). Five assessed effects on PTSD diagnosis using an independent chi-square test [34–38].

For the assessment of posttraumatic stress symptoms or reactions, the large majority of the studies ($n$ = 22, 75.9%) had sufficient power (i.e., .80) to detect a large effect size and two-thirds ($n$ = 19, 65.5%) had sufficient power to detect a moderate effect size. Only 5 (17.2%) studies had sufficient power to detect a small effect size. Among the 12 studies that reported no significant difference between the intervention and the control in effect on posttraumatic stress symptoms or reactions, 3 (25.0%), 4 (33.3%), and 11 (91.7%) did not have enough power to detect a large, medium, and small effect size, respectively. Among the five studies that tested for intervention effect on PTSD diagnosis, four, two, and none had sufficient power to detect a large, medium, and small effect size, respectively. See Table 3, which displays findings related to the statistical analyses used in the studies.

### 3.9. Efficacy of interventions

Most of the research found the intervention studied to be efficacious, with outcomes for children receiving the intervention superior to those of children in the control group ($n$ = 17, 58.6%). Interventions were not significantly different to the control in 12 (41.4%) studies. See Table 1.

Six studies (20.7%) found no significant difference between the intervention and a not well-established control [26–28,33–35]. Three studies (10.3%) found no significant difference between the intervention and a non-treatment control [31,34,39]. One study (3.4%) found no significant difference between two cognitive behavioral interventions both associated with clinically significant gains for the participating children [30]. See Table 1.

The majority of studies ($n$ = 17, 58.6%) examined cognitive behavioral interventions or included cognitive behavioral techniques. Of these 17 trials using cognitive

Table 3
Statistical and power analyses.

| Study | Intent-to-treat analysis PTS/PTSD[a] | Sample size | Statistical technique | Potential power | | |
|---|---|---|---|---|---|---|
| | | | | Small ES | Medium ES | Large ES |
| Berger et al., 2007 [32] | Y/ONA | 142 | $ANOVA_{RM\ WB}$ | .86 | 1.0 | 1.0 |
| Catani et al., 2009 [35] | Y/Y | 31 | $ANOVA_{RM\ WB}$ | .32 | .98 | 1.0 |
| | | 31 | Chi-square | .09 | .39 | .80 |
| Chemtob et al., 2002 [43] | Y/ONA | 32 | $ANOVA_{RM\ WB}$ | .33 | .98 | 1.0 |
| Chemtob et al., 2002 [26] | N/ONM | 214 | $ANOVA_{RM\ WB}$ | .96 | 1.0 | 1.0 |
| Ehntholt, Smith and Yule, 2005 [54] | NA | 26 | $ANCOVA_{B}$ | .08 | .23 | .50 |
| Gelkopf and Berger, 2009 [56] | Y/ONA | 107 | $ANOVA_{RM\ WB}$ | .75 | 1.0 | 1.0 |
| Gilboa-Schechtman et al., 2010 [36] | Y/Y | 38 | $ANOVA_{RM\ WB}$ | .34 | .97 | 1.0 |
| | | 38 | Chi-square | .09 | .46 | .87 |
| Goenjian et al., 1997 [46] | NA | 64 | $MANOVA_{RM\ WB}$ | .10 | .52 | .96 |
| Goenjian et al., 2005 [57] | NA | 63 | *t*-Test | .19 | .62 | .93 |
| Gordon et al., 2008 [41] | Y/ONM | 82 | $ANOVA_{RM\ WB}$ | .64 | 1.0 | 1.0 |
| Jaycox et al. 2010 [30] | N/ONM | 71 | Fisher's exact | 0.26[b] | – | – |
| Jordans et al., 2010[c] [19] | Y/ONM | 325 (8 clusters) | Linear mixed model | .06 | .09 | .15 |
| Kemp et al., 2010 [42] | N/ONM | 24 | $MANOVA_{RM\ WB}$ | .08 | .22 | .47 |
| Layne et al., 2001 [28] | NA | 55 | $ANOVA_{RM\ WB}$ | .47 | 1.0 | 1.0 |
| Layne et al., 2008 [29] | N/ONM | 127 | $MANOVA_{RM\ WB}$ | .20 | .80 | .99 |
| Lesmana et al., 2009 [15] | Y/ONM | 226 | $MANOVA_{B}$ | .99 | 1.0 | 1.0 |
| Ronan and Johnston, 1999 [33] | Y/ONM | 112 | $ANOVA_{RM\ BW}$ | .78 | 1.0 | 1.0 |
| Ruf et al., 2010 [40] | Y[d]/ONM | 25 | $ANOVA_{RM\ WB}$ | .29 | .97 | 1.0 |
| Sadeh et al., 2008 [14] (Study I) | NA | 62 | $ANOVA_{RM\ WB}$ | .52 | 1.0 | 1.0 |
| Sadeh et al., 2008 [14] Study II) | Y/ONM | 292 | $ANOVA_{RM\ WB}$ | .99 | 1.0 | 1.0 |
| Salloum and Overstreet, 2008 [27] | Y[d]/ONM | 45 | $ANOVA_{RM\ WB}$ | .45 | 1.0 | 1.0 |
| Scheeringa et al., 2011 [58] | N/ONA | 30 | $ANOVA_{RM\ WB}$ | .28 | .93 | 1.0 |
| Shooshtary et al., 2008 [49] | NA | 168 | $ANOVA_{RM\ WB}$ | .91 | 1.0 | 1.0 |
| Smith et al., 2007 [37] | Y/Y | 24 | $MANCOVA_{RM\ WB}$ | .08 | .22 | .47 |
| | | 24 | Chi-square | .08 | .31 | .69 |
| Stallard et al., 2006 [38] | N/Y | 132 | $Wilcoxon_{MW}$ | .09 | .35 | .79 |
| | | 158 | Chi-square | .24 | .96 | 1.0 |
| Thabet et al., 2005 [34] | NA | 111 | $ANOVA_{B}$ | .14 | .64 | .97 |
| | | | Chi-square | .14 | .82 | 1.0 |
| Tol et al., 2008[e] [20] | Y/ONM | 403 (14 clusters) | Linear mixed-effects model | .09 | 0.30 | 0.63 |
| Vijayakumar et al., 2006 [31] | NA | 114 | $Wilcoxon_{MP}$ | .17 | .84 | 1.0 |
| Wolmer et al., 2005 [39] | NA | 287 | $MANOVA_{RM\ WB}$ | .20 | .92 | .99 |

Alpha = .05 (type I error probability). Rho = .7 (correlation among measurements over time). Small, medium, and large effects sizes are based on Cohen (1988). For *t*-tests, $dz$ = .2, .5, and .8 for small, medium, and large, respectively. For ANOVA/ANCOVA/MANOVA, $f$ = .1, .25, .4; Wilcoxon, $r$ = .1, .3, .5; Chi-square, $w$ = .1, .3, .5; MP = matched pairs; MW = Mann–Whitney; RM = repeated measures; W = within groups; B = between groups.

[a] PTS = posttraumatic stress disorder symptoms or reactions; PTSD = diagnosis of posttraumatic stress disorder; ONA = outcome not analyzed (outcome was measured in the study, but no statistical tests assessing treatment effect were conducted); ONM = outcome not measured (outcome was not measured in the study); NA = not applicable.

[b] The power of the Fisher's exact test was computed regardless of the effect size.

[c] Effect size $f$ = .1, .25, .4; intra-class correlation $\rho$ = 0.219.

[d] The authors performed both an intent-to-treat and a treatment completer analysis and found no difference.

[e] Effect size $f$ = .1, .25, .4; intra-class correlation $\rho$ = 0.050.

behavioral interventions, 10 (58.8%) found the intervention superior to the control. This proportion of 58.8% was not significantly different from the proportion of studies reporting a superior intervention among trials which examined other types of interventions ($n$ = 7, 58.3%) (two-sided Fisher's exact test $p$ = 1.0).

Exposure therapy was supported in studies showing benefit relative to a not well-established treatment (time-limited psychodynamic treatment) [36] and a waitlist control [40] and not significantly different from a not well-established treatment (relaxation) in one study [35]. Ronan and Johnston [33] found an eclectic intervention with exposure not significantly different from an eclectic cognitive behavioral intervention. Relaxation was superior to a waitlist control in one study [41] and equal to exposure therapy in another [35].

The two debriefing studies failed to show benefit above that of the control conditions [34,38]. The EMDR condition was superior to a waitlist control in one study [42]. In another study, the EMDR condition was not significantly different from the waitlist condition though there was a significant reduction of symptoms from pre- to post-intervention [43].

Spiritual hypnosis was supported over a non-treatment control in one study [15]. A novel intervention encouraging children to care for a doll was supported over a non-treatment control group in two studies described in one paper [14]. See Table 1.

## 4. Discussion

Approaches used to develop clinical guidelines and those advanced to evaluate investigations of intervention effectiveness identify randomized controlled trials, followed by well-designed controlled trials without random assignment, as the strongest methodological studies conveying the strongest evidence for the treatment. Thus, selecting only controlled trials for review in this analysis of efficacy studies of child disaster interventions set a high bar for methodological rigor. The discussion focuses on the strength of the findings in light of the choice of control groups, the use of manuals and fidelity measures, the efficacy of intervention types and status of the evidence base, statistical considerations, and the clinical significance and meaningfulness of the results.

### *4.1. Efficacy of interventions relative to controlled conditions*

The use of control groups makes it possible to determine if observed changes can be attributed to the intervention rather than to unrelated factors such as the mere passage of time [44]. A hierarchy of rigor in controlled conditions clarifies the interpretation of outcomes related to intervention efficacy. First, direct comparison to no treatment or to a waitlist control is used to determine if an intervention is more efficacious than no treatment [45]. Eight studies in this review used non-treatment controls; of these, five studies showed the intervention superior and three studies yielded non-significant differences relative to the non-treatment condition. Ten of the 12 studies using waitlist comparison showed the intervention superior to the control condition. These studies using non-treatment or waitlist controls provided evidence that interventions in general were efficacious. It is possible, however, that children in non-treatment or waitlist controls have a lesser expectation of recovery skewing results in favor of the intervention condition. Smith and colleagues [37] found that children in a waitlist comparison group evidenced considerable improvement supporting the importance of natural recovery, the benefit of monitoring, and/or the influence of an expectation of benefit [44] in children's progress post event. Thus, use of waitlist controls may be ethically viable in disaster settings. Goenjian and colleagues [46], however, found that children in the comparison group who did not receive treatment after a devastating earthquake experienced worsening of posttraumatic stress and depressive symptoms raising potential ethical concerns about withholding treatment. Thus, additional research evaluating the benefits and costs of waitlist controls is warranted.

Superiority to a placebo control is evidence that improvement is the result of the intervention rather than simply the effect of expectations or of simply receiving attention [44,45]. The only study using a placebo comparison group found no significant difference between a "debriefing" intervention and an unstructured discussion unrelated to the children's trauma [38] suggesting that receiving professional attention [38] or the expectation of improvement [44] alone may enhance recovery from disaster experiences. The statistical power in this study was insufficient to detect a medium difference in posttraumatic stress symptoms and a small difference in PTSD, however, making it unclear if the failure to find a significant difference was simply due to the small sample size [38].

Comparison with another intervention—established treatment and not well-established treatment—controls for processes that are independent of treatment and those that are common to all treatment and provides support that one treatment is more efficacious than another [45]. Comparison with a rival intervention offers the greatest scientific rigor by providing evidence for a specific mechanism of effect [45]. Only two studies found an intervention superior to another intervention [29,36] while seven studies found no significant difference between the study intervention and other interventions [26–28,30,33–35]. Six [27,28,30,33–35] of the seven studies that did not find a significant difference between the intervention being tested and the control intervention did not have sufficient statistical power to detect a small difference. Thus, it is unclear if the experimental and control interventions were equivalent or if the failure to establish a significant difference was the result of the insufficient sample size. Gilboa-Schechtman and colleagues [36] found exposure superior to time-limited psychodynamic therapy after treatment and at 6-month follow-up but not at 17-month follow-up, and Layne and colleagues [29] found the use of a cognitive behavioral school-based intervention along with a psychoeducation and skills intervention superior to the psychoeducation and skills intervention alone. These findings fail to establish the efficacy of any specific intervention approach.

An additive or constructive treatment design is one in which children in both the treatment and control conditions receive the same active intervention and those in the treatment condition receive an additional component [44]. Using this design in their 2008 study, Layne's team [29] showed that the delivery of an eclectic cognitive behavioral intervention focused on trauma and grief provided additional benefit to children all of whom received a classroom psychoeducation and skills intervention. One could consider their earlier study, which compared partial and full implementation of the intervention, an example of this design as well [28].

Nock and colleagues [44] refer to comparisons in the ways in which interventions are delivered as parametric treatment. Jaycox and colleagues [30] failed to establish superiority between two similar eclectic cognitive behavioral

interventions—one with individual and conjoint parent/child sessions delivered in a community clinic and one delivered in group sessions, with one to three individual sessions, at school sites. Supporting the importance of accessibility of services, more children completed the school-based intervention than the intervention delivered in the community clinic. Chemtob and colleagues [26] and Salloum and Overstreet [27] used a parametric treatment design in their studies comparing interventions delivered to children individually and in groups; in both studies, the investigators failed to find superiority of either delivery approach.

### 4.2. Manuals and fidelity measures

Fidelity measures enhance the ability to administer an intervention as it was intended to be implemented across various participant and provider groups [10,12]. Treatment manuals are commonly used to convey the principles and procedures for implementation. They vary in the degree to which they provide explicit instructions for sessions and in the flexibility they allow in application [47]. Manuals have the added advantage of enhancing and refining the clinical skills of local practitioners who tend to have a range of experience working with children in the disaster context; thus, manuals likely facilitate rapid community-based responses. The analysis of fidelity used in the studies reported here included strategies beyond manualization. For example, Gilboa-Schechtman and colleagues [36] used approaches like reviewing the treatment manual protocol in group supervision, watching and rating videotapes of treatment sessions, and conducting adherence ratings of therapists to ensure that fidelity to a treatment protocol was maintained. These approaches have the added benefit of strengthening the skill base of local mental health and school professionals to respond to future disasters if needed.

### 4.3. Efficacy of intervention approaches and intervention components

The results of this analysis suggest that the interventions were superior to no treatment, but the studies to date are too few to definitively establish the comparative efficacy of any one treatment approach. The most commonly studied interventions used cognitive behavioral approaches. The proportion of cognitive behavioral interventions found to be superior to a control condition was not significantly different than the proportion of non cognitive behavioral interventions that were superior to a control condition. This is consistent with the conclusion of Gillies and colleagues [48] in their Cochrane Review on the treatment of PTSD in children that, while there is "fair evidence for the effectiveness of psychological therapies," especially cognitive behavioral interventions, there is "no clear evidence" that any one therapy is superior to others (p. 21).

Second in the number of investigations, four studies examined exposure techniques as either the treatment or control condition [33,35,36,40]. Evaluating the same narrative exposure intervention, KIDNET, Ruf and colleagues [40] found the intervention superior to a waitlist control and Catani and colleagues [35] found it not significantly different from a meditation–relaxation intervention. Ronan and Johnston [33] found no significant difference between an exposure intervention and an eclectic cognitive behavioral intervention. As noted above, Gilboa-Schechtman and colleagues [36] found prolonged exposure superior to time-limited psychodynamic therapy after treatment and at 6-month follow-up, but both interventions were efficacious and there was no significant difference in the two interventions at 17-month follow-up. Thus, the composite evidence failed to reveal superiority of exposure therapy over other interventions.

Because of the controversy regarding the use of debriefing in adults, the two studies using this intervention in children are of interest. Thabet's team [34] found no significant difference in outcomes for three conditions—group debriefing, teacher psychoeducation, and no treatment; neither treatment was determined to be efficacious. Stallard and colleagues [38] found significant improvement in children receiving both their individual debriefing intervention and an unstructured discussion unrelated to the child's traumatic experience with no significant difference in outcomes between the two conditions and no evidence that the debriefing intervention was harmful. Shooshtary and colleagues [49] demonstrated efficacy of a four-session eclectic cognitive behavioral group intervention which used debriefing as well as other techniques for adolescents exposed to a massive earthquake in Iran. Unfortunately, these investigators did not determine the efficacy of any specific component of the intervention; thus, the results do not contribute to the debate about the efficacy of debriefing in children.

In general, studies have not sought to identify the therapeutic component(s) responsible for improved outcomes instead measuring outcomes associated with treatment in general rather than any single component of the intervention used. While not designed with that in mind, the two studies by Layne and colleagues [28,29] have contributed to an understanding of the importance of dismantling interventions to identify the components responsible for therapeutic benefit. The next generation of studies will need more complex and sophisticated research designs to determine the "active" ingredient in interventions. Furthermore, studies have not addressed common factors of interventions that may be associated with outcome. These include, for example, the therapeutic relationship, the expectation of therapeutic success, the process of confronting or facing the problem, the opportunity to experience mastery or control over the problem, and attribution of therapeutic success or failure [50].

### 4.4. Statistical analysis, methodology, and power

A majority of the randomized studies conducted an intent-to-treat analysis of the collected data. This analytical approach preserves equality between treatment groups

created by randomization, regarding known and unknown characteristics of the participants that may be associated with treatment effectiveness. Hence, treatment groups remain comparable and the effect of the intervention can be assessed independently from these characteristics [16].

All of the 29 studies evaluated for this report clearly described the statistical methods used and the majority offered sufficient power (greater than .80) to detect large and moderate effect sizes. Few studies, however, had sufficient power to detect small effects [14,15,26,32,49]. Of the 12 studies that did not find a significant difference between the intervention and the control, only 1 (8.3%) had enough power to detect a small difference [26]. The main factors affecting power are sample size and the population effect size and outcome measurement reliability. Power increases as reliability increases because of corresponding reductions in random error in statistical tests [51]. It is recommended that researchers (1) calculate the sample size needed to detect weak, medium, and strong effects (as per the selected statistical test for treatment effects); (2) offer reasoned estimates of the expected effect size (e.g., based on prior similar research); (3) strive to achieve the critical sample size needed to detect the expected effect size at $\alpha = .05$ and $\beta = .20$ (i.e., power = .80); and (4) ensure that their outcome measures are as reliable as practically feasible. These calculations and rationales should be included in the study report to offer a suitable context for the reader in judging the merits of the study as a source of generalizable knowledge.

Detecting small effects may not be feasible in many studies because it may require the assessment of hundreds or thousands of participants. In addition to meta-analyses, researchers have suggested that pooling data across studies [52] or conducting multicenter studies [3] may be beneficial for addressing the power issue. Simple pooling of data, however, is susceptible to the Simpson's paradox, the reversal of the treatment effect observed in the individual studies [53]. On an individual basis, however, researchers studying child disaster interventions, like all researchers, struggle with the dilemma of how to cost-effectively detect meaningful differences in a particular study. Meaningful differences can involve changes in two standard deviations, reductions from the diagnostic range to the non-diagnostic range (i.e., "caseness"), reductions beyond the confidence interval for expected change, or a 25% reduction in presenting problems.

### 4.5. Clinically significant and meaningful results

Establishing statistically significant results in intervention studies is insufficient in assessing their benefit; outcomes related to an intervention also must be meaningful. For example, in their study of refugee children from war-torn countries, Ehntholt and colleagues [54] noted that while those receiving an eclectic cognitive behavioral intervention showed statistically significant better outcomes compared to a waitlist control, the improvements were clinically modest and were not maintained at follow-up. Additionally, the failure to find statistically significant differences does not mean that interventions were "equivalent." Two treatments are said to be equivalent if the difference in their therapeutic effects is significantly smaller than a pre-defined clinically meaningful difference. Thus, the results of a trial reporting no significant difference between the intervention and control groups, but not having sufficient statistical power to detect a difference below an admitted clinically meaningful difference, should not be interpreted as evidence that the two conditions are equivalent rather that there was insufficient power to detect a difference if one existed. Determining what constitutes meaningful results can be challenging especially given the range of conditions and reactions child disaster interventions are intended to address and the populations they are designed to serve. In studies of clinical interventions delivered to children whose experiences suggest the potential for pathological reactions, a clinically meaningful outcome may be assessed in terms of diagnostic status or symptom severity.

Foa and Meadows [13] emphasize the importance of using diagnostic status and a threshold of symptom severity and inclusion criteria in studies of intervention effectiveness. Smith and colleagues [37] used a 4-week symptom-monitoring baseline assessment procedure before initiating treatment for children with PTSD after single-incident trauma. Given that the focus of many disaster interventions is not on clinical populations, however, using diagnosis or symptom severity is not universally appropriate. For example, disaster interventions are commonly delivered to children in schools to provide psychoeducation and support and to normalize reactions in those who would not be expected to suffer psychopathology in relation to their disaster experiences. Thus using diagnosis for inclusion or for measuring outcome, which might improve the interpretation of results in studies of children with psychiatric conditions, is not appropriate for these community-based studies. For example, studies of school-based interventions which have not found a decrease in psychiatric symptoms have found improvement on indicators of psychological difficulty, social behavior, well-being [19], and functioning [39]. Thus, future studies may be enhanced by including a greater range of outcomes beyond psychiatric symptoms. Nonetheless, while it may be unclear what constitutes meaningful change—as opposed to statistically significant change—in the children in these nonclinical studies, it is clear that children who experience distress and nonclinical reactions warrant attention. One appropriate focus in the next generation of studies will be to identify tools that measure and establish meaningful change.

Some disaster-exposed children will require traditional mental health treatment beyond the approaches used in the studies reviewed for this report. The American Academy of Child and Adolescent Psychiatry Practice Parameter on PTSD [9] recommended psychodynamic trauma-focused psychotherapies and psychoactive medications for children with PTSD. Thus, another line of future research should

identify ways to enhance post-disaster interventions for children with persistent symptoms using, for example, individual psychotherapy, family therapy, psychoactive medication, and/or social support.

### 4.6. Limitations

Unfortunately, a lack of uniformity in the information provided across studies and constraints imposed by the samples as well as the need for parsimony limited the content examined in this analysis. For example, the analysis did not address a number of issues that may be important in clinical and community applications including, for example, the type and location of the disaster, child specific issues (e.g., gender, age, pre-existing conditions, prior trauma), characteristics of the interventions (e.g., population being served, number and length of sessions), intervention delivery (e.g., time since the disaster, site of service delivery), attrition in the samples, and methodological and design issues (e.g., use and length of follow-up). The review included only studies that examined PTSD and/or PTSD symptoms or reactions as outcomes and it included studies for which PTSD or PTSD symptoms or reactions were not the only, or even primary, outcome of interest. Some results were skewed for this reason.

## 5. Conclusions

While the scientific studies needed to establish an evidence base for child disaster mental health interventions are accumulating, the field is yet to fully mature. Nevertheless, the identification of this number of controlled trials and the quality of the methods used in many of the studies is impressive especially given the chaotic post-disaster environments in which many of the interventions were implemented and the studies were conducted. These studies suggest that while challenging, it is ethically viable to conduct methodologically rigorous intervention research.

Only cognitive behavioral approaches have been widely researched and there appears to be support for their use. The tentative conclusion is that the child disaster interventions studied in general are beneficial, but the research has been inadequate to determine whether some interventions are superior to other interventions, and the evidence that these interventions provide benefit beyond natural recovery is only preliminary. Moreover, it is unclear what constitutes meaningful change in relation to treatment especially for the numerous interventions administered to general populations of children exposed to a disaster. Finally, the results are inconclusive about which children, what settings, and what approaches are most likely to be beneficial.

Future studies are needed to identify the best child disaster interventions. The field will benefit by continuing to strive for conducting controlled studies that meet the seven gold standards proposed by Foa and Meadows [13] with minor modifications. Since many disaster interventions include public health interventions for nonclinical populations, diagnostic status may not be as important as other clearly defined target symptoms and problems. Conveying more information in reports about exclusion as well as inclusion criteria, the qualification and training of assessors and those who provide the interventions, inter-rater reliability when more than one rater is used, rater drift, power analyses, and assignment to treatment (e.g., random assignment) are easily attainable goals. Addressing treatment adherence concerns in child disaster interventions also simultaneously benefits both science and practice. As in all intervention studies, the ultimate cost–benefit of what standards are feasible will depend on the particular cultural and community disaster recovery needs with the active involvement of community stakeholders in decision making. Researchers should consider communicating explicit discussion of these decisions in future reports to facilitate future research endeavors in this area, including measuring stakeholders' responses to such research participation [55].

## Acknowledgment of support

This work was funded in part by the Substance Abuse and Mental Health Services Administration (1 U79 SM57278) which established the Terrorism and Disaster Center (TDC) at the University of Oklahoma Health Sciences Center. TDC is a partner in the National Child Traumatic Stress Network (NCTSN).

## References

[1] Feeny NC, Treadwell KRH, Foa EB, March J. Posttraumatic stress disorder in youth: a critical review of the cognitive and behavioral treatment outcome literature. Prof Psychol Res Pr 2004;35(5):466-76, http://dx.doi.org/10.1037/0735-7028.35.5.466.

[2] Silverman WK, Ortiz CD, Viswesvaran C. Evidence-based psychosocial treatments for children and adolescents exposed to trauma events. J Clin Child Adol Psychol 2008;37(1):156-83, http://dx.doi.org/10.1080/15374410701818293.

[3] Stallard P. Psychological interventions for post-traumatic reactions in children and young people: a review of randomised controlled trials. Clin Psychol Rev 2006;26:895-911, http://dx.doi.org/10.1016/j.cpr.2005.09.005.

[4] Wethington HR, Han RA, Fuqua-Whitley DS, Sipe TA, Crosby AE, Johnson RL, et al. The effectiveness of interventions to reduce psychological harm form traumatic events among children and adolescents: a systematic review. Am J Prev Med 2008;35(3):287-313, http://dx.doi.org/10.1016/j.amepre.2008.06.024.

[5] Cohen JA, Mannarino AP, Gibson LE, Cozza SJ, Brymer MJ, Murray L. Interventions for children and adolescents following disasters. Interventions following mass violence and disasters: Strategies for mental health practice. New York, NY: The Guilford Press; 2006. p. 227-56.

[6] Adler-Nevo G, Manassis K. Psychosocial treatment of pediatric posttraumatic stress disorder: the neglected field of single-incident trauma. Depression and Anxiety 2005;22:177-89, http://dx.doi.org/10.1002/da.20123.

[7] Persson TJ, Rousseau C. School-based interventions for minors in war-exposed countries: a review of targeted and general programmes. Torture 2009;19(2):88-101.

[8] Rolfsnes ES, Idsoe T. School-based intervention programs for PTSD symptoms: a review and meta-analysis. J Traumatic, Stress 2011;24(2):155-65, http://dx.doi.org/10.1002/jts.20622.
[9] American Academy of Child and Adolescent Psychiatry. Practice parameter for the assessment and treatment of children and adolescents with posttraumatic stress disorders. J Am Acad Child Adol Psychiat 2010;49(4):414-30, http://dx.doi.org/10.1016/j.jaac.2009.12.020.
[10] Foa EB, Keane TM, Friedman MJ, Cohen JA. Introduction. In: Foa EB, Keane TM, Friedman MJ, & Cohen JA, editors. Effective treatments for PTSD: practice guidelines from the International Society for Traumatic Stress Studies. New York, NY: The Guilford Press; 2009. p. 1-20.
[11] Ursano RJ, Bell C, Eth S, Friedman M, Norwood A, Pfefferbaum B, et al. Practice guideline for the treatment of patients with acute stress disorder and posttraumatic stress disorder. Am J Psychiat 2004;161(11, Suppl):3-31, http://dx.doi.org/10.1176/appi.books.9780890423363.52257.
[12] Foa EB, Keane TM, Friedman MJ. Guidelines for treatment of PTSD. J Traumatic, Stress 2000;13(4):539-88, http://dx.doi.org/10.1023/A:1007802031411.
[13] Foa EB, Meadows EA. Psychosocial treatments for posttraumatic stress disorder: a critical review. Ann Rev Psychol 1997;48:449-80, http://dx.doi.org/10.1146/annurev.psych.48.1.449.
[14] Sadeh A, Hen-Gal S, Tikotzky L. Young children's reactions to war-related stress: a survey and assessment of an innovative intervention. Pediatrics 2008;121:46-53, http://dx.doi.org/10.1542/peds.2007-1348.
[15] Lesmana CBJ, Suryani LK, Jensen GD, Tiliopoulos N. A spiritual-hypnosis assisted treatment of children with PTSD after the 2002 Bali terrorist attack. Am J Clin Hypnosis 2009;52:23-34, http://dx.doi.org/10.1080/00029157.2009.10401689.
[16] Newell JD. Intention-to-treat analysis: implications for quantitative and qualitative research. Int J Epidemiol 1992;21(5):837-41, http://dx.doi.org/10.1093/ije/21.5.837.
[17] Wertz RT. Intention to treat: once randomized always analyzed. Clin Aphasiology 1995;23:57-64.
[18] Erdfelder E, Faul F, Buchner A. GPOWER: a general power analysis program. Behavior Research Methods Instruments, & Computers 1996;28:1-11, http://dx.doi.org/10.3758/BF03203630.
[19] Jordans MJD, Komproe IH, Tol WA, Kohrt BA, Luitel NP, Macy RD, et al. Evaluation of a classroom-based psychosocial intervention in conflict-affected Nepal: a cluster randomized controlled trial. J Child Psychol Psychiat 2010;51:818-26, http://dx.doi.org/10.1111/j.1469-7610.2010.02209.x.
[20] Tol W, Komproe IH, Susanty D, Jordans MJD, Macy RD, De Jong JTVM. School-based mental health intervention for children affected by political violence in Indonesia. JAMA 2008;300:655-62, http://dx.doi.org/10.1001/jama.300.6.655.
[21] Raudenbush SW, Spybrook J, Congdon R, Liu X, Martinez A, Bloom H, Hill C. Optimal Design Software for Multi-level and Longitudinal Research. Version 3.01. 2011. [Software]. Available from: , http://www.wtgrantfoundation.org/resources/consultation-service-and-optimal-design
[22] Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Erlbaum; 1988.
[23] de Ayala RJ, Vonderharr-Carlson DJ, Kim D. Assessing the reliability of the Beck Anxiety Inventory scores. Edu Psychol Meas 2005;65:742-56, http://dx.doi.org/10.1177/0013164405278557.
[24] Schuerger JM, Zarrella KL, Hotz AS. Factors that influence the temporal stability of personality by questionnaire. J Pers Soc Psychol 1989;56:777-83, http://dx.doi.org/10.1037/0022-3514.56.5.777.
[25] Yin P, Fan X. Assessing the reliability of Beck Depression Inventory scores. Educ Psychol Meas 2000;60:201-23, http://dx.doi.org/10.1177/00131640021970466.
[26] Chemtob CM, Nakashima JP, Hamada RS. Psychosocial intervention for postdisaster trauma symptoms in elementary school children: a controlled community field study. Arch Pediat Adol Med 2002;156:211-6, http://dx.doi.org/10.1001/archpedi.156.3.211.
[27] Salloum A, Overstreet S. Evaluation of individual and group grief and trauma interventions for children post disaster. J Clin Child Adol Psychol 2008;37:495-507, http://dx.doi.org/10.1080/15374410802148194.
[28] Layne CM, Pynoos RS, Saltzman WR, Arslanagic B, Black M, Savjak N, et al. Trauma/grief-focused group psychotherapy: school-based postwar intervention with traumatized Bosnian adolescents. Group Dynamic: Theory, Research, and Practice 2001;5:277-90, http://dx.doi.org/10.1037/1089-2699.5.4.277.
[29] Layne CM, Saltzman WR, Poppleton L, Burlingame GM, Pasalic A, Durakovic E, et al. Effectiveness of a school-based group psychotherapy program for war-exposed adolescents: a randomized controlled trial. J Am Acad Child Adol Psychiat 2008;47:1048-62, http://dx.doi.org/10.1097/CHI.0b013e31817eecae.
[30] Jaycox LH, Cohen JA, Mannarino AP, Walker DW, Langley AK, Gegenheimer KL, et al. Children's mental health care following hurricane Katrina: a field trial of trauma-focused psychotherapies. J Traumatic Stress 2010;23:223-31, http://dx.doi.org/10.1002/jts.20518.
[31] Vijayakumar L, Kannank GK, Kumar G, Devarajan P. Do all children need intervention after exposure to tsunami? Int Rev Psychiat 2006;18:515-22, http://dx.doi.org/10.1080/09540260601039876.
[32] Berger R, Pat-Horenczyk R, Gelkopf M. School-based intervention for prevention and treatment of elementary-students' terror-related distress in Israel: a quasi-randomized controlled trial. J Traumatic Stress 2007;20(4):541-51, http://dx.doi.org/10.1002/jts.20225.
[33] Ronan KR, Johnston DM. Behaviourally-based interventions for children following volcanic eruptions: an evaluation of effectiveness. Disaster Prevention and Management 1999;8:169-76, http://dx.doi.org/10.1007/s00787-005-0466-7.
[34] Thabet AA, Vostanis P, Karim K. Group crisis intervention for children during ongoing war conflict. Eur Child Adol Psychiat 2005;14:262-9.
[35] Catani C, Kohiladevy M, Ruf M, Schauer E, Elbert T, Neuner F. Treating children traumatized by war and tsunami: a comparison between exposure therapy and meditation–relaxation in North-East Sri Lanka. BMC Psychiatry 2009;9:1-11, http://dx.doi.org/10.1186/1471-244X-9-22.
[36] Gilboa-Schechtman E, Foa EB, Shafran N, Aderka IM, Powers MB, Rachamim L, et al. Prolonged exposure versus dynamic therapy for adolescent PTSD: a pilot randomized controlled trial. J Am Acad ChildAdol Psychiat 2010;49:1034-42, http://dx.doi.org/10.1016/j.jaac.2010.07.014.
[37] Smith P, Yule W, Perrin S, Tranah T, Dalgleish T, Clark DM. Cognitive-behavioral therapy for PTSD in children and adolescents: a preliminary randomized controlled trial. J Am Acad Child Adol Psychiat 2007;46(8):1051-61, http://dx.doi.org/10.1097/CHI.0b013e318067e288.
[38] Stallard P, Velleman R, Salter E, Howse I, Yule W, Taylor G. A randomised controlled trial to determine the effectiveness of an early intervention with children involved in road traffic accidents. J Child Psychol Psychiat 2006;47:127-34, http://dx.doi.org/10.1111/j.1469-7610.2005.01459.x.
[39] Wolmer L, Laor N, Dedeoglu C, Siev J, Yazgan Y. Teacher-mediated intervention after disaster: a controlled three-year follow-up of children's functioning. J Child Psychol Psychiat 2005;46(11):1161-8, http://dx.doi.org/10.1111/j.1469-7610.2005.00416.x.
[40] Ruf M, Schauer M, Neuner F, Catani C, Schauer E, Elbert T. Narrative exposure therapy for 7- to 16-year-olds: a randomized controlled trial with traumatized refugee children. J Traumatic Stress 2010;23:437-45, http://dx.doi.org/10.1002/jts.20548.
[41] Gordon JS, Staples JK, Blyta A, Bytyqi M, Wilson AT. Treatment of posttraumatic stress disorder in postwar Kosovar adolescents using mind-body skills groups: a randomized controlled trial. J Clin Psychiat 2008;69:1469-76, http://dx.doi.org/10.4088/JCP.v69n0915.
[42] Kemp M, Drummond P, McDermott B. A wait-list controlled pilot study of eye movement desensitization and reprocessing (EMDR) for children with post-traumatic stress disorder (PTSD) symptoms from motor vehicle accidents. Clin Child Psychol Psychiat 2010;15:5-25, http://dx.doi.org/10.1177/1359104509339086.

[43] Chemtob CM, Nakashima JP, Carlson JG. Brief treatment for elementary school children with disaster-related posttraumatic stress disorder: A field study. J Clin Psychol 2002;58:99-112, http://dx.doi.org/10.1002/jclp.1131.
[44] Nock MK, Janis IB, Wedig MM. Research designs. In: Nezu AM, & Maguth-Nazu C, editors. Evidence-based outcome research: a practical guide to conducting randomized controlled trials for psychological interventions. New York, NY: Oxford University Press; 2008. p. 201-18.
[45] Chambless DL, Hollon SD. Defining empirically supported therapies. J Consult Clin Psychol 1998;66:7-18, http://dx.doi.org/10.1037/0022-006X.66.1.7.
[46] Goenjian AK, Karayan I, Pynoos RS, Minassian D, Najarian LM, Steinberg AM, et al. Outcome of psychotherapy among early adolescents after trauma. Am J Psychiat 1997;154:536-42.
[47] Chambless DL, Ollendick TH. Empirically supported psychological interventions: controversies and evidence. Ann Rev Psychol 2001;52:685-716, http://dx.doi.org/10.1146/annurev.psych.52.1.685.
[48] Gillies D, Taylor F, Gray C, OBrien L, DAbrew N. Psychological therapies for the treatment of post-traumatic stress disorder in children and adolescents (review). Cochrane Database of Systematic Reviews 2012;12 Art. No.: CD006726. http://dx.doi.org/10.1002/14651858.CD006726.pub2.
[49] Shooshtary MH, Panaghi L, Moghadam JA. Outcome of cognitive behavioral therapy in adolescents after natural disaster. J Adolescent Health 2008;42:466-72, http://dx.doi.org/10.1016/j.jadohealth.2007.09.011.
[50] Weinberger J. Common factors aren't so common: the common factors dilemma. Clin Psychol Sci Pr 1995;2(1):45-69, http://dx.doi.org/10.1111/j.1468-2850.1995.tb00024.x.
[51] Humphreys LG, Drasgow F. Some comments on the relation between reliability and statistical power. Appl Psychol Measurement 1989;13:419-25.
[52] Resick PA, Monson CM, Gutner C. Psychosocial treatments for PTSD. In: Friedman MJ, Keane TM, & Resnick PA, editors. Handbook of PTSD: science and practice. New York, NY: Guilford Press; 2007. p. 330-58.
[53] Bravata DM, Olkin I. Simple pooling versus combining in meta-analysis. Eval Health Prof 2001;24:218-30, http://dx.doi.org/10.1177/01632780122034885.
[54] Ehntholt KA, Smith PA, Yule W. School-based cognitive-behavioural therapy group intervention for refugee children who have experienced war-related trauma. Clin Child Psychol Psychiat 2005;10:235-50, http://dx.doi.org/10.1177/1359104505051214.
[55] Newman E, Kaloupek DG. The risks and benefits of participating in trauma-focused research studies. J Traumatic Stress 2004;17(5):383-94, http://dx.doi.org/10.1023/B:JOTS.0000048951.02568.3a.
[56] Gelkopf M, Berger R. A school-based, teacher-mediated prevention program (ERASE-Stress) for reducing terror-related traumatic reactions in Israeli youth: a quasi-randomized controlled trial. J Child Psychol Psychiat 2009;50:962-71, http://dx.doi.org/10.1111/j.1469-7610.2008.02021.x.
[57] Goenjian AK, Walling D, Steinberg AM, Karayan I, Najarian LM, Pynoos R. A prospective study of posttraumatic stress and depressive reactions among treated and untreated adolescents 5 years after a catastrophic disaster. Am J Psychiat 2005;162:2302-8, http://dx.doi.org/10.1176/appi.ajp.162.12.2302.
[58] Scheeringa MS, Weems CF, Cohen JA, Amaya-Jackson L, Guthrie D. Trauma-focused cognitive-behavioral therapy for posttraumatic stress disorder in three through six year-old children: a randomized clinical trial. J Child Psychol Psychiat 2011;52:853-60, http://dx.doi.org/10.1111/j.1469-7610.2010.02354.x.