Comparative Effectiveness Research Review Disposition of Comments Report

Research Review Title: *Psychological Treatments and Pharmacological Treatments for Adults With Posttraumatic Stress Disorder (PTSD)*

Draft review available for public comment from May 29, 2012 through June 26, 2012.

Research Review Citation: Jonas DE, Cusack K, Forneris CA, Wilkins TM, Sonis J, Middleton JC, Feltner C, Meredith D, Cavanaugh J, Brownley KA, Olmsted KR, Greenblatt A, Weil A, Gaynes BN. Psychological and Pharmacological Treatments for Adults With Posttraumatic Stress Disorder (PTSD). Comparative Effectiveness Review No. 92. (Prepared by the RTI International–University of North Carolina Evidence-based Practice Center under Contract No. 290-2007-10056-I.) AHRQ Publication No. 13-EHC011-EF. Rockville, MD: Agency for Healthcare Research and Quality; April 2013. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Comments to Research Review

The Effective Health Care (EHC) Program encourages the public to participate in the development of its research projects. Each comparative effectiveness research review is posted to the EHC Program Web site in draft form for public comment for a 4-week period. Comments can be submitted via the EHC Program Web site, mail or email. At the conclusion of the public comment period, authors use the commentators' submissions and comments to revise the draft comparative effectiveness research review.

Comments on draft reviews and the authors' responses to the comments are posted for public viewing on the EHC Program Web site approximately 3 months after the final research review is published. Comments are not edited for spelling, grammar, or other content errors. Each comment is listed with the name and affiliation of the commentator, if this information is provided. Commentators are not required to provide their names or affiliations in order to submit suggestions or comments.

The tables below include the responses by the authors of the review to each comment that was submitted for this draft review. The responses to comments in this disposition report are those of the authors, who are responsible for its contents, and do not necessarily represent the views of the Agency for Healthcare Research and Quality.

Source: http://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1435
Published Online: April 3, 2013

Comment Number	Commentator & Affiliation	Section	Comment	Response
1	Peer Reviewer #1	General	This is a “tour de force”, a review very carefully and comprehensively done, on a topic of great and growing clinical and policy importance. Not only does the review provide strong answers for key questions, but also sets forth what areas are not yet well-enough researched. As a result, this is information useful for clinicians and clinical policy makers, as well as guidance for clinical researchers and funding agencies to directions for future research.	Thank you
2	Peer Reviewer #1	General	What I’m about to say next does not gainsay the above. I did have some methodological concerns. Moreover, some of these concerns other researchers/reviewers/statisticians may not be agreement on. So, for the consideration of the author(s):	See below for responses to the comments related to methods referenced here.
3	Peer Reviewer #2	General	I think the report is clinically meaningful and the key questions are well formulated.	Thank you
4	Peer Reviewer #3	General	This report clearly represents an important contribution to the literature on PTSD treatment effectiveness. The report has the potential to be highly clinically relevant given its topic and scope, coupled with AHRQ’s reputation. My comments are offered here and to the authors in the spirit of enhancing the utility and acceptance of the findings. The rigor and completeness of the report are particular strengths. The target population and audience are explicitly defined and the key questions are appropriate and clearly stated.	Thank you. We have addressed each of the referenced comments below. We appreciate that the comments have improved the report and will likely enhance the utility and acceptance of the findings.
5	TEP Reviewer #5	General	I read the executive report in detail and then skimmed most of the rest of the report-concentrating on tables and figures. The report is clinically meaningful. The target population is explicitly defined. Not sure what "audience" is referred to. This report should be useful to PTSD researchers as well as to Policy makers- both at NIH- to fund head to head trials of psych vs. pharm treatments. It should also be a useful synthesis of material for members of professional societies and researchers. It is really useful especially for those not directly in the field as it really captures the hard data on different therapies. It also captures the stark reality that we don't know whether behavioral therapies are better than pharmacological therapies. In addition, it is hard to know whether combining them will produce better outcomes than individual therapy types.	Thank you.
6	TEP Reviewer #6	General	This is an extremely thorough and well written report. The discussion includes consideration of clinical factors that may affect application of the findings and this makes the document more helpful from a clinical standpoint than a typical systematic review (e.g., Cochrane review). The authors have done a good job of developing appropriate, explicitly phrased key questions.	Thank you

Comment Number	Commentator & Affiliation	Section	Comment	Response
7	Anonymous Public Reviewer #2	General	In the ES, on pg. 127, and elsewhere it would be important to note that "absence of evidence is not evidence of absence." That is, there are some models that have not yet been sufficiently tested. The phrase "insufficient evidence" may be misinterpreted to mean that those treatments don't work.	We have made sure to clearly define the meaning of an "insufficient" SOE grade in the Methods section. We have also added the point, as suggested, in the Discussion of the full report. New text (using the SOE definition of insufficient): "When we have graded evidence as insufficient, it indicates that evidence is either unavailable or does not permit estimation of an effect. It does not indicate that a treatment has been proven to lack efficacy."
8	Anonymous Public Reviewer #2	Executive Summary	ES2: "The therapies are delivered predominantly to individuals, but they can also be conducted in a group setting.10, 11 " Should be reworded as "a few have been studied also in group format". As written, it is stating that all of them can be done in group format—but many don't have a single study nor treatment manual for their use in groups.	We have revised the wording to avoid the implication that all of them can be done in group format. It now reads: "The therapies are delivered predominantly to individuals; some can also be conducted in a group setting." We changed the corresponding text in the full report.
9	Anonymous Public Reviewer #2	Executive Summary	ES2: "…identify trauma-focused psychological treatments." The term "trauma-focused" is not defined. Moreover, it's a problematic term—sometimes meaning "any treatment intended to address trauma" and other times meaning "exposure-based" models (i.e., intensive models to explore the past). The term should be clearly defined on first use and then used consistently throughout. Perhaps the term "trauma-specific treatments" would be more apt.	The uses of the term (trauma-focused) throughout our report are limited to instances when we're describing other reports (e.g., guidelines from various organizations) and we are not referring to exposure therapies. Instead, it is the more broad term to describe psychotherapies that are treating PTSD by addressing the trauma. We have clarified the meaning we intend by defining the term when it first shows up in both the ES and in the report.

Comment Number	Commentator & Affiliation	Section	Comment	Response
10	Anonymous Public Reviewer #2	Executive Summary	ES8: "Evidence of moderate strength supports greater effectiveness (1) for exposure therapy than for coping skills for achieving loss of PTSD diagnosis and …" This is problematic. Relaxation (which is one of the main treatments classified as "coping skills" treatment) is not usually considered a bona fide coping skills model. It was usually used as a basic comparison condition that is not generally seen as a PTSD treatment per se. Moreover, relaxation is known to be triggering for some PTSD patients. Also many of the "coping skills" studies (as defined in this draft) were really intended more as comparison treatments rather than strong coping skills approaches in their own right. It seems highly premature to draw such a strong conclusion about coping skills approaches at this point. Cross-reference: pg 24 (main document).	We have revised our approach to the coping skills section and we no longer lump the various "coping skills" comparators with each other. Instead, we now have conducted separate analyses for SIT, relaxation, etc. Following this approach, we no longer make any broad conclusions about coping skills as an entire group. We make our conclusions about SIT, relaxation, etc.
11	Anonymous Public Reviewer #2	Executive Summary	ES11: "For psychological treatments, the vast majority of studies reported no information about adverse effects ..." This is an extremely important issue and it's terrific to see this being raised. There should be mention that there are anecdotal reports in the literature of adverse effects with some PTSD treatments.	Thank you, we agree that this is an important issue to raise. We are unaware of the anecdotal reports mentioned. We don't want to propagate beliefs that are not based on good evidence (i.e., anecdotes); rather, we think this issue should be raised as one requiring future study (as we've done in the report).

Comment Number	Commentator & Affiliation	Section	Comment	Response
12	Anonymous Public Reviewer #2	Executive Summary	ES13: "Given the findings, the magnitude of benefit and SOE found for exposure therapy support its use as a first-line treatment for PTSD. However, other factors must be considered in selecting a treatment for PTSD, including patient preference, access to treatment, and clinical judgment about the appropriateness of an intervention that guide clinical decisionmaking. For example, a majority of the studies reviewed in this report excluded patients with presenting issues such as substance dependence or suicidality. Most clinicians would agree that stabilization of these issues should occur prior to initiating trauma-focused therapy ." This is very well written and an extremely important point. It should go at the end of the first paragraph of the Discussion as it tempers that main finding reported there. Moreover, it is key to add in some of the other major exclusionary criteria, i.e., "a majority of the studies reviewed in this report excluded patients with presenting issues such as substance dependence or suicidality as well as bipolar disorder, psychotic disorders, homelessness, current domestic violence, self-harm, and sometimes other Axis I or Axis II disorders." In addition, in the Appendices, there should be clear identification of exclusionary criteria. Just for your reference, examples of such exclusionary criteria are as follows: Schnurr et al. 2003: Excluded patients with current or lifetime psychotic disorder, mania, or bipolar disorder; current major depression with psychotic features, alcohol or substance dependence; unwillingness to refrain from substance use at treatment or work; significant cognitive impairment; severe cardiovascular disorder. Schnurr et al. 2007: Excluded patients with substance dependence in remission for less than 3 months; current psychotic symptoms, mania, or bipolar disorder; prominent suicidal or homicidal ideation; significant cognitive impairment; self-mutilation within the previous 6 months; involvement in a violent relationship.	Thank you, we agree that it is an important point. We have added several additional paragraphs after the one mentioned here by the reviewer to the Applicability section of the full report. The new text aims to address the exclusions made by various studies---those that the reviewer highlights as well as others clinicians may be most interested in. The new text is now the last 6 paragraphs of the Applicability section of the full report and includes quantification of the percentages of included trials of psychological treatments and of pharmacological treatments that set various exclusion criteria. We have also added a reference to this additional information in the paragraph the reviewer mentioned here from ES 13. We have not added the suggested underlined information to the ES, as it is not quite accurate and really oversimplifies the information.

Comment Number	Commentator & Affiliation	Section	Comment	Response
13	Anonymous Public Reviewer #2	Executive Summary	ES14: "The included studies assessing efficacious treatments generally enrolled subjects from outpatient settings. " This point also is extremely important and would be very helpful to see at the end of the first paragraph of the Discussion of the ES. Essentially the most severe PTSD patients—those on inpatient units, in residential treatment, in day programs—are left out of the literature by and large. Many clinicians believe these are exactly the kinds of patients who get worse if given certain intense PTSD treatments prematurely.	Thank you, we agree that this is an important point. We prefer to keep this information where it is currently located (in the Applicability section of the Discussion), as we feel this is the appropriate location, rather than to move it to the end of the first paragraph of the Discussion of the ES as suggested.
14	Anonymous Public Reviewer #2	Executive Summary	ES16: "Future studies could focus on comparisons between (1) the psychological treatments with the best evidence of efficacy, " This has the potential for unintended negative consequences such as prematurely closing off the field from many important studies of treatments that do not yet have the "best evidence" accruing to them, but which may be more powerful, more generalizable, less costly, or otherwise potentially good treatments. Granting agencies or policy-makers read this and determine that they should not fund "lesser" treatments, which can hamper the development of new treatments. Also the "future research" section should have a bullet-point list of many of the key issues raised prior in the ES (e.g., the need for inclusion of much broader patient samples, the need for reporting of costs of treatments, studies of non-outpatient samples, etc).	We agree that this is an important point and that another potential area for future research could involve assessment of potentially beneficial treatments that don't yet have the best evidence supporting their use, and may be more available, etc. We have expanded the future research section to include the following in our Table of evidence gaps for future research and in the ES: "Future studies could evaluate promising therapies that have some evidence suggesting possible efficacy or could evaluate new therapies that may be applicable to broader populations or to specific populations (e.g., those with particular comorbid conditions)."

Comment Number	Commentator & Affiliation	Section	Comment	Response
15	Brown, Peter Public Reviewer	Executive Summary	Excerpt from the abstract of my dissertation. I am submitting this paper under copyright and it should be cited per the APA manual. Please see PDF for full version. TRAUMA RESEARCH AND TREATMENT OF COMBAT VETERANS: AN EVIDENCE-BASED INTEGRATIVE LITERATURE REVIEW Abstract The mainstream treatments for Post Traumatic Stress Disorder (PTSD) are Cognitive Behav00ioral and Prolonged Exposure Therapies (CBT & PE). These closely studied evidence- based treatments also show high relapse, dropout, and failure rates of up to half of those treated (Bryant, R., et al., 2008, p. 555). While not as well researched and harder to measure in terms of the gold standard in Evidence Based Practice of Psychology (EBPP), studies of "alternative" treatments and their methods, yield different and interesting evidence. Using the standards espoused by EBPP alongside alternative movements, this study examined modalities used in veterans' treatment. A guiding question was "What can the field of trauma studies learn from a systematic and comparative review of the research and treatment of combat veterans suffering the sequelae of trauma?" ii Included in this integrative literature review—which generates a critique and theoretical synthesis of a body of literature (Torraco, R., 2005, p. 356)—were peer-reviewed studies from 2006-2010. The participating studies consisted largely of Veterans Administration (VA)-funded, CBT/PE treatments, with an average of over 32 patients per participating study, of approximately 13 weeks duration, and where 20% of patients avoided treatment, 25% dropped out, and 30% failed treatment altogether. Concept matrix analysis of data included distillation of essential statements further reflecting poor tolerability, dropout, failure, and an inability to maintain symptom reductions (75% of studies). Authors tended to overstate positive effects while omitting adequate examination of study design and construct validity, leading to dearth bias, defined as scarcity of evidence hiding behind citations. From this integrative review of the literature a reconceptualization and agenda for future research emerged. The reconceptualization stems from the usefulness of hybridized efficacy and effectiveness research, self-reflection and bracketing, and more accounting for dearth bias. The future agenda recommends practitioners use concept matrices as iii research and practice tools, conduct more common factors research, and develop more clinical practice-based evidence. Especially as related to knowledge evaluation, increased accountability, and system-wide change, these recommendations can assist the spread of more diverse and useful EBPP, to help relieve some of the pain of the traumatized combat veteran.	Thank you for submitting this dissertation, but it does not meet eligibility criteria. It is labeled as a submitted (i.e. draft) dissertation and is not a final/approved dissertation. It covers some of the literature from 2006 to 2010. We would consider this to be a narrative review. We have hand-searched the references included in this work to make sure our searches didn't miss any studies. We did not find any additional studies that our searches missed from reviewing this. The dissertation included 15 articles in the analysis and almost all are case series (including case series of drumming and dancing, which were not within the scope of our review).
16	Najavits, Lisa Public Reviewer	Executive Summary	Please note I am trying to attach several documents but it appears possible only to upload one each time. I will upload several now.	Thank you. We respond to each of the related comments below.

Comment Number	Commentator & Affiliation	Section	Comment	Response
17	Peer Reviewer #1	Introduction	Lifetime prevalence is a very misleading statistic, by and large, only used in psychiatry. Lifetime prevalence for those at a fixed age is actually the incidence by that age among those who survive to that age. When you mix subjects at different ages, you are not only taking a weighted average of those incidences at different ages, but the weights depend on the age distribution of the sample which varies from one study to another, and, if the disorder (like PTSD) is itself associated with decrease in survival, the incidences mixed are from different populations as well. It would be a boon to psychiatry if lifetime prevalences were removed from consideration, as it has been in other areas of medicine.	We agree that lifetime prevalence can be misleading, but we don't agree that this information should be completely removed. It just needs to be reported along with current prevalence. This is why we also report current (12-month) prevalence in the same sentence with lifetime prevalence.
18	Peer Reviewer #2	Introduction	The Introduction is well done.	Thank you
19	Peer Reviewer #2	Introduction	I have a few concerns: a) The "cognitive restructuring" category (page 3) is misleading, if not erroneous. Cognitive processing Therapy has both a cognitive restructuring as well as an exposure component. (In fact, the IOM, erroneously classified CPT as an "exposure therapy". This misclassification is carried through the entire report. CPT is different than the approach used by Ehlers.	We have revised the cognitive restructuring section to clarify based on this comment and others. We now call the section "Cognitive Therapy" (as suggested by peer reviewer #3) as an overarching term, which we now define in the intro. The section includes studies of CPT, cognitive restructuring, and other cognitive therapies. Most importantly, we no longer lump these all together in our analyses and we report results separately for these various types of therapies. The reviewers indicated that this was especially important for CPT.
20	Peer Reviewer #2	Introduction	b.) Table 2 does not mention guanfacine (an alpha-2 agonist) , about which 2 RCTs have been published	We have added guanfacine to Table 2.
21	Peer Reviewer #2	Introduction	Brief eclectic psychotherapy is NOT "a general class of therapies" (page 3) but a manualized and discrete psycho-therapeeutic approach.	We have revised and expanded the description of BEP to clarify. We no longer state that it is a general class of therapies and we make it clear that it is a manualized treatment.

Comment Number	Commentator & Affiliation	Section	Comment	Response
22	Peer Reviewer #3	Introduction	The literature review of the treatments reads as if written by people who do not know the literature well. It is scholarly and comprehensive, but reflects key misunderstandings that need to be corrected before the report is finalized. For example, Brief Eclectic Psychotherapy is described as a general class of therapies rather than a specific intervention. The treatment is in fact a specific branded type of therapy that was developed by Berthold Gersons. It draws from a variety of theoretical approaches, including psychodynamic and cognitive-behavioral, but it is a very specific 16-session protocol that "…combines cognitive-behavioral and psychodynamic approaches, and a farewell ritual at the end of the treatment in a single treatment method and also devotes attention to the patients' partner and work-related problems. BEP is a manualized psychotherapy for PTSD patients who have experienced a wide range of traumas (Gersons, Carlier, & Olff, 2004)." (as cited in Lindauer et al., Journal of Traumatic Stress, 2005). There is a copyrighted manual for BEP as well. This state of affairs differs markedly from the description of BEP on p. 3.	As described in the response to the previous comment, we have revised and expanded the description of BEP to clarify.
23	Peer Reviewer #3	Introduction	Similarly, Cognitive Processing Therapy is described as a type of cognitive restructuring but in fact it is a specific branded protocol, like BEP. It also has other elements that are incorrectly described under the heading of cognitive restructuring; in fact, it includes written exposure, which is why it was classified as a type of exposure therapy in the 2008 IOM report. CPT should be categorized as cognitive restructuring.	We have revised the description of CPT so that it is no longer described as a type of cognitive restructuring. Based on the beginning part of this comment and other comments from Peer Reviewer #3 (and other reviewers), we realize that the last sentence of this comment intended to say that it should *not* be categorized as cognitive restructuring. As stated elsewhere, we no longer lump CPT with anything else in our analyses. Thus results and conclusions are presented for CPT by itself.

Comment Number	Commentator & Affiliation	Section	Comment	Response
24	Peer Reviewer #3	Introduction	One other serious problem is the classification of coping skills therapy. There are two specific issues. One is that, as defined, it is a catch-all that includes a range of psychotherapeutic and not psychotherapeutic active treatments and comparison treatments. There is no clinical or rational basis for combining relaxation training (a control treatment) with stress inoculation training (a fist-line treatment in practice-guidelines around the world). Consequently, any overall effect for this category does not have sufficient scientific meaning or clinical relevance. Second, not all of these treatments are cognitive-behavioral therapy. A CBT protocol might include relaxation training but relaxation is not exclusively CBT.	We have revised our analyses so that we don't combine any of these various things we've deemed coping skills. We present results/conclusions separately for SIT and for relaxation. We have added text to note this in the introduction as well as to make sure readers are aware of the points raised here by the reviewer.
25	Peer Reviewer #3	Introduction	The introduction does little to reconcile the controversies around the existing guidelines and in fact makes it seem as if controversy is more widespread than is actually the case. There is very strong agreement about psychotherapy, with the exception of the IOM report, which erroneously classified CPT as an exposure therapy (and thereby weakened the cognitive category) and judged the EMDR findings as insufficient for methodological reasons. There is more divergence in the pharmacotherapy ratings across guidelines, particularly in the VA/DoD, IOM, and NICE ratings. It would help readers to have a more accurate picture of the state of affairs.	We have completely revised the section of the introduction. We have softened the language about controversy and added some more specifics about the various guidelines. Of note, the intention of the introduction is not to reconcile the controversies---the intention is to set the stage for an evidence review. Further, a detailed review of all the guidelines is not the intention of this report so we did not want to expand this text too much in a report that is already very large, but we have expanded it some to clarify and to address the reviewers points.
26	TEP Reviewer #5	Introduction	Well written.	Thank you
27	TEP Reviewer #6	Introduction	The introduction was well-written and provides a good background for the remainder of the report.	Thank you
28	Anonymous Reviewer #1	Introduction	This is an ambitious and comprehensive project. Unfortunately, the introductory description of EMDR therapy is inaccurate. It appears to be an erroneous rendition of "EMD" that has not been used since 1994.	Thank you We have revised the introductory description of EMDR as suggested. The new

Comment Number	Commentator & Affiliation	Section	Comment	Response
			Specifically, in EMDR, the patient is not asked to imagine the event as if it is in the present. During the Assessment phase, the patient is initially asked to identify image, negative cognition, emotion and physical sensation, along with a desired cognition and numerical baseline measures. During memory processing, the patient is not asked to repeatedly return to the memory or to rate distress. An associative process is engaged and tracked by the clinician. The positive cognition is not evoked until the end of treatment. The description of the theory does not correspond to current research (see list below). Three early research studies reported positive effects in the treatment of a single trauma within three sessions. However, treatment was not restricted to this in clinical practice. Given length restrictions, the following is an accurate description of the treatment: **EMDR therapy** is an eight phase treatment combining brief exposures to aspects of the traumatic event with concurrent induction of saccadic eye movements. The latter are theorized to both interfere with working memory and elicit an orienting response, which lower emotional arousal so that the trauma can be resolved. The patient is initially instructed to identify imaginal, cognitive and somatic elements of the traumatic memory. The clinician then asks the patient to access the memory while focusing on rapid movements of the clinician's fingers. After approximately 30 back and forth eye movements, the clinician asks the patient to report any associations that may have emerged. The clinician follows standardized procedures to monitor and guide the patient's associative process during sequential sets of eye movements. Although early studies of EMDR evaluated 1 to 3 sessions[1], current standards consist of 8 to 12 weekly 90-minute sessions.[2] 1. Rothbaum, B. O. (1997). A controlled study of eye movement desensitization and reprocessing in the treatment of post-traumatic stress disordered sexual assault victims. *Bulletin of the Menninger Clinic, 61,* 317-334 2. Shapiro, F. (2001). *Eye movement desensitization and reprocessing: Basic principles, protocols and procedures* (2nd ed.). New York: Guilford Press. It might also be useful to note the differences between CBT and EMDR treatments. In EMDR therapy, there is no detailed description of the event and no homework. In the research included in these guidelines comparing the two forms of treatment, the EMDR condition used no homework, compared to approximately 50 hours in the CBT condition. As noted by Rothbaum et al. (2005), *"An interesting potential clinical implication is that EMDR seemed to do equally well in the main despite less exposure and no homework." (p. 614)* Research supporting the theories that the eye movements disrupt working memory and elicit an orienting response:	version clarifies the points made here by the reviewer. We have hand-searched the references included by this reviewer to make sure our searches didn't miss any studies. We did not find any additional studies that our searches missed from reviewing this. (the list did not result in additional studies being added to the evidence on EMDR).

Comment Number	Commentator & Affiliation	Section	Comment	Response
			Andrade, J., Kavanagh, D., & Baddeley, A. (1997). Eye-movements and visual imagery: A working memory approach to the treatment of post-traumatic stress disorder. *British Journal of Clinical Psychology, 36,* 209-223. Barrowcliff, A.L., Gray, N.S., Freeman, T.C.A., & MacCulloch, M.J. (2004). Eye-movements reduce the vividness, emotional valence and electrodermal arousal associated with negative autobiographical memories. *Journal of Forensic Psychiatry and Psychology, 15,* 325-345. Barrowcliff, A.L., Gray, N.S., MacCulloch, S., Freeman, T. C.A., & MacCulloch, M.J. (2003). Horizontal rhythmical eye-movements consistently diminish the arousal provoked by auditory stimuli. *British Journal of Clinical Psychology, 42,* 289-302. Christman, S. D., Garvey, K. J., Propper, R. E., & Phaneuf, K. A. (2003). Bilateral eye movements enhance the retrieval of episodic memories. *Neuropsychology. 17*, 221-229. Elofsson, U.O.E., von Scheele, B., Theorell, T., & Sondergaard, H.P. (2008). Physiological correlates of eye movement desensitization and reprocessing. *Journal of Anxiety Disorders, 22,* 622-634. Engelhard, I.M., van den Hout, M.A., Janssen, W.C., & van der Beek, J. (2010). Eye movements reduce vividness and emotionality of ''flashforwards.'' *Behaviour Research and Therapy, 48,* 442–447. Engelhard, I.M., et al. (2011). Reducing vividness and emotional intensity of recurrent "flashforwards" by taxing working memory: An analogue study. *Journal of Anxiety Disorders 25,* 599–603. Gunter, R.W. & Bodner, G.E. (2008). How eye movements affect unpleasant memories: Support for a working-memory account. *Behaviour Research and Therapy 46,* 913– 931. Hornsveld, H. K., Landwehr, F., Stein, W., Stomp, M., Smeets, S., & van den Hout, M. A. (2010). Emotionality of loss-related memories is reduced after recall plus eye movements but not after recall plus music or recall only. *Journal of EMDR Practice and Research*, *4*, 106-112. Kavanagh, D. J., Freese, S., Andrade, J., & May, J. (2001). Effects of visuospatial tasks on desensitization to emotive memories. *British Journal of Clinical Psychology, 40*, 267-280. Kuiken, D., Bears, M., Miall, D., & Smith, L. (2001-2002). Eye movement desensitization reprocessing facilitates attentional orienting. *Imagination, Cognition and Personality, 21, (1),* 3-20. Lilley, S.A., Andrade, J., Graham Turpin, G.,Sabin-Farrell, R. & Emily A. Holmes, E.A. (2009). Visuospatial working memory interference with recollections of trauma. *British Journal of Clinical Psychology, 48,* 309–321. Maxfield, L., Melnyk, W.T. & Hayman, C.A. G. (2008). A working memory explanation for the effects of eye movements in EMDR. *Journal of EMDR Practice and*	

Comment Number	Commentator & Affiliation	Section	Comment	Response
			Research, 2, 247-261. Sack, M., Hofmann, A., Wizelman, L., & Lempa, W. (2008). Psychophysiological changes during EMDR and treatment outcome. *Journal of EMDR Practice and Research, 2,* 239-246. Sack, M., Lempa, W. Steinmetz, A., Lamprecht, F. & Hofmann, A. (2008). Alterations in autonomic tone during trauma exposure using eye movement desensitization and reprocessing (EMDR) - results of a preliminary investigation. *Journal of Anxiety Disorders, 22,* 1264-1271. Schubert, S.J., Lee, C.W. & Drummond, P.D. (2011). The efficacy and psychophysiological correlates of dual-attention tasks in eye movement desensitization and reprocessing (EMDR). *Journal of Anxiety Disorders, 25,* 1-11. Van den Hout, M., Muris, P., Salemink, E., & Kindt, M. (2001). Autobiographical memories become less vivid and emotional after eye movements. *British Journal of Clinical Psychology, 40,* 121-130. van den Hout, M., et al. (2011). EMDR: Eye movements superior to beeps in taxing working memory and reducing vividness of recollections. *Behaviour Research and Therapy, 49,* 92-98.	
29	Anonymous Public Reviewer #2	Introduction	4: In the example given within the definition of "coping skills therapy", the Seeking Safety treatment should be included as it is a coping skills treatment.	This therapy was one of the challenging ones to categorize. We have discussed this with our team and with experts in the field and have decided to categorize it with the “Other” interventions, rather than with the coping skills interventions. Most importantly, regardless of where it is categorized, the results for this intervention are presented without combining them with any other interventions. Of note, no other reviewers made this comment (suggesting that they felt it’s categorization to be appropriate). Although the reviewer is correct that coping skills are

Comment Number	Commentator & Affiliation	Section	Comment	Response
				taught as part of the intervention, the Seeking Safety treatment is a more comprehensive intervention than the others included in the Coping Skills section, and it is targeted for women with comorbid PTSD and substance abuse. Other features that lead to our decision to categorize it with the "other" group: 1) that it is manualized; 2) that it incorporates education on substance use disorders and skills to prevent drug use; 3) that it teaches cognitive restructuring of maladaptive thoughts associated with substance use and trauma symptoms; and 4) a includes a focus on building a healthy support network through the development of effective communication. We could exclude item #1.…

Comment Number	Commentator & Affiliation	Section	Comment	Response
30	Anonymous Public Reviewer #2	Introduction	5: "Furthermore, patient preferences need to be incorporated into shared decisionmaking about treatment because they can influence treatment adherence and therapeutic response. " This is a very key statement—good to see it. In addition, though, there need to be other caveats, such as lack of readiness for some treatments (this simply must be made clear repeatedly—one can do a lot of harm by pushing patients to do intense PTSD treatments too early, while they have vulnerability factors such as current homelessness, domestic violence, etc). Also, to mention here are workforce issues and cost of treatments (as noted earlier), and also prior non-response or iatrogenic effects. Many patients are re-entered into the same few treatments over and over (the so-called "first line" treatments) when they didn't benefit prior. Overall, this document emphasizes internal over external validity —emphasizing certain treatments that have more evidence, but which still are clearly lacking in evidence that relate to vast swaths of the PTSD population. These "first line" treatments, for example, don't yet have a single published RCT in any substance dependent patient sample. Yet substance use disorders are one of the most common psychiatric diagnoses in the US population. Many clinicians, policy makers, and entities will rely on this document. There are already concerns in the field of these "first line" treatments being done with all patients—even when there is no evidence for some subpopulations (such as those named earlier, who have been excluded from research trials).	Thank you, we agree that this is an important point. As suggested in this comment, we have added information to the report to further clarify how many studies set various exclusion criteria. See responses to other comments from this reviewer as many of them are related to this point. Our revisions include 6 new paragraphs in the Applicability section of the report.
31	Brown, Peter Public Reviewer	Introduction	Please see PDF	See response to initial comment from this reviewer.
32	Najavits, Lisa Public Reviewer	Introduction	N/A	No response required.
33	Peer Reviewer #1	Methods	What I'm about to say next does not gainsay the above. I did have some methodological concerns. Moreover, some of these concerns other researchers/reviewers/statisticians may not be agreement on. So, for the consideration of the author(s): For statistics like time to remission, time to return to work, etc. either the analysis should be based on survival curves (more on that later), or the time point at which analysis is done must be stated, e.g., whether or not remission occurs within 6 months of diagnosis or 3 months or 1 year. The effect size will change in such cases depending on the follow-up time. There is too much emphasis here on statistical significance. After all, the strength of meta-analysis is the focus on effect sizes. It would be a major improvement if it were stated 'a priori' what effect size would be the threshold of clinical significance. Then any RCT that did not have adequate power (at least 50% power) to detect any effect size greater than that threshold should be	As the reviewer thought might be the case, we do not agree with much of this and with most of the reviewer's subsequent comments about the analysis methods. In addition, our methods are consistent with those recommended by the EPC Methods Guide. There is no consensus on the approach suggested by the reviewer. As evidenced by the Methods Guide and by comments from

Comment Number	Commentator & Affiliation	Section	Comment	Response
			excluded from the meta-analysis, along with those that for other reasons (attrition, lack of blindness, etc.) were already so excluded. Studies with too small a sample size often have other more covert problems as well. Even if not, removing under-powered studies would fundamentally solve the "file drawer" problem. I notice with some misgiving the RCTs you have included with, say, 12 in the treatment group and 10 in the control group!	other reviewers, the approach used in this review is a valid approach. Regarding survival curves, we do not have primary individual patient data to conduct survival curve analyses for time to remission, etc. The data available allow calculation of the risk difference (or we could calculate RR or OR, but we feel RD is most appropriate for the dichotomous outcomes in this report); the data from individual studies are generally the proportion of people achieving remission etc. We have added the time points (study duration) to the forest plots or the relevant text when conducting analyses for loss of PTSD diagnosis, for example, so that readers can see the duration over which subjects achieved the remission (the durations were very similar in the studies we pooled data from; often data from 8 to 12 weeks). We disagree with the reviewer's comments about generally excluding small RCTs—in fact, one of the reasons meta-analysis is often recommended is for situations when there are a number of studies, each of which may be underpowered to detect a small effect (or even a medium effect).

Comment Number	Commentator & Affiliation	Section	Comment	Response
34	Peer Reviewer #1	Methods	My major problem had to do with effect sizes. I strongly, strongly disagree with presenting information on mean differences, ignoring the within groups variances and distributions. If the mean difference between treatment and control were, say, 10 points on some familiar scale, and the two distributions were both normal with standard deviation equal to 1, that would mean almost no overlap between the two populations. A major and important treatment efficacy. On the other hand, if it were 50, there is almost complete overlap, and no treatment efficacy. And if the variances were different, or the distribution skewed or long-tailed, reporting the mean difference is even less interpretable.	We disagree with the reviewer's opinion on this matter. Our methods are supported by the EPC methods manual. In addition, we have calculated and presented both WMDs and SMDs for the most important continuous outcomes (PTSD symptoms) since some readers will be more familiar with (and more comfortable with) one or the other. See the following related comment and response.
35	Peer Reviewer #1	Methods	Moreover, you keep switching from one effect size to another: from mean differences to Cohen's d (much better since it takes within group variance into consideration, but even then, misleading for skewed or long-tailed distributions), to risk difference, etc. It is impossible either to interpret many of the individual results, and even more impossible to compare results with different outcomes measures on different scales.	We had previously reported most of the psychological intervention meta-analyses for PTSD symptoms with SMDs (because several different outcome measures were used in studies) and the pharmacological ones with WMDs (because there was less variation in measures, with many using CAPS). We have run many additional analyses and we now also report SMDs for the main pharmacological outcomes (PTSD symptoms) to make it easier for readers. Of note, this comment about WMDs and SMDs only applies to continuous outcomes. We have run many additional analyses and now report results for Cohen's d (SMDs) in addition to WMDs for many more analyses, including all of the PTSD symptom reduction

Comment Number	Commentator & Affiliation	Section	Comment	Response
				analyses. Both measures, WMD and SMD, have utility. The WMD allows the results to remain in the units of the measure (e.g., CAPS score) which is often meaningful to clinicians. The SMD allows the results to be presented in the same units when various outcome measures were used; it also allows results from different analyses to be presented in the same way, and perhaps making them more comparable if different outcome measures were used in studies of different interventions.
36	Peer Reviewer #1	Methods	A much better choice would be to report Success Rate Difference (SRD)=Prob(T>C)-Prob(T<C), there "T>C" means that a randomly chosen patient in the T group has a clinically preferable response to a randomly chosen patient in the C group(Hsu, 2004; Kraemer & Kupfer, 2006). For a binary outcome that is the measure you used: the risk difference. Also NNT=1/SRD, and you did report NNT in many, but not all, cases. NNT is great for interpretability, especially for clinicians and policy makers, but is a misery to use in computations. Hence SRD rather than NNT. However, SRD can be computed for any outcome measure valid for a RCT. If Cohen's d assumptions are satisfied, SRD=2normsdist(d/sqrt(2))-1, where normsdist is the standard normal cumulative distribution function, and sqrt is the square root. If you're comparing two normal distributions with unequal variances, to compute d, you would use the average of the two groups' variances in the formula above rather than the pooled standard deviation. For ordinal data, SRD=2U/(mn)-1, where U is the Mann-Whitney U-statistic. For survival data, see (Altman & Andersen, 1999). In short, if you used SRD, you could use the same effect size, one that had precise clinical meaning regardless of what the scale of the outcome was, one that is invariant under any monotonic transformation. Confidence intervals could be obtained using bootstrap methods. 5. Using one common effect size for all possible outcome measures would also help with #3 above. Cohen's d of .5, which he called "moderate" is often suggested as the threshold of clinical significance (corresponding to SRD=.28 and NNT=3.6). (This is not even a suggestion, but an illustration. The choice of threshold would be yours based on your knowledge of the field.) If you found a pooled effect size with a 95% confidence interval contained completely between -.28 and +.28 (often non statistically significant	We disagree. SRD is not a metric that most readers are familiar with and it would make the report substantially less meaningful to the audience that will access this information. Importantly, changing to SRD would not change any conclusions—it would just make the report less accessible and less clear for clinicians, policymakers, and others that will use the information. We agree with benefits of NNT for interpretability and clinical importance—and thus we report it when possible. We have followed the EPC methods guidance for our

Comment Number	Commentator & Affiliation	Section	Comment	Response
			at the 5% level), then those two treatments should be considered “clinically equivalent”. That would be true even if such a pooled effect size were statistically significant (e.g., the confidence contained between 0 and +.28). Using a common effect size, and indicating the thresholds on your “tree” plots would then distinguish between two treatments that are statistically significantly different , but likely not clinically significantly different, and would gain information from non-significant results that indicate clinical equivalences. You could even move further to discuss clinical superiority (confidence interval completely above +.28 or completely below -.28), or clinical non-inferiority (equivalence plus clinical superiority). 6. This meta-analysis also suffers all the problems of multiple outcome measures. When several different outcomes are reported, and they do not come to the same conclusion on efficacy, it is difficult to know what to make of the results. Are the outcomes positively correlated with each other, and the result simply due to inadequate power for one or the other? Or bad luck? It could even be that both outcome measures indicate the same outcome, but are measured on different scales or with different reliabilities. Or do some patients have major benefit on outcome X, and other patients on outcome Y? Then you may see little benefit on either X or Y separately, even though every single patient has a benefit. I don’t know what can be done about this, but it is a problem. Altman, D.G., & Andersen, K. (1999). Calculating the number needed to treat for trials where the outcome is time to an event. British Medical Journal, 319, 1492-1495. Hsu, L.M. (2004). Biases of Success Rate Differences Shown in Binomial Effect Size Displays. Psychological Bulletin, 9(2), 183-197. Kraemer, H.C., & Kupfer, D. J. (2006). Size of Treatment Effects and their Importance to Clinical Research and Practice. Biological Psychiatry, 59(11), 990-996.	analyses.
37	Peer Reviewer #2	Methods	My biggest objection to the search strategy is that the National center for PTSD's PILOTS bibliographic database (Published International Literature on Traumatic Stress - accessible at <www.ptsd.va.gov>- which has 46,000 citations and is the largest and most comprehensive source in the field) was not utilized. Otherwise the methodology was sound and appropriate.	We have conducted additional searches to include the PILOTS database. This added three low or medium risk of bias publications that were companions with articles already included in our report (each generally reporting a few additional outcomes for a trial). We also identified 3 studies that were rated high risk of bias, but otherwise met inclusion criteria.

Comment Number	Commentator & Affiliation	Section	Comment	Response
38	Peer Reviewer #3	Methods	In general, the methods are well-described but there are a few minor exceptions. (1) The authors did not search the PILTOS database, which includes the world's literature on trauma and PTSD. The search strategy may have identified all of the key articles but any comprehensive search on PTSD treatment should include the primary database devoted to the topic.	As described in the previous comment response, we have conducted additional searches to include the PILOTS database.
39	Peer Reviewer #3	Methods	(2) There is a lack of justification for the exclusion of the findings on CAM and somatic treatments like rTMS.	Due to the already large scope of this review, it was important that we focus on the most clinically relevant interventions. During the topic development and refinement process, CAM interventions were considered, and the general consensus was that such interventions were of lesser interest than other interventions and that there stakeholders thought we were less likely to find sufficient reliable evidence to synthesize on those interventions. We added to the methods some text explaining reasons for not including CAM. We have included a sentence in our limitations mentioning that we did not review literature on complementary and alternative medicine treatments.
40	Peer Reviewer #3	Methods	(3) The Impact of Event Scale is incorrectly titled Impact of Events Scale.	Thank you for catching this. We have corrected it throughout the entire report (removing the "s" from Events)
41	Peer Reviewer #3	Methods	A major exception is that the specifics of how medium versus low risk categories were determined is difficult to understand, even using the detailed information provided in the appendix. Once there, I failed to understand how some of the resulting determinations were made. For example, a study by Schnurr et al. (2003) was judged to be low risk and a study by Schnurr et al. (2007) was judged to be medium risk. Both studies had reasonably high dropout and higher dropout in the treatment vs. control	We have provided text in the Methods section that explains our process for making this determination. The information is in the 5 paragraphs under Risk of Bias Assessment of

Comment Number	Commentator & Affiliation	Section	Comment	Response
			arm. (There are also two listing fro the Schnurr et al., 2003 study). Both studies had blinded assessment, the same outcomes, adequate randomization strategies, etc. Both studies used ITT analysis and robust methods for handling missing data, but the 2007 paper (which used multiple imputation to include all randomized participants) was arguably better than the 2003 paper (which included data on only those participants who provided outcome data, roughly 90% of the randomized sample). The text on p. 14 is clear on how high risk is determined. The bottom line is that readers should not have to refer to a detailed Appendix in order to understand an important distinction like this, and if they do, the information should be clearer.	Individual Studies. We have added some revisions to this section that should make the process clear. We include that: "…We determined the risk of bias rating via appraisal of responses to all 12 questions assessing the various types of bias listed above. We did not use a quantitative approach (e.g., adding up how many favorable or unfavorable responses were given), but we did require favorable responses to at least 10 questions to give a low risk of bias rating, with any unfavorable responses being of relatively minor concern (e.g., lack of provider masking in studies of psychological interventions, which is generally not considered possible)." These two studies were ones that we have discussed with our full team. The 2007 study did not have a favorable response to at least 10 questions (as required for a low risk of bias rating). The main difference between these studies that lead to the 2007 study being medium risk of bias was the overall attrition (about 30%) and the differential attrition (17% difference between two of the groups); whereas the 2003 study had attrition around just

Comment Number	Commentator & Affiliation	Section	Comment	Response
				10%. We see that there were 2 rows for the Schnurr 2003 study and we have fixed this---there is now just one row for that study. Regarding the point readers should not have to refer to a detailed Appendix in order to understand an important distinction like this (about low vs. medium risk of bias ratings), we disagree. First, there is a copious amount of information in the report and we think the best way to handle all of the detailed risk of bias information is in an appendix. Our experience over the past 15 years with EPC work has taught us that most readers prefer this approach. Second, the distinction is really not a terribly important one for this particular report. Differences between low and medium risk of bias ratings did not end up influencing our conclusions significantly. Both low and medium risk of bias studies contributed equally to our meta-analyses.
42	Peer Reviewer #3	Methods	Another major exception is how some studies are described and/or classified. The studies included in Table 8 might be better described as cognitive therapy.	See also responses to the related comments from this reviewer. As suggested, we now describe Table 8 (and this section) as studies of cognitive therapy.

Comment Number	Commentator & Affiliation	Section	Comment	Response
43	Peer Reviewer #3	Methods	The CBT coping skills category needs to be refined as indicated above.	We have addressed that comment and our approach to the coping skills section above. See previous response.
44	Peer Reviewer #3	Methods	Moreover, I could not find where the authors describe how they specifically operationalized distinctions between cognitive restructuring and CBT mixed. I think readers would seriously question the categorization of Ehler's 2005 study of cognitive therapy as CBT-mixed; her treatment is one of the purest cognitive therapies around. Similarly the CBT in Mueser et al.'s study is a type of simplified cognitive restructuring, certainly it is no more missed than CPT.	We have re-reviewed the Ehlers 2005 and the Mueser studies. We agree that Mueser should be categorized as cognitive therapy and we have moved it to that section. However, we do not agree about Ehlers 2005. We categorized the earlier study from Ehlers (2003) as cognitive therapy, and she is well known for the development of this cognitive therapy. But, her 2005 study clearly uses a different intervention that includes taking people to the site of the trauma (in vivo exposure), and thus seems to clearly go beyond cognitive therapy. Perhaps the reviewer is thinking of the Ehlers 2003 study (or mistakenly thinking that the two studies used the same intervention b/c Ehlers is famous for developing a particular CT, described in Ehlers 2003). Regarding operationalizing the CBT-mixed category, we have added text to the Psychological Interventions section of the report to address this, explaining how studies were placed in this category.

Comment Number	Commentator & Affiliation	Section	Comment	Response
45	Peer Reviewer #3	Methods	A table specifying how each of the psychotherapy categories was defined would be helpful.	We provide the relevant information in the Introduction. We have revised and clarified the categorization and definitions as described in responses to other comments from this reviewer and from Reviewer #2, including the previous comment. We provide the information using text (section titled "Psychological Interventions"), rather than a Table, as we feel that the information describing all of the interventions and definitions was fairly extensive, and would make a very long table. We think it lends itself better to text than to a Table.
46	Peer Reviewer #3	Methods	Another clarifying table would be how many studies within each category yielded data on each of the outcomes. I know this can be gleaned from looking at multiple tables but that is exactly the point—it must be gleaned form looking at multiple tables and there is little synthesis to help readers.	We have included (in the detailed synthesis section of the report) text that describes the number of studies that reported each of the relevant outcomes, with references to those studies. In addition, our Appendix of Evidence Tables shows the detailed data for each study, so readers can see the relevant outcomes that each reported. Further, the Forest plots provide yet another illustration of the studies that contributed to quantitative syntheses for each intervention and outcome.
47	Peer Reviewer #3	Methods	Waitlist controlled studies should be distinguished from studies using a nonspecific treatment such as Present-Centered Therapy or treatment as usual as a control. There is a big difference is the control provided by waitlist and nonspecific treatment. Present-Centered Therapy, for example, includes the nonspecific elements of Prolonged Exposure and even has (non-trauma-focused) homework. Scientifically, a treated	We have re-run all of our analyses for the psychological interventions, now stratifying them by comparator. Our main analyses now include studies

Comment Number	Commentator & Affiliation	Section	Comment	Response
			control group controls for more threats to validity than a waitlist control group does. Pooling them is mixing apples and oranges.	that used a waitlist comparison (also sometimes referred to as minimal attention) as well as those with a usual care or treatment as usual comparator. But, we have stratified the analyses to show a pooled estimate separately for each of those groups (one for waitlist and one for usual care) as well as an overall pooled estimate so that readers can see if and when it makes a difference. We have revised our methods section, results, and conclusions to reflect these new analyses (which did not change our conclusions). We have added text to the Methods and in some other places to make this clear to readers. The usual care or treatment as usual arms were often not well described, and comprise a heterogeneous group. Many of them appear to be essentially equivalent to no treatment, whereas some others (albeit few) involved engagement in some activity that could potentially be beneficial). Similarly, the waitlist comparator groups were heterogeneous in that some of them may essentially be equivalent to usual care groups (again with some groups allowed to receive other interventions and others probably getting no treatment).

Comment Number	Commentator & Affiliation	Section	Comment	Response
				As suggested, we have separated the present-centered therapy comparators as well as supportive control comparator, which are both qualitatively different than waitlist or usual care. We now only bring those studies in for sensitivity analyses.
48	Peer Reviewer #3	Methods	The authors need to better define how they standardized definitions of loss of diagnosis and remission, which vary substantially across trials.	We have added text to the Methods about these outcomes and how they were conceptualized; and more text to the "Limitations of the evidence base" section—which is really the key issue here as the reviewer importantly recognized. New text in the Limitations section: "Heterogeneity of outcome measures used in the included studies also posed some challenges. For example, many different measures of PTSD symptoms were used (e.g., CAPS, DTS, IES). In addition, some measures have several versions such as the CAPS, which has evolved over the past decades into its current form.[26] It was often unclear which version of a measure a study used. In addition, the definitions of loss of PTSD diagnosis and remission were somewhat heterogeneous. They were assessed using several

Comment Number	Commentator & Affiliation	Section	Comment	Response
				different instruments across the studies. In addition, complete explanations of the approach to assessing remission or loss of PTSD diagnosis were not always provided. For example, it was sometimes implied, but not explicitly stated, that loss of diagnosis was determined by assessment of DSM diagnostic criteria or by using a CAPS score (or another scale, such as the PSS) cutoff indicative of PTSD diagnosis. In addition, many studies did not clearly report specific score cutoffs used to define loss of diagnosis or remission when reporting the results."
49	Peer Reviewer #3	Methods	Also, "good end-state function" is not a measure of function at all; it is merely an index of improvement in symptoms.	Thank you, we agree. There were 2 mentions of end-state function in the report. We have corrected the misplacement and moved all of the information about end-state functioning (which was indeed measured by symptom improvement measures) to the PTSD symptoms sections.
50	Peer Reviewer #3	Methods	Lastly, by categorizing the SF-36 and SF-12 as measuring of quality of life, the authors make it appear that few studies measured functioning. I am well aware the SFs are described as measuring health-related quality of life but that is in the context of a model that conceptualizes functioning as an indicator of QoL. It is quite clear when one looks at the subscales and items that the SF does measure functioning. The distinction of whether a measure is measuring quality of life or functioning should not be based on a title, especially given conceptualization of functioning as an aspect of quality of life. My recommendation would be to use the broader model of quality of life or at least combine the categories. As presented, the categorization results in a presentation of the data that is misleading.	"Quality of life" is a broader concept than "disability and functional impairment" as generally conceptualized and as conceptualized in this report. The quoted phrases here are the 2 names of the categories we aimed to find evidence on. We developed these through a 6 month topic development and topic refinement process that

Comment Number	Commentator & Affiliation	Section	Comment	Response
				included input from many people. We have added the following to the Methods section to address this issue: “Some outcome measures may encompass aspects of both “quality of life” and “disability and functional impairment”. We considered quality of life to be a broader concept and disability and functional impairment to be one component of quality of life. We included data from general quality of life measures, such as the SF-36 and SF-12, in the quality of life section. We did not also include data from component subscales of the SF-36 or SF-12 in the disability and functional impairment section. We only included validated measures that specifically address disability and functional impairment in that section.” We do not agree with combining the categories. We think that would be less appropriate as they are not the same thing.

Comment Number	Commentator & Affiliation	Section	Comment	Response
51	TEP Reviewer #5	Methods	The criteria are justifiable and the strategies were explicitly stated and logical. The one area that was touched upon but not expanded was the question of side effects for the pharmacological therapies. It would have been useful to draw about studies done for different purposes (e.g. major depressive disroder or anxiety disorders) to eval for side effects of the meds.	Thank you. We agree that it is possible that studies enrolling subjects with major depressive disorder or anxiety could possibly yield useful information on side effects of pharmacological therapies. We include the following in the Limitations: "For harms, useful information could possibly have been provided by studies conducted in other populations (i.e., those without PTSD). For example, many studies of some of the medications reviewed in this report enrolled patients with depression. Such studies could provide important information about adverse effects of the medications."
52	TEP Reviewer #6	Methods	The inclusion and exclusion criteria are clearly stated and justifiable. The constraints on included studies do shift the consideration of the evidence towards RCTs and larger industry sponsored studies but this is typical of systematic reviews that focus on high quality evidence. The outcome measures seem appropriate as are the statistical measures (though my personal knowledge of network meta-analytic techniques in minimal).	Thank you.
53	TEP Reviewer #6	Methods	Perhaps the biggest problem (which is common to all such reviews) is that older medications are systematically downplayed due to shifts in the state-of-the-art of clinical trial design and the fact that older off-patent medications are no longer of commercial interest to pharmaceutical manufacturers. Thus, medications such as tricyclic antidepressants will rarely be given any reasonable consideration since most studies are from an era when FDA standards, trial sizes, statistical methodologies, etc. were quite different. This introduces a systematic bias against such drugs. This has implications in terms of the cost of care to society (if older cheaper drugs are equally effective but de-emphasized) and implications for patients (if older drugs are the only ones feasible for them through Wal-mart $4 program, for example).	There are two issues raised here—(1) that older off-patent medications may be less likely to be studied because they are no longer of interest to pharmaceutical manufacturers and (2) that studies of older medications were conducted under different methodological standards and may be more likely to be rated high risk of bias—and thus less likely to have insufficient evidence supporting their efficacy.

Comment Number	Commentator & Affiliation	Section	Comment	Response
				Regarding the first, we can only review evidence that does exist. Speculating what would be found from studies that were never conducted because there is less interest in funding them is purely hypothetical. Regarding the second, this may be true (that older studies are more likely to be rated high risk of bias) for some conditions. However, it does not appear to have influenced the conclusions for this pharmaceutical literature for PTSD. For older medications to be systematically downplayed in our review, this would presumably happen if studies of older medications were more likely to be rated high risk of bias (and excluded from the main analyses). However, we conducted sensitivity analyses using all studies (regardless of risk of bias rating)—and the sensitivity analyses did not change results significantly. For tricyclic antidepressants specifically, we found only one study each for amitryptiline (N=62), imipramine (N=42), and desipramine (N=27). Regardless of risk of bias rating of these 3 studies, none of the three medications has sufficient evidence supporting efficacy.

Comment Number	Commentator & Affiliation	Section	Comment	Response
54	Brown, Peter Public Reviewer	Methods	Please see PDF	See response to initial comment from this reviewer.
55	Najavits, Lisa Public Reviewer	Methods	N/A	No response required.
56	Peer Reviewer #1	Results	Well done.	Thank you
57	Peer Reviewer #2	Results	Given all the data and comprehensiveness of the report, it is a very difficult document to read. I thought that the graphic displays of results were excellent and very comprehensible.	We realize that the full technical report is very dense with data. There were a large number of interventions, comparisons, and outcomes included in the report. We have worked on condensing some parts to improve readability. However, the full report will likely remain somewhat difficult to read. We hope that the Executive Summary is much easier for readers—and seems to be based on the comments of the reviewers. In addition, we have submitted an article to a peer reviewed journal, similar to the Executive Summary, that is a much more condensed version of the most important information.
58	Peer Reviewer #2	Results	The pharmacotherapy section (and subsequent sections) were excellent. Having placebo controls for most of the studies made the exposition much easier than with psychotherapy and allowed you to make important statistical comparisons.	Thank you

Comment Number	Commentator & Affiliation	Section	Comment	Response
59	Peer Reviewer #2	Results	The psychotherapy section is much more challenging and could be improved. First, I recognize the difficulty of lumping these all of these diverse treatments into a few categories but CPT should probably stand on its own.	We have responded to the various specific comments from reviewers (including this reviewer) to improve the psychotherapy section. Regarding CPT, we no longer lump it with other interventions in our analyses (the finding stand on their own), within the section now titled Cognitive Therapy (per suggestion of Peer Reviewer #3).
60	Peer Reviewer #2	Results	I don't think you did a good job clarifying the importance of the various psychotherapy comparison groups (e.g. Wait List, TAU, another active treatment) in terms of how to interpret the results from the various RCTs under discussion. This is an extremely important omission in the present document that should be carried through the entire document.	Please see similar comment above from peer reviewer #3 and our response. In short, we have made extensive revisions to address this, including re-running all of our analyses for the psychological interventions to account for the various comparison groups. We now stratify the analyses by comparison group.
61	Peer Reviewer #2	Results	There have been a few psychotherapy RCTs with children, especially trauma-focussed CBT. It would have been great if this research could also have been included.	AHRQ has commissioned a separate report focused on children.

Comment Number	Commentator & Affiliation	Section	Comment	Response
62	Peer Reviewer #3	Results	Overall, readers have to either take a lot on faith or dig in to detailed tables in order to understand important details. The Pharmacotherapy section is easier than the psychotherapy section to read. The psychotherapy results are very difficult in places, thanks t the otherwise helpful detail, but perhaps the authors could try to make the results more user-friendly while retaining the richness of the information. For example, maybe it would be helpful for the Key Points summaries on pp. 25-26 to include references to where reader can see a complete list of the included studies, e.g., where is there a list of the 19 Mixed CBT trials?	We have responded to the various specific comments from reviewers (including this reviewer) to improve the psychotherapy section. Regarding the suggestion for the Key Points to include references to where readers can see a complete list of the included studies, we include the following in the paragraph just prior to the Key Points: "The findings in these key points are primarily based on meta-analyses of the trials that we rated low or medium risk of bias described later in the detailed synthesis sections of the chapter. Those trials are cited in the detailed synthesis and related tables." Thus, one simply needs to look at the related detailed section of the report and the related forest plot to find the list of relevant studies (we did not add references to each study whenever calling out a number of studies in the Key Points because it would increase the density of the text greatly).
63	Peer Reviewer #3	Results	Many of my concerns about the results stem from the methods of classifying the studies, so I have little to add here except that the results might change if the studies were reclassified as recommended.	We have revised the classifications as described in the responses to previous comments from this reviewer and we have updated the results.

Comment Number	Commentator & Affiliation	Section	Comment	Response
64	Peer Reviewer #3	Results	There are also little details that seem puzzling, e.g., in Table 13, duration is listed as NR for the Schnurr et al., but the paper states that the treatment lasted 10 sessions and specifies the follow-up intervals at posttreatment, 3 months, and 6 months. Follow-up is reported for only some of these studies but many of them clearly reported follow-up.	Regarding the duration for Schnurr, we have corrected this in the Table--adding "10 weeks (3 and 6 months)". We have also checked again to see if there were any other studies that reported follow-up and updated the table with any changes.
65	Peer Reviewer #3	Results	Other details seemed puzzling across tables, e.g., so few reports of functioning measures. For example, the PE study by Foa et al. 2005 included the Social Adjustment Scale and the one by Schnurr et al. 2007 included the SF-36. My comments in e. below indicate that I figured out the reasons for what seemed like an error actually resulted from the categorization of measures of functioning vs. QoL. This may be difficult for readers to grasp without going into the Appendix tables.	It is a bit surprising that there were so few measures of functioning (and QoL), but that actually turns out to be true for many reviews of treatments for various conditions. Regarding the Foa 2005 study and the Social Adjustment Scale, we had those details in our evidence tables, in the appendix, but had not included them in the main report. We have now added some description of that data to the Results section. Regarding the SF-36, see responses to other comments from this reviewer about SF-36. We have added text to the Methods about our rationale and approach for categorizing various measures in the QOL and the disability and functional impairment sections.

Comment Number	Commentator & Affiliation	Section	Comment	Response
66	Peer Reviewer #3	Results	The medication findings are easier to follow, and the network meta-analysis is especially helpful. I wondered if this approach would be possible at least for the wait-list controlled studies, where there is the greater homogeneity in the comparator than in the other psychotherapy studies.	Thank you. We don't think the approach is the best way to assess the psychotherapy studies. It is certainly possible, but there are a number of potential problems that lead us to the decision not to do a network meta-analysis for the psychotherapy studies. These included the greater degree of heterogeneity across populations enrolled in the studies and greater variation in the outcome measures used as well as the type of data reported.
67	TEP Reviewer #5	Results	I generally skimmed through some of the detail, and looked mostly at the tables which appeared adquate and descriptive.	Thank you
68	TEP Reviewer #6	Results	The level of detail seems reasonable. There does seem to be a fair amount of duplication in the various sections of the report (especially the executive summary and body of the report). The figures and tables do a nice job of delineating key comparisons. I am not aware of specific studies that should have been included/excluded.	Thank you. The Executive Summary and the body of the report are intentionally redundant, as the Executive Summary is intended to be a stand-alone document.
69	Anonymous Public Reviewer #2	Results	22: See the various comments I've listed in relation to the ES. These major conclusion on pg. 22 should be tempered immediately with some of the limitations of the literature (see my ES comments). People reading may simply stop at the end of this page and not read the relevant caveats and issues that need to be stated here.	See responses to other comments from this reviewer. We have added 6 paragraphs to the Applicability section of the report to address her concerns about applicability.
70	Anonymous Public Reviewer #2	Results	25: As stated earlier, Seeking Safety should be listed as a coping skills model in the "coping skills" paragraph. It should not be classified as an "other therapy" (on pg 26) as its focus is coping skills (see www.seekingsafety.org).	See response to the related (earlier) comment.
71	Anonymous Public Reviewer #2	Results	25-26: Key points: should list percentage of studies that were by or with the treatment developers themselves. This is a known bias ("allegiance" or as Luborksy called it, the "wildcard effect"), and for some models the developer was almost always an author or collaborator on the studies of their own model.	We have added a paragraph to the "Limitations of the Evidence Base" section of the report to mention this. Of note, it is very difficult to determine if and when allegiance bias has an influence on results. Further, it is difficult to ascertain the allegiance of

Comment Number	Commentator & Affiliation	Section	Comment	Response
				investigators that were not the developer of any particular therapy, but were trained by someone with a particular allegiance (and thus learned certain biases). But, it is appropriate to mention this as a possible limitation of some of the evidence for certain psychological interventions. New paragraph: “One criticism of psychological treatment trials has been the possibility of “allegiance bias”- referring to the potential for contamination or distortion of results due to the investigators’ theoretical perspective or treatment preferences .{Luborsky, 1999} One marker of allegiance to a treatment preference is when the developer of the method is a primary author in the study of that method. For some of the psychological therapy interventions, it appeared that the developer of the methods was an author on the majority of studies, such as narrative exposure therapy{Neuner, 2004#762}{Neuner, 2008#353}{Neuner, 2010#211}and brief eclectic psychotherapy.{Gersons, 2000#1043}{Lindouer, 2005#659} For the purposes of this review, exploring allegiance was not emphasized during critical review of the included studies and it is unclear what effect, if

Comment Number	Commentator & Affiliation	Section	Comment	Response
				any, this has on the overall validity of the results." We also mention in our research gaps/future research section that future comparative studies should perhaps be conducted by investigators free of conflict/those without a particular allegiance.
72	Anonymous Public Reviewer #2	Results	63: "Three trials assessed the efficacy of a short-term manualized cognitive behavior treatment for women with PTSD and substance abuse called Seeking Safety…" a) Seeking Safety (SS) was designed for both genders. This is stated in the SS book, and numerous article on the model. The original paper on the model in the mid-1990s described it as a model for women and many of its studies were on women. However, by the time the book was published in 2002, it was designed for both genders and it has been implemented with males as well as females for many years successfully. b) Some RCT-relevant studies on SS are missing, as follows. - Boden et al., 2012. This is a recent RCT trial on men veterans, showing significantly greater reduction in drug use for Seeking Safety compared to TAU, as well as well as greater attendance, satisfaction, and active coping. Note also that that trial was, however, underpowered (they did not attain the necessary sample based on power analysis). . - Several Hien et al. papers related to the 2009 study. d) There are also some other studies that are not covered that may be relevant: - Lynch et al., 2011 (compares SS to waitlist) - Desai et al., 2008, 2009 (the 2009 cite makes important corrections to the 2008 paper). - Gatz et al. (2007) e) Given the paucity of research studies reporting on adverse events, it is worth reporting that SS had extremely low rates of adverse events in the Hien et al. study (Killeen et al., 2008). Cross-reference: pg 116. f) It may be worth noting that various SS trials were conducted in group format, they did not have the exclusionary criteria of most PTSD treatment trials, and did not have the intensive selection/training of clinicians that other studies had.	a) We have changed "for women" to "for people" since it was designed for both genders. b) through f) Regarding the studies on SS listed in this comment, our searches identified the studies listed in the comment, but many of the studies did not meet inclusion criteria. Several of these were excluded because the enrolled population did not meet our criteria because they did not enroll a population of subjects with a DSM diagnosis of PTSD. They allowed a lot of people with sub-threshold PTSD to be enrolled or did not clearly describe inclusion criteria to ensure that all, or the vast majority of, subjects have PTSD. Below, we provide the disposition for each of the articles mentioned. Boden---this was captured by our update search. It does meet inclusion criteria and we have added it to the report.

Comment Number	Commentator & Affiliation	Section	Comment	Response
				Hien 2009 is included in our review, as are all of the relevant related papers we identified that met inclusion criteria. Lynch— this study was excluded for the wrong population (as described above). Desai—excluded for the wrong population (as described above). Gatz—we cannot locate any potentially relevant study by Gatz et al., and we did not receive a PDF for a Gatz article (as we did for most of the other articles mentioned in this reviewer's comments). Killeen— excluded for the wrong population (as described above); enrolled those with either full or subthreshold PTSD.

Comment Number	Commentator & Affiliation	Section	Comment	Response
73	Anonymous Public Reviewer #2	Results	64: Table 20 states the Hien et al. 2004 is "unclear" as to sample sizes per arm; those sample sizes are in fact stated in the Hien article (pg. 1428)—n=41 SS, n=34 RP, n=32 TAU.	The numbers reported by the reviewer here are the numbers analyzed. The article does not report the numbers randomized to each intervention group (which is what we are using in Table 20 and the other tables of characteristics of studies). It tells us that 96 participants were randomized to the active treatment groups (but not how many were randomized to each of those groups) and that 32 subjects formed the community comparison (TAU) group. We have changed "unclear" to "32" for the TAU group in the table, but have left it as "unclear" for the two active treatment groups. We have added a footnote to the Table to include the numbers analyzed mentioned here by the reviewer and to explain what is reported by the article.
74	Anonymous Public Reviewer #2	Results	66-67: There are numerous omissions with regard to Seeking Safety studies. a) The Hien et al. 2009 study actually found SS superior to the comparison condition on numerous secondary outcomes: on substance use among heavy drug users [Hien, 2009 #2865], among alcohol users [Hien, in press #2672], in therapeutic alliance [Ruglass, 2012 #2930], and in HIV risk reduction [Hien, 2010 #2666]. Moreover, as noted earlier, SS was found to be a very safe treatment, as evidenced by low adverse events [Killeen, 2008 #2885].	Each of these studies was identified by our searches. We have looked at them again (because of these comments) and explain the disposition of each of these articles below. Regarding Hien et al., 2009, this study was already included in our draft report and we already included the substance use outcomes. The relevant paragraph was (and still is) in the Seeking Safety results section titled Prevention or Reduction of

Comment Number	Commentator & Affiliation	Section	Comment	Response
				Comorbid Medical or Psychiatric Conditions. In the draft text, the Hien 2009 study was reference 127. The text includes: "…seeking safety trials reported outcome data on substance use or abuse and found no between group differences for the active treatment arms in the respective studies. One study sample comprised incarcerated women with no access to substances,[128] and two studies enrolled those in community-based substance use or abuse treatment programs.[126, 127] Substance use outcome measures included abstinence,[127, 128] and substance use severity.[126, 128] … Abstinence rates were not significantly different for seeking safety and Women's Health Education (WHE) at 12-month followup.127" Ruglass 2012 was identified by our searches. The study is a secondary data analysis looking at the association between therapeutic alliance and treatment outcomes. It does not meet eligibility criteria, based on study design. Hien 2010, related to the comment here about HIV risk reduction, was excluded for not reporting an eligible outcome. The study examined

Comment Number	Commentator & Affiliation	Section	Comment	Response
				the impact of two group therapy interventions on the reduction of unprotected sexual occasions. None of our specified outcomes of interest were reported by this article. It reports sexual risk behaviors, but no outcomes directly measuring prevention or reduction of a comorbid medical condition (which is the only outcome category this would have been close to fitting in). For Killeen 2008, a full DSM PTSD diagnosis was not required for participation in the trial. To be eligible, participants needed to have had at least one traumatic event in their lifetime and to have met DSM-IV criteria for either full or subthreshold PTSD. The percentage of participants with a PTSD diagnosis (as opposed to those with subthreshold PTSD) was not provided by the authors. Thus, like other similar articles, it was excluded based on our population criteria.

Comment Number	Commentator & Affiliation	Section	Comment	Response
75	Anonymous Public Reviewer #2	Results	b) The Hien et al. 2009 study had various methodology limitations: (i) it used less than half of the SS model (only 12 of the 25 SS sessions/topics); (ii) it allowed patients in both conditions to be in other trauma-specific treatments but without identifying which patients were in such during the trial or how this may have affected results; (iii) a large proportion of patients were either abstinent from substances at baseline or had very low levels of use (an average of 1 day in the prior month), which did not allow a correctly powered test of SUD.	We agree that this study has some limitations that limit what we can conclude about seeking safety overall as an intervention. Of note, we had already concluded that strength of evidence was insufficient for all of the outcomes for seeking safety, so further consideration of these points does not change our conclusions—of insufficient evidence. Further studies would be needed to determine whether seeking safety is efficacious for improving the outcomes of interest in this report. We have added some text to the characteristics of studies description to mention the important points here---that this study used less than half of the SS model and that a large proportion of patients were either abstinent from substances at baseline or had very low levels of use—and that these things could bias results to the null.
76	Anonymous Public Reviewer #2	Results	c) The Zlotnick et al. 2009 study was a pilot study insufficiently powered to test the difference between treatment conditions; it had no end-of-treatment assessment; and no blind assessments at all; it also had a mandatory dose of TAU that was very large compared to the voluntary SS dose.	We have added the point to the characteristics of studies that this study was a pilot study (N=49) that may have been underpowered. The lack of masking was factored in to our risk of bias assessment of the study, and that information is included in the Appendix describing risk of bias assessments. We did not

Comment Number	Commentator & Affiliation	Section	Comment	Response
				include all of those details in the body of the report as it would substantially lengthen an already large report. Regarding the point about TAU---we agree with the reviewer's point here; and the report accurately reflects the comparison made in the study, but we have added text to make the description more detailed. Although the study calls it TAU in many places, we considered this a head to head study of two interventions because TAU is a clear intervention, a relapse prevention control—all subjects were incarcerated and were enrolled in a 28-bed residential substance use treatment program in a minimum security wing; thus we considered this a head-to-head study with an active relapse prevention comparator group. Like the previous study, we now include a mention that the design/comparison (with the voluntary SS dose and the large mandatory TAU dose) could bias results toward the null.

Comment Number	Commentator & Affiliation	Section	Comment	Response
77	Anonymous Public Reviewer #2	Results	127: "Of note, safety seeking was developed to target substance use/abuse …" The correct term is "substance use disorder."	We have revised this as suggested to "substance use disorders" (made it plural to fit in the sentence appropriately). We also changed it to "seeking safety" rather than "safety seeking" to be consistent with wording elsewhere in the report.
78	Anonymous Public Reviewer #2	Results	136: The first paragraph has the same issue noted earlier—i.e., the need to add to the end of the paragraph the mitigating context as detailed in earlier comments (that these trials excluded important severe and/or complex populations, etc.)	See responses to other related comments about this issue. We have added a lot of text to the Applicability section to address the reviewers comments about these issues.
79	Anonymous Public Reviewer #2	Results	137: "Similarly, we did not find evidence to confirm or refute whether treatments are more or less efficacious for many other subgroups, including gender groups, racial or ethnic minorities, refugees, first responders, disaster victims, or for those with coexisting conditions, different PTSD symptoms, complex PTSD, exposure to childhood trauma, repeat victimization, or different levels of severity at presentation. Although many studies did not exclude subjects in these subgroups (e.g., those with a history of multiple past traumas, service connected disability, or coexisting psychiatric conditions such as depression), studies generally did not report whether interventions were efficacious for such subjects either." [emphasis added] The vast majority of PTSD trials have excluded patients with clinically important and common comorbidities (e.g., per Bradley et al., 2005, 67% of trials excluded those with SUD). The paragraph above appears inaccurate with regard to co-occuring disorders (COD). The only COD that mentioned in studies, by and large, was depression, which has major overlap with PTSD symptoms. The vast majority of PTSD RCTs excluded important CODs (e.g., bipolar, psychosis, SUD, and some Axis II disorders), as well as other vulnerabilities (as stated earlier—e.g., homelessness, domestic violence, etc.).	We have added several additional paragraphs after the one mentioned here by the reviewer. Of note, the text that the reviewer has underlined (or coexisting psychiatric conditions) is followed by "such as depression". The new text aims to address the other conditions---those that the reviewer highlights as well as others clinicians may be most interested in. The new text is now the last 6 paragraphs of the Applicability section of the full report and includes quantification of the percentages of included trials of psychological treatments and of pharmacological treatments that set various exclusion criteria.
80	Brown, Peter Public Reviewer	Results	Please see PDF	See response to initial comment from this reviewer.

Comment Number	Commentator & Affiliation	Section	Comment	Response
81	Najavits, Lisa Public Reviewer	Results	N/A	No response required.
82	Peer Reviewer #1	Discussion	This I have no way to judge.	No response required.
83	Peer Reviewer #2	Discussion	The answer is "Yes" to all of these questions.	Thank you
84	Peer Reviewer #3	Discussion	The discussion and conclusions are reasonable (given caveats that relate to issues noted above and in the following section). The implications for research are reasonable.	Thank you

Comment Number	Commentator & Affiliation	Section	Comment	Response
85	Peer Reviewer #3	Discussion	However, the statements in the discussion on p. 128 and on p. 130 that evidence for other outcomes was generally insufficient does not follow from the data reported, in which almost all studies reported on depression and/or anxiety symptoms. Few if any studies reported all of the additional outcomes, but virtually all studies reported more than PTSD. As indicated above, the data on functioning appear to be lacking perhaps because of the way the SF-12 and SF-36 are categorized as measures of QoL, but overall, the picture is much more complete than indicated in the discussion.	The sentences referred to by this comment are in fact supported by the data reported. The text states: "Similarly, evidence for improving other outcomes of interest—anxiety symptoms, quality of life, disability or functional impairment, or return to work or active duty—was generally insufficient (often with no trials reporting those outcomes)." The reviewer missed that this text does not mention depression symptoms, as the reviewer seemed to think it did based on this comment. Evidence on depression symptoms was described in the text prior to this sentence (and the evidence was sufficient—and we do have numerous conclusions about depression symptoms within this paragraph and the surrounding paragraphs). For example, "The outcomes included in the table are those most commonly reported—PTSD symptoms, loss of PTSD diagnosis, and depression symptoms." And the table has a row for depression symptoms for each intervention---and only one of those was graded insufficient (for narrative exposure therapy).
86	Peer Reviewer #3	Discussion	Lastly, it is my understanding that the IOM will be issuing a report on PTSD treatment in mid-July. The authors should discuss any differences between their report and the	The new IOM report has just been recently released, after

Comment Number	Commentator & Affiliation	Section	Comment	Response
			new IOM report. The discussion of differences from the 2008 report is helpful.	we completed all of our update searches and were nearing final edits for this report. It is approximately a 400 page book. The task of the Committee was to conduct a study of ongoing efforts in the treatment of PTSD. One of the Committee members, Johnathan Davidson, was on our TEP. There are 2 planned phases. Of note, they were not tasked with reviewing the evidence on effective therapies, or updating guidelines, or making conclusions about various treatments. Their focus was quite different from ours and is very different from the focus of the group that wrote the 2008 IOM report. The first phase (initial report) has been completed and is described in the report. In phase 1, the Committee task was to collect data from the DoD and the VA on programs and methods available for prevention, screening, diagnosis, treatment and rehabilitation of PTSD. Additionally, the Committee was tasked with considering the status of studies and clinical trials involving innovative treatments of PTSD. The second phase (not yet completed) will focus on

Comment Number	Commentator & Affiliation	Section	Comment	Response
				analyzing data received and determining success rates for the various programs. There is a chapter on Treatment (Chapter 7), but it is not based on a systematic review. They note in the report (pg 21) that "…although the committee considered many studies in this report, it did not systematically review and rank each study." It is essentially a narrative review that describes available treatments---much of the chapter is similar to the text that we have written in the introduction that describes/defines the various categories. Some of it also provides narrative review of published trials.
87	TEP Reviewer #5	Discussion	In particular Table 52 is quite useful to summarize the needed findings. It is a confusing literature, and this is particularly helpful aspect of this document.	Thank you
88	TEP Reviewer #6	Discussion	The discussion does a good job of stating key findings and describing the limitations of the review. To identify limitations in the studies, the reader has to go to the details of the tables. Even there, it is difficult to determine the specific reasons that individual studies are rated as having a particular degree of bias. This tends to be a limitation of the AHRQ/GRADE method and not this document, per se.	Thank you. There is an extensive amount of information describing the limitations of specific studies. Thus, as we often do, we have put much of that information in an appendix describing our risk of bias assessments. However, we do provide some summary of this information in the report in the Methods and Discussion sections.

Comment Number	Commentator & Affiliation	Section	Comment	Response
89	TEP Reviewer #6	Discussion	By focusing on measured efficacy/effectiveness or harms, AHRQ funded reviews do not always consider or discuss other factors that may appropriately influence medication choice. Thus, the conclusions necessarily focus more on the medications for which strongest evidence exists currently despite the fact that a number of these medications have clear problems that are distinct from other medications in the same class. For example, among the SSRIs, paroxetine is likely to be most problematic in pregnancy and it is generally viewed as having the most problems with withdrawal symptoms and anticholinergic side effects. It's short half-life can create difficulties with patient adherence and it also has significant issues with drug-drug interactions (as does fluoxetine). These sorts of issues are not incorporated in the discussion but could realistically influence medication choice, particularly in some patient groups.	We appreciate the point here, and how making clinical decisions should often consider information beyond what is included in our report, but the intention of this report is to conduct a comparative effectiveness review and not to provide clinical practice guidelines. Evidence about the use of SSRIs in pregnant patients with PTSD was not identified in any of our included studies and we are unable to address whether it's use is more problematic in pregnancy than other medications that might be considered to treat pregnant women with PTSD. Also, we added the following to the limitations section: "Our review did not include an assessment of some factors important for clinical decision making, such as adherence or interactions with other therapies that could influence real world effectiveness of treatments."

Comment Number	Commentator & Affiliation	Section	Comment	Response
90	TEP Reviewer #6	Discussion	In addition, at least from a theoretical standpoint, other unstudied or understudied SSRIs might have comparable benefits but this concept does not seem to be incorporated into the future research section.	We have expanded our future research section to include a point to cover unstudied or understudied medications. In our Table in that section, we include "Future studies could evaluate medications that have some evidence (often from one or two small trials) suggesting possible efficacy (e.g., prazosin, olanzapine, mirtazapine) or medications that have not yet been studied with some theoretical basis to support their potential efficacy."
91	TEP Reviewer #6	Discussion	In other respects, the future research section does seem able to be translated into specific studies for future research but it seems overly constrained by the current evidence. It may also be worth emphasizing that these "gaps" relate only to the key questions addressed by this report and that these suggestions would not eliminate a wide range of PTSD related research gaps that fall outside of these specific key questions.	Good point. We have added text to make sure this is clear, as suggested. New text: "Of note, these gaps relate only to the key questions addressed by this report and the gaps identified here should not eliminate a wide range of potentially important PTSD-related research that falls outside of the scope of our key questions."

Comment Number	Commentator & Affiliation	Section	Comment	Response
92	TEP Reviewer #6	Discussion	Finally, in the sections that discuss effects of treatment on comorbid conditions, especially psychiatric symptoms, the focus is on the PTSD first and the concomitant condition secondarily. However, a particular treatment might have a much stronger evidence for an effect in the comorbid disorder alone, which could also influence decisions about treatment choice. For example, as a clinician, I wouldn't have any qualms about using a treatment for which multiple good studies showed a strong effect in depression alone and a moderate effect in PTSD alone, even if no studies specifically treated patients with PTSD and measured effects on depressive symptoms. In fact, I might consider this as better evidence than a single small study that showed moderate effects on PTSD and smaller effects on depressive symptoms in PTSD patients.	Review of evidence and conclusions about interventions that address the comorbid condition alone, from studies of people that do not also have PTSD, is beyond the scope of this report. The clinicians on our team appreciate the intentions of this comment, and it is certainly part of routine clinical thinking to think about such possibilities, but this report is not the place clinicians should look for such information (on how effective some of these treatments are for various comorbid conditions alone---and that would be quite an extensive undertaking with such a large list of treatments in this report and a great number of potential conditions).
93	Brown, Peter Public Reviewer	Discussion	Please see PDF	See response to initial comment from this reviewer.
94	Najavits, Lisa Public Reviewer	Discussion	N/A	No response required.
95	Anonymous Public Reviewer #2	Discussion (Conclusion)	This section really appears to need a broader focus. It needs to state key limitations of the literature (e.g., the exclusion of many of the patients with PTSD, such as those with substance use disorders, psychosis, bipolar, cognitive impairment, domestic violence, homelessness, inpatients, etc).	See response to related comments from this reviewer above. We have addressed this in the applicability section, with the addition of 6 new paragraphs.
96	Anonymous Public Reviewer #2	Discussion (Conclusion)	It also needs to state that there are many other issues to address in future research including workforce issues (e.g., the intensive training and required professional degrees required of clinicians), cost (e.g., that research needs to address cost, not just outcomes).	See response to the related comment—4 comments after this one from this reviewer.

Comment Number	Commentator & Affiliation	Section	Comment	Response
97	Anonymous Public Reviewer #2	Discussion (Conclusion)	The vast majority of RCTs had the treatment developer as an author or collaborator. Investigator allegiance to their treatment is a known bias that should be mentioned in Conclusions.	See related comment and response from this same reviewer. We have added text to the report about allegiance and potential bias.
98	Anonymous Public Reviewer #2	Discussion (Conclusion)	Cross reference: the same set of issues apply to pg. 141 (the table of future research) which entirely omits the issues above.	See responses to related comments above
99	Anonymous Public Reviewer #2	Discussion (Conclusion)	Other issues (no page numbers): There are various other truly important issues that are not addressed, as follows. a) The draft does not address patients being *paid* for each session attended. This is a known issue with various RCT studies, for example, where the patient is paid at each session of therapy. This should be documented.	We have carefully considered this suggestion and discussed it with our research team. In reviewing this literature, when subjects were paid (which was not something mentioned by very many studies) it was generally for attending/completing assessment sessions as a strategy to optimize follow-up and minimize missing data (and not for completing each therapy session). This is similar to the trial literature for most fields. We do not feel this warrants mention in the report or that it introduces any clearly discernible bias, since both/all arms are generally being paid for completing assessments. We could understand how payments might introduce bias if they were only provided to one arm, but this does not appear to be the case.

Comment Number	Commentator & Affiliation	Section	Comment	Response
100	Anonymous Public Reviewer #2	Discussion (Conclusion)	b) The draft does not address the high cost and level of training required of some models. For example, the cost to train a PE or EMDR therapist per the requirements of the treatment developers are many thousands of dollars when one includes not just the lengthy required days of training, but also the costly consultation with a professional that they require. Also, some statement as to the required clinician qualifications is needed—e.g., many require the clinician to have an advanced degree in mental health (e.g., social work, psychology), which can limit their ultimate applicability.	We did not review any detailed information about costs or about requirements for training of various types of therapy providers as that was not within the scope of our Key Questions. However, we agree that these are potentially important issues and we have included the following in the Discussion: "Access to and availability of treatments may vary for individuals and by geography. For example, among all the potential psychological treatments for PTSD, the VA offers prolonged exposure therapy and cognitive processing therapy for its patients.[196] Many people with PTSD never seek or receive treatment—this may be due to symptoms of the disorder itself (e.g., avoidance, anxiety), particular patient characteristics that increase or decrease the likelihood of seeking treatment (e.g., age, marital status, race, comorbidities), lack of availability of treatments, stigma, costs, transportation, unfamiliarity with accessing treatment.[197-199]

Comment Number	Commentator & Affiliation	Section	Comment	Response
101	Anonymous Public Reviewer #2	Discussion (Conclusion)	c) There is no mention of the impact of external, uncontrolled treatments patients were receiving while in these RCTs. This is a huge and undocumented issue in most RCTs. It is rare for a patient with PTSD to have no other treatments (whether medication or psychotherapy or self-help groups, all of which can have powerful effects). Thus most of these RCTs are essentially testing the impact of treatments in the context of other, undocumented treatments. Most RCTs never report nor never measure it.	We agree that this could theoretically be important. The challenge in reviewing the evidence is that it is often just an unknown, with little or no information about external, uncontrolled treatments that subjects might be receiving. As such, it is very much speculation to assume that most patients in trials are receiving other treatments (especially when access to treatments is quite limited and many people with PTSD go untreated). In addition, many trials mention in their inclusion criteria that subjects were not allowed to start new medications, and had to be on stable doses of any potentially relevant medications throughout the trial.
102	Brown, Peter	References	Please see PDF	See response to initial comment from this reviewer.
103	Najavits,Lisa Public Reviewer	References	N/A	No response required.
104	Anonymous Public Reviewer #1	References	Contain numerous typos. Examples: For #2, title should be "Treatment of Posttraumatic Stress Disorder: An Assessment of the Evidence" (see http://www.nap.edu/catalog.php?record_id=11955). For #34, authors should be Atkins D, Chang S, Gartlehner G, et al. (see http://www.effectivehealthcare.ahrq.gov/ehc/products/272/603/Methods%20Guide--Atkins--01-03-2011KM.pdf). For #154, page numbers should be 1158-65 and PMID should be 17015818 (see http://www.ncbi.nlm.nih.gov/pubmed/17015818). Someone should thoroughly re-check all references.	Thank you, we have fixed these.
105	Brown, Peter Public Reviewer	Tables	Please see PDF	See response to initial comment from this reviewer.
106	Najavits, Lisa Public Reviewer	Tables	N/A	No response required.

Comment Number	Commentator & Affiliation	Section	Comment	Response
107	Brown, Peter Public Reviewer	Figures	Please see PDF	See response to initial comment from this reviewer.
108	Najavits, Lisa Public Reviewer	Figures	N/A	No response required.
109	Brown, Peter Public Reviewer	Appendixes	Please see PDF	See response to initial comment from this reviewer.
110	Najavits, Lisa Public Reviewer	Appendixes	N/A	No response required.
111	Peer Reviewer #1	Clarity and Usability	Yes	Thank you
112	Peer Reviewer #2	Clarity and Usability	Yes	Thank you
113	Peer Reviewer #3	Clarity and Usability	I believe that the recoding suggested for the psychotherapy studies is necessary before any research or policy recommendations can be made. The eccentric and difficult to understanding coding distinctions are likely to impair the report's credibility.	Thank you, we have done this as suggested. See responses to previous comments from this reviewer.
114	Peer Reviewer #3	Clarity and Usability	I suspect that some readers will take great exception to the excluded studies, particularly because some such as Foa's PE study, are seen a key pieces of evidence in other reviews and practice guidelines. Coupled with the decisions about the psychotherapy coding, some of which even seem like mistakes, it will be easy for readers to dismiss the solid work in the report. This is why it is so crucial to revise the psychotherapy section to be more in line with conventional thinking. Many people dismissed the IOM report because of concerns about the way exposure and cognitive therapies were coded. It would be a shame to have this report suffer the same fate.	As above, we have fixed this. Regarding the excluded studies (the reviewer is referring to the high risk of bias studies), this is part of the reason to conduct sensitivity analyses by bringing those high risk of bias in. Of note, the findings were essentially unchanged. For PE (prolonged exposure) specifically, the analyses without the study result in the same conclusion as the analyses with the study, and we make this clear in the report.

Comment Number	Commentator & Affiliation	Section	Comment	Response
115	Peer Reviewer #3	Clarity and Usability	I have commented above only just how difficult it is to read the report in a way other than taking it at face value. This report is an ambitious and important project, one with great current and even greater potential value. More effort should be put into presenting the data in a way that requires less effort for readers to understand what was done and to evaluate the validity of the claims.	Thank you, we have revised the report to improve the presentation of the data to require less effort for readers; see responses to the specific comments above. We realize that the full report is a large document with copious information; the executive summary provides a more condensed/readable version. We also hope to have a related journal publication to provide a more concise version that takes less effort for readers; it has been submitted and is under review.
116	TEP Reviewer #5	Clarity and Usability	It is really excellent. I enjoyed reading it and see that it will be a classic for scholars and practictioners.	Thank you
117	TEP Reviewer #6	Clarity and Usability	The structure and text of the report are clear and it is well organized for its purpose as a systematic review.	Thank you
118	TEP Reviewer #6	Clarity and Usability	The conclusions can be used to inform policy/practice decisions but it should be made clear that the huge number of gaps in the evidence base should lead decision makers to be quite cautious in adopting these conclusions at face value (e.g., give everyone paroxetine) without considering the complexities of the available evidence, the lack of information about many similar treatments, the lack of information about the effective ingredients of effective treatments (e.g., psychotherapies, EMDR) and aspects of treatment that were not covered by this review.	We agree that this is very important. In our revised discussion, we feel that it provides a balanced description of the findings as well as the limitations of the evidence and gaps in the evidence.