Comparative Effectiveness Review Review Disposition of Comments Report

Research Review Title:
Psychosocial and Pharmacologic Interventions for Disruptive Behavior in Children and Adolescents

Draft review available for public comment from December 20, 2014 to January 26, 2015

Research Review Citation: Epstein R, Fonnesbeck C, Williamson E, Kuhn T, Lindegren ML, Rizzone K, Krishnaswami S, Sathe N, Wright G, Raj M, Ficzere K, Ness G, Potter S, McPheeters M. Psychosocial and Pharmacologic Interventions for Disruptive Behavior in Children and Adolescents. Comparative Effectiveness Review No. 154. (Prepared by the Vanderbilt Evidence-based Practice Center under Contract No. HHSA 290-2012-00009I.) AHRQ Publication No. 15(16)-EHC019-EF. Rockville, MD: Agency for Healthcare Research and Quality; October 2015. www.effectivehealthcare.ahrq.gov.

Comments to Research Review

The Effective Health Care (EHC) Program encourages the public to participate in the development of its research projects. Each 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 E-mail. At the conclusion of the public comment period, authors use the commentators' submissions and comments to revise the draft 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://www.effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=2133
Published Online: October 19, 2015

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #2	Executive Summary	Note a minor edit needed on page ES-8: the section is titled "Preschool Children," but covers interventions for all three age groups.	Thank you. We have corrected the section heading.
Peer Reviewer #2	Executive Summary	No comment about harm from psychosocial interventions	We agree that there needs to be comment about harm from psychosocial interventions. These studies did not look at harms. We have added text to this section and in other places within the main report to note this.
Peer Reviewer #3	Executive Summary	Double check the last sentence on ES16 and top line on ES-17--seems a bit awkward.	Thank you. We reworked the sentence in the Executive Summary and a similar sentence in the Applicability section.
Peer Reviewer #4	Executive Summary	Detail nicely the search methodology and reasons why some types of studies were excluded.	Thank you for your comments.
Peer Reviewer #4	Executive Summary	The diagnostic criteria for each study is reasonably clear and outcome measures are the state-of-the art. Nice that so many researchers in this field are using the same outcome instruments so that outcome are comparable. The statistical methods are appropriate for this type of literature review.	Thank you.
Peer Reviewer #4	Executive Summary	The Key Questions are well word, explicitly stated, relevant and appropriate. Also they are pretty comprehensive. The search and rating methodology matches the Key questions very well in following chapters.	Thank you.
Peer Reviewer #4	Executive Summary	The search system is very comprehensive including meeting abstracts and other unpublished literature.	Thank you.
Peer Reviewer #4	Executive Summary	The search team went beyond what is generally done in the search strategies for reviews in that they also looked for unpublished results, particularly those that were of non-statistically significant results. They also searched conference abstracts and called the researchers of these psychosocial intervention to ask about unpublished results or those with non-significant results. This appears to lead to a very thorough search strategy.	Thank you.
Peer Reviewer #4	Executive Summary	The study limitations are clear. Nice that they accept parent reported outcomes because of the high cost of direct observations with independent raters.	Thank you for your comments.
Peer Reviewer #4	Executive Summary	However, their conclusion that the multicomponent programs have greater effects is my finding from my research on SFP from the very first NIDA RCT in 1982 to 1986 that tested the three different components—parenting training only, PT plus child skills training and PT + CT plus family skills training also. SFP combined with a child skills training program, produced effect sizes that were additive of the experimental conditions with one or the other only.	Thank you for your comments.

Source: http://www.effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=2133
Published Online: October 19, 2015

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #4	Executive Summary	I am not familiar with the type of Bayesian multivariate, mixed treatment statistical meta-analysis they conducted so hard to determine if their results are accurate.	To aid interpretation, we added several descriptive sentences in the "Data Synthesis, Synthesizing Results" subsection that succinctly describes the type of meta-analysis (MA) being used (multivariate network MA).
Peer Reviewer #4	Executive Summary	I have also found that the greater the baseline severity of conduct disorders in the children, the larger the effect sizes of my program, SFP as there is more to improve. Also that involvement of father in one RCT of SFP in Thailand found that you get larger effect sizes if fathers are involved. Also depression in mothers is associated with reduced recognition of improvements in the children.	Thank you for your comments.
Peer Reviewer #4	Executive Summary	It is of interest that there were no studies identified that answered KQ3 or 4. My impression is that those researchers who have developed psychosocial interventions are not interested in psychotropic medicines.	Thank you for your comments.
Peer Reviewer #4	Executive Summary	The areas of needed improvement in these studies and their publication clarity on methods of random assignment and other research issues is clear and should be highlighted for researchers in this field.	Thank you. We agree.
Peer Reviewer #4	Executive Summary	The authors make clear the limitations of their inclusion criteria and state that they omitted any preventive studies of high-risk populations because Medicaid and other insurance funders would not fund them. The new Affordable Care Act is supposed to fund preventive services but so far none of the evidence-based family prevention interventions has been tested in clinical settings, A team of these family researchers lead by Dr. William Beardsley and Dr. Hendricks Brown are working to get federal funding to test the EBP family interventions with clinical populations but it will still take several years to get the results needed to get the US Preventive Services Task Force to approve them for funding.	Thank you for your comments. We did not exclude preventive studies on the basis of funding, but because our review was focused on individuals who met a clinical threshold for a DBD. We have changed the text to be clearer.
Peer Reviewer #4	Executive Summary	The fact that they found about half (N=22) of the studies conducted in countries other than the USA is somewhat surprising. However, the Australia studies are mostly to Triple P by Matt Sanders and his research team. The others in other countries are mostly dissemination studies including the program developers of existing USA developed parenting and other family interventions. What is not clear is how much these interventions were culturally adapted for the new culture rather than just translated into the new language.	We agree that evaluations of cultural adaptations of interventions are important. That issue was not a focus on this review.
Peer Reviewer #4	Executive Summary	The potential for conflict of interest is mentioned since most of these studies are conducted by the program developer or associates. I like their terminology of “intellectual descendent” of the original program developers as one of the contributors to high “conflict of interest” bias potential. Efficacy studies with RCTs (Phase 3 NIH studies) are generally	Thank you. We agree that more real world studies of these interventions are needed. We note this in the future research needs section

Commentator & Affiliation	Section	Comment	Response
		only awarded to the program developers who are generally university professors. More Phase 4 Dissemination studies in the real world are needed of these interventions to demonstrate clinical effectiveness.	
Peer Reviewer #4	Executive Summary	They mention Functional Family Therapy but later is does not appear in any of the reviewed studies and I wondered why. Maybe because FFT is conducted mostly in court settings but with teens who have committed offenses.	Appendix A includes detailed information about our search strategy. Functional Family Therapy was a term included in our search.
Peer Reviewer #4	Executive Summary	This review says that included studies were those that provide treatment for DBDs even without a diagnosis, yet it seems only those studies that screened for DBDs were included in later chapter.	Although we did not require included studies to include only children with a formal DBD diagnosis, we did require included studies to include children with either a formal DBD diagnosis or scoring above threshold on a validated measure. We did not otherwise include children at-risk or primary prevention.
Public Comment (APA)	Executive Summary	Puerto Rico is listed as if it were not part of the United States, it is also listed this way elsewhere in report where the studies are categorized by location. Puerto Rico is part of the U.S. (a territory) so consider re-classifying studies and listing accordingly or consider adding a brief note explaining why it is being classified as not part of the U.S. (i.e., cultural differences, etc.).	Thank you. We have elected to report the study numbers for Puerto Rico separately from the studies conducted in the United States due to cultural differences. We have added a note to the methods section about this.
Public Comment (APA)	Executive Summary	While the Incredible Years intervention could certainly be used with the parent-only intervention component, it is commonly used in clinical practice as a multi-component intervention (i.e., with both child and parent components). Consider listing a different intervention as an example of a parent-only intervention (i.e., Triple P).	Thank you. We made the revision to include Triple P.
Public Comment (APA)	Executive Summary	There is a caption for pre-school children but pages appear to be missing the captions for school age children and adolescents.	Thank you. We have corrected the section heading.
KI #1	Executive Summary	You state that "there is some evidence suggesting that improved parenting practices at least partially mediate intervention effectiveness…" You may find the attached recent review (published in December, 2014) to be relevant in helping you to qualify/support this statement. There is some support, but not overwhelming, that parenting practices are the mechanism of change in parent management training.	Thank you for the reference to which we now cite in the Discussion section of KQ6.
KI #1	Executive Summary	Why is the following title on page ES-8: Preschool Children? Under this section, there are studies reported for both school age children and for adolescents.	Thank you. This has been corrected.
KI #1	Executive Summary	The number of studies in the preschool, school age, and adolescent age groups do not agree across pages vi, ES-7 & -8, and ES-15. For example, on ES-7, there are 19 studies noted for preschoolers whereas in the Table on page ES-15, there are 18 studies. Similarly, on page ES-8, there are 24	Thank you. Please note these corrections and clarifications refer to the numbers reported in the draft version reflecting the initial literature search. With the updated literature search

Commentator & Affiliation	Section	Comment	Response
		studies for preschool children but only 23 in the Table on ES-15.	conducted during peer review, the numbers of studies reported in the final report will be different. For the preschool category, in our initial literature search, there were 19 studies (18 RCTs and 1 nonrandomized controlled study). We have clarified these numbers in the table. For the school age category, there were 24 studies (23 RCTs and 1 nonrandomized controlled study). We have clarified these numbers in the table. For the Adolescent category, we have corrected the numbers to reflect the 12 studies (11 RCTs and 1 nonrandomized controlled study).
TEP Reviewer #1	Executive Summary	For the most part, the executive summary is nicely laid out and the key points are provided.	Thank you.
TEP Reviewer #1	Executive Summary	Discussion of the harms of pharmacologic interventions, in particular, (page ES-11 and beginning on page 78) should acknowledge that the studies in the literature are not designed nor powered to assess harms. Research is generally focused on detecting the benefits and so studies are powered for this endpoint - not a harm. The extent of evidence of harm that can be gleaned from the literature is clearly limited in randomized controlled trials, which comprise the majority of the studies in this review. In order to provide a fair and balanced assessment of harm, it would be advisable to note that the most that any RCT can detect are the common and likely less severe adverse events. The small samples and short follow up are insufficient to detect rare events and are thus a major methodological limitation. This issue is raised later in the report, however, a naive reader of this report could mis-interpret the lack of evidence from RCTs as an indication of no harm. As duly noted in this review, RCTs do not translate into real-world settings - and it is also important for interpreting the risks. [NOTE: This comment is repeated in the section, "Discussion"]	We have added text about each of these issues to the Discussion in the main report and the Executive Summary.
TEP Reviewer #1	Executive Summary	It might be useful to state up front that there are none of the pharmacologic interventions have an approved indication for DBD. Thus, all of the studies are indeed investigating off-label use.	We have added a statement to this effect.
TEP Reviewer #1	Executive Summary	Pages ES-8 and ES-9 of the executive summary does need to be revised for clarity. The evidence is presented for psychosocial interventions but it is all under the sub-heading of preschool children, when in fact the text that follows reviews the evidence for specific interventions in preschool children as well as studies targeting school age and adolescent children. The main report does a better job of separating this information. The executive	Thank you. We corrected the section heading.

Commentator & Affiliation	Section	Comment	Response
		summary could be improved by a similar demarcation of the evidence.	
TEP Reviewer #4	Executive Summary	In the adolescent category, include a line for parent only interventions. (Even though there are none, the same is done for child only interventions in the Preschool category.)	Thank you. We have added a row for parent only interventions under the adolescent category in the strength of evidence table.
TEP Reviewer #4	Executive Summary	Delete the heading "Preschool Children"	We have corrected this. We have added a subheading for "School Age" and "Adolescents" to this section for consistency.
TEP Reviewer #4	Executive Summary	Change "combine" to combined"	Thank you. We have corrected to read "Effectiveness of combined psychosocial and pharmacologic interventions compared with individual interventions"
TEP Reviewer #4	Executive Summary	In the Executive Summary and elsewhere throughout the document, it is stated that studies of children with ADHD were excluded unless the specific focus of treatment was on disruptive behavior. However, in the section on Inclusion and Exclusion Criteria (ES-4, lines 23-27), it is stated categorically that such studies would be excluded.	Thank you. We have revised the inclusion and exclusion criteria section to clarify our decisions for studies of children with ADHD or ADHD-related disruptive behaviors.
Peer Reviewer #4	Abstract	I think it is unfortunate that a statistical analysis such as used by the Cochrane Reviews or statistical meta-analyses using Cohen's d was not conducted since all the relevant studies were located. That would help the clinicians and policy makers to determine the average effect size of parent-only, youth-only or multiple interventions. I found in one RTC and combining child only and family interventions had roughly an additive impact in improving clinical outcomes.	We believe that the effect sizes reported serve as a functional equivalent of Cohens d, since they are estimates of the effect in standard deviations, relative to treatment as usual/control. Moreover, our effect sizes are accompanied by credible intervals that incorporate system-wide uncertainty in the associated estimate, rather than just a simple point estimate.
Peer Reviewer #4	Abstract	It would be good to clarify the definition of children with disruptive disorder as only studies that included children with a specific cut off score on a standardized test of diagnosing disruptive behaviors.	We agree but word limits for the abstract preclude us from making this change.
Peer Reviewer #4	Abstract	It would be nice in the abstract to specify how the "strength of the evidence" of the studies were determine even if in just a few more words. Was any existing system used such as that by the Cochrane Reviews that include meta-analyses or the CDC's Guides to Clinical Treatments or Preventive Services? Does strength mean the effect sizes or the quality of the study? Later you mention that "Two senior investigators graded the body of evidence for key intervention/outcome pairs using methods based on the Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Good to add this is an AHRQ approved system.	Due to word limits, we have added "in accordance with the EPC Methods Guide".
Peer Reviewer #5	Abstract	And some minor issues and typos: : "… categorized AS interventions …"	This has been corrected to read, "Psychosocial interventions were categorized as interventions..."

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #3	Introduction	Appropriately acknowledges limitations of prior reviews and guidelines.	Thank you.
Peer Reviewer #3	Introduction	PICOTS, Analytic Framework clearly presented.	Thank you for your comments.
Peer Reviewer #5	Introduction	And some minor issues and typos: “DBDs are often present in the absence of a specific DBD diagnosis”. Do the authors mean “Disruptive Behaviour symptoms”? or just “Disruptive behaviours”?	Thank you. We have revised the statement to read, "DBD symptoms are often present in the absence of a specific DBD diagnosis."
Public Comment (APA)	Introduction	Typo in sentence, “Combination therapy with antipsychotics and stimulants is commonly for patients with attention deficit hyperactivity disorder (ADHD) comorbid with DBD or aggression…”	We have inserted the word, "used" to correct the sentence to read, "Combination therapy with antipsychotics and stimulants is commonly used for patients with…"
TEP Reviewer #1	Introduction	The diagnostic criteria are appropriate - with a small caveat. The background (page 1) notes a 1-6% prevalence of ODD and 1-4% prevalence of CD - on page 8 the included studies had to focus on the DBD and this had to be the primary problem. Exclusion of studies where DBD were evaluated as symptoms or comorbidities is important to discuss in relation to who the 'included' studies really represent. There is a lot of overlap between ADHD and DBD and so it is likely that a lot of information was excluded based on this criterion. This is noted in the limitations. However, it would be useful to know if the selection of studies would likely misrepresent the 1-6% and 1-4% prevalence of ODD and CD, respectively.	We revised this text in the Executive Summary and in the Introduction of the Main Report.
TEP Reviewer #2	Introduction	Good overview. Clear statement on what will be presented.	Thank you.
TEP Reviewer #3	Introduction	Well stated overall. PLease consider the comments included in the attachment.	Thank you for your comments.
TEP Reviewer #5	Introduction	The introduction set up the review well but was overlay technical and not well written to appeal to a public policy audience.	Thank you for your comments. An Executive Summary is included and we will submit the work for publication as a manuscript in a peer review journal, both of which may be less technical than the full report.
Peer Reviewer #1	Methods	It is still not clear to me why targeted (indicated) prevention studies are not included, in cases where the screening and inclusion of children was based on ratings of children's aggressive and/or conduct problem behaviors. The IOM conceptualizes treatment and prevention as being on a continuum; findings from indicated prevention studies can inform treatment research (interventions such as IY and Coping Power have been used in targeted prevention studies as well as in treatment stdies. The lack of attention to indicated prevention should be clearly noted in general (and when the effects of these programs with substantial prevention research history are described), and described as a limitation later in the report.	We agree and think that topic is amenable to future review. However, it was outside the scope of the current project. The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of systematic reviews to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. AHRQ expects that these systematic reviews will be helpful to health plans, providers, purchasers, government programs, and the health care system as a whole. For more

Commentator & Affiliation	Section	Comment	Response
			information about AHRQ EPC systematic reviews or to nominate a topic, see www.effectivehealthcare.ahrq.gov/reference/purpose.cfm
Peer Reviewer #2	Methods	Inclusion of assessment bias in accord with Cochrane standards is also a strength.	Thank you.
Peer Reviewer #2	Methods	A neglected topic which was dismissed is prevention. The example given for Early Risers was eliminated due to implementation cost. In fact, the cost for training an Early Riser parent advocate is $5000 to $8,000, in contrast to PMT-O at $25,000 per clinician (pp. 13-14, shin/content//SMA11-4634CDDVD/SelectingEBPS-IDBD.pdf) and $1,170,000 year one cost for an agency (http://www.blueprintsprograms.com/programCosts.php?pid=c837307a9a2ad 4d08ca61a4f1bd848ba3d6890fc). Assuming the intent was to exclude prevention, an alternative rationale could be added, although Triple P is a multi-level prevention and intervention model.	We agree and think that topic is amenable to future review. However, it was outside the scope of the current project. The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of systematic reviews to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. AHRQ expects that these systematic reviews will be helpful to health plans, providers, purchasers, government programs, and the health care system as a whole. . For more information about AHRQ EPC systematic reviews or to nominate a topic, see www.effectivehealthcare.ahrq.gov/reference/purpose.cfm
Peer Reviewer #2	Methods	This reviewer is curious about the unstated rationale for the 1994 start date as significant research on DBD treatment was published in the late 1980s and early 1990s. For example, earlier research by Henggeler, et al., on MST, and Chamberlain, et al., on PMT-O for treatment foster care, were not included, nor was Chamberlain's later research.	As described in the Methods section, "Eligible studies were not limited to intervention timing or duration of followup, but we limited the search to studies published in or after 1994. We conducted a preliminary screening of records retrieved from a search with no limits to the publication year. We screened approximately 1500 records published 20 or more years ago, and found that the study populations were inadequately described and poorly characterized, rendering a large number of the older studies unusable for this review. In order to include studies of patients meeting the population criteria for this review, the team agreed to limit the retrieval of primary study data to those studies published in or after 1994, as this date cutoff aligns with the

Commentator & Affiliation	Section	Comment	Response
			availability of the DSM-IV." Additionally, it is our intention that our review of prior reviews likely covers many of these articles. We added a sentence to this effect in the Executive Summary.
Peer Reviewer #2	Methods	A concern, although not intentional by AHRQ or the authors, is that the identified interventions could get preferential policy treatment when a number of others were similarly studied with solid outcomes. At a minimum, a statement about how interventions were selected and others were not and why needs to be added to Methods. If a consensus process of experts was utilized, acknowledge with information about that process.	The literature search strategy (and list of named psychosocial interventions and pharmacologic agents) was presented to the technical expert panel for feedback. Appendix A includes detailed information about our search strategy.
Peer Reviewer #2	Methods	While acknowledging that the primary focus of the review was to assess differential outcomes for the parent-child, parent only, and child only approaches, and recognizing that the intervention choices were excellent, what is puzzling is how decisions about which interventions were selected is not apparent.	The literature search strategy included broad terms for psychosocial interventions as well as a list of named psychosocial interventions. We shared the protocol, which included the literature search strategy and specific named interventions, with technical experts for feedback. We developed inclusion and exclusion criteria (described fully in the Eligibility section in the Methods Chapter of the Full Report) to select papers that evaluated one or more eligible interventions. Appendix A includes detailed information about our search strategy and Appendix B includes the screening forms that were used to determine study eligibility.
Peer Reviewer #3	Methods	Exclusion of studies conducted only in a hospital or a systems level intervention well justified.	Thank you for your comments.
Peer Reviewer #3	Methods	Methods are clearly described. Search strategies are explicitly described and comprehensive.	Thank you for your comments.
Peer Reviewer #4	Methods	Nice that their date extraction forms for each study were deposited into an international database for replication or other studies as this would be a useful database for meta-analysis.	Thank you for your comments.
Peer Reviewer #4	Methods	They also asked Key Informants about current or unpublished research which contributes to the completeness of their review.	Thank you for your comments.
Public Comment (APA)	Methods	Add period and delete comma after sentence, “All measurement instruments shared the same study arm treatment effect in our model, “	We have corrected this.
TEP Reviewer #1	Methods	The report does a very nice job of describing the inclusion and exclusion criteria - and the justification is sound. The figures provide a clear layout of the process.	Thank you.

Commentator & Affiliation	Section	Comment	Response
TEP Reviewer #1	Methods	To be clear, a study that focused primarily on the management of DBD but the sample could have comorbid ADHD was included in the review. But a study that focused primarily on the ADHD and the sample also had a DBD was not included in the review. However, in reality the samples could be the same from a clinical perspective. This is something I am certain the authors considered but in my read of this report it was not explicitly evident. Someone could challenge that qualitatively the study samples may be very similar - but the focus of the analysis, and thus how the study was powered, is what drives the decision.	We agree with the reviewer's concern. We tried to focus on the outcomes for which studies appeared to have been powered. We have reworked text throughout the document in an attempt to make our approach more clear.
TEP Reviewer #2	Methods	Overall, methods appear reasonable and thorough.	Thank you for your comments.
TEP Reviewer #2	Methods	The proper terminology is "boxed warning," not "black box warning." This should be changed wherever there is a reference to these warnings.	We have corrected this.
TEP Reviewer #2	Methods	With regard to the 6000+ publications excluded at abstract screening: it might be helpful to provide a hyperlink to the Appendix describing the screen criteria.	The Effective Health Care Program posts the final report and associated appendixes on their website. We do not include links within the report text.
TEP Reviewer #3	Methods	I am not a researcher. That said, i though that the inclusion/exclusion criteria were well thought out and all methods of analyzing the studies very appropriate. I do not usually see such thoughtfulness in examining articles, e.g. in meta-reviews	Thank you for your comments.
TEP Reviewer #4	Methods	Throughout the document – there are contradictory statements about whether only two or three of the pharmacological studies were not supported by the pharmaceutical companies (e.g., ES- 17, line 18; page 117, line 34).	Thank you we have corrected throughout the document for consistency and added the funding source(s) to the KQ2 study characteristics table.
TEP Reviewer #4	Methods	The definition of DBD provided in the box is essentially the DSM definition of Conduct Disorder. It does not include behaviors associated with ODD or Intermittent Explosive Disorder.	Thank you. This definition was reviewed with our Key Informants, finalized with the study protocol and cannot be changed at this time. However, we note and understand your comment.
TEP Reviewer #4	Methods	Throughout the document, there is inconsistent use of either “preschool/school age/adolescent” versus “pre-kindergarten/pre-teen child/teenage.”	Thank you. We have standardized the categorizations to: “preschool”, “school age”, and “teenage”.
TEP Reviewer #5	Methods	The search strategies were well justified but overly influenced by RCT evaluations. The degree to which these evaluations reflect community practice were not evaluated which limited the public health significance of the findings.	The review includes RCTs and non-RCTs. Nevertheless, we share the reviewers concern that the literature may not reflect community practice and note this in the Future Research Needs section.
Peer Reviewer #1	Results	Although it is noted whether the comparison groups are TAU or wait-list or placebo, these distinctions are not noted wheb the strength of findings are discussed (I would expect strongest effects with wait-list control designs,	We agree that it is reasonable to expect larger effects for active treatments compared to waitlist control or treatment as usual.

Commentator & Affiliation	Section	Comment	Response
		and weakest with reasonable TAU)	However, we combined treatment as usual and waitlist control into one group for several reasons. First, there is inconsistent use of the terms across studies with some studies provided little description of the control condition. This leads us to be unsure that they are correctly classified across studies with little information to "double check." Second, doing so would reduce the number of arms per group in our meta-analysis thus, reducing statistical power.
Peer Reviewer #2	Results	Conclusions about psychosocial and psychopharmacological have clear implications for practice, research, and policy. The very important finding and message to clinicians is that including parent and child is superior to parent or child only in treatment truly matters for youth outcomes. The major message from the psychopharmacology studies will be to inform clinicians about benefits and risks of medications and medication combinations, and for investigators, a critical need for more research. Further, all six research questions were addressed responsibly to the extent that sufficient data were available, and when not so, limitations were acknowledged, except for no comment about harm from psychosocial interventions (KQ5.PES-11). [NOTE: COMMENT ON KQ5 LIMITATIONS IS INCLUDED AND ADDRESSED SEPARATELY]	We agree that there must be a comment about harm from psychosocial interventions. The studies did not examine this. We now explicitly state this in the Executive Summary and Discussion section of the Main Report.
Peer Reviewer #3	Results	Results for K1 clearly and concisely described.	Thank you for your comments.
Peer Reviewer #3	Results	Only one study rated as having a low risk of bias (Findling et al 2000): consider adding a comment about why this study was rated as low risk of bias. All studies examining the efficacy of antipsychotics were funded by pharma or all authors served on the speaker bureau, so what set this study apart from the others? study design rated as good.	We have added a statement that all elements of the risk of bias assessment were determined in our review to be of low risk of bias. This study is particularly well described in the manuscript and all elements of the risk of bias tool are clearly addressed.
Peer Reviewer #3	Results	The finding that "other FDA review documents available did not include pediatric data" is an important one. Consider a comment that would share with the reader how this could happen.	This is reflection that many drugs are used off label.
Peer Reviewer #4	Results	Good to add the name of these programs and the authors as a check of the references produced authors I have not heard of except Maldonado	It is unclear to us what the reviewer is referring to in this comment. Apologies.
Peer Reviewer #4	Results	Not sure what the n=25 studies left and the bottom for inclusion in the meta-analysis represents as not described on page 44.	We note. "We identified a subset of studies (n=30) from KQ1 to contribute data to the network meta-analysis." Please note, that this number is different from the number in the draft report due to the literature search

Commentator & Affiliation	Section	Comment	Response
			update.
Peer Reviewer #4	Results	Nice that Carolyn Webster Stratton has tested IY with both the parenting and child components. Glad to see that the combined intervention had the greatest improvement as found in SFP and other studies.	Thank you for your comments.
Peer Reviewer #4	Results	Nice to see that McMahon's Helping the Noncomplient Child now has a smart phone video application. Too bad his sample size is so small. This program is the basis of the family interaction component of SFP and we now also have a totally video version but want to get it also on smartphones.	Thank you for your comments.
Peer Reviewer #4	Results	The characteristics of the studies are by and large clearly described and contributes to the length of the document.	Thank you for your comments.
Peer Reviewer #4	Results	This figure is a good way to show how they started with 6868 records and the flow of exclusion.	Thank you for your comments.
Peer Reviewer #4	Results	Amazing that half of the studies (n=22) were conducted outside the USA but mostly by the original program developed in an effort to promote international dissemination or by request of other government to implement these family EBPs in their country. My concern is with the lack of cultural adaptation done on these EBPs or in any case any description in their publications. I just sent to the publisher –Academic Press, chapters by all of these family intervention researchers and wrote one chapter on cultural adaptation. So I have reviewed their original articles and also their chapters. Your review also does not mention as a moderator cultural adaptation, but then I didn't find any of these EBPs has had degree of cultural adaptation as a variable in RCTs or quasi-experimental studies. We found in 5 different 5 year phase in studies that cultural adaptation didn't improve outcomes but did increase by recruitment rates and reduce attrition by an average of 40%	We agree that cultural adaptation is an important issue. It was not a focus of this review. We added a statement to this effect to the Future Research Needs section.
Peer Reviewer #4	Results	Are these Cohen's d effect sizes? If so specify and also mention that d = .14 is a small effect size so not much improvement. It would be good to calculate the effect sizes for all of the interventions to compare which are most effective.	In the second paragraph of the section "Bayesian Meta-Analysis of Psychosocial Interventions", we have inserted a sentence explicitly stating the interpretation of our effect sizes. They are in terms of standard deviation changes in expected scores, relative to treatment as usual/control. A score of -1 is therefore a 1 SD reduction in expected score. The point estimates of the effect sizes for the three classes of non-control interventions are all roughly in the -1 to -1.4 range.
Peer Reviewer #4	Results	Good to mention which level or type (clinic, workbook or online) of Triple P had significant outcomes. Later I think you do clarify this.	Table 8 in KQ1 results indicates the type and level of Triple P and the comparison of

Commentator & Affiliation	Section	Comment	Response
			behavior outcomes between Triple P groups.
Peer Reviewer #4	Results	Good to specify more clearly that this type of meta-analysis is.	We have added several descriptive sentences in the "Data Synthesis, Synthesizing Results" subsection that succinctly describes the type of meta-analysis being used (multivariate network meta-analysis).
Peer Reviewer #4	Results	Good to specify the average length of the therapist calls in Triple P.	We have indicated the frequency of therapy sessions across the Triple P studies, and added the mean duration and range of telephone calls as reported in the publication by Connell, et al. (1997).
Peer Reviewer #4	Results	Interesting that mothers rated significant improvements in their children but the fathers didn't. Any speculation by the authors or the reviewers as to why?	We did not speculate about this difference.
Peer Reviewer #4	Results	There are other family and parenting prevention intervention RCT studies for high risk children with conduct disorders that are not included because of the narrow inclusion criteria. I just hope that policy makers will understand that these few psychosocial parenting only, child only or multicomponent interventions are not the only effective programs for reducing conduct and disruptive disorders that are very similar in dosage and content but do not have inclusion criteria targeting only youth with diagnosed behavioral disorders, but recruit more broadly in schools and communities to avoid to stigma of 'therapy' to have a larger public health impact at a lower cost since many of these are not individual therapy but family group skills training.	Thank you for your comment. We agree.
Peer Reviewer #4	Results	Yes, the amount of detail in the results section is appropriate but as I mentioned above under General Comments, clinicians and policy makers are likely to skip over these and just want the 'bottom line" in an Executive Summary as not very interested in the research methodology as are researchers or reviewers like myself.	An Executive Summary is included. We will also submit a manuscript length version for peer review publication.
Peer Reviewer #4	Results	Good and clearer that you added the name of the treatment intervention to the table in addition to the author.	Thank you.
Peer Reviewer #5	Results	About the pharmacological interventions: on page 19, second paragraph, the risperidone RCT's are described as "small (20-355 participants)". I believe the study with 335 participants is the Reyes et al (2006) study, correct? A pharmacological study with more than 300 participants is not considered small in mental health. This study is almost 10 times larger than most of the other pharmacological ones here listed. Should it receive more attention? It is less likely to be underpowered for their main analyses, when contrasted with the other studies.	Yes, correct. One of the three studies of risperidone included in this review randomized 355 participants and should be described as a large study rather than a small study. The methods account for study size in the strength of evidence assessments.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #5	Results	- Moreover, shouldn't more attention be paid to risperidone, since it is the most likely choice for prescribers? Significant info is provided on risperidone for KQ5 (risk of harm), but not much on KQ2 (efficacy), even though is the antipsychotic with the largest number of RCT's.	We have added information to the results for KQ2 and findings for KQ5 to address the studies of risperidone.
Public Comment (APA)	Results	I would also like to have seem more on CPS – now called Collaborative and Proactive Solutions by its developer, Ross Greene. We have a study now in press comparing CPS to Barkley's variant of PMT – in fairness though the reviewers would not have had access to it. We can make it available if such is appropriate. In the final analysis, it is really hard to keep up with this literature as studies are coming out each month in various journals – and their review is only through June of 2014 – fair enough!!	Thank you. We have conducted a literature update since Peer Review and included seven additional studies of psychosocial interventions for KQ1.
Public Comment (APA)	Results	I think there are two seminal review papers that include references to the many studies published before 2004 and many thereafter. Sheila Eyberg was an author on both – see below. Note the catch title of the Brestan and Eyberg paper published in 1998 – they take us back 29 years, 82 studies, and 5.272 kids! Also, the edited book by Murrihy, Kidman, and Ollendick (2010) contains chapters by many of these early authors (or their followers) that update this earlier work and build upon it. 1) Brestan, E., & Eyberg, S.M. (1998). Effective psychosocial treatments for conduct-disordered children and adolescents: 29 years, 82 studies, and 5272 kids. Journal of Clinical Child Psychology, 27, 179-188.; 2) Eyberg, S. M., Nelson, M. M., & Boggs, S. R. (2008). Evidence-based psychosocial treatments for children and adolescents with disruptive behavior. Journal of Clinical Child and Adolescent Psychology, 37(1), 215 - 237.; 3) Murrihy, R.C., Kidman, A. D., & Ollendick, T.H. (Eds.) (2010). Handbook of clinical assessment and treatment of conduct problems in youth. New York: Springer Publications.	We have updated the review of existing systematic reviews and include the referenced publications.
Public Comment (APA)	Results	Harms and burdens (KQ 5): Not surprisingly, there is very little information on harms of psychosocial interventions and there are a few places where the authors need to state that. [See below]	We agree and have added explicit mention of this to the Executive Summary and to KQ5 Discussion section.
Public Comment (APA)	Results	Harms of psychosocial or pharmacologic interventions, P. ES-11. No discussion regarding harms of psychosocial interventions in this section. It is stated on page ES-13 that the reason for this is because it isn't addressed in the literature but a sentence stating that would be helpful in the beginning of this section as well. It appears that the "loss to follow up in several studies" were all psychopharm studies but again that needs clarification.	Thank you. We added a sentence about harms of psychosocial interventions to the KQ5 section of the Executive Summary and clarified that the loss to followup is in pharmacologic studies.
Public Comment (APA)	Results	Much detail about psychopharm. Again need to note there is nothing on psychosocial intervention harms.	We added a sentence to note an absence of harms reporting for the psychosocial interventions in KQ5.
Public Comment	Results	We acknowledge that harms/burdens issue is tough one since most studies	Thank you. Yes, we agree and have added a

Commentator & Affiliation	Section	Comment	Response
(APA)		simply do not comment on harm. Most of the harm cited from pharm studies is based on what the reviewers call "package insert data" not on studies with DBD specifically. I think this is okay and worthwhile, however, since I do not believe any of these meds have FDA approval for DBDs (other than those for ADHD). It would be good to comment on the absence of harm noted is not an indication that harm might not be present – even for the psychosocial treatments.	statement to address that issue in the overview section for KQ5.
Public Comment (APA)	Results	Consider adding a brief explanatory paragraph about network meta-analyses when discussing this analysis in the body of the review. Many psychologists and likely other end-user clinicians are not familiar with this statistical technique and adding more explanation would make the review more user-friendly.	We have added several descriptive sentences in the "Data Synthesis, Synthesizing Results" subsection that will hopefully clarify, in broad terms, what network meta-analysis is all about, and why we are using it here.
Public Comment (APA)	Results	Second sentence- indicted should be indicated.	We have corrected this.
Public Comment (APA)	Results	For the network meta-analysis figures, consider using more distinct colors for the blues- the fill color for the control looks very similar to the color for multicomponent (i.e., see Figure 3).	We have re-generated the figures with colors having better contrast, selected using colorbrewer2.org
Public Comment (APA)	Results	Although I really liked the review – and thought it was comprehensive – I was surprised that they decided to delimit the review from 1994 – June, 2014. In doing so, they did not include many, many studies conducted by the likes of Alan Kazdin, Jerry Patterson, Russ Barkley, Rex Forehand, BobMcMahon and others for PMT (primarily with parents only) and Sheila Eyberg (PCIT) quite a bit of early work by Scott Henggeler and colleagues on MST and Jim Alexander, Tom Sexton, and others on Functional Family Therapy and its derivatives (I know, BTW, that Jim and Tom represent different aspects of FFT. Still, a review needs to start somewhere – it was just unclear to me why 2004. As noted, many, many studies are not included.	As described in the Methods section, "Eligible studies were not limited to intervention timing or duration of followup, but we limited the search to studies published in or after 1994. We conducted a preliminary screening of records retrieved from a search with no limits to the publication year. We screened approximately 1500 records published 20 or more years ago, and found that the study populations were inadequately described and poorly characterized, rendering a large number of the older studies unusable for this review. In order to include studies of patients meeting the population criteria for this review, the team agreed to limit the retrieval of primary study data to those studies published in or after 1994, as this date cutoff aligns with the availability of the DSM-IV.15" Additionally, it is our intention that our review of prior reviews likely covers many of these articles. We added a sentence to this effect in the Executive Summary.
Public Comment	Results	That said, I did like the outline and approach – breaking the review down	We agree that the potential for

Commentator & Affiliation	Section	Comment	Response
(APA)		developmentally (preschool, school age, adolescents) was welcome as was the partitioning of studies into child alone, parent alone, and multicomponent practices (though the delineation is not always clear – for example many PMT programs include child components – though they are largely parent only. The reviewers might provide additional commentary on this decision and the "reliability" of their assignments.	misclassification error is important to be aware of with this approach. Although we did not formally calculate the "reliability" of our assignments, the only assignment about which we received a question was PCIT. To test the impact of the classification of PCIT, we re-ran the analysis with PCIT classified as parent only and compared the results. The difference was nominal, with the multicomponent intervention shifting downward in value somewhat, but not changing the general result. We added text discussing this to the report.
Public Comment (APA)	Results	I would suggest there are some serious gaps in the studies included in the review. I will declare a bias but, there are at least 2 (possibly 4) Functional Family Therapy studies within the time frame and seemingly within the inclusion criteria.	Appendix A describes our search terms. Functional Family Therapy was included in our search. As described in the Methods section, "Eligible studies were not limited to intervention timing or duration of followup, but we limited the search to studies published in or after 1994. We conducted a preliminary screening of records retrieved from a search with no limits to the publication year. We screened approximately 1500 records published 20 or more years ago, and found that the study populations were inadequately described and poorly characterized, rendering a large number of the older studies unusable for this review. In order to include studies of patients meeting the population criteria for this review, the team agreed to limit the retrieval of primary study data to those studies published in or after 1994, as this date cutoff aligns with the availability of the DSM-IV." Additionally, it is our intention that our review of prior reviews likely covers many of these articles. We added a sentence to this effect in the Executive Summary.
Public Comment (APA)	Results	Beginning the review in 1994 biases those treatment interventions that have a longer history and were, thus, validated prior to this date. There are a number of organizations (Center for Violence Prevention: University of Colorado) that has done extensive work in validating some of these earlier,	We added a table summarizing existing systematic reviews (Appendix F). The Methods section describes how we incorporated existing systematic reviews, "We located recent

Commentator & Affiliation	Section	Comment	Response
		yet still effective and valid, programs. I would suggest their review.	reviews published between 2005 and 2014 and evaluated each for relevance using the review PICOTS (Appendix B). We summarized review data from relevant psychosocial and pharmacologic interventions in the discussion to put our findings in context of existing evidence. For the systematic reviews reporting harms, we, and assessed quality using AMSTAR and summarized the findings in KQ5."
Public Comment (APA)	Results	I was very impressed with the inclusiveness of the review – save my comments about the early studies that occurred prior to 1994.	Thank you.
Public Comment (APA)	Results	Most importantly, some of the primary interventions for the most difficult clinical problems in this areas were developed and the early studies occurred prior to the invitation of this review. This will be critical to include or we will get a distorted view of the sum total of the research.	We have updated the review of existing systematic reviews and include the referenced publications.
Public Comment (APA)	Results	I also think that variants of models is not a critical element—at least in some of the interventions with such variants, there is no clinical difference. My read of this and the literature would suggest that there is evidence to be able to talk about youth/problem severity and model adherence as moderators. Those would be important in developing actionable recommendations	We described evidence about baseline severity as a potential moderator in KQ6a and model adherence in KQ6d.
Public Comment (APA)	Results	There are a number of systematic reviews that would identify a whole range of these studies	We have updated the review of existing systematic reviews and include the referenced publications.
Public Comment (APA)	Results	Additional key articles prior to 1994 that should be included are:(I have a bit of a bias here…so use what you like of this. I know FFT studies well, but not those of other models). I might suggest targeting intervention programs that fit the criteria of having a long history (FFT, MDFT, MST/PMT etc) the following key articles/reviews.	Appendix A describes our search. Functional Family Therapy was included as a search term. As described in the Methods section, "Eligible studies were not limited to intervention timing or duration of followup, but we limited the search to studies published in or after 1994. We conducted a preliminary screening of records retrieved from a search with no limits to the publication year. We screened approximately 1500 records published 20 or more years ago, and found that the study populations were inadequately described and poorly characterized, rendering a large number of the older studies unusable for this review. In order to include studies of patients meeting the population criteria for this review, the team

Commentator & Affiliation	Section	Comment	Response
			agreed to limit the retrieval of primary study data to those studies published in or after 1994, as this date cutoff aligns with the availability of the DSM-IV." Additionally, it is our intention that our review of prior reviews likely covers many of these articles. We added a sentence to this effect in the Executive Summary.
Public Comment (APA)	Results	It is impressive however there are some studies (of FFT) not included. I would like to know more about the specific criteria for search.	Appendix A describes our search. Functional Family Therapy was included as a search term. As described in the Methods section, "Eligible studies were not limited to intervention timing or duration of followup, but we limited the search to studies published in or after 1994. We conducted a preliminary screening of records retrieved from a search with no limits to the publication year. We screened approximately 1500 records published 20 or more years ago, and found that the study populations were inadequately described and poorly characterized, rendering a large number of the older studies unusable for this review. In order to include studies of patients meeting the population criteria for this review, the team agreed to limit the retrieval of primary study data to those studies published in or after 1994, as this date cutoff aligns with the availability of the DSM-IV." Additionally, it is our intention that our review of prior reviews likely covers many of these articles. We added a sentence to this effect in the Executive Summary.
Public Comment (APA)	Results	Also, I seem to recall one study on Triple P being conducted in the US – the review states that studies have only been conducted in Australia by the developer (Matt Sanders). The study I am thinking of was conducted in South Carolina with Ron Prinz as lead author (Matt Saunders is a co-author, along with some others – I will try to track it down).	Thank you. This is a U.S. trial but did not meet the criteria for inclusion in this review. The Triple P trial reported by RJ Prinz was a population-level trial that did not target children with DBD. We have clarified our statement in the report to clarify that all the Triple P studies that met the criteria for inclusion in our report were conducted outside

Commentator & Affiliation	Section	Comment	Response
			of the United States.
Public Comment (Cynthia Russo)	Results	Consider referencing the most current RISPERDAL® USPI (April 2014) The Risperdal (risperidone) 2007 Package is referenced (Ref 181) rather than the current package insert date April 2014. (Ref 182)	We now reference the 2007 and the current (2014) package insert for Risperidone.
Public Comment (Cynthia Russo)	Results	Consider citing all Warnings & Precautions in the RISPERDAL USPI: "Adverse events referenced in the warnings/precautions section of the package insert include: neuroleptic malignant syndrome, tardive dyskinesia, hyperglycemia, dyslipidemia, weight gain, hyperprolactinemia, orthostatic hypotension, leukopenia, neutropenia, agranulocytosis, cognitive/motor impairment, and seizures." Not all adverse events in the warnings/precaution section of the prescribing information are included in this paragraph. Please refer to the Risperdal Prescribing Information for the complete list of warnings.	We have updated the adverse events to match those listed in the current (2014) FDA approved package insert for Risperidone.
Reviewer 10 (SRC)	Results	The incidence of sedation in pediatric patients with schizophrenia receiving 4-6 mg/day of risperidone in the 6 week trial was 12% (and not 2%). (Page 84): "Common adverse reported in the pediatric population…." Is also referenced with the 2007 Package Insert rather than the current Package Insert - April 2014 (Ref 181). In this section sedation is reported as a 2% incidence rather than 12%.	Thank you. This has been corrected.
TEP Reviewer #1	Results	The results section is presented very clearly and in a well-organized manner. The tables are clear and they provide sufficient information to stand alone. The figures and graphs are equally well designed. It is impressive that the magnitude of the information was represented in easy to read tables. The Appendix is also very helpful and I found it quite useful as I reviewed the report.	Thank you.
TEP Reviewer #2	Results	Overall, appears quite exhaustive.	Thank you for your comments.
TEP Reviewer #2	Results	Note that "normal release" is not the appropriate terminology. Should be "immediate release."	We have corrected from "normal release" to "immediate release".
TEP Reviewer #2	Results	While I appreciate the inclusion of non-DBD safety data for completeness sake, it may be worth noting somewhere that rates of adverse reactions often vary widely depending on the population being treated.	We have added a statement to this effect.
TEP Reviewer #3	Results	Also, could we create a 'head to head' analysis of sorts, even if it's just to compare the level of significance of the findings of psychosocial interventions as opposed to pharmacological interventions? What many readers may not know is that psychosocial interventions compare so favorably to, or more favorably than, pharmacological interventions.	Because the studies did not compare pharmacologic and psychosocial treatments directly, we do not think it is appropriate to do so. Moreover, the decision about what treatment to use is very complex and involves more than head-to-head comparisons of the treatment approaches. Any direct comparison not drawn from a well-conducted trial would likely suffer from significant bias, including

Commentator & Affiliation	Section	Comment	Response
			confounding by indication.
TEP Reviewer #3	Results	I think that the amount of detail is excessive, but it also depends upon the needs of the reader. Researchers may need this level of detail, while other stakeholders are less likely to desire this level of detail.	An Executive Summary is included. We will also submit a manuscript length version for peer review publication.
TEP Reviewer #4	Results	I may be reading this incorrectly, but the information presented in this paragraph does not seem to map onto Table 24.	Thanks for spotting this. These typos were corrected, and the values should indeed match.
TEP Reviewer #4	Results	The Lavigne study is included in the preschool section, but because the average age was 6 years it should be in the School-age section.	Thank you. The text was incorrect. The age of the children in the study was between 3 and 6.11 years. We have changed the text from "All children were between the ages of 6 and 6.11 years," to "All children were between the ages of 3 and 6.11 years," We categorized this study as Preschool Age.
TEP Reviewer #4	Results	Change "Dose" to "Distress"	This has been corrected.
TEP Reviewer #4	Results	Change "Helping the Non-Compliant Child" to "Helping the Noncompliant Child" throughout.	This has been corrected throughout.
TEP Reviewer #4	Results	Change "PMOT" to "PTMO"	This has been corrected.
TEP Reviewer #4	Results	For this study, report CBCL total problems score is reported in the text, but CBCL externalizing score is reported on Table 17 (page 48, line 19).	We have rechecked and corrected. The authors reported outcomes for CBCL internalizing and externalizing. We report outcomes for CBCL, Externalizing subscale in the text and in the table.
TEP Reviewer #4	Results	Under the "Age group" heading, the listing is "School, preschool, adolescent." Change this to "Preschool, school, adolescent."	We have reordered the age groupings with the table.
TEP Reviewer #4	Results	In various places in the document it is stated that the samples were "overwhelmingly male." As the data on this table show, more than 25% of the participants were female. Language should be modified accordingly.	We have changed the sentence to read, "As in the current review, participants in studies included in prior systematic reviews were mostly male and typically Caucasian,"
TEP Reviewer #4	Results	The following study is relevant to this section:Beauchaine, T. P., Webster-Stratton, C., & Reid, M. J. (2005). Mediators, moderators, and predictors of 1-year outcomes among children treated for early-onset conduct problems: A latent growth curve analysis. Journal of Consulting and Clinical Psychology, 73, 371-388.	We have added this as a relevant study for KQ6.
TEP Reviewer #4	Results	The level of detail in describing individual studies varies across sections. For example, the child component section provides the mean age of each sample and the means and standard deviations of baseline and posttreatment scores. The parent section does not. The adolescent section does provide the appropriate level of detail.	We have made edits throughout for consistency (e.g., added mean age of study population and organized sections) for consistency. However, in some instances, the level of detail does vary as a function of the number of studies and complexity of the study reports for individual sections.

Commentator & Affiliation	Section	Comment	Response
TEP Reviewer #5	Results	The studies were explicitly described but the the tables and figures were not helpful in displaying results clearly and succinctly.	We made a number of changes throughout the document in response to peer review - including to tables and figures - that we hope improve the readability of the entire document including the tables and figures.
Peer Reviewer #1	Discussion	More emphasis should be placed on the reliance on parent ratings in these studies, rather than data from other sources. it could be noted that a limitation of these studies is that there is limited data from less biased sources such as teachers during follow-up periods and from independent observers.	We address this issue in the Limitations section with the following text "Third, the field lacks consensus on the most important outcomes, and as such, there are few studies that measure similar outcomes for synthesis. Methodologically, outcomes such as direct observation by a blinded and independent observer are arguably the most valid. However, direct observations can be expensive and are not always logistically feasible. From the perspective of patient-centered outcomes research, we believe that there is a strong argument in this literature to be made in favor of the importance of parent reported outcomes. However, most of the studied interventions included a parent component either alone or in combination with other components. Blinding was not always feasible and when parent reported outcomes were included, multiple measures of similar constructs were used within and across studies."
Peer Reviewer #2	Discussion	Since costs, although relevant, are not included in the research questions that might be specified in limitations.	We agree that cost is important. Costs were not specified within the scope of this review. This affects applicability and we added text to that section to reflect this concern.
Peer Reviewer #2	Discussion	Would like to have seen some research policy recommendations	We revised the Implications for Clinical and Policy Decision-making section.
Peer Reviewer #3	Discussion	Gaps in research are explicitly stated.	Thank you for your comments.
Peer Reviewer #3	Discussion	Consider adding a comment re: substantial high and moderate bias for the psychopharm studies in the conclusion--this is an important finding. on p 109, the team does an excellent job in last para of K2 results clearly describing this bias.	We have added text to the conclusion.
Peer Reviewer #3	Discussion	Excellent point re: polypharmacy. However, is the only limitation difficulty assessing applicability in highly complex cases? This is a thorny issue. A	We agree and have added the following: "Polypharmacy with two or more antipsychotic

Commentator & Affiliation	Section	Comment	Response
		national indicator of poor quality is use of 2 antipsychotic medications without documented rationale. Is the point that there is little research to assess the safety or efficacy of combined medication treatment, particularly for two ap's? Need for improved documentation of clinical rationale and use these data to examine "practice-based evidence" of combined medication safety? The following para (In reality...) does a nice job describing how restricting the literature review to DBD treatment has implications for the findings. To preserve conciseness, an option might be to add a reference or two re: common concern for polypharm and this commonly used quality measure.	drugs is a commonly used indicator of poor quality care although it clearly occurs. A better understanding of the prevalence, circumstances, and implications of polypharmacy is needed."
Peer Reviewer #4	Discussion	The limitations of the review/studies are described adequately including the need or more independent research, replication studies, and longer term follow-up in the discussion.	Thank you for your comments.
Peer Reviewer #4	Discussion	Also stressing increasing positive and respectful interventions and learning good discipline and communication techniques.	Thank you for your comments.
Peer Reviewer #4	Discussion	Could be more complete.	We hope our efforts to address the body of comments about this section satisfy the need for additional information.
Peer Reviewer #4	Discussion	I do feel that the investigators did omit important literature from the selective and indicated prevention literature as there is a very fine line between what is prevention and treatment. Prevention for adolescent problems such as substance abuse and delinquency are often conducted with children with conduct disorder since they are the highest risk population.	We agree that this literature may be amenable to review, but it was outside the scope of the current review. Regarding the current review, we also agree that there is a fine line between prevention and treatment. We limited our review to studies examining children with a formal DBD diagnosis or who scored above threshold on a validated measure.
Peer Reviewer #4	Discussion	What is not clearly stated is that practice time of the parent and child together makes the multicomponent intervention even more effective. This was also found in the CDC meta-analysis of Kaminski, et al. that determined the aspects of these family interventions that make them effective. Having practice time together with the therapist coaching them was a key factor in effectiveness.	We did not examine the specific issue of "child and parent practice time together." We agree that this could be one of the things that distinguishes multicomponent interventions from the other broad intervention categories we examined.
KI #1	Discussion	The primary concern I have is that I do not think that age of child is adequately covered and, as a result, the conclusions are misleading. Please see the Table on the following page which is based on some of the numbers reported in the text. Please note that these numbers may not be correct as I noted in number 1 above, the numbers are not consistent in various places. Nevertheless, the percents that I note indicate that there are clear trends depending on the age of the child for whether there is parent only or parent plus other components to the intervention. It seems	We agree that this is an important issue but are not entirely clear to which section of text the comment refers. The comment is identified as targeting the Discussion section but references p. vi. Thus, we assume this means the comment is directed towards the Structured Abstract in general and the Conclusions section of the Structured Abstract more specifically.

Commentator & Affiliation	Section	Comment	Response
		inappropriate to me to conclude on page vi that aparent component in combination with other intervention components are most likely to be effective at reducing problem behaviors than psychosocial interventions that include only a parent component or a child component. If one looks at the studies (I know that these do not represent studies necessarily with supportive outcomes but generally most studies did find support), it is evident that primarily (70%) parent-only studies were done with younger children whereas by school age less than a majority of studies (39%) had been done with parents only and by adolescents none of the studies had been done with adolescents only. Similarly, when we consider parent plus other components being implemented as the intervention, this increased from 30% at preschool to 57% at school age to 83% at adolescence. To conclude that a parent component in combination with other interventions is most likely to be effective at reducing problem behaviors is really not the case: It depends on the age of the child. I would propose that the conclusions be modified to indicate something such as the following: "Once children reach school age, psychosocial interventions for children with disruptive behavior disorders that include a parent component in combination with other intervention components are most likely to be effective at reducing problem behaviors than psychosocial interventions that include only a parent component or a child component. In the preschool years, only a parent component is most commonly implemented and necessary." You may not agree fully with my conclusion but I do think something needs to be inserted to indicate that interventions at the preschool age are done primarily through parents. Otherwise, we are going to have practitioners who are trying to do interventions that involve both parent and child components with preschool children and child components really are not important at this age.	We revised that paragraph to read, "Across all age groups, psychosocial interventions for children with disruptive behavior disorders that include a parent component in combination with other intervention components are most likely to be more effective at reducing disruptive child behaviors than psychosocial interventions that include only a parent or only child component. Each of these broad intervention categories appear to perform better than control conditions. In the preschool years, a parent component is essential. Small studies of antipsychotics and stimulants report positive effects in the very short term. The most commonly reported outcomes are parent-reported outcomes. Long term and functional outcomes were not consistently reported. There was variability in the duration of long-term followup and functional outcomes reported." Regarding how we addressed child age more generally, we organized the qualitative synthesis by age group to highlight the issue of age. Unequal distribution of intervention types across age groups is clearly described in the Executive Summary and in the Findings and Discussion sections. Though we considered age-by-treatment interactions, there was not enough balance among the age and treatment combinations to include them in the final model, and we present meta-analysis results by age group.
TEP Reviewer #1	Discussion	The report includes an adequate overview of the limitations as well as the limitations specific to KQ1-KQ4.	Thank you.
TEP Reviewer #1	Discussion	Discussion of the harms of pharmacologic interventions, in particular, (page ES-11 and beginning on page 78) should acknowledge that the studies in the literature are not designed nor powered to assess harms. Research is generally focused on detecting the benefits and so studies are powered for this endpoint - not a harm. The extent of evidence of harm that can be gleaned from the literature is clearly limited in randomized controlled trials, which comprise the majority of the studies in this review. In order to	Thank you. We have included an explicit statement to address this issue in the Executive Summary, in the KQ5 Overview and in the Discussion.

Commentator & Affiliation	Section	Comment	Response
		provide a fair and balanced assessment of harm, it would be advisable to note that the most that any RCT can detect are the common and likely less severe adverse events. The small samples and short follow up are insufficient to detect rare events and are thus a major methodological limitation. This issue is raised later in the report, however, a naive reader of this report could mis-interpret the lack of evidence from RCTs as an indication of no harm. As duly noted in this review, RCTs do not translate into real-world settings - and it is also important for interpreting the risks. [Note: This comment is repreated in the section, "Executive Summary"]	
TEP Reviewer #1	Discussion	One issue that was not adequately addressed is the need to explore evidence outside of the RCT design. Observational studies have their set of limitations, however, the methods for observational research are advancing and this should not be discounted. The reality is that any evidence of the long-term effectiveness and safety of interventions, pharmacologic in particular, will require large longitudinal cohort studies. This could be applied to the future research needs section of the report. There is inadequate funding for prospective studies to address the portability and safety of interventions in real-world practice setting.	The review includes RCTs and non-RCTs. Nevertheless, we share the reviewers concern that the literature may not reflect community practice and note this in the Future Research Needs section.
TEP Reviewer #2	Discussion	Appropriate conclusions based on evidence presented.	Thank you for your comments.
TEP Reviewer #3	Discussion	Highlight the following additional key findings and key questions as important for future study or future consideration (by each of the different stakeholders). For example, in my reading of the article, the following findings have great significance: 1) The size of the evidence base for the efficacy of psychosocial interventions is larger than the size of the evidence base for pharmacological interventions. Should clinicians therefore be recommending psychosocial interventions more often than pharmacological interventions? I do not believe that this is usually the case. Also, should clinicians not recommend psychosocial recommendations that specifically have a parent component? 2) The efficacy of pharmacological interventions is very short term. Only long-term efficacy is meaningful. Pharmacological studies assess efficacy over a shorter term than psychosocial interventions. This finding has important implications for the clinician as well as for third party payers. The major cost of DBD's lies in their chronicity. Patients, clinicians, and policymakers should be very interested in studies that reduce long-term costs. This means reducing symptoms over the long term, and/or using an intervention that will not be costly over the long term. What should clinicians be recommending? Can we start to measure outcomes over a longer time interval, either in clinical settings, or in research studies? 3) Not all studies measure functional outcomes (improvement in function or skills-building). Some only measure	We agree. To address the part of this comment focusing on different stakeholders, we have reworked the Implications for Clinical and Policy Decision-making section of the Discussion. Other comments are addressed in the limitations and future research needs sections.

Commentator & Affiliation	Section	Comment	Response
		behavioral outcomes (symptom reduction).The authors point out an important distinction in outcomes measures, which is that some studies measure 'Behavioral outcomes" while other studies measure "functional outcomes." "Behavioral outcomes" are focused mostly on symptom-reduction. "Functional outcomes" are focused mainly on skills- building. This seemingly simple distinction has very significant implications for treatment efficacy, patient satisfaction, and long- term cost.Improved functional outcomes are dependent upon specific teaching or training in the functional skills that the child may need to develop. Reduction in symptoms does not automatically result in improved functional outcomes. Focusing upon function and upon improving functional outcome sometimes has the beneficial effect of reducing symptoms 4) The quality of research is inconsistent. Research quality varies and is influenced by bias (inadequate randomization; bias due to funding source; treatment fidelity of psychosocial interventions, amongst others). It may be worth listing separately all of the studies deemed 'high quality."	
TEP Reviewer #3	Discussion	I think that the implications of the study could be amplified. I provided a list of what I think are the major implications in the attachment.	No comment needed.
TEP Reviewer #3	Discussion	Should the authors limit the "future research" section to formal research studies, e.g. RCT's? Or, can we also consider research to be other aspects of research, such as research to develop clinical practice guidelines, research into costs and cost-benefit ratios, etc. In other words, can the authors expand/amplify the relevance of the research question as they apply to clinical practices, policies, and implementation by third party payers?	We agree and added, "Additionally, studies examining implementation of these interventions in real-world community practice are also needed." to this section.
TEP Reviewer #4	Discussion	You summarize 12 reviews re effectiveness of psychosocial interventions. There are several additional meta-analyses that are relevant to this section of the document	We have updated the review of existing systematic reviews and include the referenced publications.
TEP Reviewer #5	Discussion	The findings were clearly stated but the limitations of the review did not include concerns as to whether the studies reflected community practice. The literature on the dissemination of the parenting practices was not well reflected. The implications for future research was especially disappointing. The need for research that bridges research and practice was not discussed.	We agree and added, "Additionally, studies examining implementation of these interventions in real-world community practice are needed." to this section.
Peer Reviewer #1	General	The report is clinically meaningful and adequate	Thank you.
Peer Reviewer #2	General	This review represents a solid contribution, clear summary, and implications for future research. It is a very responsible review of psychosocial and medication interventions for the treatment of youth with disruptive behavior disorders. Overall, it is impressively presented in text and tables. It is comprehensive, encompassing a larger literature since the early 1990s, including national and international studies and relevant meta-	Thank you.

Commentator & Affiliation	Section	Comment	Response
		analyses.	
Peer Reviewer #3	General	Overall, this is a clearly written, well organized report. Key questions are explicitly stated. With strong scholarship, it identifies substantial gaps in the research base.	Thank you.
Peer Reviewer #4	General	My field is research o and reviews of evidence-based family interventions to reduce child disruptive behaviors and prevent later higher treatment costs of mental disorders, substance abuse, delinquency as well as current child maltreatment and foster care. Sad that only 71 unique studies of interventions effects, of which 58 assessed psychosocial interventions and 13 assessed pharmacologic interventions were found. One would think in 10 years there would be more. Glad to hear that the multicomponent interventions had the best outcomes as I have also found this to be true in my research as well as seen in the research literature.	Thank you. We agree it is surprising that a larger literature was not identified. It is also important to note that we used relatively stringent study inclusion and exclusion criteria.
Peer Reviewer #4	General	The report is definitely clinically useful and meaningful. However, clinicians will not be interested in all the detail about the scientific search and rating methodology, but this is needed. Maybe an Executive Summary for policy makers and clinicians would be helpful to more general readers.	Thank you. An Executive Summary is included. We will also submit a manuscript length version for publication in a peer review journal.
Peer Reviewer #4	General	There are many more studies, even RCTs of psychosocial interventions, including parent training, family skills training and children's skills training for youth with disruptive behaviors. I think it would be good to clarify for the readers that only those programs specifically targeting as an indicated population youth with diagnostic scores over 127 on the Eyeburg or CBCL were included and that there are other psychosocial interventions that have large effect sizes in reducing children's disruptive behaviours, but they were not included because they were selective or universal skills training interventions without explicit diagnostic cutoff scores for inclusion in your study.	We agree it is important to make study inclusion and exclusion criteria clear. We have clarified the study inclusion and exclusion criteria in the methods sections of the Executive Summary and Main Report.
Peer Reviewer #5	General	This report is a thorough review of psychological or pharmacological interventions for disruptive behaviours in children. As expected for an AHRQ review, this report presents an analytic framework and search strategy that are thorough and well done, and the methodology is sound.	Thank you.
Peer Reviewer #5	General	However, I believe there are certain aspects of the report that could be further explored and/or expanded, to increase its clinical usefulness:	No comment needed.
Peer Reviewer #5	General	The definition of what constitute "disruptive behaviours" is quite broad in some of the reviewed studies. For instance, on Analytic Framework, page 12, many of the behavioural outcomes listed are very non-specific and have a certain degree of overlap, e.g., aggressive behaviour, violent behaviour and fighting. How to address this overlap? It would have been helpful if the authors could explore a bit more these behavioural constructs.	Thank you. This definition and analytic framework was reviewed with our Key Informants, finalized with the study protocol and cannot be changed at this time. However, we note and understand your comment.
Peer Reviewer #5	General	For most of the studies here reviewed, the primary outcomes were parent	We agree that the literature's reliance on

Commentator & Affiliation	Section	Comment	Response
		reports of child disruptive behaviours. However, there is very little comment on the lack of blinding for this measure in psychosocial interventions when compared to pharmacological ones. It is much more difficult to blind a parent who is receiving an active intervention (since this seemed to be the key factor in increasing the likelihood of effectiveness for the psychosocial interventions) when compared to a placebo pill taken by the child. Isn't it possible that the psychosocial treatments have inflated rates of response, compared to pharmacological trials, due to the lack of blinding? There are no comments on this potential bias on the report. There were no comparisons of efficacy when measures were done by an independent, blinded clinician when compared to a parent report. For instance, the author state on page 42 that, "We considered the feasibility of blinding in psychosocial studies and did not downgrade where it would have been impossible" - is this based on the decision that parental assessment of child behaviour is the outcome measure of interest? This leaves the reader with the impression that a tautological argument may be happening here; i.e., if the main outcome measure elected by the report is a parental assessment of the child's behaviour, these are naturally hard to be effectively blinded, thus these studies were not downgraded. Would the results be different if the main outcome measure for the psychosocial interventions was a teacher or a clinician report, i.e., a less biased and more likely to be blinded assessment of child's behaviour?	parent reported outcomes is a limitation and have added an explicit statement to address this in the "Limitations of this Review" section of the Discussion.
Public Comment (APA)	General	Great job addressing the questions initially posed in the proposal. This is a very nice review and will be very helpful. Reviewers described it as "refreshing to read" and found it overall easy to read.	Thank you for your comments.
Public Comment (APA)	General	Consider adding evidence profiles as an appendix to the report. This will allow guideline developers to make more immediate use of the review without additional intermediate steps. We will email an example evidence profile directly to EPC.	We have added evidence profile tables for KQ1 to the Appendix.
Public Comment (APA)	General	Outcomes that would be important to include [in Evidence Profiles] are: Reductions in youth behavior problems (as evidenced by: recividism/if in a justice setting, conduct problems symptoms); Improvement in mental health functioning; Improvements in family functioning (a primary risk factor for youth behavior problems); Effects on the family and society; Stable home placements (for those at risk of out of home placement); Impairment; Adverse events/harms; Quality of life of the young persons, their peers, and their families.	The outcomes included in the quantitative synthesize were based on the three most prevalent reports of group baseline and end of treatment means and standard deviations. Because there were a number of outcomes and various measures, we were unable to synthesize each. We do however, include all the behavioral and functional outcomes, as well as moderating variables, prespecified in the Analytic Framework in a data file in the systematic review data repository (SRDR) at:

Commentator & Affiliation	Section	Comment	Response
			http://srdr.ahrq.gov/.
Public Comment (APA)	General	Outcomes: These are far-reaching problems and the symptoms are only the tip of the proverbial icebergs. These are some of the most impairing and distressing childhood disorders.	We agree and hope this review helps prompt additional research.
KI #1	General	To say the least, this is a thorough and comprehensive report. In addition, it is well written. Congratulations on an excellent job!	Thank you.
KI #2	General	Thank you for the opportunity to review the report. It was very well done and written. I have no comments for revision.	Thank you.
TEP Reviewer #1	General	This comparative effectiveness review is a comprehensive report of the existing evidence for the management of disruptive behavior disorders among children and adolescents (preschool up to age 18). This is likely to be clinically significant given the extent of these problems in outpatient settings, both formally diagnosed as well as un-diagnosed problematic aggressive and irritable behaviors. The key questions are appropriate and they are presented in a logical format. The intended target audience and how this review should be used by the different groups is nicely described.	Thank you.
TEP Reviewer #2	General	Clinically meaningful. It's too bad that including a broader population would likely either render the report unwieldy, or make it too difficult to parse out the effects of treatment on comorbidity vs. the disruptive behaviors themselves. Nonetheless, the report is still quite valuable. The question just remains how generalizable the conclusions are when considering DBDs more broadly.	Thank you. We agree that a review of an at-risk population is an important topic for future research.
TEP Reviewer #3	General	Consider reducing the length of the article. The article is very long. The level of detail offered may be important for the researcher audience, but might be perceived as excessive for the clinician, policymaker, and/or third party payer audience. You will lose part of your audience if the article is too long, and/or of certain members of your audience need to search too deeply for the information that is most pertinent to them.	We agree. An Executive Summary is included. We will also submit a manuscript length version for publication in a peer review journal.
TEP Reviewer #4	General	There is a significant misclassification that likely significantly changes some of the conclusions of the review. Parent-Child Interaction Therapy is classified as a multicomponent intervention (e.g., pp. 28-31, and elsewhere throughout the document). This is completely incorrect – PCIT is a parenting only intervention, derived from the same "ancestor" (Constance Hanf) as Helping the Noncompliant Child, the Incredible Years, and Defiant Children (see Reitman & McMahon, 2013, Cognitive and Behavioral Practice, 20, 106-116).	The potential for misclassification error is always something we need to be aware of with this approach. Fortunately, it is relatively easy to run the model under any classification scheme that we wish. So, we re-ran the analysis with PCIT classified as parent only and compared the results. The difference was nominal, with the multicomponent intervention shifting downward in value somewhat, but not changing the general result. We added text discussing this to the report in several sections, notably in the Findings,

Commentator & Affiliation	Section	Comment	Response
			Bayesian analysis chapter.
TEP Reviewer #5	General	The population is well defined but I did not think that the audience for this review was clear. The questions were well described but based on the population for which this review was intended, it is not clear to me whether these questions were the most relevant.	The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of systematic reviews to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. AHRQ expects that these systematic reviews will be helpful to health plans, providers, purchasers, government programs, and the health care system as a whole. For more information about AHRQ EPC systematic reviews, see www.effectivehealthcare.ahrq.gov/reference/purpose.cfm
Peer Reviewer #1	Clarity and Usability	The report is certainly lengthy, but it is adequately structured given that inherent limitation.	Thank you.
Peer Reviewer #2	Clarity and Usability	In summary, this report has the power to influence policy, clinical practice, and research on interventions for youth with DBDs, and if appropriate, adding recommendations for federal research would be a future contribution. Major credit to AHRQ for taking on this project, to AHRQ staff who worked diligently on it, and who can utilize the findings to pursue federal policy options for increasing availability of effective treatment for a critically important clinical population. The very careful, competent, and thorough work of the Vanderbilt investigators is fully appreciated.	Thank you.
Peer Reviewer #2	Clarity and Usability	The interventions selected for inclusion have a solid research base and are currently being widely implemented and disseminated. The careful selection of interventions for three age groups (preschool, school age, and adolescents) is an important strength.	Thank you.
Peer Reviewer #2	Clarity and Usability	This raises a question about how published reviews of evidencebased practices for DBDs were utilized. In addition to the specific citations above, key examples where such information is available include: (1) the SAMHSA Evidence-Based Practices KIT: http://store.samhsa.gov/product/Interventions-for-Disruptive- Behavior-Disorders-Evidence-Based-Practices-EBP-KIT/SMA11- 4634CD-DVD, (2) National Registry of Evidence-Based Programs and Practices (NREPP): http://www.nrepp.samhsa.gov/ (3) California Evidence-Based Clearinghouse for Child Welfare (CEBC) http://www.cebc4cw.org/, (4) Office of Juvenile Justice and Delinquency Prevention (OJJDP) Model Programs Guide, http://www.ojjdp.gov/mpg/Program, and (5)	We added a table summarizing existing systematic reviews (Appendix F). The Methods section describes how we incorporated existing systematic reviews, "We located recent reviews published between 2005 and 2014 and evaluated each for relevance using the review PICOTS (Appendix B). We summarized review data from relevant psychosocial and pharmacologic interventions in the discussion to put our findings in context of existing evidence. For the systematic reviews reporting

Commentator & Affiliation	Section	Comment	Response
		Blueprints for Healthy Youth Development, http://www.blueprintsprograms.com/, hosted by the Center for the Study and Prevention of Violence (CSPV), at the Institute of Behavior Science, University of Colorado Boulder and funded by the Annie E. Casey Foundation. There are many strong candidates in those documents. Several were referenced, e.g., Kazdin and others, which seem to meet inclusion criteria.	harms, we assessed quality using AMSTAR and summarized the findings in KQ5."
Peer Reviewer #3	Clarity and Usability	Well organized, clearly written. This report strongly supports the need for more research and to place priority on funding studies with low risk of bias for both psychosocial and pharmacologic treatments.	Thank you for your comments.
Peer Reviewer #4	Clarity and Usability	The report is well structured and organized around the Key Questions. The conclusions can be used to inform policy and practice, but more easily with an Executive Summary that they are more likely to actually read and understand.	Thank you. An Executive Summary is included. We will also submit a manuscript length version for publication in a peer review journal.
TEP Reviewer #1	Clarity and Usability	The report is very nicely structured - which made it easy to review such a large document. The key points are noted in the Executive Summary and then at the beginning of each sub-section in the results. This makes it especially easy to review the highlights and then target the detailed description for the depth of information. Future research needs is a very good summary and clinical and policy implications are concise but effective.	Thank you.
TEP Reviewer #2	Clarity and Usability	At times, it was easy to get bogged down in all the details. But, that's to be expected in a report as exhaustive as this.	Thank you for your comments.
TEP Reviewer #2	Clarity and Usability	Conclusions can certainly be used to guide practice decisions. I hope they will encourage more insurers to cover psychosocial interventions.	Thank you for your comments.
TEP Reviewer #3	Clarity and Usability	The figures/tables were especially useful to the clinician.	Thank you.
TEP Reviewer #3	Clarity and Usability	This is a very well structured report and it is well organized. THe level of detail sometimes distracts from the overall organizational scheme. I am impressed by how well it is organized.	Thank you for your comments.
TEP Reviewer #3	Clarity and Usability	Consider labeling or identifying all of the specific audiences of the article. The article implicitly addresses several stakeholders. The article would be more likely to reach it's stakeholders or audience members if those stakeholders are identified explicitly, and if the needs of each stakeholder is addressed more explicitly. As examples, the stakeholders include: a. Parents and their children b. Clinicians c. Researchers d. Policymakers and third party payers.	We agree. To address this, we have reworked the Implications for Clinical and Policy Decision-making section of the Discussion.
TEP Reviewer #3	Clarity and Usability	Consider presenting your findings and key questions with accompanying text that explains its relevance to key audience(s) or stakeholders. For example, the following subtitles could be used in the article, or could be presented in a side bar or text box: 1) Key questions and/or findings from	We agree. To address this, we have reworked the Implications for Clinical and Policy Decision-making section of the Discussion.

Commentator & Affiliation	Section	Comment	Response
		this review that are important for parents and their children. For example: Parents want to have a positive relationship with their child, and to see their child participate successfully in daily routines at home, in social interactions, and at school. Parents and children would like to know which treatment is most likely to confer the best long-term outcome in these domains. Which treatment is the most likely to assure a positive outcome in these areas? 2) Key questions and/or findings from this review that are important in the creation of clinical practice guidelines. For example, clinicians will want to know answers to the following types of questions:i. How do I choose/recommend the right treatment for each child? ii.How do I choose/recommend the right treatment based on parent characteristics?iii. How will the family access those treatments? 3) Key questions and/or findings from this review that are important for policy makers and third party payers. For example, third party payers will want to know: i. What is the cost of each treatment ii. What is the long-term benefit of each treatment? iii. Which treatments are the least costly over the long term? (e.g. consider articles that include a measure of functional outcome) iv. How much money should be allocated for which types of treatments? 4) Key questions and/or findings that are important for researchers. Researchers should consider questions such as i. How can I create a study that shows true efficacy? (non- biased; adequately randomized; uses a treatment that is likely to show efficacy) ii. How can I create a study that measures comparable outcomes to studies already published? (use of an appropriate outcomes measure) iii. How can I create a study that is clinically meaningful? (use an outcomes measure that is clinically meaningful,e.g. functional outcomes)	
TEP Reviewer #3	Clarity and Usability	Figures/tables for the pharmacological studies would be welcome, to create a 'head-to-head' comparison of sorts between pharmacological and psychosocial intervention results.	We include summary tables in the KQ2 section similar to those included in KQ1. However, the tables and figures associated with the meta-analysis for KQ1 cannot be duplicated for KQ2 as we did not conduct a meta-analysis for KQ2.
TEP Reviewer #3	Clarity and Usability	The authors provide excess detail for a clinician and probably also for a policymaker, though the level of detail may be welcome by researchers.	An Executive Summary is included. We will also submit a manuscript length version for peer review publication.
TEP Reviewer #5	Clarity and Usability	I thought that the results were described too technically which obscured the clinical and public health significance of the results.	An Executive Summary is included. We will also submit a manuscript length version for peer review publication.
Peer Reviewer #3	Tables and Figures	Figure 1: excellent documentation of excluded studies and rationale	Thank you.
Peer Reviewer #3	Tables and	Table 3: very helpful	Thank you.

Commentator & Affiliation	Section	Comment	Response
	Figures		
TEP Reviewer #3	Tables and Figures	The tables for the psychosocial interventions were very helpful, and the graphs even more so. Is there any way to create similar graphs for the pharmacological results?	We did not conduct a meta-analysis for KQ2. Thus, tables and graphs like those included for the meta-analysis for KQ1 cannot be included. Due to the small number of studies on medication use for DBDs in children, we did not use a formal statistical approach to assess the possibility of publication bias, as it would be unlikely to be informative.
TEP Reviewer #1	Appendixes	Of note, on page G-3 of the Appendix - I really liked that the full references for the excluded studies were included and the notation with the exclusion code was very good. I did note that some of the studies (reference #20, #68, #136) had an exclusion code of X-7a but there is no X-7a in the Appendix G.Reasons for Exclusion table. This also occurred for reference #213 where X-4a is listed as the code but this is not a code in the table. There may have been others that I missed so a careful review is warranted.	Thank you. This has been corrected. This was an oversight. We have added the description of all exclusion codes to the table.