Comparative Effectiveness Review Disposition of Comments Report

Research Review Title: *Therapies for Children With Autism Spectrum Disorder—Behavioral Interventions Update*

Draft review available for public comment from January 14, 2014 to February 11, 2014.

Research Review Citation: Weitlauf AS, McPheeters ML, Peters B, Sathe N, Travis R, Aiello R, Williamson E, Veenstra-VanderWeele J, Krishnaswami S, Jerome R, Warren Z. Therapies for Children With Autism Spectrum Disorder: Behavioral Interventions Update. Comparative Effectiveness Review No. 137. (Prepared by the Vanderbilt Evidence-based Practice Center under Contract No. 290-2012-00009-I.) AHRQ Publication No. 14-EHC036-EF. Rockville, MD: Agency for Healthcare Research and Quality; August 2014. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Comments to Research Review

The Effective Health Care (EHC) Program encourages the public to participate in the development of its research projects. Each research review is posted to the EHC Program Web site in draft form for public comment for a 4-week period. Comments can be submitted via the EHC Program Web site, mail or email. At the conclusion of the public comment period, authors use the commentators' submissions and comments to revise the draft 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/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1945
Published Online: August 6, 2014

Commentator & Affiliation	Section	Comment	Response
TEP Reviewer #4	Abstract	Was there an age range and a time period?	We restricted to studies in children ages 0-12 and studies published from 2000 to the present. This information is presented in the Methods section.
TEP Reviewer #4	Abstract	I'd suggests including the number of studies reviewed originally, so that the reader knows what the total is.	This information is included in Figure 2.
TEP Reviewer #4	Abstract	One extra period and one missing period!	Thank you catching this error. We have corrected it.
Peer reviewer #1	Appendices	The appendices are comprehensive and I could not identify any relevant studies that have been omitted.	Thank you for your comment.
TEP Reviewer #1	Clarity & Usability	The report is written clearly. In its current form, it is usable by doctoral-level professionals but would be inaccessible to most others.	We hope to get this information to multiple audiences through publication in the peer reviewed literature and also via the Eisenberg Center, an arm of the Effective Health Care program that develops materials for clinicians and families.
TEP Reviewer #2	Clarity & Usability	Report well structured.	Thank you for your comment.
Peer Reviewer #3	Clarity and Usability	Clarity and Usability: As mentioned earlier, the report is well organized and thorough, but repetitive in many places due to the inherent structure/order of the key questions.	We have attempted to streamline presentation of results that may repeat across categories.
Peer Reviewer #3	Clarity and Usability	Main points are very clear.	Thank you for your comment.
Peer Reviewer #3	Clarity and Usability	Conclusions are important and will likely inform policy and practice decisions to the extent that a review like this can do so.	Thank you for your comment.
Peer Reviewer #3	Clarity and Usability	Since there will be newly emerging findings published regularly, the evidence may change and some disclaimer to this effect could be made.	This disclaimer is made as part of the SOE assessment
Peer Reviewer #3	Clarity and Usability	Likewise, it seems important to reiterate what was within the scope of this review and what was not (excluded) so that policy makers do not interpret the absence of some information as a lack of an evidence base.	It would be beyond the scope of this review to catalog every intervention that was not included. Rather, we have tried to be clear that this is a focused review and not a comprehensive update of every treatment for ASD.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #3	Clarity and Usability	Individual differences cannot be overstated; these groups studies apply to group averages not to specific individuals. Thus, decisions effecting individual patients are still best made by the clinicians, who hopefully will take the research evidence into account.	We agree.
Peer Reviewer #3	Clarity and Usability	The authors should be commended. This was an impressive undertaking - thanks for letting me serve as a reviewer.	Thank you for your comment.
Peer Reviewer #4	Clarity and Usability	Clarity and Usability: The reprot was easy to follow and read dispite the length. The executive summary and the body of the report were clear. The information as mentioned above was very useful for policy makers and system administrators.	Thank you for your comment.
Peer reviewer #1	Conclusion	e. Discussion/ Conclusion: The major findings of the study are clearly stated. Thus, although there has been a considerable growth in the number of well controlled studies over the last 3-4 years, suggesting the potential effectiveness of several different approaches, research in this area continues to suffer from significant limitations	Thank you for your comment.
Peer reviewer #1	Conclusion	Among the most important of these are the lack of long-term follow-up studies and the lack of evidence of generalisation beyond the training context. There is more work taking place now on moderators or mediators of intervention although many more studies in this area are needed in order to enhance understanding of what interventions are most effective for which particular types of children.	Thank you for your comment.
Public Reviewer #2 (Association of Professional Behavioral Analysts)	Conclusions	With some caveats, we concur with the authors' conclusion (p. v and elsewhere) that the evidence from such studies shows that "Young children receiving high intensity applied behavior analysis-based early intervention over extended timeframes commonly displayed substantial improvement in cognitive functioning and language skills relative to community controls," and with some of the descriptions of the limitations of that research. The same goes for the conclusions about the reviewed studies on parent training, social skills interventions, play/interaction-based interventions, joint attention interventions, and cognitive behavior therapy (CBT) for anxiety.	Thank you for your comments.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #2 (Association of Professional Behavioral Analysts)	Conclusions	On pp. ES-16--ES-18 and elsewhere, the authors note that their conclusions should be tempered by the facts that their review included a limited subset of studies on "behavioral" interventions for ASD, and that those studies evaluated multiple different treatment approaches.	We feel that our conclusions are accurate and do not consider that they should be tempered by our inclusion criteria. Rather, as we note, other studies may provide additional information also important in making individual treatment decisions.
Peer reviewer #2	Discussion	I was surprised that the play/interaction studies did not also get a moderate for improvement in language - I am thinking of the Kasari et al 2008 study mainly but there is a hint of this in others I think...	We considered the strength of the evidence low given the fair quality of the studies assessing this outcome and imprecision of the outcome measurement given the small number of studies and sample size.
Public reviewer #1 (Richard G Kensinger, MSW)	Discussion	Again reflects our clinical reality & I trust psycho-social interventions far more than the use of potent psychotropics	Thank you for your comment.
TEP Reviewer #2	Discussion/ Conclusion	Future research well stated.	Thank you for your comment.
TEP Reviewer #3	Discussion/ Conclusion	Discussion/ Conclusion: I likely missing something, but why isn't there an assessment of the strength of the evidence for KQ 2-5 and 7?	Per EPC methods, strength of evidence is assessed for a limited number of outcomes, usually those related to effectiveness. We have noted in the Methods section of the report that we assessed strength of the evidence only for Key Question 1—outcomes of intervention because that is the primary effectiveness question.
TEP Reviewer #3	Discussion/ Conclusion	In the Applicability section, the comment is made on page 84, line 6, that PICOS information is available in Appendix G. When reading this section, I stopped, went to Appendix G and hoped that here lay some answers to KQ2 and wondered why these answers wouldn't have been incorporated into the main text...but the answer on modifiers in not there. Too bad. Perhaps a softer comment on "...MAY support translation of our findings..." wouldn't have gotten my hopes up. (This is a very minor point I'm making here...just hopeful for answers.) Use of the word "may" does appear, and I think this is very appropriate, on page 85, line 9.	Data on applicability are meant to help assess how generalizable the findings of studies included in the review are to children with ASD in the general population. We have attempted to clarify the purpose of this information.
TEP Reviewer #3	Discussion/ Conclusion	The discussion regarding the lack of outcome measures is nicely presented on page 88, lines6-22.	Thank you for your comment.
TEP Reviewer #3	Discussion/ Conclusion	The conclusion is a well summarized section and nicely written.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
TEP Reviewer #3	Discussion/ Conclusion	Clarity and Usability: Yes to all these questions.	Thank you for your comment.
TEP Reviewer #3	Discussion/ Conclusion	I think that it OK that there wasn't a lot of research to answer all the key questions. I'd like to think that some of the open issues identified by the 2011 report helped to shape the research that was reviewed in this report and that this report will go on to influence, and especially close gaps for future systematic and comparative reviews.	Thank you for your comment.
Peer Reviewer #3	Discussion/ Conclusion	Discussion/ Conclusion: In general, the discussion and conclusions seem on target. The authors did a thorough job discussing limitations of the review and challenges of the specific studies /outcomes for the most part.	Thank you for your comment.
Peer Reviewer #3	Discussion/ Conclusion	One concern was with the discussion of the moderators - it's not clear that most of the studies had the power to fully test moderators yet often the conclusion is that if moderators were "not significant" then the treatment appears to be effective for a range of child characteristics. This seems too generous. These issues are particularly salient in the tables since you have limited space and can't expand on the reason for null findings on moderators here.	We have added a comment that most studies are not designed or powered to assess moderators. We agree that this is an issue in this literature that makes it very difficult to answer this important question.
Peer Reviewer #3	Discussion/ Conclusion	In future research/gaps section (p. 87), there should be some statement on the need for a systematic review of the categories that were not included in this review.	This is an important point and has been added.
Peer Reviewer #4	Discussion/ Conclusion	Discussion/ Conclusion: The implication of the major findings were clear.	Thank you for your comment.
Peer Reviewer #4	Discussion/ Conclusion	Limitation were described adequately.	Thank you for your comment.
TEP Reviewer #1	Discussion/ Conclusion	The conclusions are consistent with the data, and the discussion of findings is helpful.	Thank you for your comment.
TEP Reviewer #1	Discussion/ Conclusion	It is perhaps worth noting that many or most of the recent RCTs are relatively small, often enrolling about 20 participants per group, with the intention of serving as pilot studies to prepare for larger trials; these larger trials will be needed to address the unanswered questions about moderators, "dose," intervention method, etc.	We have added a statement about sample size concerns to the discussion of the limitations of the evidence base.

Commentator & Affiliation	Section	Comment	Response
TEP Reviewer #1	Discussion/ Conclusion	Another point that may be worth making is that investigators are beginning to evaluate services as delivered by community providers; these evaluations have sometimes yielded null findings, the interpretation of which is complicated by pre-existing differences between groups (Boyd) or low fidelity of intervention (Mandell). However, there are also some successes such as the studies on training early childhood special educators to implement interventions for enhancing social communication.	Thank you for this comment. We have noted the need for more of these kinds of studies in the future research section of the report.
TEP Reviewer #2	Discussion/ Conclusion	the implications are appropriately stated; limitations described.	Thank you for your comment.
Peer reviewer #2	Executive summary	page v - Add to Conclusions of Abstract (if space permits) the need to identify effective elements/components of change to final sentence (as in terms of accessibility/equity this could have far-reaching consequences).	We have added this idea.
Peer reviewer #2	Executive summary	page ES-8 - Most of the studies purporting to examine moderator effects have not done the proper analysis which is an interaction/regression within an RCT design.	Yes, we have noted that few studies are designed or powered to properly assess moderator effects.
Peer reviewer #2	Executive summary	Another methods point for research is to clarify primary/secondary outcomes (in advance).	Reporting of primary outcomes in advance has improved over our last review; however, we agree that this is an important point and have added it to the future research section.
TEP Reviewer #3	Executive summary	Also of note, ES-7, the paragraph on play/interactive approaches doesn't speak to the strength of the evidence, which the other paragraphs do very nicely.	We have added the SOE to this paragraph.
Peer reviewer #2	Executive summary/Discussion	page ES-17 - Research gaps. In addition to my point on [page v] - the need to identify effective elements/components of change to final sentence I would also ad d the need to identify and test potential mediators of change. Understanding mechanisms is fundamental to refining and improving treatments in psychological therapies and the autism field has not done this well to date (whatever some may claim).	We have added a statement about the need to understand mediators to the future research section.
Peer reviewer #1	General	I found this an excellent report. It is comprehensive, thorough, and written clearly and concisely. It is not a totally new review, being an update on the authors' previous systematic review of the literature published in 2011.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Peer reviewer #1	General	It is pleasing to note that the quality of research has improved considerably since the previous review ; however, as the authors highlight, research in this area continues to suffer from many methodological and practical limitations (e.g. lack of information on generalisation to non-trained settings; very limited follow-up periods; limited analysis of moderators and mediators of treatment)	Thank you for your comment.
Peer reviewer #1	General	The target population is well described. The focus is on interventions for children with autism ages 2-12 and very young children (<2 years) at possible risk of ASD	Thank you for your comment.
Peer reviewer #1	General	However , the intended audience could be specified more clearly	The Uses of This Report section in the Introduction contains information about intended audiences.
Peer reviewer #1	General	f. Clarity and Usability: As is evident from my comments above, the report is comprehensive, well structured and organised. The conclusions pertain mainly to the need for much greater methodological rigour in child autism intervention research. Until this can be achieved it remains difficult to make informed decisions about the relative value of different interventions,	Thank you for your comment. We hope that the review will promote improvements in the rigor and reporting of ASD research.
Peer reviewer #2	General	Global comment: Given that there are now N=37 RCTs to be reviewed I do not see the need to retain the nonrandomized comparative studies. This I am guessing is an EHC wide methods issue but given the number of trials conducted and the much lower weight one gives essentially naturalistic observational this makes me for one less confident of the report's conclusions than I would be had they been removed.	While we agree that RCTs, when well conducted, are best for answering questions of causal inference, especially in a relatively small body of literature, observational studies can provide important context and pragmatic information about effectiveness. In the original review, most of the studies used a non-comparative design, so we consider the addition of comparative studies as substantially increasing our ability to draw conclusions about the body of evidence. You are correct that observational studies carry relatively less weight in assessing the strength of the evidence overall, but they do provide some information that was helpful, including beginning to look at moderators of effect.

Commentator & Affiliation	Section	Comment	Response
Peer reviewer #2	General	I do wonder if the shorthand use of 'ABA' or 'behavioral' approaches will make some readers think of the more 'traditional' Lovaas style ABA and not ESDM (which the authors always describe as behavioral/developmental or behavioral-relationship based) or LEAP which are somewhat different.	We understand that certain terms, although extremely accurate, often have less accurate connotations for lay use/understanding. In this capacity, our utilization of the heading and term "Early Intensive Developmental and Behavioral Intervention (EIDBI)" is meant to represent an inclusive grouping of ABA-based interventions that are delivered from several different contexts (i.e., along the developmental to behavioral spectrum). Rather that utilize a new and likely substantially confusing acronym (EIDBI) we chose to refer to these as 'ABA-based.' In order to further minimize confusion, we refer to traditional "UCLA/Lovaas" variants specifically as a subheading within this category, along with other variants such as the Early Start Denver Model.
Peer reviewer #2	General	I enjoyed reading this very competent and useful update and will look forward to seeing it in print.	Thank you for your comment.
TEP Reviewer #1	General	Given the recent increase in the quality and quantity of studies on behavioral interventions for children with autism spectrum disorder (ASD), this review is timely and important. Perhaps benefiting from experience with the previous review and the richer literature to evaluate, the update has a more logical organization, as well as clearer criteria for evaluating individual studies and rating the overall strength of evidence for each category of interventions.	Thank you for your comment.
TEP Reviewer #1	General	As before, the questions that the review addresses are stated concisely and are clinically relevant. Procedures for conducting the literature search, including or excluding reports, and extracting data are described well. Methods and findings from individual studies appear accurate. The conclusions are consistent with the evidence.	Thank you for your comment.
TEP Reviewer #2	General	The report is clinically meaningful; the authors explicitly described the characteristics of the target population for each question and the literature supporting or refuting effectiveness. The key questions are appropriate and explicitly stated.	Thank you for your comment.
TEP Reviewer #3	General	General Comments: The report was clear and meaningful.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
TEP Reviewer #3	General	ES-5 has a nice explanation of terms for no-researchers, although ES-6 gets a bit "wonky" at the bottom.	We have attempted to revise the executive summary to present clear, actionable information for clinicians, policy makers, and other uses of the report.
TEP Reviewer #3	General	Also terms like "manualized" are not part of the comment vernacular and might not be easy to understand for many readers. My suggestion is to write the Exec Sum so that many different readers can absorb the info given. By this I mean that the ES in particular will be read by policy makers, advocates, parents, etc. I'm not asking to dumb it down; I'm thinking more of plain English or at least a few less terms of art/research terms so that the information can be accurately understood and applied. For the full document (page 1-89) I would not change the tone. These pages will be read most commonly by folks who are familiar with the field and should have the specificity that comes with terms of art and precise language.	We have worked on clarifying the Executive Summary.
Peer Reviewer #3	General	General Comments: This is a well done systematic review and the report is important to practitioners, researchers, and policy makers working in the area of ASD.	Thank you for your comment.
Peer Reviewer #3	General	The target audience is not explicitly defined anywhere in the report that I could find. This could be added in the scope (ES-1)/introduction sections.	The Uses of This Report section in the Introduction contains information about audience. We have also added a brief summary to the Executive Summary.
Peer Reviewer #3	General	Key questions seem appropriate for this type of review and are well explicated. However, they are somewhat redundant since the information used to answer some questions is later used to answer parts of subsequent questions (e.g., Kq1, Kq2, Kq7). Thus, the report is well organized but lengthy and cumbersome.	We follow the organizational structure of EPC reports and although we agree that the report is long, we hope that the executive summary and subsequent manuscripts will provide readers with shorter and more digestible documents while still allowing readers the opportunity to read the detail.
Public Reviewer #3 (Magellan)	General	As reflected in this paragraph from the conclusion of the draft report, there exists a range of intervention services and a continuous need to determine what works best for specific children with ASD given their individualized needs and unique circumstances. While the evidence points to intensive early intervention being the best course of treatment, there are many mediating factors that limit our ability to fully understand the lasting effects of these services. Some considerations include the following:	Thank you for your comment. We agree that the evidence suggests that different treatment approaches will work for different children.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #3 (Magellan)	General	**Who exactly is providing the service?** While the Behavior Analyst Certification Board (BACB) has made strides in creating a better regulated work force, treatment providers are often a varied group of individuals assessing and treating individuals with ASD. As such, it can be difficult to assess the level of qualifications and training of frontline staff.	We agree that provider is an important factor and included information about providers in our evidence tables. We did not find studies addressing provider qualifications as mediators. We have noted the lack of information on providers in the Limitations of the Evidence Base and Research Gaps and Needs sections of the report.
Public Reviewer #3 (Magellan)	General	**Multiple methodologies.** Within a given treatment plan, multiple methodologies may be implemented into a program (e.g., Floortime Therapy, Discrete Trial Training, Visual Strategies, and Pivotal Response Treatment). Some of these methodologies may be more beneficial to the overall success of the child than others. However, it is not always clear which is having the greatest impact on the individual.	We agree and have noted this as a weakness of the literature base.
Public Reviewer #3 (Magellan)	General	**Family engagement.** The overall involvement of the family is a key determining factor in how well a child progresses, but can be problematic in gaining relevant data. A parent that is highly involved, but not following the treatment plan, can often have a more negative effect than a parent that is not highly involved in the program. Also, socioeconomic factors and cultural nuances can have an impact in the level of involvement a parent is able or willing to contribute to his/her child's treatment.	We agree and note that there is a lack of information in the published literature on the role of the family outside the specific parent training studies.
Public Reviewer #3 (Magellan)	General	**Other services.** A child may be receiving other types of treatment that can impact overall development and progress. This treatment may be in the form of services in an Individual Education Program, Speech Therapy, Occupational Therapy, medications, and other biomedical interventions, such as diets, supplements, and others. Once again, this can create difficulty in determining what is having the most effect.	We agree and that is why we assessed the degree to which concomitant interventions were held consistent in the quality assessment process.
Public Reviewer #3 (Magellan)	General	**Study setting.** Studies can also present challenges when conducted in the home community setting. If a study is conducted in a clinical setting, it may be able to limit many of the above mediating factors when collecting data. However, it also limits its ability to translate into the "real world" when the same therapy is eventually provided outside of the clinical setting. Additionally, there is limited data collected on the importance and impact of strong parent training programs.	As noted, we assessed generalizability/applicability for this reason and considered the setting of interventions in that assessment.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #3 (Magellan)	General	We highlight the following factors as potential focal points for improving quality of care for individuals with ASD.	Thank you for these comments. They appear to be related to your organization's processes and do not require changes to the report.
Public Reviewer #3 (Magellan)	General	**Paraprofessional certification.** Requiring providers to certify their paraprofessional workforce will allow for better quality control, particularly given that such workers often work in the homes and communities of individual consumers. The new guidelines of the BACB for the Registered Behavior Technician creates a universal training and certification process that allows for better trained staff to be providing services. Additionally, this process may reduce staff turnover and allow for more consistent and beneficial services for the clients, theoretically facilitating more effective care and potentially shortening the time spent in treatment.	Thank you for these comments. They appear to be related to your organization's processes and do not require changes to the report.
Public Reviewer #3 (Magellan)	General	**BACB guidelines.** It is important for organizations to ensure that their providers are strictly following the guidelines of the BACB when they are providing services to individuals. The use of comprehensive reporting tools will help to ensure that only the most critical and socially significant issues are being addressed in programming. Access to treatment plans will ensure that clients are getting the services that were intended for them based on the authorization provided.	Thank you for these comments. They appear to be related to your organization's processes and do not require changes to the report.
Public Reviewer #3 (Magellan)	General	**Audits.** With so many providers providing varying levels of service and incorporating diverse methodologies, it is imperative that providers be held to appropriate standards of care in the delivery of treatment interventions for individuals with ASD.	Thank you for these comments. They appear to be related to your organization's processes and do not require changes to the report.
Public Reviewer #3 (Magellan)	General	In conclusion, treatment for ASD is continually evolving – and that is appropriate given increasing experience, available technologies, and our expanding knowledge base. ASD is such a varying condition, that it makes studying the treatment process and results a challenge. While one child may be able to make gains to live and function appropriately in everyday life, others may end up in residential care for their lifespan. Each of these children may have started in similar places and had similar treatment plans and teams, yet ended up in very different places	Thank you for these comments. They appear to be related to your organization's processes and do not require changes to the report.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #3 (Magellan)	General	Given this fluidity and complexity and in the best interests of individuals with ASD, we believe that treatment providers at every level should be held to appropriate standards that ensure that individuals with ASD receive the most effective interventions for their needs.	Thank you for these comments. They appear to be related to your organization's processes and do not require changes to the report.
Public Reviewer #4 (Association for Behavioral Health and Wellness)	General	The following are a set of research questions on behavior therapies for Autism Spectrum Disorders (ASD) compiled by ABHW and its member companies. The questions are focused on what we collectively see as the current gaps in knowledge; answers to these questions are important to assure effective delivery of services in the provider community	Thank you for these suggestions. We agree that there are a number of gaps existing in the research and have noted this in the research gaps section of the report.
Public Reviewer #4 (Association for Behavioral Health and Wellness)	General	What treatments for ASD are likely to be effective for what subpopulations (i.e. age, gender, IQ)? Are there subpopulations with differential responses to interventions?	We note in the Research Gaps section of the review that "a critical area for further research is understanding which children are likely to benefit from particular interventions"
Public Reviewer #4 (Association for Behavioral Health and Wellness)	General	What plans are there for definitive, transparent and broad communication around knowledge of safety concerns (particular around unsafe therapies like chelation and secretin therapy)?	While we did not review non-behavioral therapies such as secretin in the current update, we note that our prior comprehensive review (published 2011) does discuss harms of such treatments. We note in the current review that the behavioral studies addressed did not report harms. The broader communication of these issues comes through publication of the report on the AHRQ website, publication of associated manuscripts as well as development of materials for patients, providers and policy makers. It should be noted, however, that these materials are based on what is in the reviewed in the report only.
Public Reviewer #4 (Association for Behavioral Health and Wellness)	General	Are there any identifiable changes early in the treatment phase that predict treatment outcomes	We note in the Research Gaps section of the review that "our understanding of early indicators of treatment response are extremely limited, such that it is not realistic to implement evidence-based changes in intervention based on assessing children's responses"
Public Reviewer #4 (Association for Behavioral Health and Wellness)	General	What is the evidence that effects measured at the end of the treatment phase predict long-term functional outcomes?	We note in the Research Gaps section of the report that little data are available to address this question.
Public Reviewer #4 (Association for Behavioral Health and Wellness)	General	What evidence supports specific components of treatment as driving outcomes, either within a single treatment or across treatments?	We note that in the Research Gaps section of the review "little data on whether specific treatment components drive effectiveness exists"

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #4 (Association for Behavioral Health and Wellness)	General	What are the modifiers of outcome for different treatments or approaches?	We note in the Research Gaps section of the review that "To date, studies have failed to characterize adequately the characteristics of interventions (or the children receiving them) in a manner that helps clarify why certain children show more positive responses than others. It is simpler to identify the characteristics of those children who show at most a minimal benefit from a particular treatment, but most existing studies also fail to adequately describe this population."
Public Reviewer #4 (Association for Behavioral Health and Wellness)	General	Intensive behavioral therapy appears to work for some children; will we ever know with reasonable certainty if treatment approaches should be viewed as a "one size fits all" model?	We note in the Research Gaps section of the review that "To date, studies have failed to characterize adequately the characteristics of interventions (or the children receiving them) in a manner that helps clarify why certain children show more positive responses than others. It is simpler to identify the characteristics of those children who show at most a minimal benefit from a particular treatment, but most existing studies also fail to adequately describe this population."
Public Reviewer #4 (Association for Behavioral Health and Wellness)	General	In general, is there a threshold IQ for which Applied Behavioral Analysis (ABA) is effective? What is the expected trajectory of change for ABA? How do you define the outcome?	We note in the Results section of the report that findings related to the effects of IQ on treatment outcomes are limited. We have noted the need for more research on understanding modifiers and predictors of outcome in the Research Gaps section of the report. We have similarly noted the need to standardize outcomes.
Public Reviewer #4 (Association for Behavioral Health and Wellness)	General	Is there an ideal age at which to begin ABA?	We note in the Results section of the report that findings related to the effects of age on treatment outcomes are limited. We have noted the need for more research on understanding modifiers and predictors of outcome in the Research Gaps section of the report. We have similarly noted the need to standardize outcomes.
Public Reviewer #4 (Association for Behavioral Health and Wellness)	General	What is the effect of ABA and how long will the gains last? What is needed to sustain the gains?	We note that our ability to isolate specific "effective ingredients" of typically multi-component ASD interventions is limited in the Research Gaps section of the report. We also comment specifically on the need to understand how to maintain gains.
Public Reviewer #4 (Association for Behavioral Health and Wellness)	General	In general, is there a threshold IQ for which the Denver Model (and any other evidence based intensive behavioral therapies) is effective? What is the expected trajectory of change for the Denver Model (and any other evidence based intensive behavioral therapies)? How do you define the outcome?	We note in the Results section of the report that findings related to the effects of IQ on treatment outcomes are limited. We have noted the need for more research on understanding modifiers and predictors of outcome in the Research Gaps section of the report. We have similarly noted the need to standardize outcomes.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #4 (Association for Behavioral Health and Wellness)	General	Is there an ideal age at which to begin the Denver Model (and any other evidence based intensive behavioral therapies)?	We note in the Results section of the report that findings related to the effects of age on treatment outcomes are limited. We have noted the need for more research on understanding modifiers and predictors of outcome in the Research Gaps section of the report. We have similarly noted the need to standardize outcomes.
Public Reviewer #4 (Association for Behavioral Health and Wellness)	General	What is the effect of Denver Model (and any other evidence based intensive behavioral therapies) and how long will the gains last? What is needed to sustain the gains?	We outline findings of studies addressing ESDM and other behavioral interventions in the Results section of the report and in the strength of the evidence tables. We note in the Research Gaps section of the report that more research is needed to understand the durability of gains and approaches to maintaining them.
Public Reviewer #4 (Association for Behavioral Health and Wellness)	General	For what characteristics is ABA, the Denver Model, and other evidence based treatments not effective?	We note that "To date, studies have failed to characterize adequately the characteristics of interventions (or the children receiving them) in a manner that helps clarify why certain children show more positive responses than others. It is simpler to identify the characteristics of those children who show at most a minimal benefit from a particular treatment, but most existing studies also fail to adequately describe this population."
Public Reviewer #4 (Association for Behavioral Health and Wellness)	General	Is there any evidence or tool that helps define the level of intensity of treatment needed?	We note that "To date, studies have failed to characterize adequately the characteristics of interventions (or the children receiving them) in a manner that helps clarify why certain children show more positive responses than others. It is simpler to identify the characteristics of those children who show at most a minimal benefit from a particular treatment, but most existing studies also fail to adequately describe this population."
Public Reviewer #4 (Association for Behavioral Health and Wellness)	General	Are there tools that you recommend to measure the effectiveness of treatment?	We discuss issues with outcome measurement in the Limitations and Research Gaps sections and note that "Research on appropriate methods for capturing meaningful change will be critical to advancing our understanding of behavioral interventions. In addition, although more studies are reporting primary and secondary outcome measures determined *a priori*, continued improvements in reporting will benefit the field.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #4 (Association for Behavioral Health and Wellness)	General	Other Research Needs Given the status of autism behavioral treatment research and the barriers to research in this population, what other research types, other than randomized control trials and matched group cohort studies, would AHRQ consider valid to test and demonstrate treatment effectiveness (e.g. single case studies, large multi-site observational studies)?	Large, well-conducted cohort studies with comparisons can be very useful in establishing effectiveness of interventions. To date, the EPC program has not used single case studies or single subject research designs to assess effectiveness at the population level.
Public Reviewer #2 (Association of Professional Behavioral Analysts)	General	They also make some astute observations about some of the methodological and other limitations of the reviewed studies, and perceived gaps in knowledge regarding behavioral interventions for ASD. We would like to see that section of the report expanded to acknowledge that some of those gaps are apparent rather than real, and are artifacts of constraints that resulted in the reviewers excluding most of the large body of behavior analytic research on interventions for ASD.	We have not expanded this section as recommended because we feel that the gaps noted in our review, as they pertain to establishing effectiveness data, are accurate and "real."
Public Reviewer #2 (Association of Professional Behavioral Analysts)	General	Because of that serious omission, the report conveys the dangerously inaccurate message that there are no efficacious interventions for building skills and reducing behaviors that affect the safety, health, and overall functioning in people with ASD who are over the age of 7 or have substantial delays in cognitive and language skills.	We disagree with this assessment of the conclusions of the report. As noted, this report is an update on one section of a prior report. Please see the 2011 report for a complete assessment of a broader range of treatments.
Public Reviewer #2 (Association of Professional Behavioral Analysts)	General	Even more than such an acknowledgement, we would like to see the report revised to include a thorough review of all of the relevant research, as well as recommendations for alternative research methods to address some of the needs the authors point out.	As noted in the report, we reviewed all relevant research for establishing comparative effectiveness in the 2011 report. This report is an update of a component of that report. We are confident that we have included all of the relevant research to answer our research questions.
Public Reviewer #2 (Association of Professional Behavioral Analysts)	General	Some specific suggested remedies for some of the limitations noted in the report are offered next.	Thank you.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #2 (Association of Professional Behavioral Analysts)	General	Expand the Review Scope and Process The Recommendation on p. ES-16 that readers take into account evidence from studies using single-case research designs (SCRDs) is well-founded, as the large majority of studies on applied behavior analysis (ABA) interventions for ASD employed such designs. We respectfully disagree, however, with the authors' assertions that that research addresses only "focused questions of short-term efficacy in individual children" and lacks evidence of "generalizable effects." It is true that because behavior occurs only at the level of the individual, most behavior analytic studies focus on individual behavior(s); however, many ABA studies involve more than one participant and more than one target behavior. Behavior is measured directly and repeatedly under baseline or control conditions (no treatment or treatment as usual), and with a treatment in effect (the experimental condition). Treatment procedures comprise environmental events that are arranged to precede and/or follow occurrences of the behavior close in time. Baseline and treatment phases are repeated with the same individual and/or other participants. The phases can be arranged in a variety of ways to produce designs for addressing a wide range of research and clinical questions. Graphed data are analyzed to determine if a treatment produced clinically meaningful improvement in comparison to baseline or another treatment procedure, and in many studies, whether the behavior change generalized across settings, interventionists, and time without the treatment in place. In sum ABA studies are controlled clinical trials (CCTs) in which each participant experiences control (baseline) and treatment conditions, and comparisons of those conditions are replicated	We recognize that setting a minimum of 10 participants for studies to be included effectively excluded much of the literature on behavioral interventions using single-subject designs. Because there is no separate comparison group in these studies they would be considered case reports (if only one child included) or case series (multiple children) under the rubric of the EPC study designs. Case reports and case series can have rigorous evaluation of pre- and post- measures, as well as strong characterization of the study participants, and case series that included at least 10 children were included in the review. Single-subject design studies can be helpful in assessing response to treatment in very short timeframes and under very tightly controlled circumstances, but they typically do not provide information on longer term or functional outcomes, nor are they ideal for external validity without multiple replications. They are useful in serving as demonstration projects, yielding initial evidence that an intervention merits further study, and, in the clinical environment, they can be useful in identifying whether a particular approach to treatment is likely to be helpful for a specific child. Our goal was to identify and review the best evidence for assessing the efficacy and population-level effectiveness of behavioral interventions for children with ASD, with an eye toward utility in the treatment setting. With the assistance of our technical experts, we selected a minimum sample size of 10 in order to maximize our ability to describe the state of the current literature, while balancing the need to identify studies that could be used to assess treatment effectiveness. These methodologic decisions were reported in our review protocol.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #2 (Association of Professional Behavioral Analysts)	General	Thousands of published ABA CCTs have demonstrated the efficacy of many ABA procedures – singly and in various combinations -- for building skills and reducing problem behaviors in people with ASD and other populations, in a wide range of settings. Additionally, methods have been developed for calculating statistical significance and effect sizes by comparing data from baseline and treatment phases aggregated across large numbers of participants in multiple ABA CCTs. Numerous published meta-analyses and other syntheses of data from large numbers of ABA CCTs show that many ABA interventions are very effective for changing the core symptoms of ASD and reducing behaviors that left untreated, often result in extensive use of costly healthcare and other services. Examples are eating problems, sleep problems, elopement (wandering), pica (ingesting inedible items), behaviors associated with anxiety and fear, aggression, and self-injurious behaviors. That is, the efficacy and generality of many ABA interventions for ASD have been demonstrated empirically and directly through multiple replications. We have compiled a bibliography of systematic reviews and meta-analyses of aggregated ABA studies that we will be happy to share with the report author	Please see the inclusion/exclusion criteria for this report. It was not within the scope of this report for the Effective Healthcare Program of AHRQ to review single-subject design studies. Meta-analyses of these studies are available to users of the report and may provide additional contextual information for decision makers.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #2 (Association of Professional Behavioral Analysts)	General	Several national organizations recognize the scientific validity of evidence from behavior analytic studies and include studies using SCRDs in their protocols for evaluating evidence about treatments. They include Divisions 12 and 16 of the American Psychological Association, the National Association of School Psychologists, and the U.S. Department of Education What Works Clearinghouse. The National Autism Center National Standards Project, the New Zealand Ministries of Health and Education, and the New York State Department of Health Early Intervention Program have used variations of those protocols to evaluate evidence about interventions for ASD. The foregoing groups recognize that behavior analytic research methods yield rich, precise information about treatment procedures and individual responses to treatment that cannot be derived from most studies using between groups research designs with statistical analyses of group averages and other mathematical abstractions. We urge the report authors to lobby the AHRQ to expand its review protocol to incorporate the full range of research methods that can produce credible evidence about treatment effects, including behavior analytic research methods.	Please see our responses regarding decisions about which studies to include or exclude and the context of this report.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #2 (Association of Professional Behavioral Analysts)	General	Differentiate Among "Behavioral" Interventions We agree wholeheartedly with the authors' candid statement that "...most of the body of literature categorized in this report as 'early intensive behavioral and developmental intervention' remains an eclectic grouping. This category of intervention presently groups different treatment approaches..." (p.ES-17). There are substantial differences between thorough going behavior analytic interventions and interventions derived from other conceptual frameworks that are sometimes described as "based on ABA" or incorporating some ABA procedures. We suspect that lumping several very disparate interventions together under the "behavioral" rubric contributed directly to many of the problems the authors had in drawing broad conclusions, comparing studies, identifying moderators of intervention effects, and teasing out features of effective interventions. That practice, though common, also contributes to widespread confusion among consumers, funders, and policymakers.	This comment has also been addressed in regard to previous reviewer concerns regarding categorization of interventions. As noted in the review we actively involved our stakeholder groups in order to determine the most appropriate methods for enhancing knowledge of complex intervention terms that consumers, funders, and policy makers. We understand that certain terms, although extremely accurate, often have less useful connotations for lay use/understanding. In this capacity, our utilization of the heading and term "Early Intensive Developmental and Behavioral Intervention (EIDBI)" is meant to represent an inclusive grouping of ABA-based interventions that are delivered from several different contexts (i.e., along the developmental to behavioral spectrum). Rather that utilize a new and likely substantially confusing acronym (EIDBI) we chose to refer to these as 'ABA-based.' In order to further minimize confusion, we refer to traditional "UCLA/Lovaas" variants specifically as a subheading within this category, along with other variants such as the Early Start Denver Model. We have attempted to transparently acknowledge the limits of categorization throughout and note explicit important variations in methodology wherever possible.

Commentator & Affiliation	Section	Comment	Response
Public Reviewer #2 (Association of Professional Behavioral Analysts)	General	A clearer picture could be obtained by analyzing studies of distinct intervention approaches separately. The characteristics of bona fide ABA interventions have been well-defined since 1968 (see Cooper, Heron, & Heward, 2007, pp. 16-18). Behavior analysts who oversee and study the early, intensive, comprehensive model of ABA intervention for ASD generally agree on the defining features of that model (not just the Lovaas version; see Eldevik et al., 2010, p. 384). Knowledgeable professional behavior analysts could therefore identify studies in which the intervention clearly had the defining characteristics of ABA, and could assist reviewers in evaluating those studies. It may be possible to glean the defining features of developmental, play-based, and mixed intervention models from articles, manuals, and the developers of those models so that the research on each of them can also be analyzed and described separately. A revised report that clearly distinguished among the "behavioral" interventions would be of value to many users.	We agree that it would be ideal to be able to identify key components and drivers of interventions that could be dissected from the current intervention literature in order to better understand the individualized impact of specific paradigms. Unfortunately, the current literature base did not permit such dissection in almost all cases. We also agree with the reviewer that it was important to involve experts and knowledgeable professionals in ABA as part of this process. Please note the specific experts involved in our TEP and invited reviewers of this and previous review were explicitly involved for this reason.
Peer Reviewer #4	General	General Comments: My responses are based on the usefulness of the information in the report to the implementation of therapies for children with ASD within a health care delivery system. As a system administrator, we are responsible for the delivery of medically necessary services to an insured population. Medically necessary services are defined as evidenced based services that are based on a hierarchy of credible research, (from RTCs and match co -hort studies) to generally accepted clinical practice standards established by credible subspecialty organization. The AHRQ report "Therapies for Children with ASD" serves as an important source of an independent review and ranking of the evidence that is relevant to the creation of medical policy and implementation of the service in a community network.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #4	General	With this in mind, there are many components of the report that provide the needed information to system administrators and policy makers. This includes: 1. Establishing a list of questions that are relevant to information needed for service implementation 2. Organizing the research by treatment domains and supplying quick reference information from each study reviewed for individual use 3. Ranking of the strength of the evidence to understand where there are definitive conclusion that can be drawn and where gaps still exist. 4. View to the scope expected outcomes domains, measurements and tools utilized to measure response to the interventions.	Thank you for your comment.
Peer Reviewer #4	General	Would suggest the report expand conclusion section to distinctly outline the important gaps in the research and particle research structures and methodologies that might be more efficient mechanism for obtaining such information, like case registries. Clarity around the gaps in research helps funding bodies and consumer advocate groups to facilitate a prioritize agenda of research.	We have attempted to identify and suggest important methodologies for answering unanswered questions in the report. Advocacy for specific case registries falls outside the scope of the report, but we have referenced this approach as potentially meaningful
Peer Reviewer #4	General	I would suggest gathering additional info about the necessary training of professionals in order to implement the interventions.	We agree that this is important information but outside the scope of the review.
TEP Reviewer #4	General	Article selection (page 15): States that 7 studies were not included as they were follow-ups to studies evaluated in the 2011 review. However, since of the question is long term efficacy and maintenance, it's unclear why these studies, which provided such information, were excluded. I'm sure there was a good reason, but it would be helpful to understand.	We have revised the wording of this section to clarify that 7 studies included in the current review report followup data from studies in the prior review.
TEP Reviewer #4	General	Top of Page 17. The first sentence says that "five of six RTCs identified in the literature measured anxiety symptoms." This is a very confusing statement, as it suggests that of the 51 new studies, only six were RTCs. However, this isn't true, since it is later stated that 37 of the 51 new studies were RTCs. So this needs to be corrected.	This statement referred to studies in the section of the report dealing with intervention addressing comorbid conditions, not all studies in the review. We have clarified the statement to note "Six RCTs (five good and one fair quality) of interventions addressing conditions commonly associated with ASD identified for the current update measured anxiety symptoms as a primary outcome."

Commentator & Affiliation	Section	Comment	Response
TEP Reviewer #4	General	I also suggest that the preceding paragraphs summarizing the research in other areas (e.g., early intervention, play, etc.,) include some sense of the quality of this research (e.g., how many of the studies were RTCs vs uncontrolled trials). The fact that we now that the anxiety studies included 5 RTCs is useful information and this type of information would be equally useful in the other sections.	We have added information about the number of RCTs and the quality to the executive summary paragraphs noted.
TEP Reviewer #4	General	Study Quality: Not clear how studies are determined to be "good," "Fair" "poor" quality, etc. Might be helpful to describe this at some point. P. 17. I noted that children with lesser impairments tended to do better in this study. This is an usual finding as regression toward the mean often finds that those with the most severe symptoms tend to do better.	The methods section of the full report and the quality appendix provide further details on the quality scoring methods. Word limits in the executive summary prohibit full description there but we have added a statement directing readers to the full report.
TEP Reviewer #4	General	Page 18. first sentence (repeated the word "often" twice)("often most often")	Corrected, thanks.
TEP Reviewer #4	General	KQ5: "Involved children WHO were receiving (missing the word, "who")	Corrected, thanks.
TEP Reviewer #4	General	K7: Suggested that there was a problem that no studies compared treatment to a no-treatment control group. I would think that might be considered unethical in the case of an early intervention study. It may not be a problem for a CBT study. But I would argue that comparing treatment to no treatment strongly biases toward finding positive effects. Comparing a treatment with TAU is a more valid study. So I don't know that this should be considered as a problem.	Thank you for your comment.
TEP Reviewer #4	General	Table B: It's not clear to the reader what is meant by Domain Ratings, Issues and Findings. None of the categories assessed have been defined and it's not clear how one takes a group of studies and assigns an overall rating. Similarly, it's not clear what makes a "fair" vs a "good" study design. I'm assuming that this is discussed elsewhere.	The methods for arriving at our assessments are described in the methods section of the full report.
TEP Reviewer #4	General	Page 24. Building Blocks Program description: It's not clear to the reader what the focus of the parent training was. For example, if the focus was on teaching language, then it would be concerning that no group differences were noted in language measures. I'd suggest a sentence that describes the primary focus.	We've noted that the intervention targeted social and communication skill development.

Commentator & Affiliation	Section	Comment	Response
TEP Reviewer #4	General	Page 27: ESDM description: As above, need to provide a sentence that describes the kinds of skills or behaviors that ESDM focuses upon (even though many in the field are familiar with ESDM).	We've noted that this intervention targets communication and general child development.
TEP Reviewer #4	General	Description of studies. These are all fabulous, detailed descriptions. The only problem I have is that within each section, it's not easily organized for the reader. By this, I mean that if the reader wanted to see the summary on a particular CBT study, she'd have to scan over 3-4 pages of dense writing to find the study. Would it make sense to have some subheadings or a label for each study so that it can easily be located?	We have attempted to streamline the organization while following AHRQ formatting specifications.
Peer reviewer #1	Introduction	b. Introduction: This provides a good background to the review. It summarises the findings of the authors' earlier review on this topic and the rationale for an update (i.e. the growth of publication in this area in the subsequent 3 years). The background makes clear from the outset that the review covers mainly behaviourally based interventions This may give rise to some criticism from those whose primary interest is in other types of intervention, but the authors' justification for the decision to focus the review in this way is clear, and appropriate.	As noted, the review was focused on behavioral interventions. Additional studies of behavioral interventions have the greatest potential to alter the low and insufficient strength of evidence reported in the original review and potentially affect treatment recommendations due to the number of new studies available.
TEP Reviewer #1	Introduction	Overall, the Introduction is well-written. The division of interventions into "comprehensive" and "focused," with several categories of focused interventions, is a substantial improvement over the classification system in the previous report.	Thank you for your comment.
TEP Reviewer #2	Introduction	Well done	Thank you for your comment.
TEP Reviewer #3	Introduction	Nice, brief summary. Well done.	Thank you

Commentator & Affiliation	Section	Comment	Response
TEP Reviewer #3	Introduction	Since common definitions have been a barrier to research, and the subsequent use of good research in clinical practice, would you consider expanding the "Categorization of Interventions" section with definitions for the categories of research not included, at least as you used these categories? Specifically, I think the reader may ask what are behavioral vs. psychosocial interventions; how do these interventions differ? How do behavioral interventions differ from educational interventions? Is it site of services? The outcome that the intervention is targeting? One of the outcomes mentioned in the Methods section, page 8, lists "academic skill development" and "academic engagement/attainment" as outcomes. Understanding that this may be muddy and quite challenging, I think it's important to include what behavioral interventions are NOT, especially the separation between behavioral and educations; this is hugely for policy makers, in both health care and education, and payers. I cannot underscore this enough.	As noted in the previous and this updated review a major challenge has been the categorization of interventions. Unfortunately, there is no one single accepted methodology for defining and classifying interventions for children with autism spectrum disorders. To a large extent the categorization of interventions was driven from a consumer perspective (i.e., how parents, clinicians, and systems of care might attempt to access or support intervention decisions). In this capacity categorization was not made in based on outcomes studies or examined, but rather treatment setting/context. Medical interventions, complementary and alternative interventions, and interventions primarily delivered by allied health interventions were NOT included in the behavioral review. We consciously chose the term behavioral to encompass a broad array of interventions that could be delivered from this perspective (e.g., ABA-based, CBT, Social Skills, Imitation/Joint Attention). This included behavioral interventions potentially delivered within educational systems, given that the divide between the two is not often clear, particularly at young ages. Further details of the categorization have been added to the report.
Peer Reviewer #3	Introduction	Introduction: The introduction/executive summary is succinct but could benefit from a statement of purpose and target audience. (p. 10, ES-1)	The Uses of This Report section in the Introduction contains information about audience. We have also added a brief summary to the Executive Summary.
Peer Reviewer #3	Introduction	Another issue is that the DSM-V criteria are used to define ASD in the intro section, but the studies reviewed all utilized DSM-IV for inclusion/exclusion. These changes are relevant to the key questions focusing on "commonly associated symptoms" (e.g., Kq1d) as defined by DSM-IV criteria, which are now considered "core" symptoms in DSM-V , not associated symptoms (e.g., sensory hypo/hyper reactivity). It did not appear that any studies included outcome measures specific to sensory features so perhaps this is less of an issue, but needs clarification and consistency none-the-less.	We have clarified this important point within the context of the review. It was our explicit intent to be able to capture important functional issues that have been/are understood differently across DSM-IV / DSM-5. This is, as noted by the reviewer, most relevant to sensory hypo/hyper reactivity. In our previous review we wanted to ensure this symptom domain was not neglected, and included it under the rubric of associated symptoms. In the current report (as DSM-5 now includes this a core symptom) this now represents a core symptom in the context of the review. Language has been added to clarify this so others will also understand that both reviews addressed this very important area of outcome.

Commentator & Affiliation	Section	Comment	Response
Public reviewer #1 (Richard G Kensinger, MSW)	Introduction	First objection in ES-1 is use of the term chronic" management, which as a clinician I consistently object to. I prefer terms such as persisting, episodic, life-long etc.	We have changed chronic to lifelong.
Public reviewer #1 (Richard G Kensinger, MSW)	Introduction	Second, is the lack of interfacing cost of care which I consider to be the clinical course vs. the cost of not treating it, the natural course. I know that my commercial health insurance pays nothing in this DRG category	We do not understand the comment
Peer Reviewer #4	Introduction	Introduction: The introduction was very clearly written. Given the complexity of this field it was helpful for the report to organize of the studies by the targeted outcomes.	Thank you for your comment.
Peer Reviewer #4	Introduction	The list of questions KQ1 – KQ6, (page 11) are relevant to gathering the necessary information for system administrators and policy makers. Questions KQ2c and KQ2d are exceptionally important to understanding the existence of and identification of subpopulations that respond differently to interventions.	Thank you for your comment.
Peer Reviewer #4	Introduction	Wondered whether AHRQ might consider formulating further research recommendations, such as performing a meta- analysis of existing research data to potentially identify subpopulations that respond differentially across similar interventions.	This is an excellent suggestion, and we have noted such meta-analyses as a research need in the Research Gaps section of the report
Peer reviewer #1	Key Questions	With respect to key questions, these are explicitly stated at the beginning of each section, and data relevant to each of the questions are provided.	Thank you for your comment.
Peer reviewer #1	Methods	c. Methods: The methodology used throughout the review us clearly stated. The literature search strategy; inclusion and exclusion criteria; information on study selection and data extraction are concisely presented.	Thank you for your comment.
Peer reviewer #1	Methods	However, I was not particularly convinced that the "analytic framework" (Figure A, p 12) "represents the process by which families of children with ASD make and modify treatment choices". It is unclear on what basis the authors make this claim about families' choices, although the framework provides a good template for the review itself.	We have revised this wording to clarify that the analytic framework illustrates the placement of the review's key questions within the context of treatment choice, potential outcomes, and characteristics that may affect outcomes.
TEP Reviewer #1	Methods	The description of criteria for quality assessment of individual studies and quality level for the evidence as a whole for an intervention are much clearer than before. The decision to exclude uncontrolled case series (included in the previous report) is reasonable but probably warrants a brief explanation.	Thank you for your comment. We have noted the limitations of case series in the Methods section.

Commentator & Affiliation	Section	Comment	Response
TEP Reviewer #2	Methods	Methods well described and excruciating detail provided for inclusion/ exclusion criteria. Outcome measures appropriate.	Thank you for your comment.
TEP Reviewer #3	Methods	Inc/exclusion criteria justifiable: yes, although studies with only 10 participants is a low bar, given the interest in treatments for ASD, I concur that a low bar/being more inclusive without sacrificing scientific integrity is justified	Thank you for your comment.
TEP Reviewer #3	Methods	. Search clear and logical: yes.	Thank you for your comment.
TEP Reviewer #3	Methods	Definitions for outcomes: Outcomes are discussed on page 10. I see a list of outcomes, but there are no definitions for these outcomes measures. Perhaps none are needed, but I don't say this to negate my comments in the Intro section (different definitions needed for Intro).	We extracted specific outcomes as reported in each study under the broad headings of symptom severity, cognitive skills, etc. Given the heterogeneity of individual outcomes reported in studies of ASD interventions, it is not feasible to pre-specify each individual outcome measure under these broad headings
TEP Reviewer #3	Methods	Statistical methods seem appropriate.	Thank you for your comment.
TEP Reviewer #3	Methods	I am looking right past a comment on the validity of the tools being used by each study to measure the outcomes of interest? Are all of the included studies using valid/reliable tools?	We incorporated explicit assessments of outcome measurement into our quality scoring systems. The use of valid/reliable tools was necessary for achieving higher levels of quality rating.
Peer Reviewer #3	Methods	Methods: Generally, the review was systematic and done with integrity, but there needs to be more detail provided on the inclusion and exclusion criteria.	The inclusion and exclusion criteria are described fully in the methods section of the full report.
Peer Reviewer #3	Methods	Specifically - study selection is listed in several places but it is never clear how the term "behavioral intervention" or "behavioral modality" is defined. In large part, the authors refer to the earlier review paper (2011) but this is insufficient. In some cases (e.g., p. 33 categorization of interventions such as social skills) behavioral intervention refers to the approach/type of the intervention and in other cases seems to refer to the types of outcomes targeted (e..g, p. 33 ...interventions targeting symptoms commonly associated with ASD such as anxiety) which does not refer to a behavioral intervention method per se.	We appreciate these comments. Definition and categorization of interventions represents a substantial challenge in this field. We utilized behavior intervention as an overarching term explicitly linked to the setting/context of intervention. This use excludes interventions that are primarily medical, complementary and alternative, exclusively focused on non-behavioral educational interventions, and allied health interventions. We have added language to clarify rationale and application.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #3	Methods	On this same topic, were the subcategories of behavioral interventions chosen a priori or did these emerge from the analysis? more details are needed on this decision process.	As noted, categories/subcategories were developed in consultation with our stakeholders in the original report. They were chosen without regard to outcome data, but based on general guiding principles emphasizing treatment context/setting. Again this was done with hopes of aiding the ultimate decision making processes of the stakeholders potentially utilizing the report. We transparently acknowledge the challenges of creating such categories knowing that there would be some studies and approaches were there would not be universal agreement regarding approach and inclusion/exclusion.
Peer Reviewer #3	Methods	It is not clear why some "allied health" and other categories of "medical" interventions were excluded if they used behavioral methods or behavioral measures given the definitions/categorizations above. Perhaps adding more comprehensive definitions (and examples) for categories of interventions that were excluded would be helpful.	We did not specifically exclude studies utilizing behavioral methodologies within varied settings. However, if we considered the primary interventions as linked in ways to a treatment/context setting that was best described by another category of intervention, we felt it most appropriate to ascribe it to that category.
Peer Reviewer #3	Methods	Similarly, the reason for including one study that involved Risperidone and another with Melatonin seems beyond "behavioral" as these are pharmaceuticals. Clarifying again whether the purpose of this review is on behavioral methods, or behavioral outcomes regardless of the method (or both) is needed.	This is another major challenge of categorization. If an approach also included a behavioral arm (as in the Risperidone study) or a primarily psychoeducational component (as in the Melatonin study), we felt there may be value in including a discussion of the results of that arm. In each of these cases the psychoeducation and parent coaching were viewed as active treatments that would primarily be delivered under the broad category of behavioral intervention. We do not disagree with the reviewer's apt critiques that the categorization utilized is not without substantial challenges; however, we have added clarification throughout and transparently acknowledged this limit as part of interpretation.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #3	Methods	Parent training is another general category that overlaps considerably with categories of interventions perhaps many of which were excluded. For example, some studies of coaching parents were excluded yet a study using pamphlets was included.	We included studies with at least 10 individuals with ASD in our target age range. In addition, for this review, studies had to include a treatment and comparison group. If an approach included a behavioral arm or a primarily psychoeducational component, we felt there may be value in including a discussion of the results of that arm. In this case the sleep education pamphlet was viewed as active treatment that would primarily be delivered under the broad category of behavioral intervention. We do not disagree with the reviewer's apt critiques that the categorization utilized is not without substantial challenges; however, we have added clarification throughout and transparently acknowledged this limit as part of interpretation.
Peer Reviewer #3	Methods	To sum up, the main concerns regarding the methods are largely with respect to the definitions of "behavioral" and the inclusion/exclusion criteria. The process by which these parameters were established from the outset need to be transparent. Adding sufficient detail in these decisions/processes and definitions would be very helpful to the reader to be able to judge the relevance of the review and any potential bias.	We included studies with at least 10 individuals with ASD in our target age range. In addition, for this review, studies had to include a treatment and comparison group. If an approach included a behavioral arm or a primarily psychoeducational component, we felt there may be value in including a discussion of the results of that arm. We do not disagree with the reviewer's apt critiques that the categorization utilized is not without substantial challenges; however, we have added clarification throughout and transparently acknowledged this limit as part of interpretation.
Peer Reviewer #3	Methods	Once we get past these inclusion/exclusion issues, the process for rating the studies was thorough and valid.	Thank you for your comment.
Peer Reviewer #3	Methods	On a minor note, it seemed a stretch to include studies of preschoolers in the "early" behavioral interventions. As the age of identification decreases, most clinicians think of early as in the period of early intervention (birth through two years). This was however explained, so at least it's transparent.	Thank you for this comment. What constitutes early can be construed in different ways. We adopted a more inclusive definition, reflective of prevalence numbers indicative of average age of diagnosis of four for defining what may constitute early.
Public reviewer #1 (Richard G Kensinger, MSW)	Methods	Are very solid	Thank you for your comment.
Peer Reviewer #4	Methods	Methods: The inclusion and exclusion criteria were justifiable and relevant.	Thank you for your comment.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer #4	Methods	The exclusion of single case studies continues to make sense until there is a clear way of assessing the generalizability of this type of research.	Thank you for your comment.
Public reviewer #1 (Richard G Kensinger, MSW)	References	Cannot be any better in regard to clinical investigations w high validity & reliability!	Thank you for your comment.
Peer reviewer #1	Results	In total the authors identified a large number (2193) of potential new publications, but as with similar reviews in this area, the majority were excluded leaving only just over 60 meeting inclusion criteria. Some of these were follow-ups to the previous review, resulting in 51 new studies, 37 of which were (RCTs) and 14 non-randomized trials or cohort.	Thank you for your comment.
Peer reviewer #1	Results	Factors considered in the statistical analysis of the studies are specified clearly, as are the criteria for scoring the quality of the studies and for grading evidence of effectiveness.	Thank you for your comment.
Peer reviewer #1	Results	d. Results: The detail presented in the Results sections is appropriate and provides a very adequate overview of the findings. Table B, for example, provides a helpful summary of the strength of the evidence for studies of outcomes in various domains (IQ adaptive behaviour etc.)	Thank you for your comment.
Peer reviewer #1	Results	The tables on individual studies provide succinct information on the characteristics of the study (including the source of funding) and nicely summarise key outcomes	Thank you for your comment.
Peer reviewer #1	Results	The summaries for each of the sections are also well constructed and provide the reader with a good overview of the data contained in each of the tables.	Thank you for your comment.
Peer reviewer #1	Results	I have just a minor quibble about Table 8 (p 20) Column 4 has the heading “Study design/N/Risk of bias”. However, I couldn’t see anything pertaining to risk of bias, whereas the fact that studies are rated as “good” fair” etc. is not indicated in the heading.	We have clarified the report and used “quality” consistently as opposed to risk of bias.

Commentator & Affiliation	Section	Comment	Response
TEP Reviewer #1	Results	The investigators appear to have identified almost all relevant studies, but I think several eligible reports are missing from the review (Casenhiser et al., 2011; Mandell et al., 2013; Tonge et al., 2012, all in *Autism*). A follow-up of a previously reviewed study also may qualify (Magiati et al., 2011, *Research in Autism Spectrum Disorders*).	Thank you for pointing out these references. We have added the Casenheiser study in our search update conducted while the report was in peer review. The Mandell study does not meet criteria for the review as it addresses an educational intervention. The Magiati study was not included because we considered the original paper as reporting a broad-based educational approach in our initial review, and this update focused only on studies of behavioral interventions.
TEP Reviewer #1	Results	The data extraction is highly accurate, and the ratings of study quality and overall strength of evidence appear reasonable.	Thank you for your comment.
TEP Reviewer #2	Results	The detail is almost too great.	We attempted to balance clarity with the need to explain a study's conduct and results. We have attempted to streamline our reporting of the results.
TEP Reviewer #2	Results	The tables help orient the reader.	Thank you for your comment.
TEP Reviewer #2	Results	No studies were overlooked.	Thank you for your comment.
TEP Reviewer #3	Results	Results: I think the "Key Points" and short summaries of the studies are quite helpful and capture the study well.	Thank you for your comment.
TEP Reviewer #3	Results	Why are some tables landscape and some portrait?	Some tables are landscape given the number of columns needed to display the data clearly.
TEP Reviewer #3	Results	Page 19, line 7: mean age in months? It's missing a unit of time. Also, mean age on entry into the program?	We have clarified the analysis of this study.
TEP Reviewer #3	Results	Pages 24, lines 46-57 - page 35, lines3-29: Which is the "good" quality study and which is the "fair" quality study?	We have clarified that the study noted here was reported in 2 publications.
TEP Reviewer #3	Results	Page 56: Could a "Key Points" be included summarizing all of modifiers? The result related to Modifers of Treatment Effects are so heterogeneous that readers will struggle to see the forest for the trees. Any assistance in the form of a key points would add value.	Unfortunately, there is not enough consistency in the literature to pull out overarching key points in this section.

Commentator & Affiliation	Section	Comment	Response
TEP Reviewer #3	Results	Similarly on page 62, there is no “Key Points” but a nice summary at the end is given. It would be good to have a consistent approach to presenting summarized info throughout the Results section.	We have added key points to the other key questions and made the approach consistent.
Peer Reviewer #3	Results	Results: The results are very detailed and comprehensive, and at times seemed redundant. I’m not sure if you can make it less repetitive because the information is grouped to follow the structure and order of the key questions. The problem is that the same results are presented in multiple places because they pertain to more than one question. Perhaps some sections could refer back to earlier ones.	We agree that there are some inevitable redundancies given the framework and methods utilized. We have attempted to minimize wherever possible.
Peer Reviewer #3	Results	Figures and tables were quite detailed and helpful, with the exception of table 17 (insufficient evidence by outcomes assessed). The check marks and blanks are difficult to interpret (e.g., do checks indicate that there is evidence but it’s insufficient or do blanks indicate insufficient evidence? Do blank cells indicate that there is no information available?)	We have revised this table.
Peer Reviewer #3	Results	Regarding other tables, some formatting is needed to make them easier to follow but they appear comprehensive.	We have attempted to streamline the table formatting.
Public reviewer #1 (Richard G Kensinger, MSW)	Results	Reflect the realities of this most complex & puzzling clinical subgroup when the stakes are so compelling high. We have almost no clue about what causes it. The impact is so often quite humbling	Thank you for your comment.
Peer Reviewer #4	Results	Results: The amount of detail was very helpful including the multiple tables and grids of the research studies, results, applicability to targeted domains.	Thank you for your comment.
Public reviewer #1 (Richard G Kensinger, MSW)	Tables/Figures	Very solid	Thank you for your comment.