Comparative Effectiveness Review
Number 137

Effective Health Care Program

Therapies for Children With Autism Spectrum Disorder: Behavioral Interventions Update

Executive Summary

Background

Autism spectrum disorder (ASD) is a neurodevelopmental disorder marked by impaired social communication and social interaction accompanied by atypical patterns of behavior and interest. ASD is differentiated from other developmental disorders by significant impairments in social interaction and communication, along with restrictive, repetitive, and stereotypical behaviors and activities.[1] Social communication and social interaction features include deficits in social-emotional reciprocity (e.g., deficits in joint attention, atypical social approach and response, conversational challenges, reduced sharing of interest, emotions, and affect); deficits in nonverbal communication (e.g., atypical eye contact, reduced gesture use, limited use of facial expressions in social interactions, challenges understanding nonverbal communication); and deficits in forming and maintaining relationships (e.g., diminished peer interest, challenges joining in play, difficulties adjusting behavior to social context).

ASD features of restricted repetitive patterns of behavior, interests, or activities may include stereotyped motor mannerisms, use of objects, or speech (e.g., simple motor stereotypies, repetitive play, echolalia, and formal or idiosyncratic speech); insistence on sameness, inflexible adherence to routines, or ritualized patterns of behavior (e.g., distress at small changes, rigid patterns of thought and behavior,

Effective Health Care Program

The Effective Health Care Program was initiated in 2005 to provide valid evidence about the comparative effectiveness of different medical interventions. The object is to help consumers, health care providers, and others in making informed choices among treatment alternatives. Through its Comparative Effectiveness Reviews, the program supports systematic appraisals of existing scientific evidence regarding treatments for high-priority health conditions. It also promotes and generates new scientific evidence by identifying gaps in existing scientific evidence and supporting new research. The program puts special emphasis on translating findings into a variety of useful formats for different stakeholders, including consumers.

The full report and this summary are available at **www.effectivehealthcare.ahrq.gov/reports/final.cfm**.

DEPARTMENT OF HEALTH & HUMAN SERVICES • USA

AHRQ
Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • www.ahrq.gov

Effective Health Care

performance of everyday activities in ritualistic manner); intense preoccupation with specific interests (e.g., strong attachment to objects, circumscribed or perseverative topics of interest); and sensory sensitivities or interests (e.g., hyperreactivity or hyporeactivity to pain and sensory input, sensitivity to noise, visual fascination with objects or movement).[2-4]

ASD symptoms cause impairment across many areas of functioning and are present early in life. However, impairments may not be fully evident until environmental demands exceed children's capacity. They also may be masked by learned compensatory strategies later in life. Many children with ASD may also have intellectual impairment or language impairment, and the disorder may be associated with known medical, genetic, or environmental factors.

Treatments for ASD that families pursue include behavioral, educational, medical, allied health, and complementary approaches. Individual goals for treatment vary for different children and may include combinations of therapies. For many individuals, core symptoms of ASD (impairments in communication and social interaction and restricted/repetitive behaviors and interests) may improve with intervention and over time;[5-8] however, deficits typically remain throughout the lifespan. Lifelong management—often using multiple treatment approaches—may be required to maximize functional independence and quality of life.

Scope and Key Questions

Scope of Review

This systematic review updates the behavioral intervention portion of our comprehensive review of therapies for children with ASD published in 2011.[9] ASD intervention categories overlap substantially, and it can be difficult to cleanly identify the category into which an intervention should be placed. Ultimately, we defined behavioral interventions to include early intensive behavioral and developmental interventions, social skills interventions, play/interaction-focused approaches, interventions targeting symptoms commonly associated with ASD, and other general psychosocial approaches. This behavioral category of intervention explicitly does not include primarily medical interventions, complementary and alternative interventions, allied health interventions, or educationally focused interventions unless a behavioral intervention representative of the operationalization above was included within the study design.

At the time of the 2011 review (available at www.effectivehealthcare.ahrq.gov/ehc/products/106/656/CER26_Autism_Report_04-14-2011.pdf), the strength of the evidence was considered low for the effectiveness of early intensive behavioral and developmental interventions. Positive outcomes from an early and intensive behavioral and developmental intervention were noted in cognitive performance, language skills, and adaptive behavior when the intervention was delivered over substantial intervals of time (i.e., 1–2 years). Variability in response to such approaches was tremendous, with subgroups of children who demonstrated a more modest response. The ability to describe and predict these subgroups was limited.

Some other behavioral interventions that varied widely in terms of scope, target, and intensity had demonstrated effects, but the lack of consistent data limited understanding of whether these interventions were linked to specific clinically meaningful changes in functioning. Information was similarly lacking on modifiers of effectiveness, generalization of effects outside the treatment context, components of multicomponent therapies that drive effectiveness, and predictors of treatment success.

Since the publication of the initial review in 2011, a sizable body of research has been published, particularly addressing 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 may potentially be used to update treatment recommendations due to the number of new studies available. For this reason, the current review update focuses on studies of behavioral interventions.

Key Questions

We focused this review on behavioral treatments for children ages 2–12 with ASD and children younger than age 2 at risk of a diagnosis of ASD. We synthesized evidence in the published literature to address the following Key Questions (KQs).

KQ 1: Among children ages 2–12 with ASD, what are the short- and long-term effects of available behavioral treatment approaches? Specifically—

> **KQ 1a**: What are the effects on core symptoms (e.g., social communication and interaction, restricted and repetitive behaviors) in the short term (≤6 months)?
>
> **KQ 1b**: What are the effects on commonly associated symptoms (e.g., motor, medical, mood/anxiety, irritability, and hyperactivity) in the short term (≤6 months)?

KQ 1c: What are the longer term effects (>6 months) on core symptoms (e.g., social communication and interaction, restricted and repetitive behaviors)?

KQ 1d: What are the longer term effects (>6 months) on commonly associated symptoms (e.g., motor, medical, mood/anxiety, irritability, and hyperactivity)?

KQ 2: Among children ages 2–12, what are the modifiers of outcome for different behavioral treatments or approaches?

KQ 2a: Is the effectiveness of the therapies reviewed affected by the frequency, duration, and intensity of the intervention?

KQ 2b: Is the effectiveness of the therapies reviewed affected by the training and/or experience of the individual providing the therapy?

KQ 2c: What characteristics, if any, of the child modify the effectiveness of the therapies reviewed?

KQ 2d: What characteristics, if any, of the family modify the effectiveness of the therapies reviewed?

KQ 3: Are there any identifiable changes early in the treatment phase that predict treatment outcomes?

KQ 4: What is the evidence that effects measured at the end of the treatment phase predict long-term functional outcomes?

KQ 5: What is the evidence that specific intervention effects measured in the treatment context generalize to other contexts (e.g., people, places, materials)?

KQ 6: What evidence supports specific components of behavioral treatment as driving outcomes, either within a single treatment or across treatments?

KQ 7: What evidence supports the use of a specific behavioral treatment approach in children under the age of 2 who are at high risk of developing ASD based on behavioral, medical, or genetic risk factors?

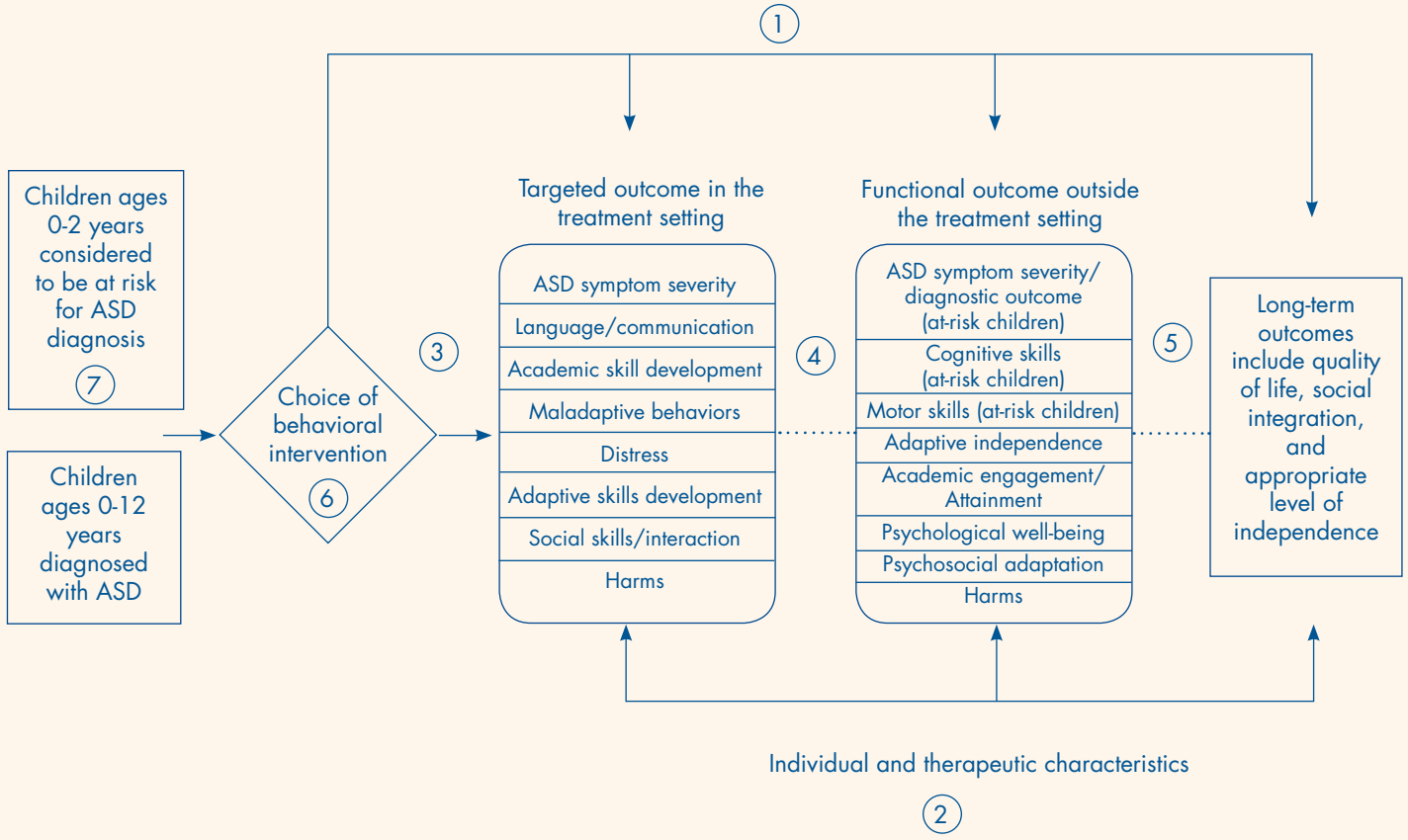
Uses of This Report

We anticipate that the report will be of value to clinicians who treat children with ASD, who can use the report to assess the evidence for different treatment strategies. In addition, this review will be of use to the National Institutes of Health, Centers for Disease Control and Prevention, Centers for Medicare & Medicaid Services, and Health Resources and Services Administration–all of which have offices or bureaus devoted to child health issues and may use the report to compare treatments and determine priorities for funding. This report can bring practitioners up to date about the current state of evidence related to behavioral interventions, and it provides an assessment of the quality of studies that aim to determine the outcomes of therapeutic options for the management of ASD. It will be of interest to families affected by ASD because of the recurring need for families and their health care providers to make the best possible decisions among numerous options. We also anticipate it will be of use to private-sector organizations concerned with ASD; the report can inform such organizations' understanding of the effectiveness of treatments and the amount and quality of evidence available. Researchers can obtain a concise analysis of the current state of knowledge related to behavioral interventions for ASD. They will be poised to pursue further investigations that are needed to understand best approaches to behavioral therapies for children with ASD.

Analytic Framework

Figure A illustrates the analytic framework for the current update. The figure illustrates the placement of the review's KQs within the context of treatment choice, potential outcomes, and characteristics that may affect outcomes. A child entering treatment may be between the ages of 0 and 2 and at risk for diagnosis of ASD or ages 0 to 12 with a diagnosis of ASD. Diagnoses may occur before age 2; thus the represented age ranges overlap.

Figure A. Analytic framework for behavioral interventions for children with ASD



ASD = autism spectrum disorder
Note: Numbers in circles represent placement of Key Questions.

Methods

Literature Search Strategy

A librarian employed search strategies provided in Appendix A of the full report to retrieve research on interventions for children with ASD. We searched MEDLINE® via the PubMed® interface, PsycINFO® (psychology and psychiatry literature), and the Educational Resources Information Clearinghouse using a combination of subject heading terms appropriate for each database and key words relevant to ASD (e.g., autism, Asperger). We limited searches to the English language and literature published since the development of the 2011 review. Our last search was conducted in December 2013. We also manually searched the reference lists of included studies and of recent narrative and systematic reviews and meta-analyses addressing ASD.

Inclusion and Exclusion Criteria

We developed criteria for inclusion and exclusion based on the patient populations, interventions, outcome measures, and types of evidence specified in the KQs and in consultation with a Technical Expert Panel. Table A summarizes criteria.

Table A. Inclusion criteria

Category	Criteria
Study population	Children ages 0–12 with ASD or 0–2 considered to be at risk for ASD based on sibling status or early developmental/behavioral vulnerabilities highly suspicious of ASD
Publication language	English only
Admissible evidence (study design and other criteria)	**Admissible designs** Randomized controlled trials, prospective and retrospective cohort studies, and nonrandomized controlled trials **Other criteria** Studies must be original research studies providing sufficient detail regarding methods and results to enable use and aggregation of the data and results. Studies must have relevant population and ≥10 participants with ASD. Studies must address 1 or more of the following for ASD: • Behavioral treatment modality • Predictors of treatment outcomes • Generalization of treatment outcomes to other contexts • Drivers of treatment outcomes Relevant outcomes must be able to be abstracted from data in the papers. Data must be presented in the aggregate (vs. individual participant data).

ASD = autism spectrum disorder

Study Selection

Two reviewers independently assessed each abstract identified for potential inclusion using an abstract review form with questions stemming from our selection criteria. If one reviewer concluded that the article could be eligible for the review based on the abstract, we retained it for full-text assessment. Two reviewers independently assessed the full text of each included study using a similar standardized form. Disagreements between reviewers were resolved by a third-party adjudicator. The group of abstract and full-text reviewers included expert clinicians and researchers and health services researchers; abstract and full-text review forms are in Appendix B of the full report.

Data Extraction

We extracted data from included studies into evidence tables that report study design, descriptions of the study populations (for applicability), description of the intervention, and baseline and outcome data on constructs of interest. Data were initially extracted by one team member and reviewed for accuracy by a second. The final evidence tables are presented in their entirety in Appendix C of the full report. For studies that were reported in the 2011 review and have followup data reported here, the evidence table for the original studies can be found in the 2011 report.[9]

Quality Assessment

We used the approach to assessing the quality of individual studies developed for the 2011 review and following methods outlined in the Agency for Healthcare Research and Quality Effective Health Care Program's "Methods Guide for Effectiveness and Comparative Effectiveness Reviews."[10] We assessed the quality of studies in domains including study design, participant ascertainment, diagnostic approach, and outcome measurement using specific questions to evaluate a study's conduct. We rated each domain individually and combined them for an overall

quality level, as described in the full report. Three levels were possible: good, fair, and poor.

Data Synthesis

We summarized all data qualitatively using evidence tables. We focused on outcomes related to core ASD symptoms (impairments in communication and social interaction and restricted/repetitive behaviors and interests); outcomes including IQ and adaptive behavior; and key symptoms in studies of interventions targeting conditions commonly associated with ASD (e.g., anxiety). For the update, we describe new comparative studies published since the original report, and we make our conclusions and assess the strength of evidence on the cumulative comparative evidence across the original report and update.

Strength of the Body of Evidence

Two senior investigators graded the entire body of evidence (i.e., studies from the 2011 review and studies identified for the current review) based on the "Methods Guide for Effectiveness and Comparative Effectiveness Reviews."[10] The team reviewed the final strength-of-evidence designation.

The assessment of the literature was done by considering how confident we were that the true effect was observed and how stable that effect is likely to be in the face of future research. Strength of evidence describes the adequacy of the current research in terms of both quantity and quality, as well as the degree to which the entire body of current research provides a consistent and precise estimate of effect. Strength of the evidence is assessed for a limited set of critical outcomes, typically those related to effectiveness of an intervention. We assessed the strength of the evidence for studies addressing KQs 1 and 7, which deal specifically with the outcomes of intervention.

We established the maximum strength of evidence possible based on criteria for each domain: study limitations, consistency in direction of the effect, directness in measuring intended outcomes, precision of effect, and reporting bias. (See the full report for further description of domains.) Then we assessed the number of studies and range of study designs for a given intervention-outcome pair and downgraded the rating when the cumulative evidence was not sufficient to justify the higher rating. The possible grades were—

- High: High confidence that the evidence reflects the true effect. Further research is unlikely to change estimates.
- Moderate: Moderate confidence that the evidence reflects the true effect. Further research may change our confidence in the estimate of effect and may change the estimate.
- Low: Low confidence that the evidence reflects the true effect. Further research is likely to change confidence in the estimate of effect and is also likely to change the estimate.
- Insufficient: Evidence is either unavailable or does not permit a conclusion.

Applicability

We assessed applicability by identifying potential population, intervention, comparator, outcome, and setting (PICOS) factors likely to affect the generalizability of results (i.e., applicability to the general population of children with ASD). For this particular review, the most likely factors that could affect applicability are the patient population (e.g., whether or not results are available to assess the utility of given interventions in target populations) and the intervention (e.g., the difficulty of applying the intervention in a nonresearch setting given available resources). We noted where data were available for specific populations and made relative assessments of applicability for intervention components in the context of resource considerations such as availability of services/programs.

Results

Article Selection

We identified 2,639 newly published citations and abstracts. (Figure 2 in the full report shows the disposition of studies.) We excluded 2,012 studies at abstract review and assessed the full text of 627 studies. Of these, 79 publications, comprising 65 unique studies, met our criteria. Eight of these studies report followup data to papers included in the 2011 review of therapies for children with ASD. The 65 new studies described in this update to add to the conclusions of the original report comprise 48 randomized controlled trials (RCTs) and 17 nonrandomized trials or cohort studies. The full report includes detailed references. Appendix E of the full report includes a list of all studies excluded at the abstract and full-text review stages.

KQ 1. Effects of Behavioral Interventions on Core and Commonly Associated Symptoms in Children With ASD

Studies of Early Intensive Behavioral and Developmental Interventions

We located 37 papers comprising 25 unique studies addressing early intensive behavioral and developmental interventions. The studies included five RCTs of good quality, six of fair quality, and one of poor quality. Individual studies using intensive University of California, Los Angeles (UCLA)/Lovaas-based interventions, the Early Start Denver Model (ESDM), the Learning Experiences and Alternate Program for Preschoolers and their Parents (LEAP) program, and eclectic variants reported improvements in outcomes for young children. Improvements were most often seen in cognitive abilities and language acquisition, with less robust and consistent improvements seen in adaptive skills, core ASD symptom severity, and social functioning.

Young children receiving high-intensity applied behavior analysis (ABA)-based interventions over extended timeframes (i.e., 8 months–2 years) displayed improvement in cognitive functioning and language skills relative to community controls (Table B). However, the magnitude of these effects varied across studies. This variation may reflect subgroups showing differential responses to particular interventions. Intervention response is likely moderated by both treatment and child factors, but exactly how these moderators function is not clear. Despite multiple studies of early intensive treatments, intervention approaches still vary substantially, which makes it difficult to tease apart what these unique treatment and child factors may be. Further, the long-term impact of these early skill improvements is not yet clear, and many studies did not follow children beyond late preschool or early school years.

Studies of high-intensity early intervention services also demonstrated improvements in children's early adaptive behavior skills, but these improvements were more variable than those found for early cognitive and language skills. Treatment effects were not consistently maintained over followup assessments across studies. Many studies measured different adaptive behavior domains (creating within-scale variability), and some evidence suggests that adaptive behavior changes may be contingent on baseline child characteristics, such as cognitive/language skills and ASD severity.

Evidence for the impact of early intensive intervention on core ASD symptoms is limited and mixed. Children's symptom severity often decreased during treatment, but these improvements often did not differ from those of children in control groups. Better quality studies reported positive effects of intervention on symptom severity, but multiple lower quality studies did not.

Since our previous review, there have been substantially more studies of well-controlled low-intensity interventions that provide parent training in bolstering social communication skills. Although parent training programs modified parenting behaviors during interactions, data were more limited about their ability to improve broad developmental skills (such as cognition, adaptive behavior, and ASD symptom severity) beyond language gains for some children. Children receiving low-intensity interventions have not demonstrated the same substantial gains in cognitive skills seen in the early intensive intervention paradigms.

Social Skills Studies

We located 13 studies addressing interventions targeting social skills, including 11 RCTs. The overall quality of studies improved in comparison with the previous review, with 2 good-quality and 10 fair-quality studies. Social skills interventions varied widely in terms of scope and intensity. A few studies replicated interventions using the Skillstreaming model, which uses a published treatment manual (i.e., is manualized) to promote a consistent approach. Other studies incorporated peer-mediated and/or group-based approaches, and still others described interventions that focused on emotion identification and Theory of Mind training. The studies also varied in intensity, with most interventions consisting of 1–2 hour sessions/week lasting approximately 4–5 weeks. However, some of the group-based approaches lasted 15–16 weeks.

Most studies reported short-term gains in either parent-rated social skills or directly tested emotion recognition. However, our confidence (strength of evidence) in that effect is low (Table B). Although we now have higher quality studies of social skills interventions that demonstrate positive effects, our ability to determine effectiveness continues to be limited by the diversity of the intervention protocols and measurement tools (i.e., no consistent outcome measures used across studies). Studies also included only participants considered "high functioning" and/or with IQ test scores >70, thus limiting generalization of results to children with more significant impairments. Maintenance and generalization of these skills beyond the intervention setting are also inconsistent,

with parent and clinician raters noting variability in performance across environments.

Play-/Interaction-Focused Studies

Since our previous review, more studies of well-controlled joint attention interventions across a range of intervention settings (e.g., clinician, parent, teacher delivered) have been published. This growing evidence base includes 11 RCTs of good and fair quality and suggests that joint attention interventions may be associated with positive outcomes for toddler and preschool children with ASD, particularly when targeting joint attention skills themselves as well as related social communication and language skills (Table B). Although joint attention intervention studies demonstrated changes within this theoretically important domain, data are more limited about their ability to improve broad developmental skills (such as cognition, adaptive behavior, and ASD symptom severity) beyond direct measures of joint attention and related communication and language gains over time.

Specific training that used naturalistic approaches to promote imitation (e.g., Reciprocal Imitation Training) was associated with some improvements, not only in imitation skills, but also potentially in other social communication skills (such as joint attention). Additionally, parent training in a variety of play-based interventions was associated with enhanced early social communication skills (e.g., joint attention, engagement, play interactions), play skills, and early language skills.

Studies of Interventions Targeting Conditions Commonly Associated With ASD

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. Five of these studies reported significantly greater improvements in anxiety symptoms in the intervention group compared with controls. Two found positive effects of cognitive behavioral therapy (CBT) on the core ASD symptom of socialization, and one reported improvements in executive function in the treatment group. The one RCT that did not find a significant benefit of CBT compared it with social recreational therapy rather than with treatment as usual or a wait-listed control group.

The studies examining the effects of CBT on anxiety had largely consistent methodologies. Six studies provided followup data reflecting treatment effects that lasted beyond the period of direct intervention. Two common factors limit the applicability of the results, however. Due to the nature of CBT, which is often language intensive and requires a certain level of reasoning skills to make abstract connections between concepts, most studies included only children with IQs much greater than 70. These studies report positive results regarding the use of CBT to treat anxiety in children with ASD (Table B). They also report some positive results in socialization, executive function, and communication; however, these results were less robust, and it is unclear in some studies if these improvements exceeded improvements related to the impact of ameliorated anxiety itself.

Additional data in the current review relate to parent training to address challenging behavior. Specifically, one fair-quality study combined a parent-training approach with risperidone. This combination significantly reduced irritability, stereotypical behaviors, and hyperactivity, and improved socialization and communication skills. However, these effects were not maintained at 1 year after treatment.

Other Behavioral Studies

Two RCTs (one fair and one poor quality) examined neurofeedback and found some improvements on parent-rated measures of communication and tests of executive function. Three fair-quality RCTs reported on sleep-focused interventions, with little positive effect of a sleep education pamphlet for parents in one, improvements in sleep quality in treatment arms (melatonin alone, melatonin + CBT) in another, and some improvements in time to fall asleep in one short-term RCT of sleep education programs for parents. One poor-quality study of parent education to mitigate feeding problems reported no significant effects.

KQ 2. Modifiers of Treatment Effects

Among the potential modifiers or moderators of early intensive ABA-based interventions, younger age at intake was associated with better outcomes for children in a limited number of studies. Greater baseline cognitive skills and higher adaptive behavior scores were associated with better outcomes across behavioral interventions, but again, these associations were not consistent. In general, children with lower symptom severity or less severe diagnoses improved more than participants with greater impairments. Many studies (e.g., social skills, CBT) restricted the range of participants' impairment at baseline (e.g., recruiting only participants with IQs >70), limiting understanding of intervention impact on broader

populations. Studies assessing parental responsiveness to children's communication typically reported better outcomes in children whose parents were more aligned with the child's communication versus those who attempted to redirect or were less synchronized. Regarding intervention-related factors, duration of treatment had an inconsistent effect. Some studies reported improved outcomes with more intervention time and others reported no association. Overall, most studies were not adequately designed or controlled to identify true moderators of treatment response.

KQ 3. Treatment Phase Changes That Predict Outcomes

The reviewed literature offers little information about what specific early changes from baseline measurements of child characteristics might predict long-term outcome and response.

KQ 4. Treatment Effects That Predict Long-Term Outcomes

Few studies assess end-of-treatment effects that may predict outcomes. Several early intensive behavioral and developmental interventions are associated with changes in outcome measures over the course of very lengthy treatments, but such outcomes usually have not been assessed beyond treatment windows. One family of studies attempted to follow young children receiving early joint attention intervention until they were school aged, but this study failed to include adequate followup of control conditions. It also involved children who were receiving many hours of uncontrolled interventions during the course of study.

KQ 5. Generalization of Treatment Effects

The majority of the social skills and behavioral intervention studies targeting associated conditions attempted to collect outcomes based on parent, self, teacher, and peer report of targeted symptoms (e.g., anxiety, externalizing behaviors, social skills, peer relations) at home, at school, and in the community. Although such ratings outside of the clinical setting may be suggestive of generalization in that they improve outcomes in the daily context/life of the child, in most cases, these outcomes are parent reported and not confirmed with direct observation. Behavioral intervention studies rarely measured outcomes beyond the intervention period, and we therefore cannot assume that effects were maintained over time.

KQ 6. Treatment Components That Drive Outcomes

We did not identify any studies meeting our inclusion criteria that addressed this question.

KQ 7. Treatment Approaches for Children Under Age 2 at Risk for Diagnosis of ASD

In the studies addressing interventions for younger children, children who received behavioral interventions seemed to improve regardless of intervention type (including the comparator interventions, which were also behavioral). None of the fair- or good-quality studies compared treatment groups with a no-treatment control group. Potential modifiers of treatment efficacy include baseline levels of object interest. Most outcome measures of adaptive functioning were based on parent report, and the effect of parental perception of treatment efficacy on perception (and report) of child functioning was generally not explored.

Discussion

Key Findings and Strength of Evidence

Since our previous review in 2011, there has been a significant increase in the quantity and quality of studies investigating behavioral interventions. These new studies add to the prior report and strengthen our ability to make conclusions about the effectiveness of behavioral interventions. Of the 45 comparative studies of behavioral interventions (29 RCTs) in the 2011 review, we considered only 2 as good quality. Among the new studies described in this current review, 19 studies are good quality, and 48 of the 65 included studies are RCTs.

Evidence from the original report and this update suggests that early behavioral and developmental intervention based on the principles of ABA delivered in an intensive (>15 hours per week) and comprehensive (i.e., addressing numerous areas of functioning) approach can positively affect a subset of children with ASD (Table B). Across approaches, children receiving early intensive behavioral and developmental interventions demonstrate improvements in cognitive, language, adaptive, and ASD impairments compared with children receiving low-intensity interventions and eclectic non–ABA-based intervention approaches.

Since our previous review, there have also been substantially more studies of well-controlled low-intensity interventions aimed at parent training for comprehensive impact on social communication skills. Although parent

training programs modified parenting behaviors during interactions, data are more limited about their ability to improve broad developmental skills (such as cognition, adaptive behavior, and ASD symptom severity) beyond short-term language gains for some children.

A growing number of studies of improved quality demonstrated positive effects of social skills interventions on at least one outcome measure, but a lack of consistency in the interventions studied and outcome measures used makes it difficult to understand specific effects of different intervention modalities.

A growing evidence base also suggests that children receiving targeted play-based interventions (e.g., joint attention, imitation, play-based interventions) demonstrate improvements in early social communication skills. Children receiving targeted joint attention packages in combination with other interventions show substantial improvements in joint attention and language skills over time. There is also evidence across a variety of play-based interventions that young children may display short-term improvements in early play, imitation, joint attention, and interaction skills. However, evidence that these short-term improvements are linked to broader indexes of change over time is not substantial.

CBT for associated conditions such as anxiety had the largest number of high-quality studies in the current review. A strong evidence base now suggests that school-aged children with average to above average intelligence and comorbid anxiety symptoms receiving manualized CBT therapy show substantial improvements in anxiety compared with wait-list controls. Table B summarizes the strength of the evidence for each category of intervention.

Table B. Strength of the evidence

Intervention	Outcome	SOE	Study Design Quality (N Participants)	Ratings for Domains Used To Assess SOE; Issues	Key Findings
Early intensive behavioral and developmental intervention: ABA based	IQ/cognitive	Moderate for positive effect	RCT: 1 good, 2 fair (360) Prospective cohort: 6 fair, 2 poor (521) nRCT; 1 good, 4 fair (170) Retrospective cohort: 1 fair, 2 poor (182)	Study limitations: Medium Consistency: Consistent Directness: Direct Precision: Precise Reporting bias: Undetected Other concerns: Approaches across studies vary substantially; it is difficult to determine the effects of these unique studies on specific groups of children.	Young children receiving high-intensity interventions display improvements in aspects of cognitive functioning. Most studies found that children in treatment and comparison groups both improved on cognitive skills, with children in early intensive behavioral interventions (target intervention) improving more than children receiving other types of services (eclectic comparators). Not all improvements were maintained at long-term followup Therefore, SOE was moderate for a positive effect relative to eclectic controls.
	Adaptive behavior	Low for positive effect	RCT: 1 good, 1 fair (76) Prospective cohort: 7 fair, 2 poor (616) nRCT: 1 good, 4 fair (170) Retrospective cohort: 1 fair, 2 poor (182)	Study limitations: Medium Consistency: Inconsistent Directness: Direct Precision: Imprecise Reporting bias: Undetected Other concerns: Behavior was always measured by parent report (Vineland Scales of Adaptive Behavior) rather than objective observation.	Most studies found that children in both treatment and control groups improved on adaptive skills. However, children in early intensive behavioral interventions improved more than children receiving other types of services. Not all group differences were maintained over long-term followup Therefore, SOE was low for a positive effect relative to eclectic controls.
	Symptom severity	Low for positive effect	RCT: 1 good, 1 fair (332) nRCT: 1 good, 1 fair (74) Prospective cohort: 4 fair, 2 poor (470) Retrospective cohort: 1 fair (142)	Study limitations: Medium Consistency: Inconsistent Directness: Direct Precision: Imprecise Reporting bias: Undetected Other concerns: Most control groups were also receiving treatments and also showed improvement, making it difficult to tease apart the effect of intervention.	There was mixed impact on symptom severity. SOE is low for a positive effect on symptom severity because 2 good-quality studies showed positive effects but multiple lower quality studies did not. More studies are needed to confirm results.

Table B. Strength of the evidence (continued)

Intervention	Outcome	SOE	Study Design Quality (N Participants)	Ratings for Domains Used To Assess SOE; Issues	Key Findings
Early intensive behavioral and developmental intervention: ABA based (continued)	Language/ communication	Moderate for positive effect	RCT: 1 good, 2 fair (360) nRCT: 1 good, 3 fair (143) Prospective cohort: 6 fair, 2 poor (616)	Study limitations: Medium Consistency: Consistent Directness: Direct Precision: Precise Reporting bias: Undetected Other concerns: Some studies measured language using direct testing, whereas others only used parent-reported measures (Vineland Scales of Adaptive Behavior).	Most studies found a positive effect of treatment on language/communication skills, although the specific domain of improvement (e.g., receptive vs. expressive language) varied across study. Some initial between-group differences disappeared at long-term followup There is moderate SOE of a positive effect on language overall.
	Social skills/ social behavior	Low for positive effect	RCT: 1 good, 1 fair (332) nRCT: 1 fair (34) Prospective cohort: 4 fair, 1 poor (406) Retrospective cohort: 1 fair (142)	Study limitations: Medium Consistency: Inconsistent Directness: Direct Precision: Imprecise Reporting bias: Undetected Other concerns: Social skills were assessed almost exclusively using parent-reported standard scores on the Vineland Scales of Adaptive Behavior.	Many studies found that treatment groups improved more than controls on measures of social skills, although a significant minority did not find any treatment effect. SOE is low for a positive effect at this time because, although positive effects were observed, they were not consistent.
Early intensive behavioral and developmental intervention: parent training	IQ/cognitive	Low for no effect	RCT: 3 fair (148) Prospective cohort: 1 good, 1 fair, 1 poor (142)	Study limitations: Medium Consistency: Inconsistent Directness: Direct Precision: Imprecise Reporting bias: Undetected Other concerns: None	Most studies of parent-implemented ABA demonstrated no improvements in IQ relative to community-based interventions; in some studies worse outcomes were reported relative to center-based treatment. SOE is low for no effect due to heterogeneity in interventions and outcomes measured.

Table B. Strength of the evidence (continued)

Intervention	Outcome	SOE	Study Design Quality (N Participants)	Ratings for Domains Used To Assess SOE; Issues	Key Findings
Early intensive behavioral and developmental intervention: parent training (continued)	Symptom severity	Low for positive effect	RCT: 3 good, 3 fair (361) Prospective cohort: 1 good, 1 fair, 2 poor, (203)	Study limitations: Low Consistency: Inconsistent Directness: Direct Precision: Imprecise Reporting bias: Undetected Other concerns: The measure of symptom severity varied across studies and was inconsistently defined.	Many studies found that treatment groups had improved ASD symptoms relative to controls.
	Language/ commun-ication	Low for positive effect	RCT: 4 good, 6 fair, 1 poor (664) nRCT: 1 poor (22) Prospective cohort: 2 good, 2 poor (176)	Study limitations: Low Consistency: Inconsistent Directness: Direct Precision: Precise Reporting bias: Undetected Other concerns: A mix of outcome measures was used—both parent reported (Vineland Scales of Adaptive Behavior) and more standardized measures such as Reynell or Mullen scales.	Parent training was associated with improvements in language (low SOE for improvements), but interventions and comparators were different across studies, as were the outcome measures. More studies are needed to confirm results.
Social skills	Social skills/ social behavior	Low for positive effect	RCT: 2 good, 11 fair, 6 poor (730) nRCT: 2 fair (45) Retrospective cohort: 1 poor (117)	Study limitations: Medium Consistency: Inconsistent Directness: Direct Precision: Precise Reporting bias: Undetected Other concerns: Interventions varied widely in terms of scope and intensity.	School-aged children diagnosed without concomitant cognitive and language deficits demonstrated short-term gains in social skills and emotion recognition. Maintenance and generalization of these skills beyond the treatment context had variable results.

Table B. Strength of the evidence (continued)

Intervention	Outcome	SOE	Study Design Quality (N Participants)	Ratings for Domains Used To Assess SOE; Issues	Key Findings
Play/interaction based interventions	Joint attention	Moderate for positive effect	RCT: 3 good, 6 fair (305)	Study limitations: Low Consistency: Consistent Directness: Indirect Precision: Precise Reporting bias: Undetected Other concerns: Children in several studies were also receiving other early intervention; disentangling results is difficult.	Selected joint attention skills consistently increased in treatment arms, but duration of effects is unclear. The SOE is lowered to moderate, as children in most studies were also receiving other early intervention and disentangling effects is difficult.
	Play skills	Low for positive effect	RCT: 3 good, 3 fair, 3 poor (265) Prospective cohort: 1 poor (12)	Study limitations: Medium Consistency: Consistent Directness: Direct Precision: Precise Reporting bias: Undetected Other concerns: Children in several studies were also receiving other early intervention; disentangling results is difficult.	Play skills increased in treatment arms but duration of effects is unclear. Imitation skills improved in treatment arms in 4 small short-term studies and in the treatment and control arms in 1 study.
	Language/ communication	Low for positive effect	RCT: 4 fair (165)	Study limitations: Medium Consistency: Consistent Directness: Direct Precision: Imprecise Reporting bias: Undetected Other concerns: Children in several studies were also receiving other early intervention; disentangling results is difficult.	Expressive, but not receptive, language skills generally increased in the treatment arms in 2 studies; prompted, but not spontaneous, communication improved in 1 study.

Table B. Strength of the evidence (continued)

Intervention	Outcome	SOE	Study Design Quality (N Participants)	Ratings for Domains Used To Assess SOE; Issues	Key Findings
Play/interaction based interventions (continued)	Social skills	Low for positive effect	RCT: 1 good, 3 fair (173)	Study limitations: Medium Consistency: Consistent Directness: Indirect Precision: Precise Reporting bias: Undetected Other concerns: Children in several studies were also receiving other early intervention; disentangling results is difficult.	Joint engagement or positive affect improved in treatment arms in 3 studies.
Interventions addressing commonly associated conditions: CBT	Anxiety	High (for positive effect in older children with at least average IQs)	RCT: 6 good, 1 fair, 2 poor (413) nRCT: 1 fair (31)	Study limitations: Low Consistency: Consistent Directness: Direct Precision: Precise Reporting bias: Undetected Other concerns: Studies included older children, typically with IQ >70.	Improvement in anxiety symptoms was greater for CBT vs. control group in 5/6 studies; study that did not show improvement compared CBT with an active treatment instead of a wait-listed control. Improvements were maintained at followup.
	Symptom severity	Low for positive effect	RCT: 2 good (81)	Study limitations: Low Consistency: Consistent Directness: Direct Precision: Imprecise Reporting bias: Undetected Other concerns: None	There was significant improvement in clinician- and parent-rated measures of anxiety severity in both studies, with improvement maintained at followup. SOE is low based on only 2 small studies.

ABA = applied behavior analysis; CBT = cognitive behavioral therapy; nRCT = nonrandomized controlled trial; RCT = randomized controlled trial; SOE = strength of evidence

Applicability

Studies of early intensive behavioral and developmental interventions were conducted primarily in preschool-age and early school-age children (i.e., typically children initially ages 1.5–7 years). The cognitive, language, and adaptive behavior profiles of participants included in these studies were generally in line with those seen in the community (i.e., typically marked by substantial impairment/delay, but with some children with more intact early cognitive/language profiles).

Often studies were conducted in highly controlled environments (e.g., university-supported intervention trials) or the methodology was not well described (i.e., nonmanualized approaches), which substantially limits their applicability to community-based settings. Even available manualized interventions require high degrees of specialization and training that make them difficult to implement in community practices.

Studies of parent training interventions and play-based interventions for preschool children often emphasized principles of ABA, in accordance with current practice recommendations for the target populations typically referred for these services. Training programs included components to improve social communication skills such as joint attention, play-based interactions, and pragmatic language approaches; interventions were conducted for approximately 1–4 hours/week, with parents trained in how to generalize these skills to other natural settings. Several programs offer manualized intervention protocols that can facilitate their use in community settings. Again, however, the number of providers in community settings who are capable of implementing these programs may be limited.

Most studies of social skills interventions targeted elementary school–aged children (6–13 years old) with few studies targeting preschool-age children, although such interventions may be important in this younger age group. Most studies also excluded children with IQs falling outside of the average range. Similarly, CBT for conditions commonly associated with ASD was targeted toward older children with generally average cognitive abilities and comorbid anxiety disorders.

Limitations of the Review Process

We limited this update to comparative studies and included only those with at least 10 individuals. Thus, we did not include data from pre-post studies or those with a very small number of children. These would include a number of single subject design studies that may be helpful for understanding focused questions of short-term efficacy in individual children and that may be useful for explicating mechanisms of action. These studies are less able to contribute to the body of evidence that we sought on population-level and generalizable effects. Users of this review may want to take those studies into account as context when applying our findings. We limited our review to English-language studies, not finding evidence that we were missing relevant research in other languages. We also did not include interventions primarily viewed as medical, educational, complementary/alternative, or allied health in nature.

Limitations of the Evidence Base

Despite improvements, the existing literature still has significant methodological concerns that in many ways continue to limit the strength of these conclusions. Evidence for the impact of intensive ABA-based interventions on cognitive, language, and adaptive skills and ASD symptoms also highlights important limitations of current treatment modalities. First, even children who demonstrate clinically significant improvements in these areas often continue to display substantial impairment in these and other areas over time. Second, not all children receiving intensive ABA-based intervention showed robust improvements in these domains. Thus, it is still challenging to predict long-term functional and adaptive outcomes on an individual level. Further, although children receiving early intensive developmental and behavioral intervention commonly display substantial improvements, the magnitude of these effects varies across studies and may indicate subgroups showing variable responses to particular interventions. Intervention response is likely moderated by both treatment and child factors.

Despite multiple studies of early intensive treatments, intervention approaches still vary substantially, which makes it difficult to tease apart what these unique treatment and child factors may be. Similarly, data on provider type and qualifications are variably reported, and the impact of provider characteristics on treatment outcomes is unclear. Study sample sizes are typically small (total numbers ranging from 11 to 284 for studies in the current review, median = 40), and some studies may be considered pilots for larger studies that may better answer questions about intervention intensity and moderators of effects. At this time, the evidence is insufficient to adequately identify and target the children who are most likely to benefit (or not benefit) from specific interventions.

Many early intervention studies found that children in all groups improved on ASD symptom measures regardless of intervention type, although the degree of improvement

was often significantly greater in the treatment group. In many studies, results were confounded by nonrandom assignment of participants, including assignment based on child characteristics (such as having the skills necessary to participate in the intervention setting) or parental preference. The latter is especially problematic when outcomes are measured by parent report, given some evidence that parental stress influences parent perceptions of child outcomes. Additionally, in most studies, both enrolled and control/wait-listed children were receiving concomitant interventions, whose magnitude was inconsistently documented and controlled for in analyses.

A remaining significant challenge to interpreting the early intensive intervention literature relates to how interventions are described and implemented. Although researchers are attempting to manualize approaches as well as operationalize and measure treatment fidelity, 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 (i.e., developmental, intensive behavioral, center based, and combinations), intensity (12 hours over 3 months vs. 30 hours over 1 week), and duration (weeks to years); varied inclusion and baseline assessment criteria; children of varying ages (intake age ranging from 18 months to 7 years); and many different outcome measurements over different periods of time (weeks to years). Manualizing intensive interventions to be delivered over the course of months and years for a heterogeneous patient population is intrinsically challenging. However, recent progress toward this end has shown that children may respond differentially to early intensive approaches.

Few studies directly compared the effects of well-controlled treatment approaches, instead comparing interventions with nonspecific "treatment as usual," which clearly lacks the level of control for expectancy bias in a placebo-controlled medication study. Additionally, little data on the practical effectiveness or feasibility of these treatments beyond research studies exist, and questions remain about whether reported findings would generalize on a larger scale within communities. Furthermore, the studies conducted have used small samples, drastically different treatment approaches and duration, and different outcome measurements. Similarly, no studies reported harms of intervention in terms of child, family, or system impact.

Although there was a fairly robust evidence base on CBT, the literature lacks head-to-head comparisons of treatment or controlled comparisons of combinations of treatments, despite the fact that most children are undergoing multiple concurrent treatments. Although the studies are well designed, the sample sizes are quite modest. Additionally, the CBT approaches were modified for children with ASD and often manualized by the study authors themselves.

Research Gaps and Needs

Given the heterogeneity of the expression of ASD across children, a critical area for further research is understanding which children are likely to benefit from particular interventions. 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. It is possible that meta-analyses of individual patient data may provide additional information for identifying subgroups of responders.

Further, our understanding of early indicators of treatment response is extremely limited, such that it is not realistic to implement evidence-based changes in intervention based on assessing children's responses. This is quite important to parents, providers, and families, as they often want to know not only when a treatment is working, but also when the lack of a robust response should lead them to pursue other treatment options. Similarly, research is lacking on the durability of treatment gains and approaches needed to maintain gains.

Currently, the evidence suggests that some children will show dramatic improvement overall, others will display robust improvement in some areas with continued areas of vulnerability in others, and still other children will show more modest responses to treatment. It is also unclear how similar groups of children would respond to differing levels of intervention intensity, approaches, and methods. Research suggests that child characteristics such as baseline cognitive, language, and adaptive skills and ASD symptoms correlate with treatment outcome regardless of intervention. However, these correlational data provide limited information to predict what treatments will work best for individual children. Intensive comprehensive intervention strategies are often, by their very nature, multicomponent, but little data exist on whether specific treatment components drive effectiveness. Also, little is known about mediators of change. Finally, intervention research often fails to collect data on pragmatic factors

related to family, culture, available resources, and stressors that are likely critical to understanding treatment response in a "real-world" context.

Measuring appropriate outcomes is a primary methodologic concern in the ASD literature. Intervention research has typically measured differing outcomes across studies, which has limited the ability to understand change within and across individual studies.[11] Many studies also used problematic methods to operationalize outcomes, doing so in terms of change on standardized measures that reference normative populations (i.e., IQ measurement, adaptive behavior scores). This may not be an appropriate or adequate method for measuring or predicting early treatment response, changes in quality of life, or long-term functional outcomes. Such measurement, while allowing for comparison with typically developing populations, may miss important information about changes that are relevant within the ASD population specifically. More simply, it is unclear that measures of cognitive ability, language, and ASD diagnostic symptoms are adequately sensitive methods for measuring symptom frequency, intensity, and impairment in children with ASD. 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.

Given that the treatment process for ASD is typically intensive and requires highly specific and well-trained individuals to deliver with fidelity, questions of feasibility and accessibility are pertinent but largely understudied. Our understanding of treatment impact and implementation would be greatly enhanced by research that explicitly evaluates which treatments have the greatest real-world impact. Similarly, evaluations of interventions delivered by community providers are important for comparing effects of such approaches with those of interventions delivered in controlled research environments. Such evaluations are complicated by the complexity of community systems and methodologic challenges, including creating similar treatment and control groups and maintaining fidelity. However, they will be increasingly valuable for scaling intervention for ASD. Also important in addressing this gap is improving our currently limited understanding of the effects of provider training and provider characteristics on outcomes of treatment.

Finally, this literature lacks studies that directly compare interventions or employ combinations of interventions (e.g., comparing medical interventions with behavioral interventions, with educational interventions, or with allied health interventions), despite the fact that most children receive multiple concurrent treatments.

Conclusions

In sum, a growing evidence base suggests that behavioral interventions are associated with positive outcomes for some children with ASD. Despite improvements in the quality of the included literature, a need remains for studies of interventions across settings and continued improvements in methodologic rigor. Substantial scientific advances are needed to enhance our understanding of which interventions are most effective for specific children with ASD and to isolate the elements or components of interventions most associated with effects.

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. Fifth ed., Washington DC: American Psychiatric Association; 2013.
2. Zwaigenbaum L, Bryson S, Lord C, et al. Clinical assessment and management of toddlers with suspected autism spectrum disorder: insights from studies of high-risk infants. Pediatrics. 2009 May;123(5):1383-91. PMID: 19403506.
3. Myers SM, Johnson CP. Management of children with autism spectrum disorders. Pediatrics. 2007 Nov;120(5):1162-82. PMID: 17967921.
4. Myers SM. Management of autism spectrum disorders in primary care. Pediatr Ann. 2009 Jan;38(1):42-9. PMID: 19213293.
5. Shattuck PT, Seltzer MM, Greenberg JS, et al. Change in autism symptoms and maladaptive behaviors in adolescents and adults with an autism spectrum disorder. J Autism Dev Disord. 2007 Oct;37(9):1735-47. PMID: 17146700.
6. McGovern CW, Sigman M. Continuity and change from early childhood to adolescence in autism. J Child Psychol Psychiatry. 2005 Apr;46(4):401-8. PMID: 15819649.
7. Fecteau S, Mottron L, Berthiaume C, et al. Developmental changes of autistic symptoms. Autism. 2003 Sep;7(3):255-68. PMID: 14516059.
8. Seltzer MM, Shattuck P, Abbeduto L, et al. Trajectory of development in adolescents and adults with autism. Ment Retard Dev Disabil Res Rev. 2004;10(4):234-47. PMID: 15666341.
9. Warren Z, Veenstra-VanderWeele J, Stone W, Bruzek JL, Nahmias AS, Foss-Feig JH, Jerome RN, Krishnaswami S, Sathe NA, Glasser AM, Surawicz T, McPheeters ML. Therapies for Children With Autism Spectrum Disorders. Comparative Effectiveness Review No. 26. (Prepared by the Vanderbilt Evidence-based Practice Center under Contract No. 290-2007-10065-I.) AHRQ Publication No. 11-EHC029-EF. Rockville, MD: Agency for Healthcare Research and Quality; April 2011. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

10. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality; January 2014. Chapters available at www.effectivehealthcare.ahrq.gov.

11. Bolte EE, Diehl JJ. Measurement tools and target symptoms/skills used to assess treatment response for individuals with autism spectrum disorder. J Autism Dev Disord. 2013 Nov;43(11): 2491-501. PMID: 23479074.

Full Report

This executive summary is part of the following document: 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.

AHRQ Pub. No. 14-EHC036-1-EF
August 2014