ECRI Institute
The Discipline of Science. The Integrity of Independence.

ECRI Institute
5200 Butler Pike
Plymouth Meeting, PA 19462
Phone: (610) 825-6000
Fax: (610) 834-1275

Comprehensive Educational and Behavioral Interventions for the Treatment of Autism Spectrum Disorders

Full In-Depth Health Care Technology Assessment (CLIN 3002)

Contract No. H94002-05-D-0003
Task Order No. 26

November 25, 2008

Prepared for:
Department of Defense
TRICARE Management Activity
Aurora, Colorado

Policy Statement

This report was prepared by ECRI Institute's Health Technology Assessment Information Service (HTAIS) under contract to TRICARE Management Activity (Contract No. H94002-05-D-0003). ECRI Institute is an independent, nonprofit health services research agency and a Collaborating Center for Health Technology Assessment of the World Health Organization. ECRI Institute has been designated an Evidence-based Practice Center (EPC) by the U.S. Agency for Healthcare Research and Quality. ECRI Institute's mission is to provide information and technical assistance to the healthcare community worldwide to support safe and cost-effective patient care. The results of ECRI Institute's research and experience are available through its publications, information systems, databases, technical assistance programs, laboratory services, seminars, and fellowships. The purpose of this technology assessment is to provide information for policy makers on the current state of knowledge on this topic. It is not intended as instruction for medical practice, or for making decisions regarding individual patients.

November 25, 2008

Ms. René Morrell
Contracting Officer's Representative
Department of Defense
TSO/TRICARE Management Activity (CMP)
16401 E. Centretech Parkway
Aurora, CO 80011-9043

Re: Contract No. H94002-05-D-0003
Delivery Order No. 26
Task Order No. 26
Full In-Depth Health Technology Assessment Report
Comprehensive Educational and Behavioral Interventions for the Treatment of Autism Spectrum Disorders

Dear Ms. Morrell:

ECRI Institute is pleased to provide the report *Comprehensive Educational and Behavioral Interventions for the Treatment of Autism Spectrum Disorders*, pursuant to the contract and delivery order cited in the subject line of this letter.

We trust you will find that this report conforms to TRICARE's specifications and meets with your satisfaction.

If we can be of further assistance or if you have any questions regarding this report, please contact me at (610) 825-6000, ext. 5337.

Sincerely,

Karen Schoelles, M.D., S.M.
Medical Director

/ldd

Enclosure

cc: V. Coates (ECRI Institute)
D. Downing (ECRI Institute)
PROJECT FILE (ECRI Institute)

Table of Contents

Tables

Figures

Summary of Findings

Autism spectrum disorders (ASDs, also known as pervasive developmental disorders), refer to a wide continuum of associated cognitive and neurobehavioral disorders, including, but not limited to, three core-defining features: impairments in socialization, impairments in verbal and nonverbal communication, and restricted and repetitive patterns of behaviors. Within the *Diagnostic and Statistical Manual of Mental Disorders-Fourth Edition (DSM-IV) and Text Revised Edition (DSM-IV-TR),* ASDs are divided into five specific diagnostic categories— autistic disorder, Asperger's disorder, Rett's disorder, childhood disintegrative disorder, and pervasive developmental disorder, not otherwise specified (PDD-NOS). Data from a population-based, multisite surveillance study conducted by the Centers for Disease Control and Prevention's Autism and Developmental Disabilities Monitoring Network (ADDM) indicated that in 2002 the prevalence of ASD in the United States per 1,000 children aged 8 years ranged from 3.3 (95% confidence interval [CI] 2.7 to 3.9) in Alabama to 10.6 (95% CI 9.5 to 11.5) in New Jersey. The overall mean prevalence was 6.6 (95% CI 6.3 to 6.8)

A number of comprehensive treatment programs are available for children with ASDs. Unlike focal or targeted interventions that aim to modify one or more specific behaviors or deficits, comprehensive programs seek to simultaneously address most or all of the symptoms of ASDs. In general, comprehensive programs use a combination of interventions which collectively target education and skill development as well as problematic behavior. Most comprehensive programs emphasize early intervention (beginning at diagnosis, which for most children occurs after the age of two) and the importance of individualizing interventions in a manner that meets the needs of each child and family. Other similarities across programs include specific curriculum content, highly supportive teaching environments and generalization strategies, highly trained staff, predictable routines, and active family involvement. Further, most comprehensive programs involve intensive hours of treatment (usually more than 15 hours per week) delivered over a long period of time (one or more years).

What distinguishes one comprehensive treatment program from another is its theoretical orientation, with some being behaviorally oriented and others being developmentally oriented. In brief, behavioral approaches, such as the University of California at Los Angeles (UCLA) Young Autism Program, uses strategies and procedures derived from applied behavior analysis (ABA) in a systematic manner to produce observable and socially significant changes in a child's behavior and skills. Developmental programs, such as the Denver Model and the Developmental, Individual-differences, Relationship model (DIR model), organize a child's environment to encourage or facilitate communicative and social interactions. Developmental programs are child-directed in that the child initiates interaction and the adult responds. In most developmentally-oriented programs, play is a primary vehicle for learning social, emotional, communicative, and cognitive skills. Other comprehensive programs, such as the Division for the Treatment and Education of Autistic and Communication Handicapped Children (DivisionTEACCH) and the Social Communication, Emotional Regulation, and Transactional Supports (SCERTS) model, are considered mixed because they incorporate both developmental and behavioral procedures.

This report addresses five key questions that pertain to the efficacy and safety of comprehensive educational and behavioral interventions for the treatment of autism spectrum disorders:

1. Does any comprehensive educational or behavioral intervention improve outcomes for children with ASD when compared to no treatment, waitlist control, or standard care (e.g., special/supported public education and/or a mix of paramedical services, such as speech and occupational therapy)?

2. Is one comprehensive educational or behavioral intervention more effective than another in improving outcomes for children with ASD?

3. Are home-based interventions of similar intensity and/or structure as comprehensive educational or behavioral interventions provided in other settings (e.g., center or clinical setting) more effective in improving outcomes for children with ASD?

4. What adverse events and harms have been reported to occur in association with the use of comprehensive educational or behavioral interventions for children with ASD?

5. What is the consensus among experts about the safety and efficacy of comprehensive educational or behavioral interventions for the treatment of children with ASD?

We based the answers to Key Questions 1, 2, 3, and 4 on a systematic review of data from clinical studies, whereas Key Question 5 is based on the expert opinion of professional societies. In answering these questions, we provide two ratings of the evidence, one for the evidence underlying our qualitative conclusions (which answer the question "Does it work?"), and one for the evidence underlying our quantitative conclusions (which answer the question "How well does it work?"). We express the ratings for evidence underlying qualitative conclusions as the strength of the evidence, and the ratings for the evidence underlying quantitative conclusions as the stability of the evidence. The following table presents the ratings we use and the definitions of each relevant term.

Table 1. Definitions of Strength and Stability of Evidence

Strength of Evidence Rating	Interpretation
Qualitative Conclusion (Direction of Effect)	
Strong Evidence	Evidence supporting the qualitative conclusion is convincing, making it highly unlikely that new evidence will lead to a change in this conclusion.
Moderate Evidence	Evidence supporting the qualitative conclusion is somewhat convincing. However, a small chance exists that new evidence will overturn or strengthen our conclusion. Regular monitoring of the relevant literature is recommended at this time.
Weak Evidence	Although some evidence supports the qualitative conclusion, this evidence is tentative and perishable. A reasonable chance exists that new evidence will overturn or strengthen our conclusions. Frequent monitoring of the relevant literature is recommended at this time.
Insufficient	The available evidence that exists is not of sufficient strength to warrant drawing an evidence-based conclusion. Frequent monitoring of the relevant literature is recommended at this time.

Strength of Evidence Rating	Interpretation
Quantitative Conclusion (Magnitude of Effect)	
High Stability	The estimate of effect size in the conclusion is stable, making it highly unlikely that the magnitude of this estimate will substantially change as a result of the publication of new evidence.
Moderate Stability	The estimate of effect size in the conclusion is somewhat stable. However, a small chance exists that the magnitude of this estimate will substantially change as a result of the publication of new evidence. Regular monitoring of the relevant literature is recommended at this time.
Low Stability	The estimate of effect size in the conclusion is likely to be unstable. A reasonable chance exists that the magnitude of this estimate will substantially change as a result of the publication of new evidence. Frequent monitoring of the relevant literature is recommended at this time.
Unstable	Estimates of the effect size are too unstable to allow a quantitative conclusion to be drawn at this time. Frequent monitoring of the relevant literature is recommended.

A summary of the findings for each of the five questions we addressed are presented below. For Key Question 1 through 3, we considered the following outcomes: cognitive/intellectual status, language/communication skills, adaptive behavior, problem behaviors, academic/developmental achievement, and parental/family well-being. The studies that met the study selection criteria for this report used various instruments to measure these outcomes.

Key Question 1: *Does any comprehensive educational or behavioral intervention improve outcomes for children with ASD when compared to no treatment, waitlist control, or standard care?*

Three non-randomized controlled studies enrolling a total of 128 children with a diagnosis of autistic disorder or PDD-NOS addressed this question. Each study compared the efficacy of early intensive behavioral intervention (EIBI) to standard care (e.g., less intensive special education). In all three studies, children in the EIBI group received 20 to 40 hours of treatment per week, whereas children in the standard care group received 15 hours or less of treatment per week. The median quality assessment score for the studies was moderate primarily due to lack of randomization, which introduces the potential of selection bias. To enhance group comparability, all three studies employed either matching strategies or used statistical methods to control for any between-group differences observed at baseline.

ECRI Institute's Conclusions for Key Question 1

- **After one year of treatment, children with ASD who receive early intensive behavioral intervention score higher on tests of IQ than children who receive standard care. Estimated effect size is a standardized mean difference (SMD) of 0.750 (95% confidence intervals [CI] 0.302 to 1.199, p <0.001), which corresponds to a between-group difference of 14.8 points in overall IQ. Strength and Stability of Evidence: Moderate.**

- **The evidence was insufficient to determine whether children with ASD who receive early intensive behavioral intervention continue to demonstrate higher scores on tests of IQ than children who receive standard care at later followup times (greater than one year).**
- **Children with ASD who receive early intensive behavioral intervention are more likely to achieve an IQ score within normal range for typically developing children (85 or higher) than children who receive standard care. The estimated size of the effect is an odds ratio of 2.616 (95% CI 1.160 to 5.902, p = 0.021). Strength of Evidence: Moderate and Stability of Evidence: Low.**
- **After one year of treatment, children with ASD who receive early intensive behavioral intervention perform more adaptive behaviors as indicated by higher scores on the Vinland Adaptive Behavior (VAB) Composite Scale than children who receive standard care. Estimated effect size is a SMD of 0.952 (95% CI 0.507 to 1.400, p <0.001), which corresponds to a between-group difference of 10.7 points. Strength and Stability of the Evidence: Moderate.**
- **The evidence was insufficient to determine whether children with ASD who receive early intensive behavioral intervention continue to perform more adaptive behaviors than children who receive standard care at later followup times (greater than one year).**
- **The evidence was insufficient to determine whether children with ASD who receive early intensive behavioral intervention perform better on tests of language and communication than children who receive standard care.**
- **For the following outcomes: problem behaviors, academic/developmental achievement, and parental/family well-being, the limited size (one study per outcome) and quality of the evidence prevented us from drawing conclusions about whether early intensive behavioral intervention was more effective than standard care in improving these outcomes for children with ASD.**

Key Question 2: *Is one comprehensive educational or behavioral intervention more effective than another in improving outcomes for children with ASD?*

Three non-randomized controlled studies enrolling a total of 109 children with a diagnosis of autistic disorder or PDD-NOS addressed this question. Each study compared the efficacy of intensive applied behavior analysis (ABA) to an intensive eclectic intervention program. In all three studies, children in the ABA group received between 20 to 40 hours of one-to-one instruction per week, and children in the eclectic group received a mix of interventions, including methods from TEACCH, ABA, and DIR, delivered at the same intensity as the ABA group. The median quality assessment score for the studies was moderate, which was mainly due to lack of randomization. All three studies did, however, attempt to enhance group comparability by either matching participants or using statistical methods to control for differences observed at baseline. In one of the three studies that addressed this question, the children were substantially older (by three years) than the children in the other two studies. Because of the large difference in age, we did not attempt to combine data from this study in any analyses with data from the other two studies.

ECRI Institute's Conclusions for Key Question 2

- **For intellectual/cognitive status, language/communication skills, and adaptive behavior, clinical differences between the studies reporting on these outcomes (children in one study significantly older than children in the other study) prevented us from drawing any conclusions about whether intensive applied behavior analysis was more effective than an intensive eclectic intervention program in improving these outcomes for children with ASD.**
- **For problem behaviors and academic/developmental the limited size (one study per outcome) and quality of the evidence prevented us from drawing conclusions about whether intensive applied behavior analysis was more effective than an intensive eclectic intervention program in improving these outcomes for children with ASD.**
- **None of the studies that addressed this Key Question reported data on parental/family well-being.**

Key Question 3: *Are home-based interventions of similar intensity and/or structure as comprehensive educational or behavioral interventions provided in other settings (e.g., center or clinical setting) more effective in improving outcomes for children with ASD?*

Our searches of the literature did not identify any studies that met our study selection criteria that directly compared one type of treatment setting to another. However, we did identify two randomized controlled trials that compared clinic-directed early intensive behavioral intervention (EIBI) delivered primarily in the home to parent-directed EIBI. In both studies, children in the clinic-directed group received greater than 30 hours of intervention per week delivered by a therapist trained in applied behavior analysis (ABA, primarily using the Lovaas method). However, the type and intensity of treatment delivered in the parent-directed group differed substantially between the studies. In one study, parents were extensively trained in delivering ABA using the Lovaas method, provided their children with about 30 hours of intervention per week, and received a moderate amount of supervision from trained professionals. In the other study, parents received five hours per week of training in methods of ABA based primarily on the principles of the Lovaas method for three to nine months, and were instructed to provide their children with five hours of instruction per week. Throughout the course of this study, children in the parent-directed group were enrolled in special education classes for 10 to 15 hours per week. Because of the differences between the studies in how treatment was delivered in the parent-directed group, we did not attempt to combine data on any of the outcomes reported in these two studies.

ECRI Institute's Conclusions for Key Question 3

- **Differences between the studies comparing clinic-directed EIBI to parent-directed EIBI in terms of how services were delivered in the parent-directed group precluded us from drawing any conclusions about whether clinic-directed EIBI is more effective than parent-directed EIBI for children with ASD.**

Key Question 4: *What adverse events and harms have been reported to occur in association with the use of comprehensive educational or behavioral interventions for children with ASD?*

- **None of the studies that met the study selection criteria for this review reported on adverse events.**

Key Question 5: *What is the consensus among experts about the safety and efficacy of comprehensive educational or behavioral interventions for the treatment of children with ASD?*

ECRI Institute's searches of the National Guideline Clearinghouse™ (NGC)™ and the Healthcare Standards database identified six treatment guidelines published between the years 2000 to present that included recommendations for the use of comprehensive educational and behavioral interventions for children with ASDs.

In general, the published guidelines recommend that treatment involving behavioral interventions, such as ABA, should be initiated when the child is young, include a minimum of 15 to 20 hours per week of one-to-one instruction, be designed to fit the needs of the individual child, and include the family in the planning and provision of services. However, there is overall agreement that the existing evidence supporting one comprehensive intervention over another is limited due to methodological flaws (e.g., lack of randomized controlled trials, small sample sizes, etc) in the published studies. Further, more research is needed to identify 1) the common effective elements of treatment programs, 2) the effects of treatment on children across the full spectrum of autism, 3) the optimal age and IQ range of children who derive the most benefit, 4) the optimal intensity and duration of treatment, and 5) whether gains on outcomes such as IQ translate to improved quality of life for children with ASDs.

ECRI Institute's Overall Conclusions

For Key Question 1, data from three non-randomized studies that compared EIBI for young children with autistic disorder or PDD-NOS to standard care were used to perform eight separate meta-analyses on the following outcomes: intellectual/cognitive functioning (as measured by tests of IQ), number of children reaching within normal levels of IQ (85 or higher), adaptive behavior, language and communication. After one year of treatment, children who received EIBI demonstrated significantly higher performance on tests of IQ and adaptive behavior compared to children who received standard care. Children who received EIBI also were more likely to reach IQ scores within normal range compared to children who received standard care. However, the evidence was insufficient (only two small studies) to determine whether these differences continued at later followup times. Similarly, for language and communication, the evidence was considered insufficient to permit a conclusion due to the small size of the studies that contributed data for these outcomes. For all other outcomes considered (problem behaviors, academic/developmental achievement, and parental/family well-being), the limited size (one study per outcome) and quality of the evidence base precluded us from drawing any conclusions.

The meta-analytic results for both measures of intellectual functioning and adaptive behavior should be interpreted with caution due to the small size of the evidence base, the moderate quality of the studies, and the variability in performance of children on these outcomes. The evidence base for each outcome consisted of only three studies enrolling a total of 128 children that compared EIBI to standard care. The quality of the studies was limited primarily because children in all three studies were not randomly allocated to the study groups. In all the studies children were assigned to one or the other treatment group based on parental preference. While the authors of all three studies tried to enhance group comparability by either matching children

on key variables and/or statistically controlling for any differences between the groups observed at baseline, these methods do not completely eliminate the potential for selection bias.

Further, when considering the results of the meta-analyses for Key Question 1, it is important to keep in mind that they are based on the overall average performance of the children in each group. This means that not all children may have benefitted equally. In an exploratory analysis, the authors of one of the studies that addressed Key Question 1 investigated variables likely to be associated with changes in IQ. In this study, by Remington et al., the authors found that the children who benefitted most from EIBI differed from the children who did not benefit as much along the following baseline characteristics: higher IQ scores, higher mental age, and higher scores on the Vineland Adaptive Behavior Scales, and less reported problem behaviors. Future studies on EIBI should focus on which children benefit the most from this intervention.

For Key Question 2, the evidence from three non-randomized controlled trials each comparing intensive ABA to a comprehensive "eclectic" intervention program was considered. However, the evidence from these studies did not permit us to draw any conclusions about whether one comprehensive intervention was more effective than another for children with ASD. This was because 1) not all of the studies reported data on the same outcomes of interest; 2) for some outcomes, only one study of moderate quality reported data on the outcome; and 3) in one of the three studies, the children were substantially older (by three years) than the children in the other two studies. This age difference, which we considered an important source of clinical heterogeneity, precluded combination of data from this study with data from the other two studies.

For Key Question 3, our searches of the literature did not identify any studies that met the study selection criteria that directly compared one treatment setting to another. However, we did identify two randomized controlled trials that compared clinic-directed early intensive behavior intervention to parent-directed EIBI. Differences between these studies, however, in terms of the level of parent training and intensity and nature of the treatment delivered to children in the parent-directed group, precluded the formation of any conclusions as to whether clinic-directed EIBI is more effective than parent-directed EIBI for children with ASD.

Preface

Organization of This Report

There are six major sections in this report: 1) *Overview*, 2) *Key Questions and Outcomes Assessed*, 3) *Methods*, 4) *Synthesis of Results*, 5) *Economic and Regulatory Issues*, and 6) *Conclusions*. In the *Overview* section, we provide background information about the health condition or illness under evaluation, including details about its epidemiology, diagnosis, and treatment. This includes background information on other procedures used for diagnosing the condition or illness, and details about the specific intervention(s) evaluated in this report. The final parts of the *Overview* section address previous systematic reviews and meta-analyses of studies of this technology. This background material supports the *Key Questions and Outcomes Assessed*. The questions were developed in consultation with TRICARE; and the section on *Key Questions* explains the rationale for each question and the type of evidence that can answer it.

The *Methods* section details how we identified and analyzed information for this report. It covers our literature searches, criteria for including studies in our analysis, evaluation of study quality, assessment of the strength of the evidence base for each question, and methods for abstracting and synthesis of clinical study results. The *Methods* section provides a synopsis of these activities. Specific details of literature searches, study quality and evidence strength measurement, and statistical approaches (understanding of which is not necessary for understanding the findings of this technology assessment) are documented in appendices.

The *Synthesis of Results* section of this report is organized by Key Question. For each question, we report the quality and quantity of the studies that provided relevant evidence. Then we summarize the results of the reported clinical studies that met our criteria for analysis. Detailed results from each included study are found in evidence tables in Appendix D. Each subsection closes with our evidence-based conclusions on the Key Question.

In the *Economic and Regulatory Issues* section, we provide information on the manufacturers of devices or technologies used in the studies analyzed for this assessment. Where available, we also provide cost information for the device. We include information on whether the technology is regulated by the U.S. Food and Drug Administration (FDA) and, if so, the status of the technology in the FDA market clearance/approval process. We provide information on health insurance coverage for the technology under evaluation. This includes a discussion of the coverage policies of Medicare, Medicaid, and other third party payers.

This report ends with a *Conclusions* section that briefly summarizes the answers to the questions addressed in it, and summarizes other important information that was presented in other sections.

Scope

This report evaluates the efficacy of comprehensive educational and behavioral interventions for the treatment of autism spectrum disorders (ASD). For this report, comprehensive educational or behavioral interventions are defined as interventions that use multiple treatment strategies to address most or all of the deficits/symptoms associated with ASD. The use of these interventions to treat conditions other than ASD is outside the scope of this report, as are other forms of treatment for ASD, such as focal interventions (e.g., discrete trial training, pivotal response training), pharmacological or dietary interventions or any other treatment that aims to have a

physiological effect, auditory or sensory integration, surgical interventions, special education, paramedical therapies (e.g., physical, occupational, or speech therapy), or alternative treatments (e.g., music therapy, massage therapy). These other forms of treatment are only considered when they are being directly compared within a study that meets the study selection criteria for this review to comprehensive educational or behavioral interventions.

In part, this report serves to update a previous report published by ECRI Institute in 1999 titled *Comprehensive Programs for the Treatment of Children with Autism*. The results and conclusions of the previous report will be discussed in this report. However, studies reviewed in the previous report will only be included as part of the evidence base for this review if they meet the study selection criteria (See the *Methods* section of this report for a detailed list of the study selection criteria) for this review. This report also serves to complement another report currently being produced by ECRI Institute titled *Educational/Behavioral Focal Interventions for the Treatment of Autism Spectrum Disorders.*

Overview

In this section, we provide background information on autism spectrum disorders and comprehensive educational and behavioral interventions used to treat ASD. Although this background information is necessary for understanding the evidence discussed later in this assessment, it is based largely upon opinion, and ECRI Institute has not critically assessed its accuracy. This section of the assessment is therefore not evidence-based, and no statement in this *Overview* section should be interpreted as an endorsement or a criticism by ECRI Institute. The "*Methods*" section begins the evidence-based section of the report.

Autism Spectrum Disorders

Autism spectrum disorders (ASDs, also known as pervasive developmental disorders) refer to a wide continuum of associated cognitive and neurobehavioral disorders, including, but not limited to, three core-defining features: impairments in socialization, impairments in verbal and nonverbal communication, and restricted and repetitive patterns of behaviors.(1) Within the *Diagnostic and Statistical Manual of Mental Disorders-Fourth Edition (DSM-IV) and Text Revised Edition (DSM-IV-TR),* ASDs are divided into five specific diagnostic categories—autistic disorder, Asperger's disorder, Rett's disorder, childhood disintegrative disorder, and pervasive developmental disorder, not otherwise specified (PDD-NOS).(2)

While all children with ASDs demonstrate similar core features, the severity of impairments, age of onset, and associations with other disorders (e.g., mental retardation, specific language delay, and epilepsy) vary considerably.(3) Further, manifestations of ASDs vary across children and within a child over time. According to a report published by the National Research Council, despite strong and consistent commonalities, there is no single behavior that is always typical of autistic disorder or of any of the other ASDs and no behavior that would automatically exclude an individual child from diagnosis of autistic spectrum disorder.(4) Below, we briefly describe each of the diagnostic categories of ASD.

Autistic Disorder

The DSM-IV criteria for the diagnosis of autistic disorder (AD) are presented in Table 2. According to the DSM-IV, the essential features of AD are the presence of "markedly abnormal or impaired development in social interaction and communication, and a markedly restricted repertoire of activities and interests."(2) To meet the criteria for AD, a child must demonstrate at least six of the symptoms listed in Table 2, with at least two coming from criterion 1 and one coming from criterion 2 through 4. Further, at least one symptom must have been present before the child's third birthday.

Table 2. The DSM-IV Diagnostic Criteria for Autistic Disorder

Criterion	Description of Symptoms
Criterion 1	Qualitative impairment in social interaction, as manifested by at least two of the following:
A	Marked impairment in the use of multiple nonverbal behaviors, such as eye-to-eye contact, facial expression, body postures, and gestures to regulate social interaction.
B	Failure to develop peer relationships appropriate to developmental level.
C	Lack of spontaneously seeking to share enjoyment, interests, or achievements with other people.
D	Lack of social or emotional reciprocity.
Criterion 2	Qualitative impairments in communication as manifested by at least one of the following:
A	Delay in, or total lack of, the development of spoken language (not accompanied by an attempt to compensate through alternative modes of communication such as gesture or mime).
B	In individuals with adequate speech, marked impairment in the ability to initiate or sustain a conversation with others.
C	Stereotyped and repetitive use of language or idiosyncratic language.
D	Lack of varied, spontaneous make-believe play or social imitative play appropriate to developmental level.
Criterion 3	Restricted repetitive and stereotyped patterns of behavior, interests, and activities, as manifested by at least one of the following:
A	Encompassing preoccupation with one or more stereotyped and restricted patterns of interest that is abnormal either in intensity or focus.
B	Apparently inflexible adherence to specific, nonfunctional routines or rituals.
C	Stereotyped and repetitive motor mannerisms.
D	Persistent preoccupation with parts of objects.
Criterion 4	The delays in normal functioning must have been manifest in at least one of the following areas with onset prior to age 3 years:
A	Social interaction.
B	Language as used in social communication.
C	Persistent preoccupation with parts of objects.
Criterion 5	The disturbance cannot be better accounted for by any other disorder.

Adapted from the *DSM-IV*.(2)

According to the DSM-IV, impairment in reciprocal social interaction is "gross" and "sustained."(2) Children with AD might display marked impairments in the use of multiple nonverbal behaviors that normally act to regulate social interaction and communication, such as eye-to-eye contact, facial expression, body postures, and gestures (Criterion 1A). They might fail to develop peer relationships appropriate to their developmental level (Criterion 1B). Children with AD might lack the normal behavior of spontaneously seeking to share enjoyment, interests, or achievements with others (Criterion 1C). For example, normal children will usually show or point out an object that they find interesting to other people, whereas a child with autistic disorder might not. Children with AD may also lack social or emotional reciprocity (Criterion 1D). For example, a child with autistic disorder might not actively participate in simple social play or games, preferring solitary activities only involving others as tools or mechanical aids to their own play.

Impairments in communication are also "gross" and "sustained" with both verbal and nonverbal skills being affected.(2) Children with AD may demonstrate a delay or a total lack of development of the spoken language (Criterion 2A). In children who are not mute, there may be an impairment in their ability to initiate or sustain a conversation with others (Criterion 2B), or they might engage in stereotyped and repetitive use of language (Criterion 2C). Children with AD may also lack varied, spontaneous make-believe play or social imitation appropriate to their developmental level (Criterion 2D).

Children with autistic disorder typically demonstrate restricted repetitive and stereotyped patterns of behavior, interests, and activities. This may manifest itself in one (or more) of four ways. There may be an all-encompassing preoccupation with one or more stereotyped and restricted patterns of interest (Criterion 3A). For example, a child with autistic disorder may appear to be preoccupied with one very narrow interest, such as collecting information about bus schedules. The child may also demonstrate an apparently inflexible adherence to a specific, nonfunctional routine or ritual (Criterion 3B) that might, for example, result in catastrophic consequences when the bus schedule is changed. Children with AD may demonstrate stereotyped and repetitive motor mannerisms, which might include clapping the hands or rocking the body back and forth (Criterion 3C). Finally, a child with AD may demonstrate a persistent preoccupation with particular parts of objects such as a button or parts of their own body (Criterion 3D).

Asperger's Disorder

The DSM-IV diagnostic criteria for Asperger's disorder are presented in Table 3. A diagnosis of Asperger's disorder applies to those children who demonstrate at least three autistic-like deficits without demonstrating a delay in language development or an important cognitive deficit.(2) Two of these deficits must manifest as impairments in sociability and one must present as impairment in the range of the individual's interests and activities. Individuals with Asperger's disorder typically have a normal IQ (>70), but are socially awkward, pedantic, and preoccupied with narrow interests, such as memorization of lists.

Table 3. The DSM-IV Diagnostic Criteria for Asperger's Disorder

Criterion	Description of Symptoms
Criterion 1	Qualitative impairment in social interaction, as manifested by at least two of the following:
A	Marked impairment in the use of multiple nonverbal behaviors, such as eye-to-eye gaze, facial expression, body postures, and gestures to regulate social interaction.
B	Failure to develop peer relationships that are appropriate to developmental level.
C	A lack of spontaneously seeking to share enjoyment, interests, or achievements with other people.
D	Lack of social or emotional reciprocity.
Criterion 2	Restricted repetitive and stereotyped patterns of behavior, interests, and activities, as manifested by at least one of the following:
A	Encompassing preoccupation with one or more stereotyped and restricted patterns of interest that is abnormal either in intensity or focus.
B	Apparently inflexible adherence to specific, nonfunctional routines or rituals.
C	Stereotyped and repetitive motor mannerisms.
D	Persistent preoccupation with parts of objects.
Criterion 3	The disturbance causes clinically significant impairment in social, occupational, or other areas of functioning.
Criterion 4	There is no clinically significant general delay in language.
Criterion 5	There is no clinically significant delay in cognitive development or in the development of age-appropriate self-help skills, adaptive behavior, and curiosity about the environment in childhood.
Criterion 6	Criteria are not met for any other disorder.

Adapted from the *DSM-IV*.(2)

The validity of Asperger's disorder as a discrete diagnostic entity distinct from high-functioning (verbal) autistic disorder remains controversial.(1) Many clinicians have used the term Asperger's disorder loosely to refer to all children with autistic disorder who show normal to high intelligence. While a consensus is beginning to emerge that the two conditions are more similar than different, the DSM-IV, as currently written, indicates that if criteria for autistic disorder are met, a diagnosis of Asperger's disorder is precluded.

Rett's Disorder

The DSM-IV criteria for Rett's disorder are presented in Table 4. Rett's disorder is found only in girls.(2) Children with Rett's disorder develop normally until approximately six months of age, when developmental delays and regression occur. Affected children typically exhibit reduced muscle tone, autistic-like behavior, stereotyped hand movements consisting mainly of wringing and waving, loss of purposeful use of the hands, a lag in brain and head growth, gait abnormalities, and seizures.

Recently, a gene was isolated on the X chromosome, MECP2, which appears responsible for most cases of Rett's disorder.(3) While boys can inherit the mutation, they display an X-linked mental retardation syndrome that does not include the symptoms associated specifically with the Rett's phenotype (e.g., unsteady gait, lack of language, lack of functional hand use, and stereotyped hand movements).

Table 4. The DSM-IV Diagnostic Criteria for Rett's Disorder

Criterion	Description of Symptoms
Criterion 1	All of the following:
A	Apparently normal prenatal and perinatal development.
B	Apparently normal psychomotor development through the first 5 months after birth.
C	Normal head circumference at birth.
Criterion 2	Onset of all the following after a period of normal development:
A	Deceleration of head growth between ages 5 and 48 months.
B	Loss of previously acquired purposeful hand skills between ages 5 and 30 months with the subsequent development of stereotyped hand movements.
C	Loss of social engagement early in the course of development.
D	Appearance of poorly coordinated gait or trunk movements.
E	Severely impaired expressive and receptive language development with severe psychomotor retardation.

Adapted from the *DSM-IV*.(2)

Childhood Disintegrative Disorder

The DSM-IV diagnostic criteria for childhood disintegrative disorder (CDD) are presented in Table 5. The diagnosis of disintegrative disorder applies to children who demonstrate normal early development, including the development of language for at least the first two years of life. Then, between the ages of two and ten years, they undergo behavioral and cognitive regression that results in severe autism and mental retardation. The period of regression typically lasts four to eight weeks and is marked by agitation and panic on the part of the child. Childhood disintegrative disorder can occur in either boys or girls, but is much more common in boys. Unlike typical autistic disorder, children with CDD display very little developmental growth after treatment and the condition continues as a chronic, severe developmental disability. Many researchers suspect that CDD is a distinct neurodegenerative disorder with a very different etiology from autistic disorder.(3)

Table 5. The DSM-IV Diagnostic Criteria for Childhood Disintegrative Disorder

Criterion	Description of Symptoms
Criterion 1	Apparently normal development for at least 2 years after birth as manifested by age appropriate verbal and nonverbal communication, social relationships, play, and adaptive behavior.
Criterion 2	Clinically significant loss of previously acquired skills (before the age of 10 years) in at least two of the following areas:
A	Expressive and receptive language.
B	Social skills and adaptive behavior.
C	Bowel and bladder control.
D	Play.
E	Motor skills.
Criterion 3	Abnormalities of functioning in at least two of the following:
A	Quantitative impairment in social interactions.
B	Quantitative impairments in communication.
C	Restricted, repetitive, and stereotyped patterns of behavior, interests, and activities, including motor stereotypies and mannerisms.
Criterion 4	The disturbance is not better accounted for by any other disorder.

Adapted from the *DSM-IV*.(2)

PDD-NOS

A DSM-IV diagnosis of PDD-NOS is applied to those children who demonstrate severe impairments in sociability, language, and range of activities and who do not meet the DSM-IV diagnostic criteria for any of the other ASDs, schizophrenia, schizotypal personality disorder, or avoidant personality disorder. PDD-NOS is a diagnosis by exclusion of the other autistic spectrum disorders. For example, a diagnosis of PDD-NOS would be given to a child who does not meet the six of possible 12 criteria for the diagnosis of autistic disorder, or who had symptom onset after the age 36 months.(1) Also, children whose symptoms are atypical or not as severe would be coded under this diagnosis.

Associated Disorders

In addition to the behavioral deficits described by the DSM-IV (and the ICD-10), children with autism spectrum disorders often present with a variety of other developmental disorders, medical conditions, and behavioral problems. The most commonly co-occurring developmental disorder is mental retardation. Approximately 75% of children diagnosed with ASD have an associated diagnosis of mental retardation (IQ <70), with roughly half of this group functioning at the range of mild to moderate mental retardation and half in the severe to profound range.(3,5) The degree of mental retardation appears to be highly correlated with the severity of autistic symptoms.

Children with ASD are also at risk for developing seizure disorders throughout the developmental period.(4) The incidence of seizures in children with ASD has been estimated to be three to 28 times higher than that found in the general population, with the prevalence being highest among those with mental retardation or motor deficits.(6,7) Seizure disorders in ASDs are of various types and may sometimes present in unusual ways, such as staring spells, cessation of activity, or aggressive escalations.(1) The most prevalent type of seizure appears to be partial complex seizures, with electrophysiological testing showing abnormalities occurring most often over the temporal lobes.(1) Other medical conditions that may co-occur with ASDs include metabolic disorders, Angelman syndrome, fragile X syndrome, and tuberous sclerosis.(1,3,4,6,7)

Other comorbid behavioral problems associated with autistic spectrum disorders include fear/phobias, sleeping and eating disturbances, Tourette syndrome and other tic disorders, hyperactivity, inattentiveness, aggressiveness, self-injurious behavior, and obsessive-compulsive behavior.(1,3,4) Abnormal responses to sensory stimuli such as loud sounds, oversensitivity to light touch, fascination with certain visual stimuli, and insensitivity to pain are also often seen in children with ASD. Additionally, disorders of mood and affect may be present, manifesting as laughing or crying for no apparent reason, lack of, or excessive fearfulness, generalized anxiety, temper tantrums, and decreased or absent emotional reaction.(3)

Diagnostic Strategies

The diagnosis of autistic spectrum disorders can be challenging. According to the National Research Council, complexities in diagnosis and evaluation relate to the range of syndrome expression in these conditions along various dimensions such as language abilities and associated mental handicap.(4) Other factors such as differential diagnosis, concerns with labeling, diagnostic terminology, and lack of expertise in assessment and diagnosis can add to the challenge. Generally, the diagnosis of ASD is carried out by a multidisciplinary team of experts,

which may include pediatricians, psychologists, psychiatrists, neurologists, speech pathologists, occupational and physical therapists, and special and general educators. Diagnosis should be based on a careful and comprehensive assessment that includes specific evaluations of language and communication skills, cognitive and adaptive functioning, sensorimotor functions, behavioral deficits, and family functioning and resources.(1) The evaluation should include measures of parental report, child observation and interactions, and clinical judgment. An expanded medical and neurological evaluation should also be conducted to assess for possible co-morbid conditions. Because there is evidence that ASDs have a genetic basis, at least in some cases, details of other family members with ASDs or other mental illnesses, such as manic depressive disorder, should also be recorded.(1)

A number of instruments have been developed to aid in the diagnosis of ASD. The most widely recognized diagnostic instruments include the parent-interview Autism Diagnostic Interview-Revised (ADI-R)(8) and the performance-based Autism Diagnostic Observation Schedule-Generic (ADOS-G).(9) Both of these instruments permit DSM-IV and ICD-10 diagnosis within the autism spectrum, with definitive threshold scores for the diagnosis of autistic disorder.(1) Other diagnostic instruments include the Childhood Autism Rating Scale (CARS), Autism Behavior Checklist, Aberrant Behavior Checklist, Parent Interview for Autism (PIA), Social Communication Questionnaire (SCQ), and the Diagnostic Interview of Social and Communication Disorders (DISCO).(1,4,10)

The symptoms of ASD are often measurable by 18 months of age.(4) While there is still some concern about the reliability and validity of early diagnosis (prior to age three), most clinicians now recognize the potential benefits of early diagnosis.(11) According to Rogers (2001), early recognition helps answer parents' questions about the nature of their child's developmental delay and the implications of this delay in the future, allows for the most appropriate treatment to be selected and delivered, and has been associated with the possibility of better outcomes.(12) Recently, several standardized tests and checklists have been developed to help assist in the early recognition and identification of children with ASD. Such instruments include the Checklist for Autism in Toddlers (CHAT) and modified version (M-CHAT), the Pervasive Developmental Screening Test-II (PDDST-II), the Screening Tool for Autism in Toddlers (STAT), and the Early Screening for Autism questionnaire. Other screening instruments have been developed for undiagnosed older verbal children, including the Australian Scale for Asperger's Disorder, the Autism Spectrum Screening Questionnaire (ASSQ), and the Gillian Asperger's Disorder Scale (GADS).(1,4)

Course and Prognosis

The onset of ASD typically occurs before age three, with the majority of children displaying developmental abnormalities within the first two years of life.(3) Although they are not always recognized at the time, a careful retrospective interview with the parents typically reveals evidence of abnormalities in social responsiveness and early communication behaviors (e.g., baby games and communication gestures). According to Ozonoff and Rogers (2003), a smaller group of children with autistic disorder display a period of normal or mostly normal development, followed by a loss of communication and social skills and onset of autism.(3) The regression generally occurs between 12 and 24 months, thus distinguishing it from childhood disintegrative disorder, in which severe regression occurs after at least two years of normal

development.(3) The causes of the regression are not yet understood. Some believe that it is influenced by environmental factors, while others contend that it is genetically influenced.

Most individuals continue to meet the criteria for ASD as teenagers and adults. Studies of adolescents and adults with autistic disorder have found that some of the symptoms that are associated with autism, such as hyperactivity, self-injurious behavior, compulsivity, and stereotypies, are exacerbated in about 35% of individuals during puberty.(13) In later adolescence and adulthood, abnormalities such as stereotyped motor movements, flat affect, generalized anxiety, and social improprieties are frequently observed, even in high-functioning individuals. In such individuals, social ineptitude and employment can also become acute problems.(13) Adults with severe autistic disorder may develop complex obsessive-compulsive rituals and abnormal speech behaviors, such as idiosyncratic usage, preservation, excessive concreteness, monotonous tone, repetitive questioning, and talking to oneself.(13)

The long-term prognosis for patients with autism, as defined by measures of social adjustment, the ability to work, and the ability to function independently, is poor.(13,14) Based on an assessment of the few available long-term followup studies, Gillberg and Nordin found that 60% to 70% of children with autistic disorder will have "a poor" or "very poor" outcome with regard to social adjustment, and only 5% to 15% of children with autism will experience a "good" outcome.(13) The best single predictor of outcome is IQ,(15-17) with an IQ of <50 at the age of five to six being a strong predictor of a poor prognosis.(13) Another predictor of a poor outcome is the lack of communicative speech at the age of five to six.(13)

Epidemiology

Data from a population-based, multisite surveillance study conducted by the Centers for Disease Control and Prevention's Autism and Developmental Disabilities Monitoring Network (ADDM) indicated that in 2002 the prevalence of ASD in the United States per 1,000 children aged 8 years ranged from 3.3 (95% confidence interval [CI] 2.7 to 3.9) in Alabama to 10.6 (95% CI 9.5 to 11.5) in New Jersey. The overall mean prevalence was 6.6 (95% CI 6.3 to 6.8).(18) To determine the prevalence of ASD, the ADDM collected data on 407,578 children from 14 different states. Children were identified as having ASD through screening and abstraction of evaluation records at health facilities and through psychoeducational evaluations for special education services. Children whose records documented behaviors consistent with the DSM-IV-TR criteria for autistic disorder, pervasive developmental disorder, not otherwise specified, or Asperger's disorder were classified as having ASD. Among the 407,578 children for which data were collected, 2,685 (0.66%) were identified as having ASD. Table 6 lists each state that was surveyed, total number of children identified as having ASD within each state, and individual state prevalence rates. To date, the ADDM's study represents the largest and most complete study on the prevalence of ASD in the United States.

Table 6. Prevalence Surveillance States and Rates of ASD in 2002

State	Total Number of Children	Total Children with ASD	Overall Rate (95% CI)
Alabama	35,472	116	3.3 (2.7 to 3.9)
Arizona	45,113	280	6.2 (5.5 to 7.0)
Arkansas	36,472	251	6.9 (6.1 to 7.8)
Colorado	11,020	65	5.9 (4.6 to 7.5)
Georgia	44,299	337	7.6 (5.5 to 8.5)
Maryland	29,722	199	6.7 (5.8 to 7.7)
Missouri	28,049	205	7.3 (6.4 to 8.4)
New Jersey	29,748	316	10.6 (9.5 to 11.9)
North Carolina	20,725	135	6.5 (5.5 to 7.7)
Pennsylvania	21,051	111	5.3 (4.4 to 6.4)
South Carolina	23,191	140	6.0 (5.1 to 7.1)
Utah	26, 108	196	7.5 (6.5 to 8.6)
West Virginia	21,472	153	7.1 (6.1 to 8.4 0
Wisconsin	35,126	181	5.2 (4.5 to 6.0)

Note: Data for this table were abstracted from the Centers for Disease Control and Prevention's (CDC) Web site (http://www.cdc.gov/mmwr/preview/mmwrhtml.htm).

In addition to overall prevalence rates, the ADDM study also provided information on demographic characteristics of children with ASD. Prevalence of ASD varied to a certain extent by race and ethnicity across states. In ten states, prevalence was higher among non-Hispanic white children than among non-Hispanic black children, but this difference was only statistically significant ($p < 0.05$) for five states.(18) In all states the prevalence was lower for Hispanic children than for non-Hispanic white and non-Hispanic black children. A consistent finding across all states was a significantly ($p < 0.001$) higher prevalence of ASDs among males than among females. Prevalence for males ranged from 5.0 per 1,000 population to 16.8 per 1,000 population, and for females the prevalence ranged from 1.4 per 1,000 population to 4.0 per population. The male-to-female ratio across the various states ranged from 3.4 to 6.5 boys to 1 girl. Finally, the median age of earliest documented ASD ranged from 49 to 66 months.

One of the limitations of the ADDM study is that prevalence rates were not provided for specific diagnostic categories of ASD. A recent survey of the epidemiological literature by Fombonne (2005) found that the reported prevalence was 13 per 10,000 for autistic disorder, 2.6 per 10,000, for Asperger's disorder, and 0.2 per 10,000 for childhood disintegrative disorder.(19) The prevalence of PDD-NOS is estimated to be 20.8 per 10,000. Of the 34 studies reviewed, 30 reported male to female ratio among children with autism. The male to female ratio varied from 1.33 to 16.0, with a mean male to female ratio of 4 to 1.(19) According to Fombonne, none of the studies identified more girls than boys with autism.

The data reported by the ADDM reflects a substantial increase in prevalence of ASD from 1/1,000 in the early 1990s to 1/152 in 2002.(3) However, there is no clear explanation for this apparent increase. The increase most likely reflects changes in the clinical definition of autism, and a greater awareness of autistic behaviors by clinicians, teachers, and parents.(3) Recent surveillance studies, such as the one conducted by the ADDM, now include children that were unlikely to have been previously considered to have autism, such as children with less severe forms of autistic disorder and children with Asperger's disorder. Similarly, children with coexisting mental retardation and autism may now have a primary diagnosis of autism rather than mental retardation. Finally, greater awareness of autism has led to more screening and availability of treatment services in schools and the community, which may also partly explain the increase in prevalence.

Pathology and Etiology

At present, the exact etiology of ASD remains unclear for most affected children. Given the range of symptoms associated with ASD and the heterogeneity of the children affected by the disorder, it is very unlikely that one single etiology will turn out to be responsible. Currently, the most widely accepted belief is that ASD is a biologically-based neurodevelopmental disorder with a strong genetic basis.(20) Evidence for a genetic basis comes from twin studies that show a high concordance for ASD in monozygotic twins and relatively small concordance in dizygotic twins.(3) The most recent studies of twins, which used standardized diagnostic measures and total population screening, found a monozygotic concordance rate of 60% for AD and 93% for the broader spectrum of social and communication deficits with sterotypies.(20) The rates for dizygotic twins were shown to be 0.0% to 5.0% for AD and 10% to 30% for the broader spectrum.

This strong decrease in risk from monozygotic twins to dizygotic twins suggests a polygenic model of inheritance.(20) Recent statistical modeling of the genetics of ASD indicates that at least three (perhaps as many as 20) gene loci contribute to the wide spectrum of symptoms. According to McPhersen et al. (2007), preliminary linkage studies have identified gene markers on chromosomes 1p, 7q, 16p, and 17p, with the highest log of odds score across studies for chromosome 7.(20) Other factors associated with an increased risk of autism are single gene defects or deletions, such as those that cause tuberous sclerosis,(21,22) phenylketonuria, fragile X, Angelman's, and Cornelia de Lange's syndromes; intrauterine exposure to rubella, thalidomide, or valproate; and herpes encephalitis.(3,20)

The common association of ASD with seizures and mental retardation suggests a neurological basis.(3,20) Neuroimaging and autopsy studies have revealed a variety of developmental brain abnormalities. According to Ozonoff et al.(2001), the findings of recent neuroimaging studies have shown deviations from normal in the volume of the hippocampus and amygdale, cerebellum, brainstem, neocortex (particularly the frontal and temporal lobes), and the cerebellar vermis (particularly lobules VI and VII).(20) Postmortem studies of a limited number of individuals with ASD most of whom also had significant mental retardation revealed increased neuronal density in the hippocampus, olivary dysplasia, scattered areas of cortical and white matter dysplasia, and other nonspecific developmental abnormalities in the brainstem and cerebellum.(20) In addition to anatomic abnormalities, quantitative abnormalities have also been found in serotonin, dopamine, opioid, and most recently, γ-aminobutyric acid neurotransmitter transport systems.

Given the phenotypic variability of ASD, even among monozygotic twins (e.g., one twin displays more severe symptoms than the other), it is unlikely that ASD is purely a genetic disorder. A number of environmental factors have been hypothesized to play a role in modulating the autism phenotype.(20) The list of factors include, but is not limited to the following—maternal exposure to mercury, pesticides and other environmental toxins, diet and nutrition, and more recently, vaccines containing the preservative called thimerosal, which is 50% mercury.(4) Research, however, assessing the association of environmental factors with ASD has been largely inconsistent.

Comprehensive Educational and Behavioral Interventions

A number of comprehensive treatment programs are available for children with autistic spectrum disorders. Unlike focal or targeted interventions that aim to modify one or more specific behaviors or deficits, comprehensive programs seek to simultaneously address most or all of the symptoms of ASDs. In general, comprehensive programs use a combination of interventions which collectively target education and skill development as well as problematic behavior.(4) Most comprehensive programs emphasize early intervention (treatment beginning at diagnosis, which for most children occurs after the age of two) and the importance of individualizing interventions in a manner that meets the needs of each child and family.(23) Other similarities include specific curriculum content, highly supportive teaching environments and generalization strategies, highly trained staff, predictable routines, and active family involvement.(23) Further, most comprehensive programs involve intensive hours of treatment (usually more than 15 hours per week) delivered over a long period of time (one or more years).

What distinguishes one comprehensive treatment program from another is its theoretical orientation, with some being behaviorally oriented and others being developmentally oriented. In brief, behavioral approaches, such as the University of California at Los Angeles (UCLA) Young Autism Program, use certain techniques or strategies collectively referred to as Applied Behavior Analysis (ABA) in a systematic manner to produce observable and socially significant changes in a child's behavior and skills.(24) Some of the techniques include chaining, or breaking a task down to its smallest parts; prompting, to encourage the child to respond appropriately; fading, or using the least intrusive prompts to bring about a desired result until prompting is no longer needed; shaping, or gradually modifying inappropriate behaviors; and, finally, providing various levels of positive or negative reinforcement depending on the difficulty of the task.

Unlike behaviorally-oriented programs, most developmental approaches do not rely on a specific set of strategies or techniques to modify behaviors or teach new skills.(4) Instead, developmental programs, such as the Denver Model and the Developmental, Individual-differences, Relationship model (DIR model), organize a child's environment to encourage or facilitate communicative and social interactions.(25) Developmental programs are child-directed in that the child initiates interaction and the adult responds. In most developmentally-oriented programs, play is a primary vehicle for learning social, emotional, communicative, and cognitive skills. Other comprehensive programs, such as TEACCH and the SCERTS model, are considered mixed because they incorporate both developmental and behavioral procedures.(4,26)

In the section below, we briefly describe some of the more widely recognized comprehensive treatment programs. Table 7 summarizes key features of each program described in this section. This section is not intended to serve as an exhaustive list of all available comprehensive programs.

Behaviorally-oriented Programs

The University of California at Los Angeles (UCLA) Young Autism Program

The UCLA Young Autism Program (also known as the Lovaas model of applied behavior analysis) was developed by O.I. Lovaas and colleagues. This early intensive behavioral intervention (EIBI) utilizes strategies and procedures derived from applied behavior analysis (ABA) to treat young children with ASD.(4,24,27-29) The program is intended for children between the ages of two to three years old and is delivered over the course of two to three years.(24) In the initial phases of the program, children receive intensive one-to-one discrete trial training for 40 hours per week, which is implemented by parents and trained therapists in the child's home. Discrete trial training involves repeating a single cycle of a behaviorally-based instruction routine several times in succession, several times a day, over several days (or even longer) until the skill is mastered.(4)

Overall, the UCLA program focuses on developing language and early cognitive skills and decreasing excessive rituals, tantrums, and aggressive behaviors.(4) The first year of intervention is aimed at teaching children to respond to basic requests, to imitate, to begin to play with toys, and to interact with family members. During the second year, the focus is on teaching expressive and early abstract language and interactive play with peers. In more advanced stages of the intervention, the child is taught at home and at school (1) early academic tasks such as beginning

reading, writing, and arithmetic; (2) socialization skills; (3) cause-effect relationships; and (4) to learn by observing other children learn.(30)

During the third year of intervention, the teaching procedures become less structured as the child progresses, and the procedures are generalized to the child's school and everyday environment. According to federal law, children at a chronological age of three years are eligible for services through their local educational system. The services provided to individual children are determined by an Individualized Education Program (IEP) team.(31) Once an appropriate placement has been decided, student aides accompany the child to school, facilitate the transition from home-based to classroom routines, and encourage interaction with other children. Lovaas recommends that the time the child attends school is increased gradually from as little as 30 minutes per day initially to full time (i.e., two to three hours per day), and the student aid is carefully phased out.(24) School hours are included in the total of 30 to 40 hours per week for Lovaas intervention. Parents, teachers, and treatment staff maintain close supervision of the child's progress in school so that success can be maximized, and the chances of the child being ostracized are minimized. This, argues Lovaas, reduces the potential bias of teaching staff in influencing a particular placement over another, which could adversely affect the child.(24)

The Douglass Developmental Center at Rutgers University

The Douglass Developmental Center (also known as the Rutgers Autism Program) is a university-based, early intensive behavioral intervention program that has been serving preschoolers with autistic spectrum disorders since 1987.(4) Douglass has a continuum of three programs including an intensive home-based intervention, a small-group segregated preschool, and an integrated preschool. Like the UCLA program, the curriculum uses ABA techniques, beginning with discrete-trial formats and shifting across the continuum to more naturalistic procedures.(4) Initial instruction is focused on teaching compliance, cognitive and communication skills, and basic social skills, as well as the elimination of serious behavior problems. The small group classroom emphasizes communication, cognitive and self-help skills, and interactive play. The integrated preschool program emphasizes communication, socialization, and pre-academic skills.(4) Children in the Douglass program receive 30 to 40 hours per week of treatment and instruction.

Project DATA (Developmentally Appropriate Treatment for Autism)

Project DATA, which started as a federally-funded model demonstration program at the University of Washington in Seattle, Washington, serves preschool children with autism and their families in inclusive and developmentally-appropriate programs.(32)The program is currently receiving funding through the Seattle public education system. Project DATA consists of five interdependent components: high quality inclusive early childhood program, extended instructional time, collaboration and coordination across services, transition support, and dependent measures. The program utilizes strategies and procedures drawn from ABA and early special education.(32) Some of the specific instructional strategies include incidental teaching, discrete trial training, the Picture Exchange Communication System (PECS), visual support strategies, and effective prompting strategies (e.g., time delay, most to least prompting, and graduated guidance). Children in this program receive 20 hours per week of classroom-based instruction that is provided through a mix of individual instruction and small groups. Social and technical support is offered to families through monthly home-based visits to assist families through problem situations, resource coordination, and monthly parent support and networking

evenings. The program is staffed by head teachers with master's level training in autism, ABA, and early childhood special education and teaching assistants who are graduate students in the special education program at the University of Washington. Additionally, Ph.D. level consultants provide training, family support, and consultation.

The LEAP Program

The Learning Experiences, An Alternative Program for Preschoolers and Parents (LEAP) is a comprehensive preschool service system that was designed to meet the needs of both children with ASD and normally developing children. This program was developed by Philip Strain and his colleagues in Pittsburgh, Pennsylvania, and includes parent training and involvement, along with five-day-a-week, three-hour-a-day classroom instruction (total intervention time is 25 hours/week).(33,34) The LEAP program does not provide intervention on a one-to-one basis. Instead, classroom instruction is provided by a master's degree level teacher and an assistant who implement the LEAP program with ten normally developing children and three to four children with autism at a time.(34) In addition to these special education teachers, a full-time speech therapist and contracted occupational and physical therapists work, directly with the children in their classrooms.(33)

School instruction is based on a method of individualized group instruction known as the Tri-I (Innovative; Integrative; Individualized) curriculum for mainstreaming (TRIIC).(34) TRIIC is a program that involves the implementation of an individualized learning program in an integrated environment. Such an approach deviates from the traditional view that one-to-one instruction is the foremost method for treatment of individuals with ASD. The TRIIC curriculum that is utilized in the LEAP program was specifically designed for the instruction of children who function at different levels. The initial component of the TRIIC program involves frequent child assessment, the results of which dictate the scope and sequence of instruction on a child-by-child basis. Upon entering the intervention program, each child is assessed using the Learning Accomplishment Profile (LAP) instrument.(35) This instrument establishes the child's developmental level based upon which objectives are determined to fall within the child's "functioning range." This assessment is repeated at three-month intervals, which allows the child's individualized curriculum to be revised as necessary.

Walden Early Childhood Programs at the Emory University School of Medicine

The Walden program has a toddler, preschool, and prekindergarten program. The toddler program is both center-and home-based, and focuses teaching children sustained engagement, functional verbal language, responsiveness to adults, tolerance and participation with typical peers, and independence in daily living.(4) The preschool and prekindergarten programs focus on language expressions, peer interaction, and academic skills training. All three components of the Walden program rely on incidental teaching based on behavioral research.

The Autism Preschool Program

This is a collaborative program based at the University of Manitoba that is staffed by a multidisciplinary team and requires collaboration between the university hospital, the provincial government, and local community resources.(30) It is similar to the Douglass Program at Rutgers in that it utilizes a variety of standard behavioral methods, such as ABA.

The Children's Unit at the State University of New York at Binghamton

This program is an intensive, short-term program for children with severe behavioral disorders. The program largely seeks to identify the factors crucial in preventing children from participating in and benefitting from services provided in the community. It primarily uses traditional applied behavior analysis techniques.(4)

Developmentally Oriented Programs

The Denver Model

The Denver Model (sometimes referred to as the Playschool Model) is a developmentally-based program that began in 1981 at the University of Colorado Health Sciences Center. The program was developed as a joint therapeutic and educational program for young children aged two to six who had been diagnosed with ASD, or other severe emotional/behavioral disorders involving deficits in communication, socialization, cognition, and the ability to adjust to change.(36) In 1998, the treatment unit was closed, and the intervention format was changed to the more natural contexts available in home and preschool environments with typical peers.(4) An individualized curriculum is developed for each child, and it is this curriculum that drives all of the child's instruction in all settings. Treatment is typically delivered for about 20 hours per week.

Focus areas of this treatment model include the development of communication and play skills, sensory activities, personal independence, and the reduction of unwanted behavior.(36) The Denver Model utilizes a "multifaceted" approach to the development of communication skills that includes several teaching elements: the development of nonverbal communication through "elicitation and shaping of natural gestures followed by conventional gestures to serve a variety of functions"; teaching motor-imitation skills that are related to language in order to teach children how to imitate other peoples' speech; and, teaching the meaning and importance of speech.(36) Play skills are taught because they are considered to be a crucial element in normal development. It is argued that children with ASD cannot benefit maximally from interactions with other children if they cannot engage in the social core of activities that preschoolers use such as play.(36)

The Denver Model views a child's sensory systems as a "crucial regulator of attention, arousal, and affect."(36) As a result, the treatment program involves the child with ASD in what it terms "sensory-social activities."(36) These are dyadic interactions that involve simple repeated social routines aimed at engaging the child as an active participant. Child independence is also valued in this program model and is encouraged through carrying out routines of daily living, independent play, and independent goal-orientated tasks that contribute to the family. Visual strategies are used as needed to support these activities to maximize independent functioning. To minimize unwanted behavior in children with ASD, the Denver Model utilizes the tools of functional behavioral analysis; communication training; positive teaching of alternative, more conventional behaviors; and redirection to provide the new behavioral strategies by which the child can achieve his or her goals.(36)

The DIR Model

The Developmental, Individual-differences, Relationship model (DIR model, also referred to as the Greenspan Floor-Time Program) is a developmentally-based, one-to-one treatment program

delivered ten to 25 hours per week. The primary intervention method used in this model is intensive interactive "floor-time" play sessions, in which an adult follows a child's lead in play and interaction.(25) The program consists of three components: home-based play sessions, individual therapies, and early education programs. The intense floor-time sessions at home are intended to help children reach what Weider and Greenspan describe as the key elements of early development: self-regulation and shared attention, engagement and relating, two-way intentional communication, problem solving, symbolic and creative use of ideas, and logical and abstract use of ideas and thinking.(25) Parents or caregivers typically participate in the floor-time sessions. The program also involves speech, occupational and physical therapists, educators, and/or psychotherapists all of whom work one-to-one with the child using specialized techniques informed by floor-time principles to deal with the child's specific challenges.(30)

Mixed Programs

The Division TEACCH Program

The Division for the Treatment and Education of Autistic and Communication Handicapped Children (Division TEACCH) was founded in the early 1970s by Eric Schopler and colleagues as a statewide program serving individuals of all ages with ASDs throughout the state of North Carolina. The administrative headquarters of the TEACCH program are in Chapel Hill, North Carolina, and there are currently nine regional centers around the state of North Carolina. Regional centers provide regular consultation and training to parents, schools, preschools, daycare centers, and other placements throughout the state.(4) The TEACCH program has gained national and international recognition, and is now being implemented in many other states and countries. For instance, the EarlyBird Programme in London, UK and the Scottish Centre for Autism preschool treatment program incorporate many aspects of the TEACCH program.(37,38)

The overall objectives of the TEACCH program are to maximize adaptation through the structured teaching of new adaptive skills; to develop environmental modifications to accommodate the child's deficits; to maintain close collaboration between teacher and parent; to provide a continuity of structured teaching throughout the life of the individual with ASD; and to prevent the development of further behavioral problems.(39) TEACCH has developed a communication curriculum that makes use of behavioral procedures, with adjustment that incorporate more naturalistic procedures along with alternative communication strategies for non-verbal children.(4) The program is typically delivered in a school or clinical setting for about 25 hours per week.

The primary technique used in the TEACCH program is structured teaching, in which environments are organized with clear, concrete, visual information. Such structuring provides the child with ASD with continuity that, it is argued, allows them to "understand where to be, what to do, and how to do it, all as independently as possible."(39) In a typical TEACCH classroom, there are clear indications of where each activity will occur to help the student learn to stay in certain areas. For example, work tasks for teaching cognitive, fine motor, eye/hand integration, and organizational skills occur at tables. Self-help skills such as toilet training, eating, washing hands, wiping tables, and hanging up coats are taught in a special self-care area. Expressive communication, receptive language, and social interaction are formally taught in another marked area but also occur as part of other activities.

SCERTS Model

The SCERTS® model prioritizes goals and implements practices that focus on enhancing Social Communication, Emotional Regulation, and Transactional Supports (SCERTS) for children with ASD and related social-communicative disabilities and their families.(26) It is based on research and practice that indicates that educational programming should focus on 1) developing spontaneous, functional communication and secure, trusting relationships with children and adults (Social Communication), 2) enhancing the ability to maintain a well-regulated emotional state for learning and interacting (Emotional Regulation), and 3) supporting children, their families, and professionals to maximize positive social experiences across home, school and community settings (Transactional Support).(26) The SCERTS® model focuses on functional skills in everyday activities across settings, and is informed by research on the unique learning style of children with ASD. It is not an exclusive approach, in that it provides a framework in which practices from other approaches may be integrated. Treatment is typically provided for 20 or more hours per week. The SCERTS® model can be used with children who exhibit a wide range of ages and developmental abilities, including both preverbal and verbal children. It also is relevant for older school-age children and adults. Particular emphasis is given to parent-professional collaboration and careful coordination across all settings and partners.

Table 7. Summarization of Key Features of Comprehensive Programs

Program	Orientation	Mean Age at Entry (range) in Months[1]	Hours Per Week	Primary Setting	Primary Teaching Procedure
UCLA Young Autism Program	Behavioral	32 (30 to 46)	20 to 40	Home	One-to-one Discrete-trial
Douglass Program at Rutgers	Behavioral	47 (32 to 74)	30 to 40	Home and school	One-to-one or small group format using discrete-trial and naturalistic training.
Project DATA	Behavioral	36 (36 to 72)	20	School	One-to-one and small group
Autism Preschool Program	Behavioral	NR	NR	Integrated day care setting, home	Group training using a variety of standard behavioral methods
Children's Unit	Behavioral	40 (13 to 57)	27.5	School	Applied behavior analysis/discrete trial training
LEAP	Behavioral	43 (30 to 60)	25	School, home	Individualized group instruction
Walden	Behavioral	30 (18 to 36)	36	School, home	Incidental teaching

Program	Orientation	Mean Age at Entry (range) in Months[1]	Hours Per Week	Primary Setting	Primary Teaching Procedure
Denver Model	Developmental	46 (24 to 60)	20	School, home, and community	Playschool curriculum
DIR Model	Developmental	36 (22 to 48)	10 to 25	Home or clinic	Mostly one-to-one floor-time therapy
TEACCH	Mixed (Developmental/behavioral)	36 (24 and up	25	School and/or clinic	Structured teaching
SCERTS	Mixed (Developmental/behavioral)	NR	>20	School, home, and community	Multiple methods used including one-to-one training, play, and naturalistic training

[1] Information presented on mean age at entry into the comprehensive program was abstracted from the National Research Councils' Report on *Educating Children with Autism*.(4)

Care Setting

As indicated in Table 7, treatment and instruction provided through the comprehensive programs discussed in this report take place in a variety of settings, including the child's home, daycare center, preschool, and community.

Staff Training for Comprehensive Programs

Most of the comprehensive programs discussed in this report have developed standardized training protocols and supervisory systems to accommodate the training and supervision needs of the staff who provide treatment and instruction.(4) For instance, the UCLA Young Autism Program has packaged both manuals and tapes to standardized staff training. Primary therapists in the UCLA program are typically undergraduates who have completed courses in learning theory and behavior modification and who have worked for a minimum of six months under supervision.(30) Supervisors typically have a Master's degree in psychology and two or more years of experience with the intervention program. In the TEACCH program, each classroom is headed by a clinical psychologist and is staffed by five to seven "psycho-educational therapists." All staff within the TEACCH program are formally trained on the instructional philosophy and techniques used within the program, and are required to have a minimum of two years experience working with individuals with ASDs and related disorders.(30)

In most of the comprehensive programs, part or all of the treatment takes place within the child's home. As such, parents are highly involved and, in some instances, are considered co-therapists. Most of the programs include a parent training component, which typically consists of both group and individual training from the professional staff. Some programs, such as the UCLA program, have developed manuals specifically for parents and teachers.

Competing/Complementary Treatment

In addition to the comprehensive programs described above, a number of other treatment options are currently available for children with ASDs. These include pharmacotherapy and complementary or alternative treatments, such as chiropractic manipulation, sensory and auditory integration, facilitated communication, hyperbaric oxygen therapy, dietary interventions; exercise; and surgical procedures. In addition, most children with ASDs are eligible to receive special education, physical, occupational, and/or speech therapy. Finally, a number of focal behaviorally- or educationally-based interventions are available to treat either the core deficits of ASDs or associated symptoms (e.g., anxiety, depression, and/or anger). Although these treatments are not the focus of this report, we briefly describe some examples of each below.[1] This list, however, is not exhaustive of all the competing/complementary treatments available for children with ASD.

Pharmacotherapy

The three major classes of psychotropic agents that have demonstrated efficacy in ASD in open label or placebo controlled trials and are widely used today include atypical neuroleptics (antipsychotics), antidepressants, and psychostimulants/alpha-adrenergic agonists.(3) Of these, the psychostimulants appear to be the most commonly used, in approximately 12% of children aged 7 through 13 with ASDs.(40)

Atypical neuroleptics, including clozapine, risperidone, quetiapine, olanzapine and ziprasidone, are increasingly being used to treat the symptoms of ASD because, on the whole, they do not produce serious side effects, especially extrapyramidal effects like tardive dyskinesia. One exception is clozapine, which does carry a risk of seizure and agranulocytosis, requiring frequent monitoring with blood tests. Of the medications in this category, risperidone has shown the most promise in reducing repetitive behavior, aggression, anxiety, depression, and irritability in individuals with ASDs in clinical trials.(3)

Among the antidepressants, selective serotonin reuptake inhibitors (SSRIs) are the preferred medication in this class, again given their superior safety profile as compared to the tricyclics. There is some evidence that fluvoxamine and sertraline are effective at reducing repetitive and/or maladaptive behaviors and aggression, but further studies are needed.(3)

Finally, the psychostimulant merthylphenidate appears to be effective in about half of ASD children who are also hyperactive or have ADHD symptoms. Other medications in this class, including Concerta, Adderall XR, Ritalin LA, Metadate CD, and Methylin ER, have been found to be efficacious in treating children with ADHD, although they have yet to be studied in children with ASD and ADHD. Clonidine and guanfacine, two alpha-adrenergic agonists, have demonstrated some efficacy in autism. In particular, for children with Fragile X and autism, clonidine appears effective in reducing tantrums and aggression.(3)

While the three medication classes listed above show efficacy in treating some symptoms often associated with ASD, other medications which were touted to alleviate the core deficits of autism (social skills, language, and cognitive function) but have since been discredited include

[1] ECRI Institute is currently producing an evidence review of focal treatments used to treat the core deficits of ASD.

naltrexone, haloperidol, propranolol, and fluvoxamine for increasing communicative language and improvements in socialization.(3,40) There is disagreement in the literature on lomotrigine's effect on ASD core deficits.(3,40) Newer medications that look promising for treating the core deficits of autism but need more research include olanzapine as well as agents designed to increase glutamatergic transmissions.(3)

Complementary and Alternative Medicine (CAM)

The National Center for Complementary and Alternative Medicine (NCCAM) recognizes five domains of complementary and alternative medicine: alternative medicine systems (e.g., Chinese medicine), mind-body interventions (e.g., meditation), body-based medicine (e.g., sensory integration therapy), biologically-based therapies (e.g., mega-vitamin therapy), and energy therapies (e.g., magnet therapy). Among individuals with ASD, it is estimated that 50%-95% have been treated with one or more of these therapies despite a lack of empirical support for these treatments.(41)

In the body-based medicine category, chiropractic manipulation is the most commonly used method, but sensory integration therapy, to compensate for brain deficits in processing sensory input, is also popular.(41) Sensory integration therapy is typically administered by occupational therapists with an emphasis on manipulation of the child's environment. Specific treatment approaches include but are not limited to trampoline jumping, wearing weighted vests, "smooshing" a child between pads or pillows, playing with textured toys. Auditory integration training, or playing acoustically modified music, is believed to reduce the volume of frequencies to which the child is hypersensitive. Facilitated communication (FC), in which a facilitator guides the hand of a nonverbal individual, to assist them in using a computer or typewriter; is believed to help abnormal motor functioning in some children with severe ASD.(41) Finally, hyperbaric oxygen therapy to decrease blood perfusion to several areas of the brain believed to be affected in ASD is another example of body-based treatments.(41)

Specifically, among the biologically-based therapies, there is a popular belief that dietary manipulation may eliminate some or all of the symptoms associated with ASDs. In particular, reduced sugar intake (Feingold diet) has been purported to reduce hyperactivity and impulsivity in these children.(41) The use of secretin, a hormone involved in the control of digestion that stimulates the secretion of pancreatic fluid, has gained significant attention.(3,41) Proponents of secretin therapy, which is usually delivered in a single dose, allege improved behavioral outcomes. However, the authors of a recent Cochrane review on secretin therapy for autism concluded that the available evidence does not show that it is effective in treating the core features of autism.(42)

Other supplementary dietary therapies include large doses of omega-3 fatty acid, ketogenic diets, and the addition of vitamin B6-magnesium complex.(41) Vitamin B6-magnesium is believed to be beneficial by many because of its role in neurotransmitter production. Further, the elimination of casein and gluten (milk and wheat proteins) from the child's diet is believed by some to prevent the manifestation of autism altogether, by altering cerebral neurotransmitter metabolism. Other highly publicized therapies in this category include chelation therapy to rid the body of excess mercury and not vaccinating one's child with the measles, mumps, rubella (MMR) vaccine.(3,41)

Exercise

Some programs emphasize the importance of physical exercise. Proponents of this method believe that stimulation of muscle activity may bring about a rewiring of the brain's neural network. Two examples of interventions that incorporate exercise as the main component of treatment are the Doman-Delacato Program and Daily Life Therapy (Higashi).(30)

Surgical

Surgery is not a treatment typically used to treat ASDs. However, because children with ASD have a higher frequency of seizure (3% to 30%) and other neurological symptoms than normal children, neurosurgery and vagal nerve stimulation to reduce or eliminate seizures has been used to treat children with ASDs.(43)

Special education, occupational, speech and physical therapy

Per the United States Government Accountability Office, in 2002, 120,000 individuals aged six to 21 were diagnosed with ASD and received services under the Individual with Disabilities Education Act (IDEA). These services entail an individualized education program (IEP) which utilizes one or more of the following: special education teachers, counselors/psychologists, and speech-, occupational-, behavioral-, and physical-therapists based on the child's unique deficits.(44)

Focal treatments

A number of focal treatments are available for children with autistic spectrum disorders. Unlike the comprehensive interventions previously described, focal interventions use only one treatment strategy to address one or more symptoms of ASDs. Many focal strategies are components of comprehensive treatment programs. Some of the better-known focal treatments include the Picture Exchange Communication System (PECS) and Social Stories.(45) Picture Exchange Communication System (PECS) is part of the ABA approach in which pictures and other symbols are used to improve functional communication skills in individuals with ASDs. Social Stories, including Comic Strip Conversations, Thinking stories and Story boxes, present everyday social situations with an emphasis on which cues in the story are most relevant to the reader and some common ways of responding to these situations.(46) Other lesser used focal treatments include Gentle Teaching, Holding Therapy, and the Option Method (also called the Son Rise Program).(45) These other treatments focus mainly on addressing problem behaviors. Finally, some focal treatments, such as cognitive behavioral therapy and anger management, are aimed at reducing secondary symptoms associated with ASD, such as anxiety, depression, anger, and sleep disorders.(47-49)

Economic and Regulatory Issues

The Individuals with Disabilities Education Act (IDEA)

The education of children with autism is governed by the Individuals with Disabilities Education Act (IDEA). IDEA is made up of both statutory laws enacted by Congress and the regulation of those laws by the Department of Education. IDEA incorporates six guiding principles. The first of which is a zero rejection policy that prohibits the exclusion of a student with a disability from free appropriate education.(31) This includes provisions governing how a child with a disability may be disciplined, limiting schools to a 10-day suspension for any violation of the school's code of conduct and up to a 45-day removal to an interim alternative education setting for serious safety threats to the child or another person. In addition, schools are not permitted to institute a change of placement if the behavior leading up to the change is a manifestation of the child's disability, unless the parent consents. When a change of placement is initiated, a behavioral intervention plan must be developed to address the problem behavior and positive behavioral interventions and supports (PBS) must be considered to remedy the situation. To ensure that the cost of some of these needed related services are covered for children with disabilities, IDEA specifies that public agencies, including state Medicaid agencies, must assume financial responsibility for services to these children.(31)

Under IDEA, a child is entitled to a nondiscriminatory evaluation (NDE), which ensures that socioeconomic status, language or other such factors do not bias the evaluation, and education in the least restrictive environment (LRE), which means that if the child can benefit from an education alongside his/her typically developing peers, that is the setting in which the child should be taught. Other IDEA principles include a policy of due process, or the rights of parents to contest any school decisions regarding the education plan of their child, and an emphasis on parent and student participation in the decision-making process. Finally, under IDEA, each child is entitled to appropriate education, or education that benefits the student and is appropriate to their individual needs. However, a free appropriate public education (FAPE) does not entitle the child to other interventions (e.g., Lovaas method, TEACCH, etc.) unless it can be shown that denial of these other interventions would constitute a denial of FAPE. As IDEA routinely uses Positive Behavioral Interventions and Supports, the burden of proof falls on parents to try to show that PBS is not beneficial to their child and that one of these other interventions would be more beneficial.(31)

Charges and Fees

Our searches of both the published and grey literature (e.g., intervention-specific Web sites) identified very little reliable information on the cost of specific behavioral interventions for the treatment of ASDs. However, some data were available for applied behavioral analysis (ABA) and parent-directed discrete trial training (DTT). In terms of healthcare utilization, based on data from 1997-2000 from three national surveys, families of children with an ASD were more likely than families of children with mental retardation to have private insurance and were found to average $2,239 on home healthcare expenditures, of which $179 was for ABA.(50)

Chasson et al. conducted a projected cost comparison study of children receiving three years of discrete trial training as compared to if those same children required a full 18 years of special education in Texas. The authors incorporated special education costs ($20,000 annually), early intensive behavioral intervention (EIBI) costs (assumed to be $22,500 annually with the parent-directed model of DTT), EIBI effectiveness (assuming a proportion of 0.28 of children who receive EIBI but fail to mainstream into regular education), population estimates of children with autism in Texas, and the expected number of years required for each type of service into the model. They found that the state of Texas could save $84,300 per child over the child's total school years. Assuming 10,000 children in Texas have autism, that is a savings of $843 million in state budgeted funds and $2.09 billion in actual funds (state funds plus local, federal and private funds).(51)

More generally, some investigators have attempted to compare the health care costs of children with ASDs versus other children. In one such study, the average annual total Medicaid expenditures for children with ASD versus those diagnosed with either mental retardation or other developmental/psychiatric disorders for the years 1994-1999 for one Pennsylvania county was found to be 3.5 times higher, or about $10,000, for the ASD children, but no breakdown by behavioral service was provided.(52) In a similar study which examined total 1993-2003 medical expenditures for a national sample of children with ASDs covered under employer-based private health insurance plans, average medical expenses for these children were between 4.1 and 6.2 times higher than for children without a diagnosis of ASD.(53) When compared with children with another mental disorder, for the year 2004, children with ASDs cost private healthcare insurers approximately $6,700 a year in total autism expenditures, surpassed only by those with a diagnosis of mental retardation (or about $10,000 per year).(54) Again, no breakdown for behavioral interventions was presented in these studies.

As the costs of programs like the Lovaas method are not routinely covered under IDEA, recently some states have taken action to remedy this coverage gap. For instance, the Nevada Autism Task Force found that less than 6.0% of ASD individuals in the state receive funding from state programs to assist with the costs of ABA and that most insurance companies, including Medicaid and Nevada check up, do not cover it. The task force is currently pressuring the Nevada Legislature to require health insurance policies and medical assistance programs to cover these costs for individuals under 21 years of age.(55) Other states are also in the processes of passing insurance reform, while some others have already done so. Table 8 below lists the insurance reform status by state.(56)

Table 8. Insurance Reform Status by State

Insurance Reform Status	States
The state has a law in place that requires private insurance to cover autism services, including ABA	Arizona, Florida, Indiana, Louisiana, Minnesota, Pennsylvania, South Carolina, Texas
The state currently has a bill seeking autism insurance reform that has been endorsed by Autism Speaks	Michigan, New Jersey, Virginia
The state is currently in the process of working on autism insurance reform, but does not yet have a bill endorsed by Autism Speaks	Kansas, Maryland, Massachusetts, Mississippi, Missouri, Nevada, New York, North Carolina, Ohio, Oklahoma, Washington, Wisconsin
The state is either in the very early stages of working on a bill or is not working on an autism insurance reform bill at all	Alabama, Alaska, Arkansas, Connecticut, Delaware, Idaho, Kentucky, Maine, Montana, Nebraska, New Hampshire, New Mexico, North Dakota, Oregon, Rhode Island, South Dakota, Tennessee, Utah, Vermont, West Virginia, Wyoming

Centers for Medicare and Medicaid Services Coverage Policy

The U.S. Centers for Medicare and Medicaid Services (CMS) does not have a national coverage policy for the use of educational or behavioral interventions for individuals with ASDs. Coverage decisions are left to the discretion of local Medicaid carriers. According to the 2008 Easter Seals Web site, only the following 16 states had a Medicaid coverage policy applicable to individuals less than 21 years of age: Arkansas, Colorado, Florida, Georgia, Illinois, Indiana, Kansas, Massachusetts, Maryland, Montana, Nebraska, South Carolina, Tennessee, Utah, Wisconsin, and Wyoming. Only 11 of the 16 states had coverage policies that specified behavioral interventions for individuals with ASDs. Table 9 below presents the behavioral services covered in the policies of the 11 states that specified some type of behavioral intervention.

Table 9. State Medicaid Coverage Policies

State	Behavioral service covered
Arkansas	Waiver for intensive early intervention individualized therapy, such a behavioral therapies, for children 3 to 10 years of age with a diagnosis of PDD, covering services up to $50,000 per year.
Colorado	Behavioral therapy services are covered at a maximum of $25,000 per year for three years or until the child's sixth birthday.
Florida	Individual with autism three and older requiring intermediate care facility for the developmentally disabled may seek services under a waiver for behavioral analysis and behavior assistant services. Massage therapy, IQ testing and psychological assessments are not covered.
Georgia	Waiver in effect covers behavioral support consultation.
Illinois	Home-based support services for children 3 to 21 years of age, with a monthly allocation not to exceed 200% of the monthly federal SSI payment. Participants may select from a range of services including behavior intervention and treatment. For those requiring residential care, behavior interventions to an annual maximum of 66 hours are covered.
Indiana	Waiver covers Applied Behavioral Analysis and behavioral support.
Kansas	Waiver covers early intensive intervention treatment through 5 years of age for a maximum of four years.
Montana	In process of developing a waiver for autistic children between 2 to 5 years of age for a maximum of three years of treatment with 20 to 25 hours per week of early intensive rehabilitation in the home by a qualified provider.
Nebraska	Waiver in process that would cover children up to nine years of age for intensive early intervention services.
South Carolina	Waiver for children 3 to 10 years of age for early intensive behavioral intervention.
Wisconsin	Waiver program covers intensive in-home treatment (although no specific behavioral intervention is listed) for children birth to 21 years of age.

More detailed information about local coverage decisions (LCD) can be found by searching the following Web site: http://www.easterseals.com/site/PageServer?pagename=ntlc8_autism_state_profiles.(57)

Third Party Payer Coverage

Table 10 below shows the current state-by-state private health insurance mandated coverage status for ASDs based on the 2008 Easter Seals Web site. This coverage may or may not include behavioral interventions. For more specific information on each state's health insurance policies, visit the following Web site: http://www.easterseals.com/site/PageServer?pagename=ntlc8_autism_state_profiles.

Table 10. State Mandated Coverage

State Insurance Coverage	States
Mental health parity law only; no specific health insurance mandate for coverage for ASDs	Alabama, Arkansas, California, Louisiana, Maine, Massachusetts, Minnesota, Missouri, Montana, Nebraska, Nevada, New Hampshire, New Jersey, New Mexico, North Carolina, North Dakota, Rhode Island, Utah, Vermont, Virginia, Washington
State insurance coverage mandate in effect	Arizona, Colorado, Connecticut, District of Columbia, Florida, Georgia, Hawaii, Illinois, Indiana, Iowa, Kentucky, Maryland, New York, Oregon, Pennsylvania, South Carolina, Texas, Tennessee
No specific policy reported that covers ASDs	Alaska, Delaware, Idaho, Kansas, Michigan, Mississippi, Ohio, Oklahoma, Puerto Rico, South Dakota, West Virginia, Wisconsin, Wyoming

In addition to the above, we searched the Web sites of 15 private third-party payers for coverage policies of comprehensive or intensive intervention programs for children with ASD (See the *Literature Search Methods* in Appendix A for a list of sites searched). Only one of the providers, Aetna, specifically indicated that such interventions are covered. Aetna's policy specifies coverage of "intensive educational interventions in which the child is engaged in systematically planned and developmentally appropriate educational activity toward identified objectives, including services rendered by a speech-language pathologist to improve communication skills." While less specific, the coverage policy of Blue Cross Blue Shield of Massachusetts indicates that early intervention is covered for children who are three years or younger and have an established, biological or environmental risk; a known disabling physical or mental condition, and four or more risk factors. The coverage policies of two other providers—Cigna and Medica—specifically indicate that early intensive intervention programs, such as ABA and the Lovaas model are not covered. The remaining providers either did not describe their coverage policy on line or did not specify whether comprehensive interventions are covered or not covered. See Table 52 in Appendix J for more information about the coverage policies of the third party payers.

Key Questions and Outcomes Assessed

For this report, we addressed the following five Key Questions:

1. Does any comprehensive educational or behavioral intervention improve outcomes for children with ASD when compared to no treatment, waitlist control, or standard care (e.g., special/supported public education and/or a mix of paramedical services, such as speech and occupational therapy)?

2. Is one comprehensive educational or behavioral intervention more effective than another in improving outcomes for children with ASD?

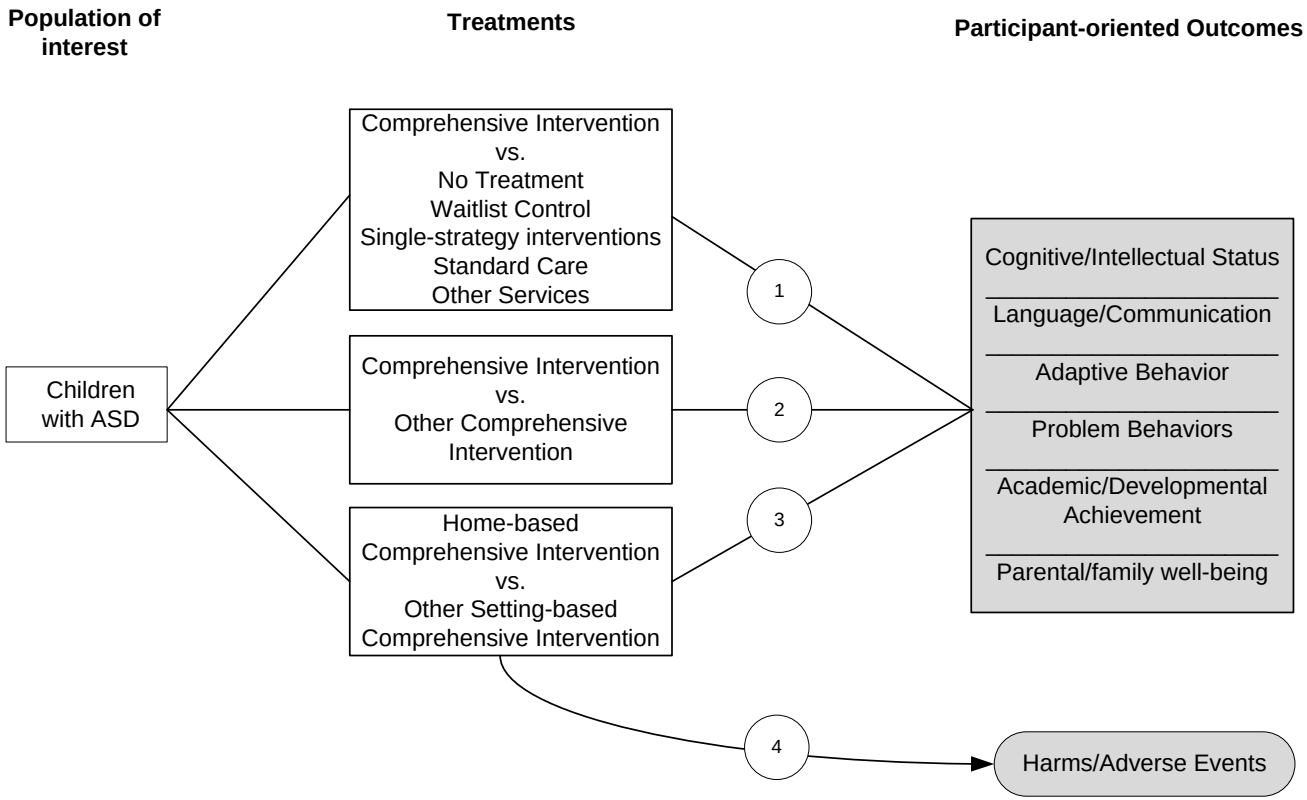
3. Are home-based interventions of similar intensity and/or structure as comprehensive educational or behavioral interventions provided in other settings (e.g., center or clinical setting) more effective in improving outcomes for children with ASD?

4. What adverse events and harms have been reported to occur in association with the use of comprehensive educational or behavioral interventions for children with ASD?

5. What is the consensus among experts about the safety and efficacy of comprehensive educational or behavioral interventions for the treatment of children with ASD?

These questions, along with the treatments and outcomes we evaluated to address these questions, are illustrated in Figure 1 below. This figure portrays the pathway of events that participants' experiences, starting from when they are first identified (the far left of the figure), to the treatments they receive, and to participant-oriented outcomes. As such, participants in the population of interest are identified and "enter" the pathway at the left of the figure. The outcomes we address are shown to the right side of the figure. Key Question 5 is not depicted in the figure because this question deals with current expert opinion on treatment for ASD and does not address participant-oriented outcomes. We address this question by summarizing pertinent information from clinical practice guidelines and consensus or position statements.

Figure 1. Analytic Framework



Note: Circled numbers, e.g., (1) denote Key Questions.

Definition of Outcomes Assessed

Below, we briefly describe the outcomes assessed in this review. The outcomes represent those that are most commonly measured in studies evaluating comprehensive interventions for children with ASD.(58) Numerous standardized instruments are available to measure these outcomes. The instruments used in the studies that met the study selection criteria for this report are listed in Table 12 and further described in Table 17 in Appendix B.

- **Cognitive/intellectual status—**measured typically, according to a review by Wolery and Garfinkle (2002), as a change in IQ score on standardized, age-appropriate IQ tests.(58) Some studies may measure this outcome as the proportion of children with ASD who after intervention reach IQ status that is considered normal for their age group (using the normative mean from the general population as the comparison).
- **Language/Communication skills**—typically measured as child's ability for verbal expression, receptive skills, and pragmatic communication (e.g., body language, turn taking, and understanding intention and interest of others). Language and communication skills are generally measured using various standardized language and non-verbal communication tests. Subscales of IQ tests may also be used.
- **Adaptive behavior** (e.g., daily living/functional skills)—refers to a person's social responsibility and independent performance of daily activities. The most widely used measure of adaptive behavior is the Vineland Adaptive Behavior Scale, which measures the following domains: communication, daily living skills, socialization, and for children under five years of age, motor skills. This is a well studied, validated instrument completed during an interview with a parent or teacher.
- **Problem behaviors—**this outcome encompasses a wide range of behaviors associated with ASD, including severe difficulty in initiating and maintaining social interactions and relationships, aggression, self-injury, use of restrictive and repetitive behaviors, also known as stereotypical behaviors. For the most part, problem behavior is measured using various validated instruments and checklists.
- **Academic/Developmental achievement**—this outcome may be measured in different ways. According to Wolery and Garfinkel (2002), it is sometimes measured as post-intervention preschool classroom placement into a regular integrated classroom, regular classroom with support and modified curriculum, or special education classroom.(58) However, the validity of classroom placement as an outcome has been questioned as it may be influenced by factors other than a child's abilities such as, according to Wolery and Garfinkel, the extent to which schools view inclusive classes as legitimate options, idiosyncratic policies, traditions, and goals of schools, and the influence of parents. Other ways this outcome may be measured include change in children's diagnostic status, reduction of autistic symptoms, or using various standardized achievement tests.
- **Parental/family well-being** (e.g., family stress, quality of life)—parents are highly involved in most comprehensive programs, and in some cases, may be the primary provider of the intervention. So, some studies may include measures assessing family outcomes.

- **Harms/Adverse Events**—According to Matson (2005), children, particularly young children, who are expected to comply to structured tasks over extensive periods of time on a daily basis are likely to experience unintended adverse events such as tantrums, noncompliance, yelling, etc.(59)

Methods

Identification of Clinical Studies

One characteristic of a good technology assessment is a systematic and comprehensive search for information. Such searches distinguish ECRI Institute's assessments from traditional literature reviews. Traditional reviews use a less rigorous approach to identifying and obtaining literature and allow a reviewer to include only articles that agree with a particular perspective, and to ignore articles that do not. Our approach precludes this potential reviewer bias because we obtained and included articles according to explicitly determined *a priori* criteria. The criteria used for this report is explained in detail below under *Study Selection.*

Electronic Database Searches

We searched 17 external and internal databases, including MEDLINE, EMBASE, and PsycINFO, for clinical trials on the use of comprehensive interventions to treat ASD. To supplement the electronic searches, we examined the bibliographies of included studies, scanned the content of new issues of selected journals, and reviewed relevant gray literature for potential additional relevant articles. Gray literature includes reports and studies produced by local government agencies, private organizations, educational facilities, and corporations that do not appear in the peer-reviewed literature. Although we examined gray literature sources to identify relevant information, we only evaluate published literature in this report. All of the databases and the detailed search strategies used in this report are presented in Appendix A.

Study Selection

We selected the studies that we considered in this report using *a priori* inclusion criteria. As mentioned above, arriving at these criteria before beginning the analysis is one way of reducing bias.

We used the following criteria to determine which studies would be included in our analysis.

Population

1. *At least 85% of children included in a study must have a primary diagnosis of ASD based on the diagnostic criteria established by the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders -IV or Text Revised edition (DSM-IV or DSM-IV-TR) or the World Health Organization's International Statistical Classification of Diseases-10th edition (ICD-10).* If less than 85% then the study must have reported outcomes separately for children who met a primary diagnosis of ASD. Studies that included children with co-morbid psychological conditions, such as depression, attention deficit/hyperactivity disorder (ADHD), obsessive compulsive disorder (OCD), were not excluded as long as these conditions were secondary to a diagnosis of ASD.
2. *Children in the studies must have been between the ages of one and eight years old at the start of treatment.* Studies that included children outside of this age range must have reported outcomes separately for those children within the specified age range. We limited the selection of studies to those that included children under the age of eight years

old because most comprehensive intervention programs are specifically designed and directed toward younger children with ASDs. Also, the specific curriculum, intervention practices, and behaviors taught within a program are likely to differ between younger and older children with ASD.(58)

Intervention

3. *Studies must have assessed the efficacy of a comprehensive educational or behavioral intervention defined as an intervention in which more than one treatment strategy is used to address most of the deficits/symptoms associated with ASD.* Studies that focused on other interventions for ASD, such as focal interventions (e.g., discrete trial training, pivotal response training), physiological or surgical interventions, special education, or paramedical services (e.g., occupational or speech therapy) were excluded from this review, unless they were being directly compared within a study to a comprehensive educational or behavioral intervention, and the study met the other inclusion criteria for this report.

Study Design

4. *Studies must have been prospective randomized or non-randomized controlled trials. Studies that employed a non-randomized design must have used methods to enhance group comparability, such as matching participants on key variables (e.g., chronological age, intellectual disability, and overall severity of ASD), or using statistical procedures to control for any differences observed between groups at baseline.* According to Matson (2006), intellectual disability, chronological age, and overall severity of ASD have been shown to be prognostic in establishing long-term outcome, regardless of the intervention.(59) Thus, studies were excluded if the authors did not demonstrate group comparability on these characteristics.
5. *Studies must have included five or more children in each the treatment and the control condition.* The results of studies with very small patient groups are often not applicable to the general population.

Outcomes

6. *All relevant outcomes must have been measured using an instrument(s) for which the properties of reliability and validity have been verified in the published literature.* However, if a study did not use a validated instrument, then the entire study was not necessarily excluded for all outcomes—only its data from instruments in which the psychometric properties were not reported in the published literature were excluded.
7. Study must have reported on at least one of the outcomes of interest for one or more of the Key Questions.
8. For all outcomes, we only considered time points for which at least 50% of the enrolled participants contributed data.

Publication Type

9. *Study must have been published in English.* Moher et al. have demonstrated that exclusion of non-English language studies from meta-analyses has little impact on the conclusions drawn.(60) Juni et al. found that non-English studies typically were of lower methodological quality and that excluding them had little effect on effect size estimates

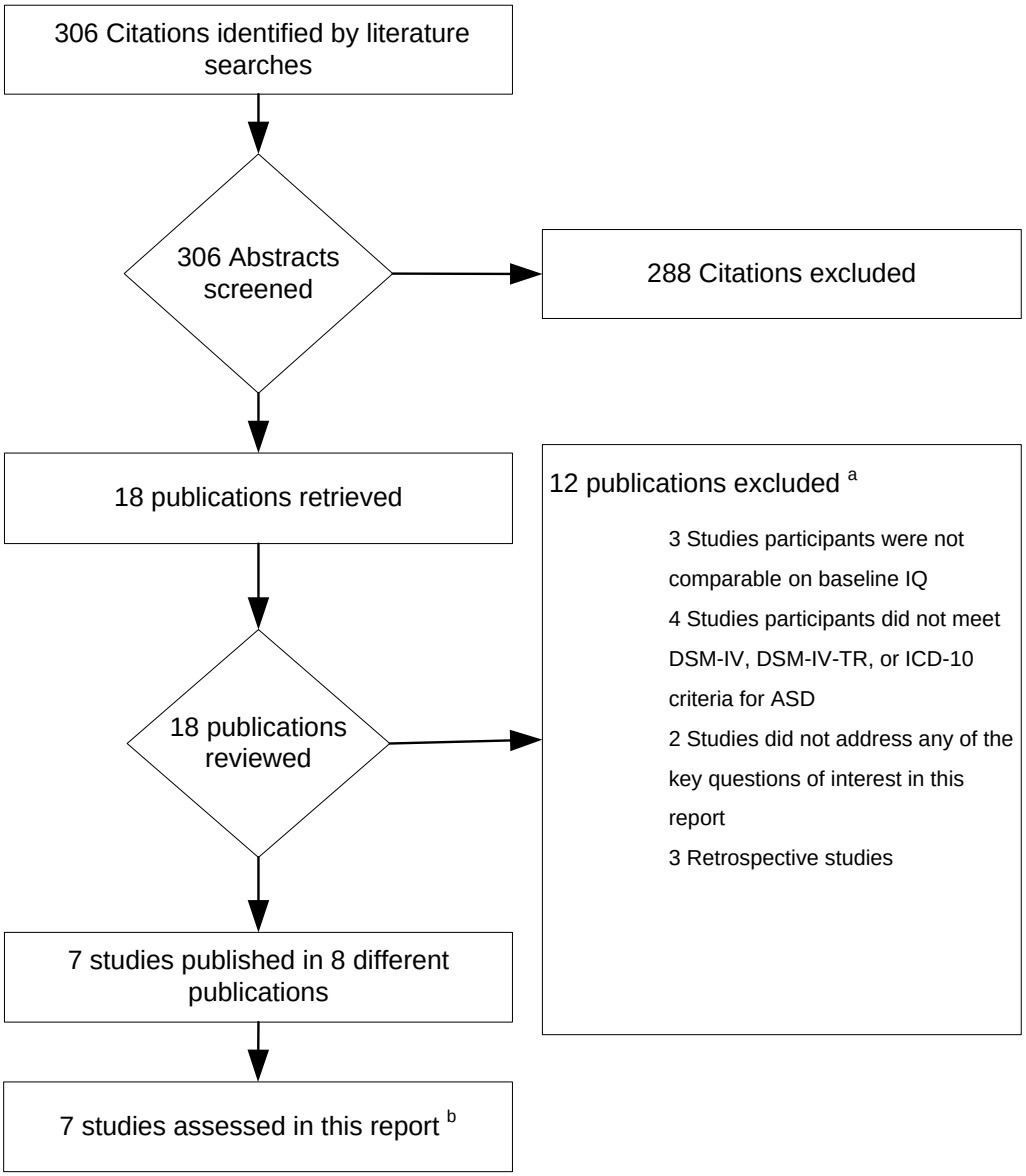
in the majority of meta-analyses they examined.(61) Although we recognize that in some situations exclusion of non-English studies could lead to bias, we believe that the few instances in which this may occur do not justify the time and cost typically necessary for translation of studies to identify those of acceptable quality for inclusion in our reviews.

10. *Study was published as a full-length article in a peer reviewed journal rather than an abstract or letter.* Published abstracts and letters do not include sufficient details about experimental methods to permit verification and evaluation of study design.(62,63) However, we included data from any abstract that reported additional outcomes from a study and patient group that had been reported in a full-length article that met all inclusion criteria.(64)
11. *When several reports from the same center were available, only outcome data from the report with the largest number of patients was included.* This is to avoid double-counting of patients. If a smaller report had provided data on an outcome that was not provided by the largest report, we included the data.

Articles Identified by Searches

Our searches identified 306 potentially relevant articles. Most articles were excluded at the abstract level because they were not clinical studies, did not address any of the Key Questions, or were case series studies. Figure 2 below provides a diagram of our study selection process. Of the 306 abstracts reviewed, 18 full-length articles were retrieved for further review. A total of 12 of the 18 studies were excluded from consideration. Studies were excluded for the following reasons: the participants in the study groups differed substantially at baseline (i.e., differences found by the authors to have an impact on study outcomes, three studies); participants did not meet the DSM-IV or DSM-IV-TR diagnostic criteria for ASD (four studies); the study was found not to address one of the Key Questions of interest in this report (two studies); or the study did not meet the study design criteria for this report (three studies). Table 16 in Appendix A lists the reasons for exclusion of all excluded studies. A total of seven studies published in eight different publications made up the evidence base for this review. Of the seven studies, three addressed Key Question 1, three addressed Key Question 2, and two addressed Key Question 3. One of the seven studies, Howard et al.(65), addressed both Key Question 1 and 2. Table 11 lists the studies included in this review and the Key Questions and outcomes addressed in each of the studies.

Figure 2. Study Attrition Diagram



[a] Table 16. Excluded
[b] Table 11. Key Questions Addressed by Included Studies

Table 11. Key Questions Addressed by Included Studies

Study	Study Design	Interventions	Number of Children	Key Question 1 (Comprehensive vs. no treatment or standard care)	Key Question 2 (Comprehensive vs. comprehensive)	Key Question 3 (Home-based vs. other setting)
Remington et al. (2007)(66)	CT	Home-based EIBI	23	✓		
		SC	21			
Eikeseth et al. (2007)(67) & Eikeseth et al. (2002)(68)[1]	CT	School-based EIBI	13		✓	
		Eclectic/developmental	12			
Zachor et al. (2006)(69)	CT	Center-based EIBI	20		✓	
		Center-based Eclectic/developmental	19			
Cohen et al. (2006)(70)	CT	Home-based EIBI	21	✓		
		SC	21			
Howard et al. (2005)(65)	CT	Mixed-setting EIBI	29	✓	✓	
		Mixed-setting Eclectic/developmental	16			
		SC	16			
Sallows &Graupner (2005)(71)	RCT	Clinic-directed EIBI	12			✓
		Parent-directed EIBI	10			
Smith et al. (2000)(72)	RCT	Clinic-directed EIBI	15			✓
		Parent-directed EIBI	13			
		Total	**261 children**	**3 studies**	**3 studies**	**2 studies**

Note: None of the included studies addressed Key Question 4. Thus, this question is not presented in the table.

[1] Studies included same participant population, but reported data for different follow-up times.

ABA Applied behavior analysis
CT Non-randomized controlled trial
EIBI Early Intensive behavioral intervention
RCT Randomized controlled trial
SC Standard care

Table 12. Outcomes Assessed and Instruments Used in Included Studies

Instrument Name	Key Question 1			Key Question 2			Key Question 3	
	Remington et al. (2007)(66)	Cohen et al. (2006)(70)	Howard et al.(65)	Eikeseth et al.(67) & Eikeseth et al.(68)[1]	Zachor et al. (2006)(69)	Howard et al.(65)	Sallows & Graupner (2005)(71)	Smith et al. (2000)(72)
Cognitive/Intellectual Status								
Bayley Scales of Infant Development (BSID)	✓	✓	✓	✓	✓	✓	✓	✓
Stanford-Binet Intelligence Tests(SBIS)	✓				✓			✓
Wechsler Intelligence Scale for Children				✓			✓	
Wechsler Preschool and Primary Intelligence Scale (WPPSI)		✓	✓	✓		✓	✓	
Language/Communication Skills								
Early Social Communication Scales	✓							
Merrill-Palmer Scale of Mental Tests (MPSMT)		✓	✓	✓		✓	✓	✓
Reynell Developmental Language Scales	✓	✓	✓	✓		✓	✓	✓
Adaptive Behavior								
Vineland Adaptive Behavior Scale(VABS)	✓	✓	✓	✓		✓	✓	✓
Problem Behaviors								
Achenbach Child Behavior Checklist				✓			✓	✓
Autism Screening Questionnaire	✓							
Developmental Behavior Checklist	✓							

Instrument Name	Key Question 1			Key Question 2			Key Question 3	
	Remington et al. (2007)(66)	Cohen et al. (2006)(70)	Howard et al.(65)	Eikeseth et al.(67) & Eikeseth et al.(68)[1]	Zachor et al. (2006)(69)	Howard et al.(65)	Sallows & Graupner (2005)(71)	Smith et al. (2000)(72)
Academic/Developmental Achievement								
Autism Diagnostic Interview-Revised							✓	
Autism Diagnostic Observation Schedule (stability of diagnosis)					✓			
Classroom placement (not an instrument)		✓						✓
Wechsler Individualized Achievement Test								✓
Woodcock-Johnson Tests of Achievement							✓	
Parent/Family Wellbeing								
Hospital Anxiety and Depression Scale	✓							
Kansas Inventory of Parental Perceptions Positive Contributions	✓							
Niosonger Child Behavior Checklist	✓							
Questionnaire on Resources and Stress-Friedrich short-form (QRS-F)	✓							

Note: The instruments listed in the table represent those for which data were collected from most or all of the study participants within a study.

[1] Same participant population

Rating the Stability and Strength of Evidence

In evaluating the stability and strength of a body of literature, we used the ECRI strength-of-evidence system (described in more detail in Appendix C).(73) This system employs decision points that collectively yield an overall category that describes the strength of the evidence for a *quantitative* estimate and *qualitative* conclusion as strong, moderate, weak, or unacceptably weak. The qualitative conclusion addresses the question, "Does it work?" The quantitative estimate addresses the question, "How well does it work?" This distinction allows an evidence base to be considered weak in terms of the quantitative estimate of effect (e.g., if estimates vary widely among studies) but strong or moderate with respect to the qualitative conclusion (e.g., if all studies nevertheless demonstrate the same direction of effect).

The system addresses five general aspects of the evidence: internal validity, quantity, consistency, robustness, and magnitude of effect.(73) Internal validity refers to the degree of potential bias in the design or conduct of studies. Quantity refers to the number of studies and the number of enrolled patients. Consistency addresses the degree of agreement among the results of available studies. Robustness involves the constancy of conclusions in the face of minor hypothetical alterations in the data. Magnitude of effect concerns the quantitative amount of benefit (or harm) that patients experience after treatment, and it is only considered in the qualitative section of the system. These concepts, and the rules we used to incorporate the concepts in this technology assessment, are described more fully in Appendix C.

Quality of Evidence

To aid in assessing the quality of each of the studies included in this assessment, we used the quality assessment instrument developed by ECRI Institute for comparative studies as shown in Appendix C. This instrument examines different factors of study design that have the potential to reduce the validity of the conclusions that can be drawn from a trial. In brief, the tools were designed so that a study attribute that, in theory, protects a study from bias receives a "Yes" response. If the study clearly does not contain that attribute it receives a "No" response. If poor reporting precludes assigning a "Yes" or "No" response for an attribute, then "NR" is recorded (NR = not reported).

To estimate the quality of an individual study, we computed a normalized score so that a perfect study received a score of 10, a study for which the answers to all items was "No" received a score of 0, and a study for which the answers to all questions was "NR" was 5.0. We then classified the overall quality of the evidence base by taking the median quality score. Quality scores were converted to categories as shown in the table below. The definitions for what constitutes low or moderate quality evidence were determined *a priori* by a committee of four ECRI Institute methodologists, and are presented in Table 13 below.

Table 13. Study Quality Categories

	Overall Quality of Evidence Base		
	Low	**Moderate**	**High**
Median Overall Quality Score of the evidence base	≤6.0	>6.0 but <8.5	≥8.5

Data Synthesis

When the evidence base included three or more studies and when 75% or more of the available study data for an outcome could be used in the analysis, we attempted to reach *quantitative* conclusions using a random-effects meta-analysis. Statistical significance was set at $p < 0.05$ and heterogeneity was determined using the I^2 statistic.(74,75) An I^2 greater than or equal to 50% was evidence of substantial heterogeneity among study results. If at least five studies were used in a meta-analysis, we performed meta-regression in an attempt to explain the heterogeneity using the permutation test p-value as described by Higgins and Thompson.(75) The following covariate variables were selected *a priori* as possible sources of heterogeneity: duration of treatment, intensity of treatment, training/experience of provider(s), fidelity/integrity of treatment, quality of study, use of blinded outcome assessors, and use of concomitant treatment. We did not attempt to obtain a quantitative summary effect estimate from an evidence base with unexplained heterogeneity. We tested summary estimates from evidence bases without substantial heterogeneity for robustness by removal and replacement of each separate study, and by performing cumulative meta-analysis by publication date (beginning with oldest and adding subsequent studies one-by-one to the most recent study). These methods are described more fully in Appendix C.

We used the standardized mean difference (SMD, also known as Cohen's d), which represents the difference between groups on a standardized scale, as the measure of effect size for all meta-analyses of continuous data. We computed baseline-adjusted SMD values using a pre-post correlation of 0.5.(76) For dichotomous outcomes, we used the odds ratio as the measure of effect size; values greater than one favored the experimental group, and values less than one favored the control group.

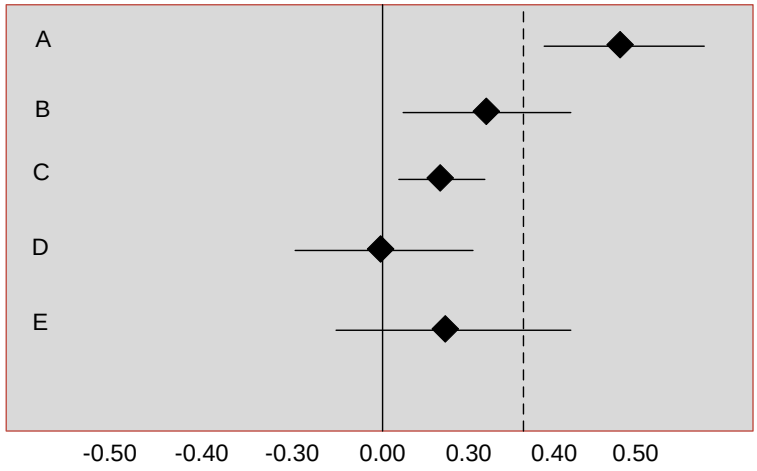
If a summary effect size could be obtained, we then determined whether or not the summary effect size estimate was informative. The summary effect size estimate was considered to be informative if it met one of the following criteria: 1) it was statistically significant or 2) it was not statistically significant and the 95% confidence intervals surrounding it did not overlap the boundaries of a clinically significant effect. A clinically significant effect refers to a quantity of change between groups (e.g., experimental and control) that is considered enough change to be meaningful. The minimally clinically significant change is the *minimal* change that is considered clinically significant.

In this report, for outcomes measured using standardized tests for which normative data were available, we defined the *minimum* level of clinical significance as a between-group difference at post-treatment of one half the standard deviation unit of the mean for typically developing children. Using the standard deviation unit as a way of determining if a change is meaningful, according to Bain and Dollaghan, is useful because it provides an estimate of a child's standing relative to a normative sample.(77) To adjust for the fact that the children in the studies assessed in this report have substantially lower levels of cognitive functioning than typically developing children, we chose a more conservative approach by using half of the standard deviation unit for that of typically developing children.

For most of the tests measuring IQ, language skills, and adaptive behavior in the studies included in this report, the normative or population mean of typically developing children is 100 with a standard deviation of 15 points.(78) Thus, for this report a *minimally* clinically significant difference between groups would be a standard deviation of 7.5. To determine if an effect size or

SMD was clinically significant based on this definition, we divided the minimum difference of 7.5 by the pooled standard deviation of the study groups at post-treatment. So, for example, if the pooled standard deviation of the study groups was 19.7, a SMD of 0.38 would be considered a *minimally* clinically significant difference, provided that the 95% confidence intervals do not overlap the boundaries of this value. Figure 3 presents examples of findings in which an effect size using the above definition is either 1) clearly clinically significant, 2) the clinical significance is unclear or inconclusive, or 3) the effect size is clearly not clinically significant.

Figure 3. Example of Findings and Associated Clinically Significant Conclusion



Dashed line = Threshold for a clinically significant difference (SMD of 0.32)

A shows that the treatment effect is both statistically and clinically important; **B** shows that the treatment effect is statistically significant, but it is unclear whether it is clinically important; **C** shows that the treatment effect is statistically significant, but too small to be clinically important; **D** shows that it is unclear if the treatment effect is significantly important, but it is clearly not clinically important; and **E** shows that it is unclear whether the treatment effect is statistically or clinically important.

For outcomes measured using instruments for which normative data were not available or were not located through our searches of the literature, we defined the *minimum* level of clinical significance as a SMD of 0.2, which corresponds to a small effect size.(79) For all dichotomous outcomes, such as change in diagnostic status, proportion of children reaching IQ scores within the normal range (85 or greater), or proportion of children placed in regular, integrated classrooms, any statistically significant difference was considered to be a clinically significant difference. See Table 14 below for the definition of the minimum difference between groups at post-treatment to be considered clinically significant for each outcome of interest addressed in this review.

When a quantitative conclusion was not possible, we entered all available data into a random effects meta-analysis to determine the robustness of a qualitative conclusion. We performed the same sensitivity analyses as described above. The data were considered robust if the summary

effect size at each iteration of adding or removing a study remained statistically significant (did not cross zero on the SMD plane) and the direction of the effect size did not change (go from positive to negative or negative to positive) during the analysis. Further details on sensitivity analyses and statistical approaches we used are described in Appendix C. All effect size estimates and meta-analyses were calculated using the Comprehensive Meta-Analysis Statistical Software Package Version 2 (Biostat/ Englewood, NJ).

Table 14. Minimum Difference for Clinical Significance

Outcome	***Minimum*** **difference between groups at post-treatment to be considered clinically significant**
Cognitive/Intellectual Status and Language/Communication Skills	One half of the standard deviation of the mean for typically developing children, which for most tests of IQ and language skills is a standard deviation of 15.(78) So, the minimum difference to be considered clinically significant for this report is a standard deviation of 7.5. For dichotomous outcomes (e.g., number of children moving into normal range on IQ scores, which is a score of 85 or greater), any statistically significant difference is considered to be clinically significant.
Adaptive Behavior	One half of the standard deviation of the mean for typically developing children, which for the Vineland Adaptive Behavior Scales is a standard deviation of 15.(80) So, the minimum difference to be considered clinically significant for this report is a standard deviation of 7.5.
Problem Behaviors	A SMD of 0.2, which corresponds to a small effect size, is considered to be clinically significant.(79)
Academic/Developmental Achievement	For continuous outcomes (i.e., achievement tests), a SMD of 0.2, which corresponds to a small effect size, is considered to be clinically significant. For dichotomous outcomes (classroom placement or change severity of symptoms), a statistically significant difference is considered to be clinically significant.
Parent/family Well-being	A SMD of 0.2, which corresponds to a small effect size, is considered to be clinically significant.(79)

Synthesis of Results

***Key Question 1*:** *Does any comprehensive educational or behavioral intervention improve outcomes for children with ASD when compared to no treatment, waitlist control, or standard care (e.g., special/supported public education and/or a mix of paramedical services, such as speech and occupational therapy)?*

- **After one year of treatment, children with ASD who receive early intensive behavioral intervention score higher on tests of IQ than children who receive standard care. Estimated effect size is a standardized mean difference (SMD) of 0.750 (95% confidence intervals [CI] 0.302 to 1.199, p <0.001), which corresponds to a between-group difference of 14.8 points in overall IQ. Strength and Stability of Evidence: Moderate.**
- **The evidence was insufficient to determine whether children with ASD who receive early intensive behavioral intervention continue to demonstrate higher scores on tests of IQ than children who receive standard care at later follow-up times (greater than one year).**
- **Children with ASD who receive early intensive behavioral intervention are more likely to achieve an IQ score within normal range for typically developing children (85 or higher) than children who receive standard care. The estimated size of the effect is an odds ratio of 2.616 (95% CI 1.160 to 5.902, p = 0.021). Strength of Evidence: Moderate and Stability of Evidence: Low.**
- **After one year of treatment, children with ASD who receive early intensive behavioral intervention perform more adaptive behaviors as indicated by higher scores on the Vinland Adaptive Behavior (VAB) Composite Scale than children who receive standard care. Estimated effect size is a SMD of 0.952 (95% CI 0.507 to 1.400, p <0.001), which corresponds to a between-group difference of 10.7 points. Strength and Stability of the Evidence: Moderate.**
- **The evidence was insufficient to determine whether children with ASD who receive early intensive behavioral intervention continue to perform more adaptive behaviors than children who receive standard care at later follow-up times (greater than one year).**
- **The evidence was insufficient to determine whether children with ASD who receive early intensive behavioral intervention perform better on tests of language and communication than children who receive standard care.**
- **For the following outcomes: problem behaviors, academic/developmental achievement, and parental/family well-being, the limited size (one study per outcome) and quality of the evidence prevented us from drawing conclusions about whether early intensive behavioral intervention was more effective than standard care in improving these outcomes for children with ASD.**

Three studies enrolling a total of 128 children with a diagnosis of autistic disorder or pervasive developmental disorder-not otherwise specified (PDD-NOS) addressed this question. Each study compared the efficacy of early intensive behavioral intervention (EIBI) to standard

care.(65,66,70) In all three studies the primary outcome was intellectual or cognitive functioning, which was measured using age-appropriate standardized tests of intelligence (IQ tests). All three studies also assessed language or communication skills using the Reynell Developmental Language Scales and adaptive behavior using the Vinland Adaptive Behavior Scale. Additionally, one study assessed problem behavior and family well-being using various validated checklists and questionnaires, and another study assessed classroom placement.

The median quality assessment score for the studies was moderate (median score 6.5, range 6.4 to 7.5). Table 22 through Table 27 in Appendix D presents the quality assessment scores for each outcome reported in each of the studies. All three studies were non-randomized controlled trials. Lack of randomization introduces the potential of selection bias, in which participants in the experimental condition differ from participants in the control condition in ways that may impact the effect of treatment. To enhance group comparability, the authors in all three studies addressing Key Question 1 either matched participants on key variables, such as IQ and chronological age(70) or used statistical procedures, such as multiple regression to assess whether any between-group differences at baseline affected study outcomes.(65,66) However, in only two of the studies the authors reported that the outcome assessors were blinded to which type of treatment (EIBI or standard care) the children received.(66,70) Finally, in only one study the authors reported assessing treatment fidelity of the providers in the EIBI group.(70)

Patient Baseline Characteristics of Included Studies

All three studies included young children who met the DSM-IV criteria for a primary diagnosis of autistic disorder or PDD-NOS. The average age of the children across the studies was 2.8 years (range 2.5 to 3.2). In the two studies that reported on gender, the majority of the children in the studies were boys (>80% in each study).(65,70) None of the children in the studies had any serious co-morbid medical conditions. The average baseline IQ score across the studies was 60.5 (SD 23.21), which indicates that most of the children had a significantly lower than average level of general intellectual functioning. The average level of intellectual functioning as measured by standardized tests of IQ among typically developing children is 100 (range 85 to 115).(78) Table 29 in Appendix E provides further information about the children who participated in the studies.

Treatment Characteristics of Included Studies

All three studies compared EIBI to standard care. In all three studies, children in the EIBI condition received 25 or more hours (range 25 to 40 hours) of treatment per week. Treatment in the EIBI condition was based on the principles of applied behavior analysis (ABA), which uses behavioral interventions, such as prompting, shaping, and the use of various reinforcements, in a systematic manner to produce positive changes in children with ASD. One of the studies specifically used the UCLA Young Autism /Lovaas model of ABA (see the section on *Comprehensive Educational and Behavioral Interventions* of this report for a description of the Lovaas model).(70)The other two studies did not report that the treatment was intended to follow any specific model of ABA.(65,66) The authors of the Remington et al. study reported that some children in the ABA group were receiving other interventions, such as speech therapy, dietary interventions, routine prescription medication, vitamin therapy, and homeopathic interventions (See Table 30 for further details). In all three studies, children in the EIBI group received one-to-one treatment primarily in the home setting. Treatment was delivered by a highly trained multidisciplinary team of therapists. Parents also participated in delivering the treatment. In all

three studies, treatment was delivered for more than one year (range 1.2 to 3 years). Table 30 provides more information about the treatment received in the EIBI condition.

Children in the standard care condition in all three studies received 15 or fewer hours of treatment per week. Treatment consisted of a mix of interventions provided through the local education system, and, as such, varied across the studies in terms of the specific interventions used. In the Remington et al. study, children received their local education standard provision of services provided in the United Kingdom, which for 52% of children included methods drawn from TEACCH, 76% of children received the Picture Exchange Communication System (PECS), and 48% received instruction in sign language or Makaton communication (see the sections on *Comprehensive Educational and Behavioral Interventions* and *Competing/Complimentary Treatments* of this report for description of these interventions).(66) None of the interventions were delivered on a one-to-one basis. In the Cohen et al. study, children in the standard care group received classroom-based special education three to five days a week for up to five hours a day plus related services such as speech, occupational and behavioral therapy.(70) Teacher to student ratio varied from one-to-one to three-to-one for children in this group. Similarly, children in the standard care group in the Howard et al. study received non-specific classroom-based special education for 15 hours per week.(65) Teacher to child ratio in this study was one-to-six. In all three studies, children in the standard care group received treatment for the same length of time as children in the EIBI group. Table 30 provides more information about the treatment received in the standard care condition.

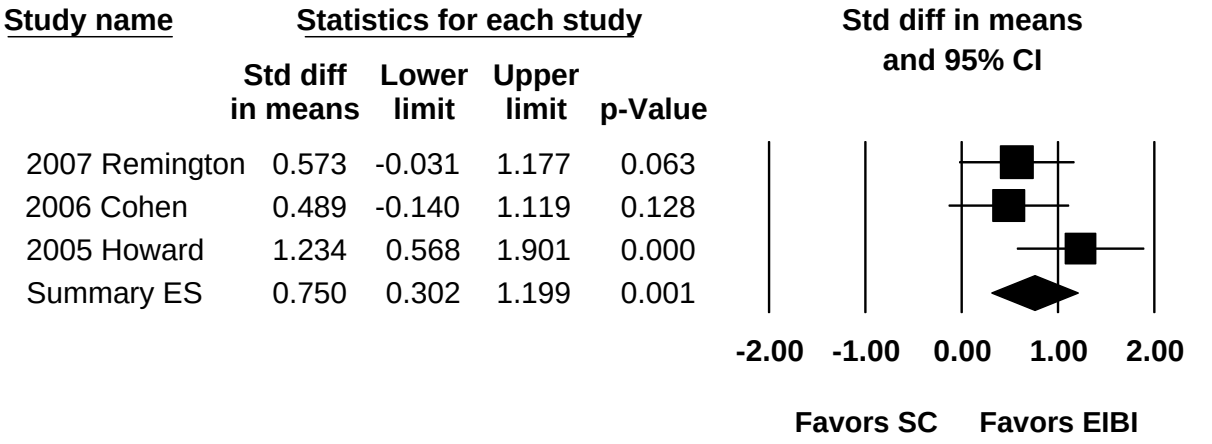
Individual Study and Meta-Analytic Results

Intellectual/Cognitive Status

All three studies that addressed Key Question 1 reported data from IQ tests at or close to one year followup and two studies reported data at later follow-up points (Remington et al. at two years and Cohen et al. at three years). The studies used either the Bayley Scales of Infant Development, Stanford Binet test, or Wechsler Primary Preschool Scales of Intelligence to measure IQ. Each of these tests has a normative mean of 100 with a standard deviation of 15 points.(78) Table 31 in Appendix F presents individual study results of the overall performance of the study groups on tests of IQ. We pooled data for this outcome into two separate random-effects meta-analyses—one using data from the first followup (about one year) and a second using data from the two studies reporting at later time points (two and three years).

Heterogeneity testing indicated that the three studies included in the meta-analysis of first follow-up data were moderately consistent (I^2 was 33.7). The estimated random-effects summary statistic was an SMD 0.750 (95% confidence intervals [CI] 0.302 to 1.199), which corresponds to a statistically significant effect. However, because the lower 95% CI is slightly less than the minimal threshold of the *minimally* clinically significant value (SMD of 0.380), it is unclear whether the effect is clinically significant. Figure 4 below presents the results of our meta-analysis. The results indicate that EIBI for young children with autistic disorder or PDD-NOS leads to higher scores on tests of IQ than standard, less intensive care, with a between-group difference of 14.8 points in overall IQ scores. However, because the quality of the studies and the size of the estimated effect were moderate, the strength and stability of the evidence supporting the results of our analysis was moderate. The results of our sensitivity analyses are presented in Figure 11 and Figure 12 in Appendix I.

Figure 4. Difference in Scores on Standardized IQ Tests at 1 Year Followup



Random Effects Meta Analysis/ I^2=33.7

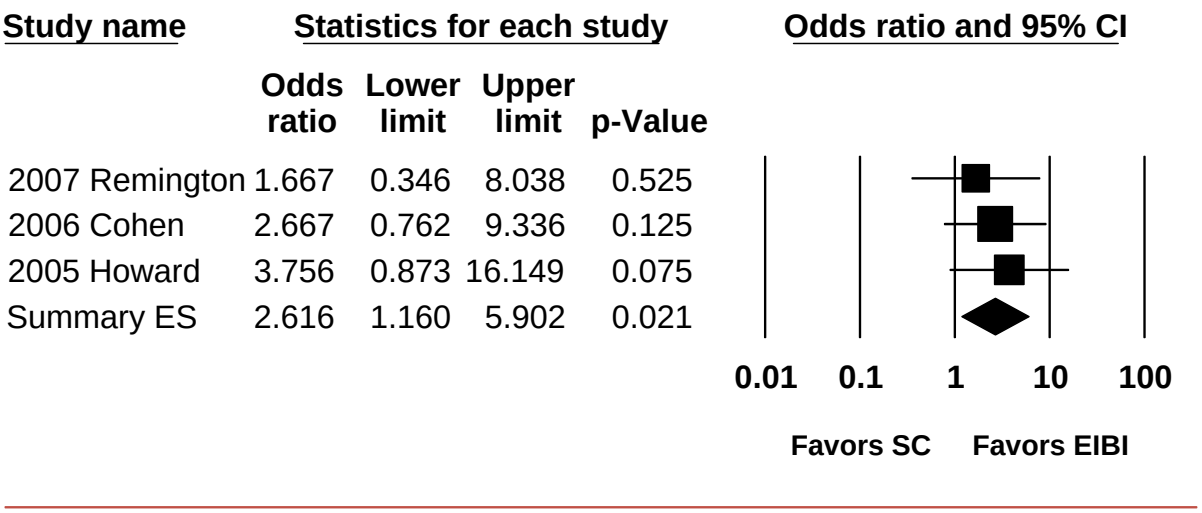
Both Remington et al. and Cohen et al. measured IQ at later follow-up points (greater than one year). At two years followup, there was no significant difference in IQ between children in the EIBI group and children in the standard care group in the Remington et al. study. Similarly, there was no significant between-group differences in the Cohen et al. study at three year's followup. We combined data from the last follow-up point in each study in a random-effects meta-analysis. However, since only two small studies contributed data to this analysis and the observed effect was not extremely large, we deemed the evidence insufficient to permit conclusions.

In addition to reporting the average IQ scores of the study groups, all three studies reported the proportion of children whose score reached the normal range for typically developing children (IQ score of 85 or higher). Data on this outcome were only reported at last followup in each study. In all three studies more children who received EIBI had IQ scores at followup that reached 85 or higher than children who received standard, less intensive care. Across the studies, the percentage of children in the EIBI group whose IQ score was 85 or higher at followup ranged from 26% to 57%, compared to 14% to 33% of children in the standard care group. Table 36 in Appendix F presents data from each study on the number of children whose IQ score reached within the normal range.

We pooled the data from the three studies on the number of children in each group whose IQ score reached 85 or higher into a random-effects meta-analysis. Consistent with the findings of the individual studies, the results of our analysis indicated that children in the EIBI group were more likely to achieve an IQ within normal range for typically developing children than children in the standard care group. The estimated size of the effect was an odds ratio of 2.616 (95% CI 1.160 to 5.902, p = 0.021), which was both statistically and clinically significant. Figure 5 below presents the results of our meta-analysis. Sensitivity testing, however, indicated that the results of our analysis were not quantitatively robust. Removal of either the Cohen et al. study or the Howard et al. study caused the lower 95% CI to overlap the minimal threshold of clinical significance for this outcome (an odds ratio of 1.0). Thus, stability of the evidence supporting the

results of our analysis was considered low (Figure 14 and Figure 15 presents the results of our sensitivity analyses). For the qualitative conclusion that there was a clinically significant difference between groups, we rated the strength of the evidence as moderate, because the analysis did not pass our qualitative robustness tests (again, removal of either the Cohen et al. study or the Howard et al. study caused the lower 95% CI to overlap the minimal threshold of clinical significance for this outcome).

Figure 5. Difference in Number of Children Achieving an IQ Score of 85 or Above at Last Followup



Language/Communication Skills

All three studies that addressed Key Question 1 measured children's verbal and nonverbal communication skills. The Reynell Developmental Language Scales were used in all three studies to measure children's verbal communication skills. The Reynell measures both receptive and expressive language among children between the ages of one and seven years old. Higher scores on this instrument indicate higher language skills. Follow-up data for this instrument, however, were only reported in two of the three studies—Cohen et al.(70) and Howard et al.(65) Table 32 (Appendix F) presents individual study results for the Reynell scales for both of these studies.

The authors of the third study reported that they were unable to obtain a score for some of the children using the Reynell Scales due to their limited verbal abilities, so raw data were incomplete.(66) Instead of reporting actual scores, the authors of this study reported the number of children within each study group who scored on the Reynell at followup. Individual study results, which are presented in Table 33 in Appendix F, indicated that more children in the EIBI group scored on both the Reynell receptive and expressive language scales at one and two year's followup than children in the standard care group.

After one year of treatment, children who received EIBI in both the Cohen et al. study and Howard et al. study demonstrated significantly higher scores on both the Reynell scale for expressive and receptive language than children who received standard care. However, only Cohen et al. reported data for further follow-up times. In this study, there were no significant between-group differences observed on either of the Reynell scales at three year's followup. We combined the data reported after one year of treatment from these studies in two separate random-effects meta-analyses—one for scores on the Reynell scale for expressive language and one on scores for receptive language. However, because there were only two combinable studies, neither of which was a multicenter study, and the observed effect was not extremely large, we deemed the evidence insufficient to permit conclusions. Figure 16 and Figure 17 in Appendix I present the results of our analyses.

As previously indicated, all three studies also measured non-verbal communication skills among children in the study groups. However, not all the studies measured this outcome using the same instrument. Non-verbal communication skills in the Remington et al.(66) study were measured using the Early Social Communication Scales, while in the Cohen et al. study(70) and the Howard et al. study(65) the Merrill-Palmer Scale was used. Table 17 describes each of these instruments, and Table 34 (Appendix F) presents individual study results on measures of non-verbal communication skills. We combined data from the first follow-up (one year) of the two studies that reported on the Merrill-Palmer scale in a random-effects meta-analysis. However, again only two studies contributed data to this analysis, neither study was a multicenter study, and the observed effect was not extremely large. Thus, we deemed the evidence insufficient to permit conclusions. The results of our analysis are presented in Figure 18 in Appendix I.

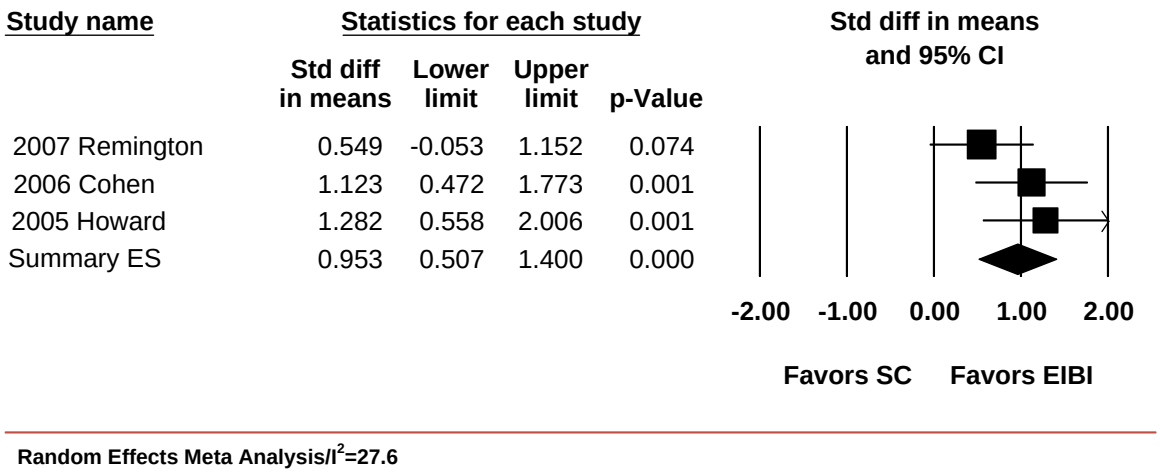
Adaptive Behavior

All three studies measured adaptive behavior using the Vinland Adaptive Behavior Scales (VABS). The VABS is administered by interviewing the child's parents, teachers, or care providers. The VABS assesses adaptive behavior across five domains—communication, daily living skills, socialization, motor skills, and maladaptive behaviors. Raw scores from each domain are converted to standard scores with a population mean of 100 and standard deviation of 15. Higher scores indicate better outcomes or performance of more adaptive behaviors. All three studies reported data on the VABS after one year of treatment, but only two of the studies reported data at further follow-up times (Remington et al. two years and Cohen et al. three years). Data were reported in a manner that allowed us to perform two separate meta-analyses using the composite score of the VABS— one using data from the one year followup and a second using data from the two studies reporting at later time points. Since not all three studies reported scores for all five subscales (or domains), we only performed a meta-analyses using data from the composite score of the VABS. Individual study results for the VABS are presented in Table 35 in Appendix F.

Heterogeneity testing indicated that the studies included in the first meta-analysis were moderately consistent (I^2 was 27.6). The estimated random-effects summary statistic was a SMD of 0.953 (95% CI 0.507 to 1.400, $p < 0.001$), which corresponds to a statistically and clinically significant between-group difference of 10.7 points. Figure 6 below presents the results of our analysis. Sensitivity testing indicated that the findings of our analysis were both quantitatively and qualitatively robust. However, because the quality of the studies supporting our conclusion was moderate, we rated the strength and stability of our finding as moderate. Figure 19 and Figure 20 in Appendix I present the results of our sensitivity analyses. Since only two small

studies contributed data to our second analysis of VABS scores at later time points and the observed effect was not extremely large, we deemed the evidence insufficient to permit conclusions. The results of this analysis are presented Figure 21 in Appendix I.

Figure 6. Meta-Analytic Results of Vinland Adaptive Behavior Composite Score at 1 Year Followup



Problem Behaviors

Only one study reported data on this outcome. Remington, et al. used various validated checklists and questionnaires to assess changes in a child's behavior problems, prosocial behaviors, and autistic behavior.(66) For all of the assessments, parents were asked to rate their child's behavior. Table 12 lists the instruments used in this study and Table 17 in Appendix B provides a brief description of the instruments. For most measures, no statistically significant differences were observed between parent ratings of children in the EIBI and those of children in the standard care group. However, at one year followup, mothers of children in the EIBI group reported significant improvement in positive social behavior on the Nisonger Behavior Checklist compared to mothers of children in the standard care group. Significant differences were no longer observed for this outcome at two year's followup. Individual results for each instrument used in this study to assess problem behaviors are presented in Table 37.

Academic/Developmental Achievement

Only Cohen et al. reported data on this outcome.(70) In this study, achievement was measured as placement into integrated regular education without assistance. The authors of the study found a statistically significant difference in the number of children who received EIBI and were placed in regular education compared to the number of children who received standard care. Specifically, six (29%) of the children who received EIBI were placed into regular education without assistance compared to zero children who received standard care.

Parental/family Well-being

Remington et al. was the only study to report data on parental/family well-being.(66) In this study various aspects of parent and family well-being were measured, including parental stress, anxiety, depression, and access to resources of support. Table 12 lists the instruments used in this

study and Table 17 in Appendix B provides a brief description of the instruments. For most measures, no statistically significant differences were observed between parents in the EIBI and those in the standard care group at one year and two year's followup. However, mothers in the EIBI group reported significantly more anxiety than mothers in the standard care group at one year followup ($p = 0.039$). This finding was no longer significant at three year's followup. Individual results for each instrument used in this study to assess parental/family stress are presented in Table 38.

Key Question 2: *Is one comprehensive educational or behavioral intervention more effective than another in improving outcomes for children with ASD?*

- **For intellectual/cognitive status, language/communication skills, and adaptive behavior, clinical differences between the studies reporting on these outcomes (children in one study significantly older than children in the other study) prevented us from drawing any conclusions about whether intensive applied behavior analysis was more effective than an intensive eclectic intervention program in improving these outcomes for children with ASD.**
- **For problem behaviors and academic/developmental the limited size (one study per outcome) and quality of the evidence prevented us from drawing conclusions about whether intensive applied behavior analysis was more effective than an intensive eclectic intervention program in improving these outcomes for children with ASD.**
- **None of the studies that addressed this Key Question reported data on parental/family well-being.**

Overall Evidence Base

Three studies, one reported in two separate publications, enrolling a total of 109 children with a diagnosis of autistic disorder or PDD-NOS addressed this question.(65,67-69) Each study compared the efficacy of intensive applied behavior analysis (ABA) to an intensive eclectic intervention program. In two of the studies, the primary outcomes were intellectual/cognitive status, language skills, and adaptive behavior.(65,67) In the third study, intellectual functioning was secondary to change in the status of the study groups on the core symptoms of autism.(69) In this study, the Autistic Disorder Observation Scale (ADOS) was used to measure change in autistic severity along the following domains: language and communication and reciprocal social interaction. Additionally, one of the three studies assessed change in problem behaviors using the Achenbach Child Behavior Checklist.

The median quality assessment score for the studies was moderate (median score 6.4, range 6.4 to 6.8). Table 22 through Table 27 in Appendix D presents the quality assessment scores for each outcome reported in each of the studies. All three studies were non-randomized controlled trials, which increases the risk of selection bias. However, in one study the children in the study groups were carefully matched for chronological age, autism severity, and cognitive functioning (IQ).(69) In the other two studies, the authors used statistical methods to enhance group comparability.(65,67) In only one of the studies the authors reported that the outcome assessors were blinded to which intervention the children received.(67) None of the authors in all three of the studies reported whether treatment fidelity was assessed.

Patient Baseline Characteristics of Included Studies

In all three studies, children met the DSM-IV criteria for a primary diagnosis of autistic disorder or PDD-NOS. In two of the studies, one by Zachor et al. and one by Howard et al., the children were young with an average age across the studies of 2.6 years (range 2.5 to 2.75).(65,69) In the third study by Eikeseth et al., which was published in two separate publications reporting on different follow-up times (one year and three years), the children were substantially older with an average age of 5.5 years (range 5.1 to 5.8).(67) Because of the large difference in age (three years), we did not attempt to combine data from the Eikeseth et al. study in any analyses with data from the other two studies.

In all three studies, the majority of the children who received treatment were boys (>80.0%). None of the children had any serious co-morbid medical problems. The average baseline IQ score across the studies was 64.6 (SD 15.2), which indicates that most of the children had a significantly lower than average level of general intellectual functioning. Table 29 in Appendix E provides further information about the children who participated in each of the studies.

Treatment Characteristics of Included Studies

All three studies compared the efficacy of intensive ABA to an intensive eclectic program. In all three studies, children in the ABA group received between 20 to 40 hours of one-to-one instruction per week. Treatment was provided by a multidisciplinary team that included therapists, teachers, aides, and parents. One of the studies, reported in separate publications, specifically used the UCLA Young Autism /Lovaas model of ABA (see the section on *Comprehensive Educational and Behavioral Interventions* of this report for a description of the Lovaas model).(67,68) The other two studies did not report that the treatment was intended to follow any specific model of ABA.(65,69) In this study, children received treatment addressing various developmental fields, such as imitation, receptive and expressive language, joint attention, non-verbal communication, pre-academic skills, play, fine motor skills, and adaptive behavior. In two studies, treatment was primarily delivered outside of the home at school or in a center.(67,69) In the third study, the treatment setting was mixed with treatment occurring in the home, school, and community.(65) In all three studies, children received treatment for at least one year (range across studies 1 to 3 years). Table 39 in Appendix G provides more information about the treatment received in the ABA condition.

Currently, there is no single definition or description of "eclectic" treatment. In the three studies addressing this question, children in the eclectic group received a mix of interventions drawn from various approaches including TEACCH, DIR, and ABA (see the section on *Comprehensive Educational and Behavioral Interventions* of this report for a description of these interventions). In all three studies, children in the eclectic group received about the same number of hours of treatment (between 20 to 40 hours per week) delivered in the same manner (mostly one-to-one instruction) for the same length of time (one or more years) as children in the ABA group.

In the study by Eikeseth et al., children in the eclectic group received the following: alternative communication (25% of children), ABA (42% of children), total communication (17% of children), sensory-motor therapy (25% of children), TEACCH (50% of children), clinical experience (50% of children), and/or some other type of intervention (67% of children).(67,68) In the study by Zachor et al., children in the eclectic group received individual therapy from various therapists, including speech and language, occupational and music therapies, and structured cognitive teaching.(69) Finally, children in the eclectic group in the Howard et al.

study received a variety of methods designed for children with autism, including discrete trial training, Picture Exchange Communication System (PECS), sensory integration, and methods drawn from the TEACCH model.(65) Table 39 in Appendix G provides more information about the treatment received in the eclectic condition.

Individual Study Results

Intellectual/Cognitive Status

All three studies reported baseline scores of each study group on tests of IQ. However, only Eikeseth et al. and Howard et al. reported scores at followup. In the Eikeseth study, follow-up scores were reported in two separate publications, one publication reporting on scores after one year of treatment(68) and the second publication reporting on scores after three years of treatment.(67) The follow-up time in Howard et al. was after 1.2 years of treatment.(65) In both of these studies IQ was measured using either the Bayley Scales of Infant Development or the Wechsler Primary Preschool Scales. However, because of the previously mentioned age difference of the children in the Eikeseth study compared to the children in the other two studies (difference of three years), no meta-analyses were preformed for this outcome. Below, we briefly describe key elements of the results of each study. Further information about the results of each study for this outcome is presented in Table 40 in Appendix G.

In the Eikeseth study, performance on tests of IQ improved for both children in the ABA group and children in the eclectic group, with no statistically significant between-group differences observed at one year or three year's followup.(67,68) In the Howard study, children in both groups demonstrated improvement from baseline to followup, but children in the ABA group had significantly higher IQ scores compared to children in the eclectic group after 1.2 years of treatment ($p < 0.001$).

Both Eikeseth et al. and Howard et al. reported on the number of children who scored within the average range of typically developing children (score of 85 or above) on tests of IQ at followup. In both studies, more children in the intensive ABA group increased their scores to within the average range than children in the intensive eclectic group. Seven of 13 children (54%) in the ABA group in the Eikeseth et al. study increased their IQ scores to within normal range at three years followup, compared to only two of 12 children (17%) in the eclectic group. The difference between the groups, however, was not statistically significant ($p = 0.06$). Similarly, in the Howard et al. study, 13 of 26 children in the ABA group increased their IQ scores to within normal range at 1.2 year's followup, compared to zero of 16 children in the eclectic group. The between-group difference in this study was statistically significant ($p = 0.04$).

Language/Communication Skills

Both Eikeseth et al. and Howard et al. measured language skills using the Reynell Developmental Language Scales.(65,68) As previously described under Key Question 1, the Reynell measures receptive and expressive language among children between the ages of one and seven years old. Higher scores indicate higher language skills. No meta-analyses were performed for this outcome due to the difference in overall age of the children in the two studies reporting on this outcome. Individual study results indicated that children in the ABA group in the Eikeseth study scored significantly higher than children in the eclectic group at one-year followup (three year follow-up data not reported) on the expressive language scale ($p = 0.004$). No significant between-group difference was demonstrated on the receptive language scale. In Howard et al, the ABA group scored significantly higher than the eclectic group on both

receptive and expressive language scales (p = 0.025 and 0.027, respectively). Children in the ABA group in this study also scored significantly higher than the eclectic group on the Merrill-Palmer scale (p = 0.020). Table 41 in Appendix G provides more information on the individual study results for this outcome.

Adaptive Behavior

Both Eikeseth et al. and Howard et al. measured adaptive behavior using the Vinland Adaptive Behavior Scales (VABS), which as previously described under Key Question 1 measures a child's communication skills, daily living skills, socialization, motor skills, and maladaptive behaviors based on an interview with the child's caregiver(s) or teacher. Higher scores indicate better outcomes or performance of more adaptive behaviors. Individual study results indicated that the overall composite score of the VABS in both studies was significantly higher for children in the ABA group than for children in the eclectic group at both one and three year followup in the Eikeseth study (p = 0.048 and p <0.001, respectively) and at 1.2 year followup in the Howard study (p = 0.003). Table 42 in Appendix G provides further study level results for each of the subscales of the VABS.

Problem Behaviors

Only one of the three studies that addressed Key Question 2 reported data on this outcome. Eikeseth et al. used the Achenbach Child Behavior Checklist (ACBC)-Teacher Edition to measure change in problem behaviors of children receiving ABA or an eclectic intervention program.(67) The ACBC is for children four to 18 years old and is completed by an adult informant, typically the child's primary caregiver or a teacher.(78) A child's behavior is rated along several categories of behavior including: aggressive behavior, anxious/depressed behavior, attention problems, delinquent rule-breaking behavior, social problems, somatic complaints, thought problems, and withdrawnness. Higher scores on this instrument reflect more problem behaviors. For the most part, children in the ABA group in the Eikeseth study did not differ from children in the eclectic group at three year's followup on most items of the scale. However, children in the ABA group displayed significantly fewer social problems and aggressive behavior than children in the eclectic group (p = 0.039 and 0.002, respectively). Table 43 in Appendix G presents individual study findings for this outcome.

Academic/Developmental Achievement

Zachor et al. assessed the stability of diagnosis of autism spectrum disorder of children in the study groups prior to the start of treatment and after one year of treatment.(69) Stability of diagnosis was assessed using the Autism Diagnostic Observation Schedule (ADOS), which is a semistructured assessment of social interaction, communication, play, and imaginative use of materials included in the test packet.(9) The ADOS is standardized in terms of the materials used, the activities presented, the examiner's introduction of activities, the hierarchical sequence of social presses provided by the examiner, and the way behaviors are coded or scored. Following the administration of the ADOS, behaviors are coded using a 0- to 3-point coding system, with a 0 indicating that the behavior is not abnormal in the way specified in the coding description and a 3 indicating that a behavior is abnormal and interferes in some way with the child's functioning.

ADOS classifications are based on specific coded behaviors that are included in a scoring algorithm using the DSM-IV diagnostic criteria, resulting in a Communication score, a Reciprocal Social Interaction score, and a Total score (a sum of the Communication and

Reciprocal Social Interactions scores). Scores are compared with an algorithm cut-off score for autistic disorder or the more broadly defined ASD in each of these areas. If the child's score meets or exceeds cut-offs in all three areas, they are considered to meet criteria for that classification on the measure.

Based on the ADOS criteria, four out of 19 (21%) children in the intensive ABA group in the Zachor et al. study no longer met the diagnostic criteria for autistic spectrum disorder after one year of treatment, compared to zero out of 18 children in the eclectic group. This difference, however, was not statistically significant ($p = 0.121$). Similarly, four out of 19 (21%) children in the ABA group changed from a diagnosis of autistic disorder to a less severe form of ASD, compared to three out of 18 (17%) children in the eclectic group. Again, the difference between groups was not statistically significant ($p = 0.734$). In both treatment groups, about 80% of children remained stable in their diagnosis from pre-to-post-treatment.

Parental/family Well-being

None of the studies that addressed this Key Question reported data on parental/family well-being.

Key Question 3: *Are home-based interventions of similar intensity and/or structure as comprehensive educational or behavioral interventions provided in other settings (e.g., center or clinical setting) more effective in improving outcomes for children with ASD?*

ECRI Institute's searches of the literature did not find any studies that met our study selection criteria that directly compared one treatment setting to another. However, we did identify two studies that compared clinic-directed early intensive behavior intervention (EIBI) delivered primarily in the home to parent-directed EIBI.

- **Differences between the studies comparing clinic-directed EIBI to parent-directed EIBI in terms of how services were delivered in the parent-directed group precluded us from drawing any conclusions about whether clinic-directed EIBI is more effective than parent-directed EIBI for children with ASD.**

Overall Evidence Base

The two studies considered under Key Question 3, one by Sallows & Graupner(71) and the other by Smith et al.(72), enrolled a total of 51 children with a diagnosis of autistic disorder or PDD-NOS. The primary outcomes in both the studies were intellectual/cognitive functioning, language and communication skills, adaptive behavior, and classroom placement. Smith et al. also reported on family satisfaction.(72) However, this outcome was measured using a non-validated instrument. Thus, we do not report data for this outcome.

The median quality assessment score for the two studies was high (median score 8.8, range 8.8 to 9.3). Table 22 through Table 27 in Appendix D presents the quality assessment scores for each outcome reported in the studies. Both studies were randomized controlled trials, however, neither study reported whether there was concealment of randomization. Both studies reported that treatment fidelity was assessed, but in only one study were all outcome assessors blinded to which treatment condition the children received.(72)

Patient Baseline Characteristics of Included Studies

Both studies comparing clinic-directed to parent-direct EIBI included young children who met the DSM-IV criteria for a primary diagnosis of autistic disorder or PDD-NOS. The average age of the children across the studies was 2.9 years (range 2.8 to 3.0). In both studies, the majority of children were boys (>80% in each study). None of the children in the studies had any serious co-morbid medical conditions. The average baseline IQ score across the studies was 51.1 (SD 10.55), which indicates that most of the children had a significantly lower than average level of general intellectual functioning. Table 29 in Appendix E provides further information about the children who participated in these studies.

Treatment Characteristics of Included Studies

Children in the clinic-directed group in both studies received intensive applied behavior analysis based primarily on methods of the Lovaas model (see the *Comprehensive Educational and Behavioral Interventions* section of this report for a description of this model). Children in both studies received one-to-one treatment delivered by a multidisciplinary team of therapist for 30 or more hours per week for the first year of treatment. In both studies, treatment was gradually decreased or phased out after about two years when children entered school.

While children in the parent-directed group in both studies received treatment that was based on the Lovaas model of ABA, the level of parent training and intensity and nature of the treatment provided to the children varied considerably between the two studies. In the study by Sallows & Graupner, parents received extensive training on the Lovaas method of ABA and were the primary providers of treatment.(71) The intensity of treatment in this study was determined by the parents, and averaged 32 hours/week the first year and 31 hours/week during the second year. Parent-directed children also received six hours of in-home supervision per month by a therapist highly trained in the methods used in the Lovaas model.

In the study by Smith et al., parents received two sessions per week of training in the methods employed by the Lovaas model totaling five hours per week, in their homes for three to nine months. At termination of training, parents were asked to implement the methods learned during training with their child for five hours per week. Throughout the course of the study, children in the parent-directed group were enrolled in special education classes for ten to 15 hours per week, with no direct instruction based on methods of the Lovaas model. Because of differences in the intensity and nature of the treatment delivered to children in the parent-directed group in the two studies, we did not attempt to combined data from these studies in any analyses. Below, we provide key elements of the results of each study. Table 44 in Appendix H provides more information about the treatments assessed in each study, and Table 45 through Table 48 presents individual study results.

Individual Study Results

In the Sallows & Graupner study the scores of children in both the clinic-directed and parent directed group improved significantly from pretreatment to follow-up on measures of IQ.(71) However, no significant between-group differences were observed. Similar results were demonstrated on measures of language and communication and adaptive behavior and in the number of children who achieved IQ scores within normal ranges (85 to 115). The authors suggest that the similarity in performance of children in the two groups at followup was likely due to children in the parent-directed group receiving roughly the same number of hours of

treatment as children in the clinic-directed group, and that, for the most part, parents in the parent-directed group were highly motivated and very supportive of each other (e.g., filling in for one another to provide treatment and actively seeking peer support for their children).

In the course of their analyses, the authors of the Sallows' study noted a bimodal distribution of pre-post scores on measures of IQ across the two study groups, indicating a group of children who showed rapid progress and a group that showed moderate progress, with no overlap between outcome distributions. As a result, the authors chose to collapse the study groups and performed subsequent analyses on the two subgroups of children—rapid learners and moderate learners. While no data are reported in this review on the subgroup analyses (as such collapsing of the two groups essentially makes this study a case series study), the overall results reported by the authors indicate that at followup the rapid learners were succeeding in regular classrooms, with fluent verbal skills and socially interacting with peers on a regular basis.

In the Smith et al. study, statistically significant differences in favor of the clinic-directed group were demonstrated for IQ ($p = 0.027$) and on the Merrill-Palmer test of non-verbal communication ($p = 0.030$). Performance on Reynell Language Development Scales and Vineland Adaptive Behavior Scales did not demonstrate any significant differences between children who received clinic-directed treatment and those who received parent-directed treatment. Parent and teacher ratings of problem behaviors using the Achenbach Child Behavior Checklist indicated little difference between the two treatment groups. However, children in the clinic-directed group were reported to have significantly less anxiety and depression (as reported by parents) and withdrawal (as reported by teachers) than children in the parent-directed group.

Key Question 4: *What adverse events and harms have been reported to occur in association with the use of comprehensive educational or behavioral interventions for children with ASD?*

- **None of the studies that met our study selection criteria for this review reported whether or not any adverse events occurred.**

Key Question 5: *What is the consensus among experts about the safety and efficacy of comprehensive educational or behavioral interventions for the treatment of children with ASD?*

ECRI Institute's searches of the National Guideline Clearinghouse™ (NGC)™ and the Healthcare Standards database identified six treatment guidelines published between the years 2000 to present that included recommendations for the use of comprehensive educational and behavioral interventions for children with ASDs. The guidelines were published by the following organizations:

- New York Department of Health Early Intervention Program 2008(81)
- Scottish Intercollegiate Network (SIGN) 2007(82)
- Canadian Pediatric Society 2004 (reaffirmed 2008)(83,84)
- Alberta Heritage Foundation for Medical Research 2001(85)
- Canadian Coordinating Office for Health Technology Assessment (CCOHTA) 2001(86)
- British Columbia Office of Health Technology Assessment 2000(87,88)

In general, the guidelines published by the organizations listed above recommend that treatment involving behavioral interventions, such as ABA, should be initiated when the child is young, include a minimum of 15 to 20 hours per week of one-to-one instruction, be designed to fit the needs of the individual child, and include the family in the planning and provision of services. However, there is overall agreement among the organizations that the existing evidence supporting one comprehensive intervention over another is limited due to methodological flaws (e.g., lack of randomized controlled trials, small sample sizes, etc.) in most of the published research. Further, more research is needed to identify 1) the common effective elements of treatment programs, 2) the effects of treatment on children across the full spectrum of autism, 3) the optimal age and IQ range of children who derive the most benefit, 4) the optimal intensity and duration of treatment, and 5) whether gains on outcomes such as IQ translate to improved quality of life for children with ASDs. More information on the objectives, treatments considered, and specific recommendations of the published guidelines reviewed for this report is provided in Table 49 in Appendix K.

Our searches of the literature also identified position and consensus statements from the following organizations/authors:

- Ministry of Health, New Zealand, 2008(89)
- Association for Science in Autism, 2007(46)
- Prior and Roberts, 2006(90)
- National Early Childhood Technical Assistance Center (NECTAC)(91)

In general, the position of these organizations is similar to that found in the published guidelines summarized above. Specifically, the position of these organizations is that treatment should start early, include a minimum of 20 hours of structured teaching per week, address multiple symptoms/deficits of children with ASDs, including deficits in language and social skills and problem behaviors, and encourage partnership between parents and professionals. Table 50 in Appendix K provides more information about the recommendations presented in the additional position and consensus statements. Finally, several states have developed practice guidelines for the provision of treatment to individuals with ASDs. Table 51 provides more information about individual state guidelines.

Findings of Other Systematic Reviews

In part, this review serves to extend the findings of a previous systematic review produced by ECRI Institute in 1999, titled *Comprehensive Programs for the Treatment of Autism*.(30) In the previous review, we addressed the following three key questions: 1) Does any comprehensive program, alone or in conjunction with any other intervention, lead to clinically significant improvements in autistic symptomology; 2) Can improvement be attributed to the particular comprehensive treatment program in question; and 3) Which comprehensive program is most effective. Most of the analyses in our previous review consisted of assessing the validity of the methods and designs of each relevant study. Only the results of studies considered to employ valid methods and of sound design were examined.

A total of six studies met the methodological criteria for inclusion in ECRI Institute's previous review. Specifically, each study included a control group that was considered comparable to the experimental group at baseline on the following variables: chronological age, IQ, and number of autistic symptoms. None of these studies, however, met the current review's study selection criteria. In five of the studies the diagnostic criteria was not based on current standards established in the DSM-IV or TR edition and one study was a retrospective case-controlled trial (See Table 16 for further details).

Four of the studies included in the previous review assessed the effectiveness of the Lovaas model of applied behavior analysis(92-94), one assessed the TEACCH program(95), and the final study assessed the Autism Preschool Program.(96) Based on the results of the individual studies, the following primary conclusions were drawn: 1) the functioning of the children described in the studies of the Lovaas program appeared to improve, but it is unclear whether any improvement can be conclusively attributed to the treatment; 2) the one study of the TEACCH program found statistically significant improvements in imitation, perception, fine motor skills, gross motor skills, and cognitive performance, but it is unclear whether these findings are clinically or practically meaningful; and 3) the language abilities of the children described in the one study of the Autism Preschool Program appeared to improve, but, again, it is unclear whether improvements were clinically meaningful. Since none of the studies included in the previous review compared one comprehensive program to another, no conclusions could be drawn about the comparative effectiveness of one program to another.

Our searches of the literature for more recent systematic reviews identified 11 reviews all published between 2000 and 2008. All of the reviews included studies that reported data on intensive or comprehensive behavioral and/or educational interventions for children with ASD. Some of the reviews also included information from studies that assessed the effectiveness of other interventions (e.g., focal interventions, pharmacological treatments)(97-100), and one of the reviews included studies that assessed treatments for children that suffered from disorders other than ASD, such as depression and attention deficit hyperactivity disorder (ADHD).(100) Since the scope of this review focuses specifically on comprehensive interventions for children with ASDs, we only report on information from the previous reviews that was specific to the interventions and population of interest in this review. Table 53 (Appendix K) presents important information about the search strategy, patient populations, methodology, results, and authors' conclusions of the previous reviews.

In general, the conclusions reached by previous systematic reviews agree with those drawn in ECRI Institute's current review. Specifically, that early intensive behavior intervention (EIBI) appears to improve intellectual functioning and adaptive behavior for some children with ASD. Similarly, most other reviewers conclude that due to the limited number and quality of studies comparing one comprehensive treatment approach to another, it is currently not possible to determine if one approach is more or less effective than another. Most reviewers agree that future studies are needed that compare one widely recognized comprehensive treatment program to another widely recognized program.

While there is general agreement in the conclusions reached, ECRI Institute's review differed from previous reviews in terms of the scope, study selection criteria, and analytical methods employed. Unlike most previous reviews, the key questions addressed in ECRI Institute's review were comparative. Our goal was to compare the efficacy of comprehensive treatment programs for children with ASD to no treatment/wait list control, standard or general care, or to another comprehensive treatment program. As such, we excluded studies that used single group designs (e.g., case series) because these types of studies do not include a comparison group. The scope of most of the other reviews we assessed was more general and descriptive in nature, making study design less important. Further, with the exception of one review, the conclusions drawn in other previous reviews were based on a narrative summarization of the findings of the included studies, with no attempt to quantitatively synthesize the results.

The one review in which the authors performed quantitative analyses focused on the efficacy of early intensive behavioral interventions based specifically on the UCLA Young Autism Project model.(101) In this review, the authors performed a meta-analysis on changes in IQ using data from 12 studies, four of which overlapped with ECRI Institute's current review (see Table 53 for a list of these studies). Other studies did not overlap with ECRI Institute's review primarily because they were either single group studies (k = 3), retrospective comparisons (k = 3), included children with a diagnosis other than ASD (k = 1), or included study groups that were not comparable at baseline (k = 1). The authors of this review did not perform any comparative analyses. Instead, they only considered data from children who received EIBI. The results of their analysis "suggest that EIBI is, on average, an effective intervention for increasing IQ scores for children with autism." However, the results of this review should be interpreted with caution because without controls any gains observed in the children cannot be solely contributed to the effects of treatment. Other factors, such as maturation and practice, may have influenced the children's performance.

Ongoing Clinical Trials

To locate recently conducted and ongoing clinical trials of comprehensive intervention programs for children with ASDs, we searched two databases: http://clinicaltrials.gov and http://www.controlled-trials.com. In addition to these two databases, we also searched the grey literature for possible ongoing studies. Our searches identified two studies that are currently enrolling participants. Important information about these trials is presented below in Table 15.

Table 15. Ongoing Clinical Trials

Clinicaltrials.gov identifier or other identifier	Sponsor	Design	Purpose	Start date	Expected completion date	Estimated Enrollment
NCT00698997	National Institutes of Health (NIH) in collaboration with University of California, Davis; University of Washington; and University of Michigan	Randomized, single blind with parallel assignment	The purpose of the project is to answer the following questions: 1) Does the Early Start Denver Model experimental intervention for toddlers with autism reduce disability associated with autism significantly more than standard community interventions?; and 2) What environmental, child, and biological characteristics mediate and moderate intervention response and outcomes at age 4.	April 2008	December 2012	108 children

Clinicaltrials.gov identifier or other identifier	Sponsor	Design	Purpose	Start date	Expected completion date	Estimated Enrollment
R324B070219 (for more information go to http://www.fpg.unc.edu/~asdtc/overview.cfm)	Institute of Education in collaboration with University of North Carolina, University of Colorado, University of Miami, and University of Minnesota	Non-randomized control trial	This 4-year, multi-site project will compare outcomes of preschool aged children with autism who receive TEACCH, LEAP, or control classrooms The geographic sites involved in this project are the states of North Carolina (primary site), Colorado, Florida, and Minnesota.	June 2008	June 2011	75 preschool classrooms

Overall Conclusions and Discussion

This review addressed five key questions pertaining to the efficacy and safety of educational and behavioral comprehensive interventions for children with ASD. In the first question, we considered evidence from clinical studies that compared comprehensive interventions to no treatment, a wait list control, or to what we considered standard care (less intensive care provided in an educational/clinical setting). In the second question, we considered evidence from studies that compared one comprehensive intervention to another. The third question was intended to compare the treatment setting of comprehensive interventions (e.g., clinical versus home or other setting), and the fourth question, for which we found no information, was intended to assess the possible harms of comprehensive treatment. Finally, the fifth question involved reviewing and summarizing the recommendations of recent clinical practice guidelines and consensus statements regarding comprehensive interventions for children with ASD. Below, we briefly discuss the main findings of our review and analyses of data (when possible) from the clinical studies that met the study selection criteria for this review.

For Key Question 1, data from three non-randomized studies that compared EIBI for young children with autistic disorder or PDD-NOS to standard care, were used to perform eight separate meta-analyses on the following outcomes: intellectual/cognitive functioning (as measured by tests of IQ), number of children reaching within normal levels of IQ (85 or higher), adaptive behavior, language and communication. After one year of treatment, children who received EIBI demonstrated significantly higher performance on tests of IQ and adaptive behavior compared to children who received standard care. Children who received EIBI also were more likely to reach IQ scores within normal range compared to children who received standard care. However, the evidence was insufficient (only two small studies) to determine whether these differences continued at later follow-up times. Similarly, for language and communication, the evidence was considered insufficient to permit a conclusion due to the small size of the studies that contributed data for these outcomes. For all other outcomes considered (problem behaviors, academic/developmental achievement, and parental/family well-being), the limited size (one study per outcome) and quality of the evidence base precluded us from drawing any conclusions.

The meta-analytic results for both measures of intellectual functioning and adaptive behavior should be interpreted with caution due to the small size of the evidence base, the moderate quality of the studies, and the variability in performance of children on these outcomes. The evidence base for each outcome consisted of only three studies enrolling a total of 128 children that compared EIBI to standard care. The quality of the studies was limited primarily because children in all three studies were not randomly allocated to the study groups. In all the studies children were assigned to one or the other treatment group based on parental preference. While the authors of all three studies tried to enhance group comparability by either matching children on key variables and/or statistically controlling for any differences between the groups observed at baseline, these methods do not completely eliminate the potential for selection bias.

Further, when considering the results of the meta-analyses for Key Question 1, it is important to keep in mind that they are based on the overall average performance of the children in each group. This means that not all children may have benefitted equally. In an exploratory analysis,

the authors of one of the studies that addressed Key Question 1 investigated variables likely to be associated with changes in IQ.(70) In this study, by Remington et al., the authors found that the children who benefitted most from EIBI differed from the children who did not benefit as much along the following baseline characteristics: higher IQ scores, higher mental age, and higher scores on the Vineland Adaptive Behavior Scales, and less reported problem behaviors. Future studies on EIBI should focus on which children benefit the most from this intervention.

For Key Question 2, the evidence from three non-randomized controlled trials each comparing intensive ABA to a comprehensive "eclectic" intervention program was considered. However, the evidence from these studies did not permit us to draw any conclusions about whether one comprehensive intervention was more effective than another for children with ASD. This was because 1) not all of the studies reported data on the same outcomes of interest; 2) for some outcomes, only one study of moderate quality reported data on the outcome; and 3) in one of the three studies, the children were substantially older (by three years) than the children in the other two studies. This age difference, which we considered an important source of clinical heterogeneity, precluded combination of data from this study with data from the other two studies.

For Key Question 3, our searches of the literature did not identify any studies that met the study selection criteria that directly compared one treatment setting to another. However, we did identify two randomized controlled trials that compared clinic-directed early intensive behavior intervention to parent-directed EIBI. Differences between these studies, however, in terms of the level of parent training and intensity and nature of the treatment delivered to children in the parent-directed group, precluded the formation of any conclusions as to whether clinic-directed EIBI is more effective than parent-directed EIBI for children with ASD.

Bibliography

1. Filipek PA, Accardo PJ, Baranek GT, Cook EH Jr, Dawson G, Gordon B, Gravel JS, Johnson CP, Kallen RJ, Levy SE, Minshew NJ, Ozonoff S, Prizant BM, Rapin I, Rogers SJ, Stone WL, Teplin S, Tuchman RF, Volkmar FR. The screening and diagnosis of autistic spectrum disorders. J Autism Dev Disord 1999 Dec;29(6):439-84.

2. American Psychiatric Association (APA), Task Force on DSM-IV. First MB, editor. Diagnostic and statistical manual of mental disorders: DSM-IV. 4th ed. Washington (DC): American Psychiatric Press, Inc; 1994 Jan. 886 p.

3. des Portes V, Hagerman RJ, Hendren RL. Pharmacotherapy. In: Ozonoff S, Rogers SJ, Hendren RL, editors. Autism spectrum disorders: a research review for practitioners. Washington (DC): American Psychiatric Association Publishing; 2003. p. 161-86.

4. National Research Council. Educating children with autism. Washington (DC): National Academies Press; 2001. 324 p.

5. Cohen DJ, Volkmar FR. Handbook of autism and pervasive developmental disorders. 2nd ed. New York (NY): John Wiley; 1997. 1092 p.

6. Volkmar FR, Nelson DS. Seizure disorders in autism. J Am Acad Child Adolesc Psychiatry 1990 Jan;29(1):127-9.

7. Tuchman RF, Rapin I, Shinnar S. Autistic and dysphasic children. II: Epilepsy. Pediatrics 1991 Dec;88(6):1219-25.

8. Lord C, Rutter M, Le Couteur A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. J Autism Dev Disord 1994 Oct;24(5):659-85.

9. Lord C, Risi S, Lambrecht L, Cook EH Jr, Leventhal BL, DiLavore PC, Pickles A, Rutter M. The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. J Autism Dev Disord 2000 Jun;30(3):205-23.

10. Schopler E, Reichler RJ, Renner BR. The childhood autism rating scale (CARS) for diagnostic screening and classification of autism. New York (NY): Irvington Publishers, Inc; 1986. 63 p.

11. Bryson SE, Rogers SJ, Fombonne E. Autism spectrum disorders: early detection, intervention, education, and psychopharmacological management. Can J Psychiatry 2003 Sep;48(8):506-16.

12. Rogers SJ. Diagnosis of autism before the age of 3. Int Rev Res Ment Retard 2000;23:1-31.

13. Nordin V, Gillberg C. The long-term course of autistic disorders: update on follow-up studies. Acta Psychiatr Scand 1998 Feb;97(2):99-108.

14. Rumsey JM, Rapoport JL, Sceery WR. Autistic children as adults: psychiatric, social, and behavioral outcomes. J Am Acad Child Psychiatry 1985 Jul;24(4):465-73.

15. Gillberg C, Steffenberg S. Outcome and prognostic factors in infantile autism and similar conditions: a population-based study of 46 cases followed through puberty. J Autism Dev Disord 1987;17(2):273-87.

16. Rutter M. Autistic children: infancy to adulthood. Semin Psychiatry 1970 Nov;2(4):435-50.

17. Venter A, Lord C, Schopler E. A follow-up study of high-functioning autistic children. J Child Psychol Psychiatry Allied Discip 1992;33(3):489-507.

18. Autism and Developmental Disabilities Monitoring Network Surveillance Year 2002, Centers for Disease Control and Prevention. Prevalence of autism spectrum disorders--autism and developmental disabilities monitoring network, 14 sites, United States, 2002. MMWR Surveill Summ 2007 Feb 9;56(1):12-28.

19. Fombonne E. Epidemiology of autistic disorder and other pervasive developmental disorders. J Clin Psychiatry 2005;66 Suppl 1:3-8.

20. Technical report: the pediatrician's role in the diagnosis and management of autistic spectrum disorder in children. Pediatrics 2001 May;107(5):E85.

21. Chess S. Autism in children with congenital rubella. J Autism Child Schizophr 1971 Jan-Mar;1(1):33-47.

22. Hunt A, Dennis J. Psychiatric disorder among children with tuberous sclerosis. Dev Med Child Neurol 1987 Apr;29(2):190-8.

23. Corsello CM. Early intervention in autism. Infants Young Child 2005 Apr-Jun;18(2):74-85.

24. Lovaas OI, Buch G. Intensive behavioral intervention with young children with autism. In: Singh NN, editors. Prevention and treatment of severe behavior problems: models and methods in developmental disabilities. Pacific Grove (CA): Brooks/Cole Publishing Co; 1997. p. 61-86.

25. Wieder S, Greenspan SI. Climbing the symbolic ladder in the DIR model through floor time/interactive play. Autism 2003 Dec;7(4):425-35.

26. Prizant BM, Wetherby AM, Rubin E, Laurent AC. The SCERTS Model: a transactional, family-centered approach to enhancing communication and socioemotional abilities of children with autism and spectrum disorder. Infants Young Child 2003 Oct-Dec;16(4):296-316.

27. Lovaas OI, Smith T. Intensive and long-term treatments for clients with destructive behaviors. In: Thompson T, Gray DB, editors. Destructive behavior in developmental disabilities: diagnosis and treatment. Thousand Oaks (CA): Sage Publications, Inc; 1994. p. 243-60. (Sage focus editions; vol. 170).

28. Lovaas OI, Ackerman AB, Alexander D, Firestone P, Perkins J, Young D. Teaching developmentally disabled children: the me book. Baltimore: University Park Press; 1981. 250 p.

29. Lovaas I, Calouri K, Jada J. The nature of behavioral treatment and research with young autistic persons. In: Gillberg C, editors. Diagnosis and treatment of autism. New York (NY): Plenum Press; 1989. p. 285-305.

30. ECRI. Comprehensive programs for the treatment of children with autism [technology assessment report]. Plymouth Meeting (PA): ECRI; 1999. 128 p.

31. Turnbull HR 3rd, Wilcox BL, Stowe MJ. A brief overview of special education law with focus on autism. J Autism Dev Disord 2002 Oct;32(5):479-93.

32. Schwartz HS, Sandall SR, McBride BJ, Boulware GL. Project DATA (developmentally appropriate treatment for autism): an inclusive school-based approach to educating young children with autism. Top Early Child Spec Educ 2004;24(3):156-68.

33. Strain PS, Kohler FW, Goldstein H. Learning experiences . . . an alternative program: peer-mediated interventions for young children with autism. In: Hibbs ED, Jensen PS, editors. Psychosocial treatments for child and adolescent disorders: empirically based strategies for clinical practice. Washington (DC): American Psychological Association; 1996. p. 573-87.

34. Hoyson M, Jamieson B, Strain PS. Individualized group instruction of normally developing and autistic-like children. The LEAP curriculum model. J Div Early Child 1984;157-72.

35. LeMay D, Griffin P, Sanford A. Learning accomplishment profile - diagnostic edition. Chapel Hill (NC): Chapel Hill Training-Outreach Project; 1977.

36. Rogers SJ, Hall T, Reaven J. The Denver model: an integrated approach to intervention for young children with autism. A guideline for parents. Denver (CO): JFK Center for Developmental Disabilities; 11 p.

37. Shields J. The NAS EarlyBird Programme: partnership with parents in early intervention. The National Autistic Society. Autism 2001 Mar;5(1):49-56.

38. Salt J, Sellars V, Shemilt J, Boyd S, Coulson T, McCool S. The Scottish Centre for Autism preschool treatment programme. I: A developmental approach to early intervention. Autism 2001 Dec;5(4):362-73.

39. Schopler E, Mesibov GB, Shigley RH, Bashford A. Chapter 4, Helping autistic children through their parents: the TEACCH mode. In: Schopler E, Mesibov G, editors. The effects of autism on the family. New York (NY): Plenum Press; 1984. p. 65-81. (Current issues in autism).

40. Aman MG, Langworthy KS. Pharmacotherapy for hyperactivity in children with autism and other pervasive developmental disorders. J Autism Dev Disord 2000 Oct;30(5):451-9.

41. Schechtman MA. Scientifically unsupported therapies in the treatment of young children with autism spectrum disorders. Pediatr Ann 2007 Aug;36(8):497-8, 500-5.

42. Williams KW, Wray JJ, Wheeler DM. Intravenous secretin for autism spectrum disorder. In: Cochrane Database of Systematic Reviews [database online]. Issue 3. Hoboken (NJ): John Wiley & Sons, Ltd.; 2005 Jul 20 [updated 2005 May 24]. [accessed 2008 Nov 21]. [Art. No.: CD003495]. Available: DOI: 10.1002/14651858.CD003495.pub2.

43. Tharp BR. Contributions of neurology. In: Ozonoff S, Rogers SJ, Hendren RL, editors. Autism spectrum disorders: a research review for practitioners. Washington (DC): American Psychiatric Association Publishing; 2003. p. 111-32.

44. Special education: children with autism. Report to the Chairman and Ranking Minority Member, Subcommittee on Human Rights and Wellness, Committee on Government Reform, House of Representatives. Washington (DC): United States Government Accountability Office (GAO); 2005 Jan. 39 p. (GAO-05-220).

45. Howlin P. The effectiveness of interventions for children with autism. J Neural Transm Suppl 2005;69:101-19.

46. Association for Science in Autism Treatment. [Web site]. Crosswicks (NJ): Association for Science in Autism Treatment (ASAT); [accessed 2008 Sep 9]. [various p]. Available: http://www.asatonline.org/.

47. Sofronoff K, Attwood T, Hinton S. A randomised controlled trial of a CBT intervention for anxiety in children with Asperger syndrome. J Child Psychol Psychiatry Allied Discip 2005 Nov;46(11):1152-60.

48. Sofronoff K, Attwood T, Hinton S, Levin I. A randomized controlled trial of a cognitive behavioural intervention for anger management in children diagnosed with Asperger syndrome. J Autism Dev Disord 2007 Aug;37(7):1203-14.

49. Weiskop S, Richdale A, Matthews J. Behavioural treatment to reduce sleep problems in children with autism or fragile X syndrome. Dev Med Child Neurol 2005 Feb;47(2):94-104.

50. Liptak GS, Stuart T, Auinger P. Health care utilization and expenditures for children with autism: data from U.S. national samples. J Autism Dev Disord 2006 Oct;36(7):871-9.

51. Chasson GS, Harris GE, Neely WJ. Cost comparison of early intensive behavioral intervention and special education for children with autism. J Child Fam Stud 2007 Jun;16(3):401-13.

52. Mandell DS, Cao J, Ittenbach R, Pinto-Martin J. Medicaid expenditures for children with autistic spectrum disorders: 1994 to 1999. J Autism Dev Disord 2006 May;36(4):475-85.

53. Shimabukuro TT, Grosse SD, Rice C. Medical expenditures for children with an autism spectrum disorder in a privately insured population. J Autism Dev Disord 2008 Mar;38(3):546-52.

54. Leslie DL, Martin A. Health care expenditures associated with autism spectrum disorders. Arch Pediatr Adolesc Med 2007 Apr;161(4):350-5.

55. Nevada Autism Task Force. 2008 report of the Nevada Autism Task Force: executive summary. Carson City (NV): Nevada Department of Health and Human Services; 2008 Jul 31. Also available: http://dhhs.nv.gov/Autism.htm.

56. Autism speaks state initiative map. [internet]. Washington (DC): Autism Votes; 2008 [updated 2008 Oct 22]; [accessed 2008 Nov 3]. Available: http://www.autismvotes.org/site/c.frKNI3PCImE/b.3909861/k.B9DF/State_Initiatives.htm.

57. State Autism Profiles. [internet]. Chicago (IL): Easter Seals; 2008 [accessed 2008 Sep 12]. [various]. Available: http://www.easterseals.com/site/PageServer?pagename=ntlc8_autism_state_profiles.

58. Wolery M, Garfinkle AN. Measures in intervention research with young children who have autism. J Autism Dev Disord 2002 Oct;32(5):463-78.

59. Matson JL. Determining treatment outcome in early intervention programs for autism spectrum disorders: a critical analysis of measurement issues in learning based interventions. Res Dev Disabil 2007 Mar-Apr;28(2):207-18.

60. Moher D, Pham B, Klassen TP, Schulz KF, Berlin JA, Jadad AR, Liberati A. What contributions do languages other than English make on the results of meta-analyses? J Clin Epidemiol 2000 Sep;53(9):964-72.

61. Juni P, Holenstein F, Sterne J, Bartlett C, Egger M. Direction and impact of language bias in meta-analyses of controlled trials: empirical study. Int J Epidemiol 2002 Feb;31(1):115-23.

62. Narine L, Yee DS, Einarson TR, Ilersich AL. Quality of abstracts of original research articles in CMAJ in 1989. CMAJ 1991 Feb 15;144(4):449-53.

63. Pitkin RM, Branagan MA, Burmeister LF. Accuracy of data in abstracts of published research articles. JAMA 1999 Mar 24-31;281(12):1110-1.

64. Taddio A, Pain T, Fassos FF, Boon H, Ilersich AL, Einarson TR. Quality of nonstructured and structured abstracts of original research articles in the British Medical Journal, the Canadian Medical Association Journal and the Journal of the American Medical Association. CMAJ 1994 May 15;150(10):1611-5.

65. Howard JS, Sparkman CR, Cohen HG, Green G, Stanislaw H. A comparison of intensive behavior analytic and eclectic treatments for young children with autism. Res Dev Disabil 2005 Jul-Aug;26(4):359-83.

66. Remington B, Hastings RP, Kovshoff H, degli Espinosa F, Jahr E, Brown T, Alsford P, Lemaic M, Ward N. Early intensive behavioral intervention: outcomes for children with autism and their parents after two years. Am J Ment Retard 2007 Nov;112(6):418-38.

67. Eikeseth S, Smith T, Jahr E, Eldevik S. Outcome for children with autism who began intensive behavioral treatment between ages 4 and 7: a comparison controlled study. Behav Modif 2007 May;31(3):264-78.

68. Eikeseth S, Smith T, Jahr E, Eldevik S. Intensive behavioral treatment at school for 4- to 7-year-old children with autism. A 1-year comparison controlled study. Behav Modif 2002 Jan;26(1):49-68.

69. Zachor DA, Ben-Itzchak E, Rabinovich AL, Lahat E. Change in autism core symptoms with intervention. Res Autism Spect Disord 2007 Oct;1(4):304-17.

70. Cohen H, Amerine-Dickens M, Smith T. Early intensive behavioral treatment: replication of the UCLA model in a community setting. J Dev Behav Pediatr 2006 Apr;27(2 Suppl):S145-55.

71. Sallows GO, Graupner TD. Intensive behavioral treatment for children with autism: four-year outcome and predictors. Am J Ment Retard 2005 Nov;110(6):417-38.

72. Smith T, Groen AD, Wynn JW. Randomized trial of intensive early intervention for children with pervasive developmental disorder.[erratum appears in Am J Ment Retard 2000 Nov;105(6):508]. Am J Ment Retard 2000 Jul;105(4):269-85.

73. Treadwell JT, Tregear SJ, Reston JT, Turkelson CM. A system for rating the stability and strength of medical evidence. BMC Med Res Methodol 2006 Oct 19;6:52. Also available: http://www.biomedcentral.com/1471-2288/6/52.

74. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002 Jun 15;21(11):1539-58.

75. Higgins JP, Thompson SG. Controlling the risk of spurious findings from meta-regression. Stat Med 2004 Jun 15;23(11):1663-82.

76. Follmann D, Elliott P, Suh I, Cutler J. Variance imputation for overviews of clinical trials with continuous response. J Clin Epidemiol 1992 Jul;45(7):769-73.

77. Bain BA, Dollaghan CA. Clinical forum: treatment efficacy. The notion of clinically significant change. Lang Speech Hear Serv Sch 1991 Oct;22(4):264-70.

78. California Departments of Education and Developmental Services, Collaborative Work Group on Autistic Spectrum Disorders. Best practices for designing and delivering effective programs for individuals with autistic spectrum disorders: recommendations of the Collaborative Work Group on Autistic Spectrum Disorders. Sacramento (CA): California Department of Education; 1997 Jul. Autism resources: assessment procedures. p. appendix 1-22. Also available: http://www.behavior-consultant.com/aut-dx-devices.htm.

79. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale (NJ): Lawrence Erlbaum Associates; 1988. 567 p.

80. Sparrow SS, Cicchetti DV. Diagnostic uses of the Vineland Adaptive Behavior Scales. J Pediatr Psychol 1985 Jun;10(2):215-25.

81. New York State Department of Health. Autism/pervasive developmental disorders: assessment and intervention for young children (Age 0-3 year). Albany (NY): New York State Department of Health, Early Intervention Program; 1999 Nov. various p. Also available: http://www.health.state.ny.us/community/infants_children/early_intervention/autism/index.htm.

82. Scottish Intercollegiate Guidelines Network (SIGN). Assessment, diagnosis and clinical interventions for children and young people with autism spectrum disorders. A national clinical guideline. Edinburgh (Scotland): Scottish Intercollegiate Guidelines Network (SIGN); 2007 Jul. 65 p. (SIGN publication; no. 98).

83. Burrows K. Early intervention for children with autism. Paediatr Child Health 2004 Apr;9(4):267-70.

84. Canadian Paediatric Society (CPS). Early intervention for children with autism. Reaffirmed February 2008. Paediatr Child Health 2004;9(4):267-70.

85. Ludwig S, Harstall C. Intensive intervention programs for children with autism. Edmonton (AL): Alberta Heritage Foundation for Medical Research; 2001 Feb. 45 p. (Health technology assessment; no. 8-B).

86. McGahan L. Behavioural interventions for preschool children with autism. Ottawa: Canadian Coordinating Office for Health Technology Assessment (CCOHTA); 2001. 88 p. (Technology report; no. 18).

87. Basset K, Green CJ, Kazanjian A. Autism and Lovaas treatment: a systematic review of effectiveness evidence. Vancouver (BC): BC Office of Health Technology Assessment, Centre for Health Services and Policy Research, University of British Columbia; 2000 Jul. 68 p.

88. Green C, Basset K, Kazanjian A. Critical appraisal of submitted cost-benefit models of 'Lovaas' early intensive behavioural intervention for autism. Vancouver (BC): BC Office of Health Technology Assessment, Centre for Health Services and Policy Research, University of British Columbia; 2000 Feb. 22 p.

89. Ministries of Health and Education. New Zealand autism spectrum disorder guideline. Wellington: New Zealand Ministry of Health; 2008 Mar 1. 312 p. Also available: http://www.moh.govt.nz/moh.nsf/pagesmh/7561/$File/asd-guideline-apr08.pdf.

90. Prior M, Roberts J. Early intervention for children with autism spectrum disorders: guidelines for best practice. Woden, ACT: Australian Government Department of Health and Ageing; 2006. 12 p.

91. Elements of effective programs. Chapel Hill (NC): National Early Childhood Technical Assistance Center (NECTAC); 2007. Also available: http://www.nectac.org/topics/autism/effecprog.asp?text=1.

92. McEachin JJ, Smith T, Lovaas OI. Long-term outcome for children with autism who received early intensive behavioral treatment. Am J Ment Retard 1993 Jan;97(4):359-72; discussion 373-91.

93. Lovaas OI. Behavioral treatment and normal educational and intellectual functioning in young autistic children. J Consult Clin Psychol 1987 Feb;55(1):3-9.

94. Sheinkopf SJ, Siegel B. Home-based behavioral treatment of young children with autism. J Autism Dev Disord 1998 Feb;28(1):15-23.

95. Ozonoff S, Cathcart K. Effectiveness of a home program intervention for young children with autism. J Autism Dev Disord 1998 Feb;28(1):25-32.

96. Jocelyn LJ, Casiro OG, Beattie D, Bow J, Kneisz J. Treatment of children with autism: a randomized controlled trial to evaluate a caregiver-based intervention program in community day-care centers. J Dev Behav Pediatr 1998 Oct;19(5):326-34.

97. Case-Smith J, Arbesman M. Evidence-based review of interventions for autism used in or of relevance to occupational therapy. Am J Occup Ther 2008 Jul-Aug;62(4):416-29.

98. Rogers SJ, Vismara LA. Evidence-based comprehensive treatments for early autism. J Clin Child Adolesc Psychol 2008 Jan-Mar;37(1):8-38.

99. Levy S, Kim AH, Olive ML. Interventions for young children with autism: a synthesis of the literature. Focus Autism Dev Disabil 2006 Spring;21(1):55-62.

100. Chorpita BF, Yim LM, Donkervoet JC, et al. Toward large-scale implementation of empirically supported treatments for children: a review and observations by the Hawaii empirical basis to services task force. Clin Psychol Sci Pract 2002;9:165-90.

101. Reichow B, Wolery M. Comprehensive dynthesis of early intensive behavioral interventions for young children with autism based on the UCLA young autism project model. J Autism Dev Disord 2008 Jun 6;Epub ahead of print.

102. Magiati I, Charman T, Howlin P. A two-year prospective follow-up study of community-based early intensive behavioural intervention and specialist nursery provision for children with autism spectrum disorders. J Child Psychol Psychiatry Allied Discip 2007 Aug;48(8):803-12.

103. Rickards AL, Walstab JE, Wright-Rossi RA, Simpson J, Reddihough DS. A randomized, controlled trial of a home-based intervention program for children with autism and developmental delay. J Dev Behav Pediatr 2007 Aug;28(4):308-16.

104. Eldevik S, Eikeseth S, Jahr E, Smith T. Effects of low-intensity behavioral treatment for children with autism and mental retardation. J Autism Dev Disord 2006 Feb;36(2):211-24.

105. Reed P, Osborne LA, Corness M. Brief report: relative effectiveness of different home-based behavioral approaches to early teaching intervention. J Autism Dev Disord 2007 Oct;37(9):1815-21.

106. Salt J, Shemilt J, Sellars V, Boyd S, Coulson T, McCool S. The Scottish Centre for Autism preschool treatment programme. II: the results of a controlled treatment outcome study. Autism 2002 Mar;6(1):33-46.

107. Panerai S, Ferrante L, Zingale M. Benefits of the treatment and education of autistic and communication handicapped children (TEACCH) programme as compared with a non-specific approach. J Intellect Disabil Res 2002 May;46(Pt 4):318-27.

108. Smith T, Eikeseth S, Klevstrand M, Lovaas OI. Intensive behavioral treatment for preschoolers with severe mental retardation and pervasive developmental disorder. Am J Ment Retard 1997 Nov;102(3):238-49.

109. Birnbrauer JS, Leach DJ. The Murdoch early intervention program after 2 years. Behav Change 1993;10:63-74.

110. Bridges LJ, Berry D, Calkins J, Margie NG, Cochran SW, Ling TJ, Zazlow MJ. Early childhood education and school readiness: Conceptual models, constructs, and measures: Profiles of early childhood measures. Washington, DC: Child Trends; 2003. 296 p. Also available: http://www.childcareresearch.org/SendPdf?resourceId=3054.

111. Richmond DM. A review and evaluation of the Reynell Developmental Language Scales. Slow Learn Child 1973 Jul;20(2):102-16.

112. Developmental behaviour checklist (DBC). [internet]. Los Angeles (CA): Western Psychological Services (WPS); 2007 Feb 12 [accessed 2008 Sep 8]. [12 p]. Available: http://www.med.monash.edu.au/spppm/research/devpsych/dbc.html.

113. Use the new WIAT--II--abbreviated (WIAT--II--A) to efficiently identify and track basic academic skills and intervention needs in children and adults. [internet]. San Antonio (TX): Pearson Education, Inc.; [accessed 2008 Sep 30]. [1 p]. Available: http://pearsonassess.com/hai/ProductLongDesc.aspx?Catalog=TPC-USCatalog&ISBN=015-8983-858&Category=AchievementAccountability.

114. Martin CR, Lewin RJ, Thompson DR. A confirmatory factor analysis of the Hospital Anxiety and Depression Scale in coronary care patients following acute myocardial infarction. Psychiatry Res 2003 Aug 30;120(1):85-94.

115. Kansas inventory of parental perceptions. [internet]. Lawrence (KS): Beach Center on Disability, University of Kansas; [accessed 2008 Sep 29]. [5 p].
Available: http://www.beachcenter.org/resource_library/beach_resource_detail_page.aspx?intResourceID=2483.

116. Armitage P, Berry G. Statistical methods in medical research. 3rd ed. Oxford, England: Blackwell Scientific; 1994. Significance tests and confidence intervals. p. 94-9.

117. Braitman LE. Confidence intervals assess both clinical significance and statistical significance [editorial]. Ann Intern Med 1991 Mar 15;114(6):515-7.

118. West S, King V, Carey TS, Lohr KN, McKoy N, Sutton SF, Lux L. Systems to rate the strength of scientific evidence. (Prepared by Research Triangle Institute - University of North Carolina Evidence-based Practice Center under Contract no. 290-97-0011). AHRQ Publication no. 02-E016. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); 2002 Apr. 199 p. (Evidence report/technology assessment; no. 47). Also available: http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=hstat1.chapter.70996.

119. Harbord R, Statistical Software Components, Boston College Department of Economics. METAREG: Stata module to perform meta-analysis regression. [software download]. Stockholm, Sweden: EconPapers; 2004 Sep 22 [accessed 2004 Nov 3]. Available: http://econpapers.hhs.se/software/bocbocode/s446201.htm.

120. Sanchez-Meca J, Marin-Martinez F, Chacon-Moscoso S. Effect-size indices for dichotomized outcomes in meta-analysis. Psychol Methods 2003 Dec;8(4):448-67.

121. Resources in Special Education (RiSE). Best practices for designing and delivering effective programs for individuals with autistic spectrum disorders. Recommendations of the Collaborative Work Group on Autistic Spectrum Disorders. Sacramento (CA): California Departments of Education and Developmental Services; 1997 Jul. 92 p. Also available: http://www.isciii.es/htdocs/centros/enfermedadesraras/pdf/aut_gcalif.pdf.

122. Autism Task Force. Autism spectrum disorder. Intervention guidance for service providers and families of young children with autism spectrum disorders. Hartford (CT): Connecticut Birth to Three System; 2008 Jan. 57 p. (Service guideline; no. 1). Also available: http://www.birth23.org/Publications/Autism%202008.pdf.

123. Pratt C, Vicker B, Davis K. Early intervention for young children with autism spectrum disorders. Recommendations for designing effective programs. Bloomington (IN): Indiana Institute on Disability and Community at Indiana University, Bloomington; 2001. 40 p. Also available: https://www.iidc.indiana.edu/.

124. Report of the MADSEC Autism Task Force. Manchester (ME): Maine Administrators of Services for Children with Disabilities (MADSEC); 2000 Feb. 67 p. Also available: http://www.madsec.org/docs/atf.htm.

125. Evans S, DiFerdinando GT Jr, Wood CA, Harrison TL, Dunning MS. Service guidelines for children with autism spectrum disorders. Trenton (NJ): New Jersey Department of Health and Senior Services; 2004 Feb. 44 p. Also available: http://www.state.nj.us/health/fhs/documents/autismguidelines.pdf.

126. NM Family Infant Toddler Program. Autism spectrum disorders. Santa Fe (NM): New Mexico Family Infant Toddler Program; 2004 Jan. 60 p. Also available: http://www.health.state.nm.us/ddsd/fit/pdf%5CAutism-Spectrum-Disorders.pdf.

127. State University of New York (SUNY) at Albany. The availability and effectiveness of programs for preschool children with autism. A report to the Board of Regents, the Governor and the New York State legislature. Albany (NY): New York State Education Department, Office of Vocational and Educational Services for Individuals with Disabilities Special Education Policy, Planning and Partnerships; 2004 Mar. various p. Also available: http://www.vesid.nysed.gov/specialed/autism/preschoolstudy.htm#west.

128. Clinical practice guideline: report of the guideline recommendations. Autism/pervasive developmental disorders: assessment and intervention for young children (Age 0-3 Years). Albany (NY): New York State Department of Health, Early Intervention Program; 1999. 322 p. Also available: http://www.health.state.ny.us/community/infants_children/early_intervention/autism/index.htm.

129. Office of Special Education. Guidelines: identifying, serving and educating children and youth with autism. Bismarck (ND): North Dakota Department of Public Instruction; 2003 Mar 17. 67 p. Also available: http://www.dpi.state.nd.us/speced/guide/autism.pdf.

130. Service guidelines for individuals with autism spectrum disorder/pervasive developmental disorder (ASD/PDD): birth through twenty-one. Columbus (OH): Ohio Developmental Disabilities Council; 103 p. Also available: http://ddc.ohio.gov/Pub/Child.htm.

131. Morris SL. Successes in serving families and infants and toddlers with autism. Olympia (WA): Washington State Department of Health and Human Services, Infant Toddler Early Intervention Program; 6 p. Also available: http://www1.dshs.wa.gov/word/adsa/iteip/SLM_Autism.doc.

132. Medicaid waivers manual. Madison (WI): Wisconsin Department of Health Services; 2007 Dec. Allowable services. p. IV1-132. Also available: http://dhs.wisconsin.gov/bdds/waivermanual/waiverch04_08.pdf#page=85.

133. Aetna. Clinical policy bulletin 0648: pervasive developmental disorders. [internet]. Hartford (CT): Aetna, Inc.; 2008 May 9 [accessed 2008 Sep 24]. [28 p]. Available: http://www.aetna.com/cpb/medical/data/600_699/0648.html.

134. Cognitive skills development and sensory integrative techniques. [internet]. Chicago (IL): American Medical Association (AMA); 2001 Dec [accessed 2008 Aug 27]. [3 p]. Available: http://coverageandpayment.mediregs.com.

135. <<97532>>Cognitive skills development - physical therapy service, for which payment may not be made. [internet]. Chicago (IL): American Medical Association (AMA); 2008 [accessed 2008 Aug 27]. [3 p]. Available: http://coverageandpayment.mediregs.com.

136. <<97533>>Sensory integration - physical therapy service, for which payment may not be made. [internet]. Chicago (IL): American Medical Association (AMA); 2008 [accessed 2008 Aug 27]. [3 p]. Available: http://coverageandpayment.mediregs.com.

137. Group evidence of coverage, point of service health benefit plan. Athens (GA): Athens Area Health Plan Select, Inc.; 2005 Nov. 52 p.

138. Medica. Chelation therapy. [internet]. Minnetonka (MN): Medica; 2005 Dec 1 [accessed 2008 Sep 24]. [3 p]. Available: http://www.medica.com.

139. Blue Cross and Blue Shield of Massachusetts. Early intervention; special needs; orthoptics training; facilitated communication. Policy #: 281. [internet]. Rockland (MA): Blue Cross and Blue Shield of Massachusetts; 2008 Jun 24 [accessed 2008 Sep 24]. [9 p]. Available: http://www.bcbsma.com.

140. Blue Cross and Blue Shield of Massachusetts. Cognitive rehabilitation; coma stimulation; sensory integration therapy. Policy #: 439. [internet]. Rockland (MA): Blue Cross and Blue Shield of Massachusetts; 2008 Jun 25 [accessed 2008 Sep 24]. [12 p]. Available: http://www.bcbsma.com.

141. Blue Cross and Blue Shield of Tennessee. Speech/language therapy, occupational therapy and physical therapy for autism. [internet]. Chattanooga (TN): Blue Cross and Blue Shield of Tennessee; 2007 Apr 12 [accessed 2008 Aug 28]. [4 p]. Available: http://www.bcbst.com.

142. Centers for Medicaid and Medicare Services. Autism spectrum disorder/pervasive developmental disorders: assessment and treatment. Coverage position number 0447. [internet]. Bloomfield (CT): CIGNA Corporation; 2007 Apr 15 [accessed 2008 Sep 24]. [26 p]. Available: http://www.cignamedicare.com.

143. HealthPartners. Pervasive development disorder. Number P006. [internet]. Minneapolis (MN): HealthPartners; 2006 Jun 1 [accessed 2008 Sep 24]. [1 p]. Available: http://www.healthpartners.com.

144. Kaiser Permanente. Kaiser Permanente traditional plan: evidence of coverage for Varian, Inc. Oakland (CA): Kaiser Foundation Health Plan, Inc.; 2004 Oct 22.

145. State of Maryland Eagle PPO plan guide. Group certificate schedule Q3. Rockville (MD): MAMSI Life and Health Insurance Company.

146. Medica. Lovaas therapy/intensive early intervention behavior therapy services (IEIBTS)/intensive behavior intervention (IBI). [internet]. Minnetonka (MN): Medica; 2008 Jun 1 [accessed 2008 Sep 24]. [2 p]. Available: http://www.medica.com.

147. Medica. Health Research Institute / Pfeiffer Treatment Center protocols. [internet]. Minnetonka (MN): Medica; 2006 Mar 1 [accessed 2008 Sep 24]. [2 p]. Available: http://www.medica.com.

148. Premera Blue Cross rescinds exclusion policy on SGD/AACs for autism. [internet]. Rockville (MD): American Speech-Language-Hearing Association (ASHA); 2008 Mar 12 [accessed 2008 Aug 28]. [3 p]. Available: http://www.asha.org/about/legislation-advocacy/2008/PremeraBlueCross.htm.

149. Regence Group. Augmentative communication devices and systems. Policy no. 52. [internet]. Portland (OR): Regence Group; 2008 Sep 1 [accessed 2008 Sep 24]. [3 p]. Available: http://www.regence.com.

150. Wellmark Blue Cross and Blue Shield of Iowa/South Dakota. Chelation therapy. Medical policy: 08.01.06. [internet]. Des Moines (IA): Wellmark Blue Cross and Blue Shield of Iowa/South Dakota; 2008 Mar [accessed 2008 Sep 24]. [3 p]. Available: http://www.wellmark.com.

151. Wellmark Blue Cross and Blue Shield of Iowa/South Dakota. Sensory integration therapy. Medical policy: 08.03.04. [internet]. Des Moines (IA): Wellmark Blue Cross and Blue Shield of Iowa/South Dakota; 2007 May [accessed 2008 Sep 24]. [2 p]. Available: http://www.wellmark.com.

152. Doughty C. What is the evidence for the effectiveness of behavioural and skill-based early intervention in young children with autism spectrum disorder (ASD)? Christchurch (NZ): New Zealand Health Technology Assessment (NZHTA); 2004. 47 p. (NZHTA tech brief; vol. 3, no. 1).

153. Finch L, Raffaele C. Developing expert practice. Intensive behavioral intervention for children with autism: a review of the evidence. Occup Ther Now 2003;5:20-3. Also available: http://www.asatonline.org.

154. Diggle T, McConachie HR, Randle VR. Parent-mediated early intervention for young children with autism spectrum disorder (Review). In: Cochrane Database of Systematic Reviews [internet]. Issue 2. Hoboken (NJ): John Wiley & Sons, Ltd.; 2002 [accessed 2008 Jul 17]. [Art. No. CD003496].

155. Ludwig S, Harstall C. Intensive intervention programs for children with autism. Edmonton (AB): Alberta Heritage Foundation for Medical Research (AHFMR); 2001. 40 p.

Appendix A. Literature Search Methods

Electronic Database Searches

The following databases have been searched for relevant information:

Name	Date limits	Platform/provider
CINAHL (Cumulative Index to Nursing and Allied Health Literature)	2000 through July 17, 2008	EBSCOhost
The Cochrane Central Register of Controlled Trials (CENTRAL)	Through 2008, Issue 3	www.thecochranelibrary.com
The Cochrane Database of Methodology Reviews (Methodology Reviews)	Through 2008, Issue 3	www.thecochranelibrary.com
The Cochrane Database of Systematic Reviews (Cochrane Reviews)	Through 2008, Issue 3	www.thecochranelibrary.com
Database of Abstracts of Reviews of Effects (DARE)	Through 2008, Issue 3	www.thecochranelibrary.com
EMBASE (Excerpta Medica)	2000 through October 8, 2008	OVID
ERIC (Education Resources Information Center)	2000 through August 29, 2008	http://eric.ed.gov/
Health Technology Assessment Database (HTA)	Through 2008, Issue 3	www.thecochranelibrary.com
Healthcare Standards	July 30, 2008	ECRI Institute
International Health Technology Assessment (IHTA)	July 30, 2008	ECRI Institute
MEDLINE	2000 through October 8, 2008	OVID
PreMEDLINE	Searched June 2, 2008	OVID
Psychology & Behavioral Sciences Collection	1998 – 2008	EBSCOhost
PsycINFO	2000 through July 31, 2008	Dialog
U.K. National Health Service Economic Evaluation Database (NHS EED)	Through 2008, Issue 3	www.thecochranelibrary.com
U.S. National Guideline Clearinghouse™ (NGC)	July 30, 2008	www.ngc.gov

Hand Searches of Journal and Nonjournal Literature

Journals and supplements maintained in ECRI Institute's collections were routinely reviewed. Nonjournal publications and conference proceedings from professional organizations, private agencies, and government agencies were also screened. Other mechanisms used to retrieve additional relevant information included review of bibliographies/reference lists from peer-reviewed and gray literature. (Gray literature consists of reports, studies, articles, and monographs produced by federal and local government agencies, private organizations, educational facilities, consulting firms, and corporations. These documents do not appear in the peer-reviewed journal literature.)

The search strategies employed combinations of freetext keywords as well as controlled vocabulary terms including (but not limited to) the following concepts. The strategy below is presented in OVID syntax; the search was simultaneously conducted across EMBASE, MEDLINE, and PsycINFO. A parallel strategy was used to search the databases comprising the Cochrane Library.

Medical Subject Headings (MeSH), EMTREE, PsycINFO and Keywords

Conventions:

OVID

$	=	truncation character (wildcard)
exp	=	"explodes" controlled vocabulary term (e.g., expands search to all more specific related terms in the vocabulary's hierarchy)
.de.	=	limit controlled vocabulary heading
.fs.	=	floating subheading
.hw.	=	limit to heading word
.md.	=	type of methodology (PsycINFO)
.mp.	=	combined search fields (default if no fields are specified)
.pt.	=	publication type
.ti.	=	limit to title
.tw.	=	limit to title and abstract fields

Dialog

?	=	truncation character (wildcard)
!	=	"explodes" controlled vocabulary term (e.g., expands search to all more specific related terms in the vocabulary's hierarchy)
/de	=	limit controlled vocabulary heading
pt=	=	publication type
/ti	=	limit to title
/ti,ab	=	limit to title and abstract fields

Topic-specific Search Terms

Concept	Controlled Vocabulary	Keywords
Pervasive developmental disorders	autism Exp autism Autistic disorder Exp child development disorders, pervasive Exp pervasive developmental disorders	Asperger$ Autis$ Disintegrative disorder$ pdd$ rett
Comprehensive treatment programs	Early intervention Early intervention education Exp behavior therapy (and keywords comprehensive or intensive)	ABA APP Applied behavio?r$ analysis Autism Preschool Program Comprehensive intervention$ Comprehensive program$ Comprehensive therapy Comprehensive training Denver model Early intervention EIBT IBT Intensive behavio?r intervention$ Intensive behavio?r program$ Intensive behavio?r therapy Intensive behavio?r training LEAP Lovaas Princeton Child Development Institute River Street Autism Program RSAP Rutgers Autism Program SCERT TEACCH

EMBASE/MEDLINE
English language, human
2000 – 2008
OVID syntax

Set Number	Concept	Search statement
1	Autism (controlled vocabulary)	Exp Child development disorders, pervasive/ or exp pervasive developmental disorders/ or exp autism/ or autism.de. or autistic disorder.de.
2	Autism (text words)	Autis$ or pdd or asperger$.tw. or rett.tw. or disintegrative disorder$
3	Combine sets	1 or 2
4	Comprehensive tx	(comprehensive or Intensive behavio?r$) adj2 (training or therapy or intervention$ or program$)
5		(early intervention or early intervention education).de.
6		Early intervention or ABA or Applied behavio?r$ analysis or EIBT or IBT
7		Autism preschool program or APP or Denver model or LEAP or Lovaas or Princeton Child Development Institute or River Street Autism Program or RSAP or Rutgers Autism Program or SCERT or TEACCH
8		Limit 3 to 3312 (behavior therapy & behavior modification)
9		3 and exp behavior therapy/
10		(8 or 9) and (comprehensive or intensive)
11	Combine sets	4 or 5 or 6 or 7 or 10
12	Combine sets	3 and 11
13	Eliminate overlap	Remove duplicates from 12
14	Limit by publication type	13 not ((letter or editorial or news or comment or note or conference paper).de. or (letter or editorial or news or comment).pt.)
15	Limit by publication type	14 and ((Randomized controlled trials or random allocation or double-blind method or single-blind method or placebos or cross-over studies or crossover procedure or cross over studies or double blind procedure or single blind procedure or placebo or latin square design or crossover design or double-blind studies or single-blind studies or triple-blind studies or random assignment or exp controlled study/ or exp clinical trial/ or exp comparative study/ or cohort analysis or follow-up studies.de. or intermethod comparison or parallel design or control group or prospective study or retrospective study or case control study or major clinical study or evaluation studies or follow-up studies).de. or random$.hw. or random$.ti. or placebo$ or ((singl$ or doubl$ or tripl$ or trebl$) and (dummy or blind or sham)) or latin square or ISRCTN$ or ACTRN$ or (NCT$ not NCT))

PsycINFO
English language, human
2000 – 2008
Dialog Syntax

Set Number	Concept	Search statement
1	Autism	s child development disorders, pervasive! or autistic disorder/de or (autis? or pdd or asperger? or disintegrative()disorder)/ti,ab
2	Therapy (controlled vocabulary)	s behavior therapy! or early intervention (education)/de
3	Therapy	s therapy! and (communication or social()skills or motor or perceptual or sensory or integrative or interpersonal or modeling)
4	Combine sets	s s2 or s3
5	Combine sets	s s1 and s4
6	Limit by publication type	s s5 not pt=(book or letter or dissert?)

Total Identified	Total Downloaded
1,085	306

Reimbursement

The following Web sites were searched for reimbursement policies:

- Aetna US Healthcare (http://www.aetnaushc.com/cpb/cpb_alpha.html)
- American Medical Association (http://coverageandpayment.mediregs.com)
- Athens Area Health Plan Select, Inc. (http://www.aahps.com/pdfs/EOCamend012006.pdf)
- Blue Cross/Blue Shield of Alabama (http://www.bcbsal.org/providers/policies/)
- Blue Cross/Blue Shield of Massachusetts (http://www.bcbsma.com/common/en_US/hresource/medcat.jsp)
- Blue Cross/Blue Shield of Tennessee (http://www.bcbst.com/providers/mpm.shtm)
- Cigna (http://www.cigna.com/health/provider/medical/procedural/coverage_positions/medical/index.html)
- Health Partners (http://www.healthpartners.com/policies/)
- Humana (https://providers.humana.com/ciinter/cihome.asp)
- Kaiser Permanente Northern California Region (www.kaiserpermenente.org)
- MAMSI Life and Health Insurance Company State of Maryland (www.mamsiunitedhealthcare.com/s/g/md/0726299-0105MD.pdf)
- Medica (http://provider.medica.com/C9/MedicalPolicies/default.aspx)
- Premera Blue Cross (http://www.ashya.org/about/legislation-advocacy/2008/PremeraBlueCross.htm)
- Regence Blue Cross/Blue Shield (http://www.regence.com/trgmedpol/)
- Wellmark Blue Cross/Blue Shield (http://www.wellmark.com/e_business/provider/medical_policies/medical_policies.asp)

We also used the Google and Vivisimo internet search engines to locate reimbursement information, using a combination of topic-specific keywords and the following search terms: (reimburs* OR coverage OR "medical policy").

Table 16. Excluded Studies

Study	Treatment(s)	Reason for Exclusion
Magiati, et al. (2007)(102)	EIBI	Study groups were not comparable at baseline. The EIBI group demonstrated significantly higher baseline IQ scores than the "eclectic" group (83.0 compared to 65.2), which the authors indicated significantly ($p < 0.001$) covaried with all posttreatment outcome variables.
Rickards, et al. (2007)(103)	Eclectic plus parent support	Study did not address one of the key questions of interest in this review.
Eldevik et al. (2006)(104)	Low intensity EIBI	Not a prospective controlled trial.
Reed et al. (2006)(105)	ABA	Study did not address one of the key questions of interest in this review.
Salt et al. (2006)(106)	Scottish Center for Autism	The study groups were not comparable at baseline. The control group demonstrated significantly higher IQ scores than the SCA group (39.43 compared to 55.67, $p < 0.05$). The study authors did not attempt to perform statistical analyses to determine if this difference effected study outcomes. Further, the majority of the waitlist control received speech therapy and demonstrated similar improvement on receptive and expressive language as the SCA group.
Panerai et al. (2002)(107)	TEACCH	The authors indicated that the two study groups demonstrated significant baseline differences on chronological age, mental age, and symptom severity, with the experimental group being at a statistical disadvantage compared to the control group. Further, the children included in this study had a diagnosis of autism and severe intellectually disability, and, in some cases, physical disabilities.
Jocelyn et al. (1998)(96)*	Community daycare with specialized support	Children in study did not meet DSM-IV or TR edition criteria for diagnosis of ASD. Study authors used DSM-III criteria.
Ozonoff & Cathcart (1998)(95)*	TEACCH	Children in study did not meet DSM-IV or TR edition criteria for diagnosis of ASD. Study authors used DSM-III criteria.
Sheinkopf & Siegal (1998)(94)*	ABA (Lovaas method)	Not a prospective controlled trial.
Smith et al. (1997)(108)	IBI	Not a prospective controlled trial and children in study did not meet DSM-IV or TR edition criteria for diagnosis of ASD/
Birnbrauer & Leach (1993)(109)	ABA (Lovaas method)	Children in study did not meet DSM-IV or TR edition criteria for diagnosis of ASD. Study authors used DSM-III criteria. Additionally, in ECRI Institute's previous review on comprehensive programs, this study was considered to have major flaws in its design (e.g., drop out rate >30%, lack of intent to treat analysis, and unknown group comparability), and was subsequently excluded from the review.
Lovaas et al. (1987)(93) & McEachlin et al. (1993)(92)* Same patient population in each study.	ABA (Lovaas method)	Children in study did not meet DSM-IV or TR edition criteria for diagnosis of ASD. Study authors used DSM-III criteria.

*Four of the last five studies were included as part of the evidence base in the previous ECRI Institute report on comprehensive treatment for autism. The results of these studies are discussed in this report in the section on *Previous Systematic Reviews*. Data from three studies were excluded from the current review because the diagnosis of children with ASD was based on DSM-III criteria which differs from the criteria in the DSM-IV(30), and one study was a retrospective case-controlled trial.

ABA Applied behavior analysis
EIBI Early intensive behavioral intervention
IBI Intensive behavioral intervention
TEACCH Treatment and Education of Autistic and Communication Handicapped Children

Appendix B. Description of Instruments Used to Measure Outcomes in Included Studies

Table 17. Name and Description of Validated Instruments

Instrument Name	Description of Instrument
Instruments Measuring Cognitive/Intellectual Status	
Bayley Scales of Infant Development (BSID)	The BSID consists of three scales (mental, motor, and behavior) that are used to assess development in young children aged one to 42 months. The test has a population mean or standard score of 100 and a standard deviation of 15 points. Higher scores indicate higher cognitive ability.(110)
Stanford-Binet Intelligence Tests(SBIS)	The SBIS is for individuals aged 2-years to adult and provides scores in 4 areas: verbal reasoning, abstract and visual reasoning, quantitative reasoning, and short-term memory. The test has a population mean or standard score of 100 and standard deviation of 16 points (subtests have a mean of 50 and a standard deviation of 8). Higher scores indicate higher cognitive ability.(110)
Wechsler Intelligence Scale for Children (WISC)	The WISC assesses the cognitive ability of children from 6.0 to 16.6 years. It contains 12 subtests which yield a Full Scale IQ, Verbal Scale IQ, and Performance Scale IQ. The population mean or standard score is 100 with a standard deviation of 15 each scale. Higher scores indicate higher cognitive ability.(78)
Wechsler Preschool and Primary Intelligence Scale (WPPSI)	The WPPSI is a frequently used intelligence test for children aged 3 to 7 years. It contains 12 subtests that measure children's visual-spatial skills, comprehensive, and vocabulary skills. For each subtest, the population mean or standard score is 10 and the standard deviation is 3. Higher scores indicate higher cognitive ability.(78)
Instruments Measuring Communication/Language Skills	
Early Social Communication Scales (ESCS)	The ESCS is a videotaped structured observation measure that requires 15 to 25 minutes to administer. It is designed to measure non-verbal communication skills in children as young as 8 months. Specifically, the ESCS measures the following behavioral areas: joint attention behaviors, behavioral requests, and social interaction behaviors. Higher scores indicate higher communication skills.(78)

Instrument Name	Description of Instrument
Merrill-Palmer Scale of Mental Tests (MPSMT)	The MPSMT is widely used as a non-verbal test instrument for measuring visual-spatial skills in children aged 16 months to six years.(78) This test is sometimes given in addition to the BSID. Higher scores indicate higher cognitive ability.
Reynell Developmental Language Scales	The Reynell is a language test designed for children one to seven years old that measures comprehension (receptive language) and expressive language.(111) Higher scores indicate higher language skills.
Instruments Measuring Adaptive Behavior	
Vineland Adaptive Behavior Scale(VABS)	The VABS comes in three forms varying in degree of detail and proposed setting. There is the Survey Form, the Expended Form, and the Classroom Edition. The VABS is administered by interviewing the child's parents, teachers, or care providers. The scales range in age from birth to 19 years. Raw scores from communication, daily living skills, socialization, motor skills, and maladaptive behaviors are converted to standard scores with a population mean of 100 and standard deviation of 15. Higher scores indicate better outcomes or performance of more adaptive behaviors.(78,80)
Instruments Measuring Problem Behavior	
Achenbach Child Behavior Checklist (ACBC)	The ACBC is for children 4 to 18 years old and is completed by an adult informant, typically the child's primary caregiver. It has two major scales—externalizing and internalizing behaviors—each of which have four subtests.(78) Higher scores reflect poorer outcomes. There is a separate version of the test developed for teachers, the Teacher Report From.
Autism Screening Questionnaire	The ASQ is a 40-item screening scale that has good discriminative validity between autistic spectrum and other disorders in children ages 4 years and older.(46) A score of 1 is given if the abnormal behavior is present, and a score of 0 is the behavior is absent. The cutoff for consideration of a diagnosis of autism is a score of 15 or higher.

Instrument Name	Description of Instrument
Developmental Behavior Checklist	The DBC is a 96-item checklist that is completed by parents or other primary carers or teachers, reporting behavioral and emotional problems in children and young people with intellectual or developmental disability over a six month period.(112) Each behavioral description is scored on a 0, 1, 2 rating scale where 0 = 'not true as far as you know', 1 = 'somewhat or sometimes true', and 2 = 'very true or often true'. The DBC can be completed in 10-15 minutes. The DBC provides scores at 3 levels: *Total behavior problem score* - an overall measure of behavioral and emotional problems. Receiver Operating Characteristic (ROC) analysis has shown that a total score of 46 or greater indicates clinically significant levels of behavioral and emotional problems. *Five sub-scales* (derived from factor analysis) - disruptive/antisocial behavior, self absorbed behavior, communications disturbance, anxiety problems and social relating problems. *Individual behavior items* - indicates the prevalence and severity of individual symptoms.
Preschool Behavior Checklist (PBCL)	The PBCL is 22-item checklist used to screen 2- to 5-year-old children for behavioral and emotional difficulties.(78) The checklist covers emotions (fears, worries, and mood), conduct, temper, activity level, concentration, social relations, speech, language, habits, wetting, and soiling. Each item lists three or four degrees of a particular behavior. Items are scored for frequency and severity. Higher scores indicate more behavioral problems.
Instruments Measuring Achievement	
Autism Diagnostic Interview-Revised	The ADI-R is a semi-structured, investigator-based interview for caregivers of children and adults for whom autism or pervasive developmental disorders is a possible diagnosis.(8) The ADI-R is recognized as one of the better standardized instruments currently available for establishing a diagnosis of autism based on the DSM-IV criteria for autism.
Autism Diagnostic Observation Schedule (stability of diagnosis)	The ADOS is a semistructured assessment of social interaction, communication, play, and imaginative use of materials included in the test packet.(9) The ADOS is standardized in terms of the materials used, the activities presented, the examiner's introduction of activities, the hierarchical sequence of social presses provided by the examiner, and the way behaviors are coded or scored. Following the administration of the ADOS, behaviors are coded using a 0- to 3-point coding system, with a 0 indicating that the behavior is not abnormal in the way specified in the coding description and a 3 indicating that a behavior is abnormal and interferes in some way with the child's functioning.

Instrument Name	Description of Instrument
Wechsler Individualized Achievement Test	The WIAT is a measure of individual achievement skills for the population targeting children, adolescents, college students, and adults, age 4 through 85.(113) The scales are reading, mathematics, written and oral language, and it also contains 9 sub-test scores. The suggested use of the WIAT is in settings such as schools, clinics, private practices and residential treatment facilities. These facilities can use the WIAT in order to assist with diagnosis, eligibility, placement, and decisions regarding interventions.Standard scores are by age or by grade, (M = 100, SD = 15) with a range of 40 to 160 and seasonal tables for pre-K - 8.
Woodcock-Johnson Tests of Achievement	This is a comprehensive, individually administered set of 27 tests that assess three areas of functioning: cognitive ability, achievement, and interest. The Test of Achievement in Part II covers 10 achievement areas such as reading, spelling, knowledge of science, etc. The battery assesses from age 3 years through adulthood. Higher scores indicate higher levels of achievement.(110)
Instruments Measuring Parent/Family Wellbeing	
Hospital Anxiety and Depression Scale	The HADS is a self-report 14-item scale which measures the states of anxiety and depression without contamination of scores by reports of physical symptomatology.(114) The scale was designed to be used in general medical outpatient populations, but is valid for use in community work (i.e., not in hospitals).
Kansas Inventory of Parental Perceptions Positive Contributions	The KIPP was developed for use among parents of children with disabilities. It measures the following areas: positive contributions, social comparisons, causal attributes, and mastery/control. Higher scores indicate more positive behaviors.(115)
Nisonger Child Behavior Checklist	The Nisonger is designed for children 3 to 16 years old, and assess problem behaviors such as conduct disorder, anxiety, hyperactivity, and stereotyped behavior.(78) It also includes a section that looks at prosocial or adaptive behavior.
Questionnaire on Resources and Stress-Friedrich short-form (QRS-F)	The QRS-F is a short form of the QRS. It is a 52-item true/false scale which measures the parents' perceptions of the impact of a developmentally delayed or chronically ill child on the family. There are 4 subscales parent and family problems; pessimism; child characteristics; physical incapacity. The total score represents perceived overall stress.(78)

Appendix C. Quality of Literature and Evidence Strength Rating

Determining the Quality of Individual Studies

For Key Questions 1, 2, and 3, we assessed the quality of each of the studies included in this assessment using a quality assessment instrument developed by ECRI Institute. This instrument examines twenty-two different factors of study design that have the potential to reduce the validity of the conclusions that can be drawn from a trial. Each question is answered with "Yes", "No" or "NR" (not reported).

Quality Checklist Items:

Comparability of Groups at Baseline

1. Were participants randomly assigned to the study's groups?
2. Did the study employ stochastic randomization?
3. Were any methods other than randomization used to make the participants in the study's groups comparable?
4. Were participants assigned to groups based on factors other than child or provider preference?
5. Were the characteristics of participants in the different study groups comparable at the time they were assigned to groups?
6. Did participants in the different study groups have similar levels of performance on all of the outcome variables at the time they were assigned to groups?
7. Was the comparison of interest prospectively planned?
8. Did ≥85% of the participants complete the study?
9. Was there a ≤15% difference in completion rates in the study's groups?
10. Were all of the study's groups concurrently treated?
11. Was compliance with treatment ≥85% in both of the study's groups?
12. Was there concealment of allocation?

Blinding

13. Were the outcome assessors blinded to the group to which the participants were assigned?

Measurement/Instrument

14. Was the outcome measure of interest objective and was it objectively measured?
15. Were the same instruments used to measure the outcomes in all of the study's groups?
16. Was the instrument used to measure the outcome standard?
17. Were the follow-up times in all of the study's relevant groups approximately equal?

Treatment

18. Was the same treatment given to all participants enrolled in the experimental group?
19. Was the same treatment given to all participants enrolled in the control group?
20. Were all of the study's groups treated at the same center?
21. Was the treatment provider's adherence to the intervention protocol (treatment fidelity) assessed?

Investigator Bias

22. Was the funding for this study derived from a source that does not have a financial or proprietary interest in its results?

We scored the quality for each outcome/timepoint by coding +1 for each Yes, -1 for each No, and 0 for each NR. The numbers were added, and then we transformed the total so that the best possible study would score 10 (i.e., all Yes's), and the worst possible study would score 0 (i.e., all No's). If the resulting combined score was <7, we categorized the quality as Low; if the score was ≥7, we categorized quality as Moderate. We then used these quality categories to proceed through the Strength of Evidence system, described next.

Strength-of-Evidence System

Ideally, the body of evidence to support a conclusion would be strong. Often, however, the evidence suffers from various limitations concerning the possible risk of bias in available studies, small numbers of studies and patients, and/or inconsistent effects. These limitations often mean that the strength of the evidence is only moderate, weak, or even insufficient to permit any conclusion. In order to gauge the impact of these possible limitations, we applied a formal rating system developed at ECRI Institute.(73)

Our system allows one to separate the question "is the treatment effective" (leading to a yes or no conclusion) from the question "how effective is the treatment" (leading to a quantitative conclusion with an estimate of the magnitude of effect). Thus, even if the evidence for a precise quantitative effect may not be strong, the same evidence may be strong with respect to the direction of the effect. The interpretation of the strength of the evidence for qualitative and quantitative conclusions is shown in Table 18.

Table 18. Interpretation of Different Categories of Strength of Evidence Supporting Conclusion

Strength of Evidence	Interpretation
Qualitative Conclusion (Direction of Effect)	
Strong Evidence	Evidence supporting the qualitative conclusion is convincing. It is highly unlikely that new evidence will lead to a change in this conclusion.
Moderate Evidence	Evidence supporting the qualitative conclusion is somewhat convincing. There is a small chance that new evidence will overturn or strengthen our conclusion. ECRI recommends regular monitoring of the relevant literature at this time.
Weak Evidence	Although some evidence exists to support the qualitative conclusion, this evidence is tentative and perishable. There is a reasonable chance that new evidence will overturn or strengthen our conclusions. ECRI recommends frequent monitoring of the relevant literature at this time.
Insufficient	Although some evidence exists, this evidence is not of sufficient strength to warrant drawing an evidence-based conclusion from it. ECRI recommends frequent monitoring of the relevant literature at this time.
Quantitative Conclusion (Magnitude of Effect)	
High Stability	The estimate of effect is stable. It is highly unlikely that the magnitude of this estimate will change substantially as a result of the publication of new evidence.
Moderate Stability	The estimate of effect is somewhat stable. There is a small chance that the magnitude of this estimate will change substantially as a result of the publication of new evidence. ECRI recommends regular monitoring of the relevant literature at this time.
Low Stability	The estimate of effect is likely to be unstable. There is a reasonable chance that the magnitude of this estimate will change substantially as a result of the publication of new evidence. ECRI recommends frequent monitoring of the relevant literature at this time.
Unstable	Estimates of the effect are too unstable to allow a quantitative conclusion to be drawn at this time. ECRI recommends frequent monitoring of the relevant literature.

The system employs 14 decision points (Table 19). Four of them are listed in the General section because they apply to both quantitative conclusions as well as qualitative conclusions. The other 10 apply specifically to either quantitative conclusions (Decision Points 5-9) or qualitative conclusions (Decision Points 10-14). The rest of this appendix defines these decision points and describes how we resolved them for this report. After these descriptions, the pathways for the full system appear in Figure 7 through Figure 10.

Note that we applied this system separately for each outcome of interest. This is because many aspects of the evidence (quality, consistency, etc.) can vary by outcome.

Table 19. Decision Points in the ECRI System

Category	Decision Point
General	1) What is the quality of individual studies?
	2) What is the overall quality of evidence?
	3) Could a quantitative estimate be appropriate?
	4) Are data informative?
Quantitative	5) Are data quantitatively consistent?
	6) Are data quantitatively robust?
	7) Are there sufficient data to perform meta-regression?
	8) Does meta-regression explain heterogeneity?
	9) Is the meta-regression model robust?
Qualitative	10) Are data qualitatively robust?
	11) Is meta-analysis possible?
	12) Are data qualitatively consistent?
	13) Was at least one study a multicenter study?
	14) Is the magnitude of effect extremely large?

Decision Point 1: What is the quality of individual studies?

To aid in assessing the quality of each of the studies included in this assessment, we used a quality instrument developed by ECRI Institute for controlled trials. This instrument examines different factors of study design (attributes) that have the potential to reduce the validity of the conclusions that can be drawn from a trial (see *Determining the Quality of Individual Studies* in the above section for the complete instrument). In brief, the scale was designed so that a study attribute that, in theory, protects a study from bias receives a "Yes" response. If the study clearly does not contain that attribute it receives a "No" response. If poor reporting precludes assigning a "Yes" or "No" response for an attribute, then "NR" is recorded (NR = not reported).

To assess the quality of an individual study, we computed a normalized score so that a perfect study received a score of 10, a study for which the answers to all items was "No" received a score of 0, and a study for which the answers to all questions was "NR" was 5. Quality scores

were converted to categories as shown in Table 13 (see *Methods* section of main document). The definitions for what constitutes low and moderate quality evidence were determined *a priori* by a committee of four methodologists. Because the quality was determined separately for each outcome, a study that scored as moderate quality for one outcome might score as low quality for another outcome.

Decision Point 2: What is the overall quality of evidence?

We classified the overall quality of the evidence base by taking the median quality score of the individual studies. We used the median because it is the appropriate measure of central tendency to represent the "typical" quality score, and is less sensitive to outliers than the mean. Depending on the overall quality scores for each outcome, we then followed the high, moderate, or low quality branch of the system.

Decision Point 3: Is calculating a quantitative estimate potentially appropriate?

The answer to Decision Point 3 depends upon the adequacy of reporting in available studies as well as the number of available studies. In order to permit a quantitative estimate of an effect size for a given outcome, the data for that outcome must be reported in at least three studies in a manner that allows the data to be pooled in a meta-analysis. If less than three studies are available, no quantitative estimate is usually appropriate, regardless of reporting. Another situation that does not permit a quantitative estimate is when at least three studies are relevant to the general topic, but fewer than 75% of them reported the outcome and as well as sufficient information for determination of the effect size and its dispersion, either by direct reporting from the trial or calculations based on reported information. If no quantitative estimate would be appropriate, then one moves directly to Decision Point 10 to determine whether the evidence supports a qualitative conclusion.

Decision Point 4: Are data informative?

For this decision point, we determined whether the precision of an evidence base was sufficient to permit a conclusion. Statistically significant results are informative because they mean that a treatment effect may exist. Statistically non-significant results are also potentially informative, but only if they exclude the possibility that a clinically significant treatment effect exists.

When a meta-analysis is performed, a key concern is the confidence interval around the random-effects summary statistic. If this interval is so wide that it is includes a clinically significant (or substantial) effect in one direction *and also an effect in the opposite direction*, then the evidence is inconclusive, and therefore uninformative.(116,117)

Thus, when considering the summary effect size from a meta-analysis (or the effect size from a single study), there are three ways in which the effect can be "informative":

1) The effect size is statistically significantly different from 0. This would be indicated whenever the confidence interval does not overlap 0.

2) The confidence interval is narrow enough to exclude the possibility that a *clinically significant difference* exists.

3) The confidence interval is narrow enough to exclude the possibility that a *substantial difference* exists. This possibility is included to address situations when even a very small effect can be considered "clinically significant" (e.g., a difference in mortality rates), but the effect may not be "substantial".

The second possibility requires definitions of a minimum "clinically significant difference" for each outcome. For the outcomes in this report, Table 20 lists our definitions of "clinical significance".

Table 20. Minimum Difference for Clinical Significance

Outcome	***Minimum*** **difference between groups at post-treatment to be considered clinically significant**
Cognitive/Intellectual Status and Language/Communication Skills	One half of the standard deviation of the mean for typically developing children, which for most tests of IQ and language skills is a standard deviation of 15.(78) So, the minimum difference to be considered clinically significant for this report is a standard deviation of 7.5. For dichotomous outcomes (e.g., number of children moving into normal range on IQ scores, which is a score of 85 or greater, any statistically significant difference is considered to be clinically significant.
Adaptive Behavior	One half of the standard deviation of the mean for typically developing children, which for the Vineland Adaptive Behavior Scales is a standard deviation of 15.(80) So, the minimum difference to be considered clinically significant for this report is a standard deviation of 7.5.
Problem Behaviors	A SMD of 0.2, which corresponds to a small effect size, is considered to be clinically significant.(79)
Academic/Developmental Achievement	For continuous outcomes (i.e., achievement tests), a SMD of 0.2, which corresponds to a small effect size, is considered to be clinically significant. For dichotomous outcomes (classroom placement or change severity of symptoms), a statistically significant difference is considered to be clinically significant.
Parent/family Well-being	A SMD of 0.2, which corresponds to a small effect size, is considered to be clinically significant.(79)

Note that when the evidence base consists of one or two studies, and the only usable data from one study consists of a p-value that was calculated using the wrong statistical test, then the data cannot generally be considered "informative." If, however, the study reported sufficient information for one to perform the correct test, then informativeness can be determined.

Decision Point 5: Are data quantitatively consistent?

Quantitative consistency (also referred to as lack of heterogeneity) refers to the extent to which the effect sizes of studies in an evidence base were statistically similar.(118) To measure quantitative consistency, we used Higgins and Thompson's I^2 statistic.(74) For this report, we considered an evidence base to be quantitatively consistent when I^2 <50%.

Decision Point 6: Are data quantitatively robust?

Robustness of findings refers to whether the evidence for a summary estimate is both *precise* and *stable*. A precise estimate is one for which the evidence permits a narrow confidence range for possible values of the parameter. A stable estimate is one that does not change substantially in

response to minor alterations in the analysis. In this report, we considered an estimate to be quantitatively robust if all of the following conditions were met:

1. The overall estimate is sufficiently precise
2. The estimate remains sufficiently precise after the removal of any single study
3. The estimate remains sufficiently precise after a cumulative robustness test by year

Test #1: Sufficient precision. An important component of the evidence for a summary estimate is the precision of that estimate. Specifically, we refer to the 95% confidence interval (CI) around the estimate as a measure of precision. This is an objective measure of the quantity of evidence that *simultaneously incorporates* 1) the number of studies; 2) the number of patients in those studies; and 3) within-study variability of effect sizes; and 4) between study-variability of effect sizes (because we only perform random-effects meta-analyses). An imprecise estimate is one that could easily change when future evidence becomes available (i.e., a wide confidence interval), whereas a precise estimate is unlikely to change (i.e., a narrow confidence interval).

To assess whether precision is "sufficient", we refer to the minimum difference that is considered to be clinical significant. Specifically, we defined a "sufficiently precise" estimate as one where the lower and upper confidence bounds were *each within one clinically significant difference* from the summary estimate. If not, then the evidence base is not precise enough to locate the effect within a clinically equivalent range. For example, suppose the summary effect size is 10, with a CI of 8.5 to 11.5. Further suppose that the definition of clinical significance is 2 units. This indicates that data *are* sufficiently precise to provide an estimate that is within 1 clinically significant difference, and so the estimate would pass this test. However, suppose the CI had been 7 to 13. Then the interval suggests that the true effect could be a full three units above or below the estimate of 10. Three units is greater than the minimum clinically significant difference of 2, therefore a 7 to 13 interval would fail this test.

For some variables (e.g., change in diagnosis, classroom placement) any difference at all can be considered clinically significant. For variables, such as standardized tests, we defined the magnitude of a "substantial difference" as one half the standard deviation for typically developing children (i.e., 7.5). For other continuous variables, we defined the magnitude of substantial difference as defined by Cohen.(79) Thus, if the effect size metric is SMD or Hedges' g, we defined a "substantial difference" as $d = 0.2$, or if the effect size metric is the log odds ratio, we defined a "substantial difference" as $\ln(OR) = 0.4$.

Test #2: Removal of one study at a time. The summary estimate should not depend heavily on the inclusion of any particular study in the evidence base. To test whether any one study exerts a large impact on the summary estimate, we remove it from the evidence base and recalculate the summary estimate, replace it, and then perform a series of subsequent analyses, each time removing one study and recalculating the summary estimate before replacing it and repeating with the next study. In order to pass this test, the lower and upper bounds of the 95% CI in all analyses should be within one clinically significant difference from the *all-study* summary estimate. Thus, this test produces a new set of CIs (one set of CIs for each calculation study removal), and each CI is compared to the all-study summary estimate.

Test #3: Cumulative robustness test by year. If recent studies have reported very different effect sizes from older studies, then not-yet published studies may be expected also to cast doubt

on a summary effect size. For this test, we determined whether effect sizes demonstrate a clear downward or upward trend over time. If so, the quantitative estimate was deemed not robust.

Decision Point 7: Are there sufficient data to perform meta-regression?

We required a minimum of five studies before attempting meta-regression.

Decision Point 8: Does meta-regression explain heterogeneity?

This decision point provides decision rules for the conduct of a meta-regression analysis and the interpretation of its results. The project internal review committee must determine *a priori* what methods will be used in performing a meta-regression should one be necessary. In addition, the committee must define the rules that will be used for interpretation of the findings of the meta-regression analysis. We use the permutation test for all meta-regressions.(119) This test was developed by Higgins and Thompson in attempt to control the Type I error rate for meta-regression.(75)

For this topic, we chose the following covariates as potential explanations of heterogeneity:

1. Duration of treatment (defined as less than 1 year or greater than 1 year).
2. Intensity of treatment (only applicable to key question 2, in cases where one comprehensive program was delivered for less hours than another comprehensive program).
3. Training/experience of provider (the definition of this variable may differ depending on the intervention and who is providing it (parents or therapist). If meta-regression is possible, we will need to consider what training is required and establish whether the criteria were met. So, this might be a continuous variable measured in hours or a dichotomous variable measured as provider met established training requirements or not)
4. Fidelity/integrity of treatment-when measured within a study (Yes or No).
5. Quality category of study (High or Moderate or Low)
6. Use of blinded assessors (Blinded or Not blinded)
7. Use of concomitant treatment in experimental group (medication or supplemental services versus none)

In order to determine that a given covariate "explains" the heterogeneity, the resulting I^2 must have been less than 50%, and the beta coefficient for the covariate must have been statistically significant by the permutation test.

Decision Point 9: Is the meta-regression model robust?

The purpose of Decision Point 7 is to test the robustness of any quantitative findings that may emanate from meta-regression analysis. The only necessary robustness test involves removing one study at a time to determine whether this alters the findings of the meta-regression. If removal of one study results in heterogeneity that is greater than or equal to $I^2 = 50\%$, or caused the covariate to become statistically non-significant by the permutation test, then the meta-regression model is not robust.

Decision Point 10: Are data qualitatively robust?

If the evidence base for an outcome had three or more studies, we determined whether the qualitative findings could be overturned by sensitivity analyses. We considered findings to be overturned only when a sensitivity analysis altered the conclusion (e.g., a statistically significant finding becomes non-significant as studies are added to the evidence base). The same sensitivity analyses used to test quantitative robustness were used to test qualitative robustness (except for the sufficient precision test, which does not apply to this decision point).

The system allows for several general types of qualitative conclusions:

a) A conclusion that the effect is statistically significant

b) A conclusion that the effect is clinically significant (see definition of clinical significance in Decision Point #4 above).

c) A conclusion that the effect is not clinically significant

d) A conclusion that the effect is not "substantial." (see definition of "substantial" in Decision Point #4 above)

For each of these types of conclusions, the qualitative robustness test will depend critically on a different threshold. For conclusion a, the question is whether the statistical significance of the finding is preserved across all qualitative robustness tests. In practical terms, this means that the lower bound of the 95% confidence interval must not overlap with 0 in any of the robustness tests. For conclusion b, the issue is whether the lower bound of the confidence interval stays consistently *above* the level of clinical significance across all robustness tests. For conclusion c, the issue is whether the lower bound of the confidence interval stays consistently *below* the level of clinical significance across all robustness tests. Finally, for conclusion d, the issue is whether the lower bound of the confidence interval stays consistently *below* the level of a substantial difference across all robustness tests.

Note that more than one qualitative conclusion could apply to the same outcome. For example, a treatment could be both statistically and clinically significantly better than an alternative (conclusions a and b). Or, a treatment could be statistically better than an alternative but clearly not clinically better (conclusions a and c). Conclusions b, c, and d, however, are mutually exclusive. Conclusions b and c are opposites; conclusion d only applies when the notion of "clinical significance" is inappropriate (see Decision Point #4 for further explanation).

Decision Point 11: Is meta-analysis possible?

This Decision Point is used only when the evidence base for an outcome consists of two studies.

A meta-analysis is possible if each study reports an effect size and its standard error, or if each study reports sufficient information for the reader to calculate these values. Note that meta-analysis is never appropriate if two studies have statistically significant effect sizes in opposite directions.

Decision Point 12: Are data qualitatively consistent?

This Decision Point is used only when the evidence base for an outcome consists of two studies. Table 21 depicts several situations using confidence intervals. For each situation, one can ask three questions:

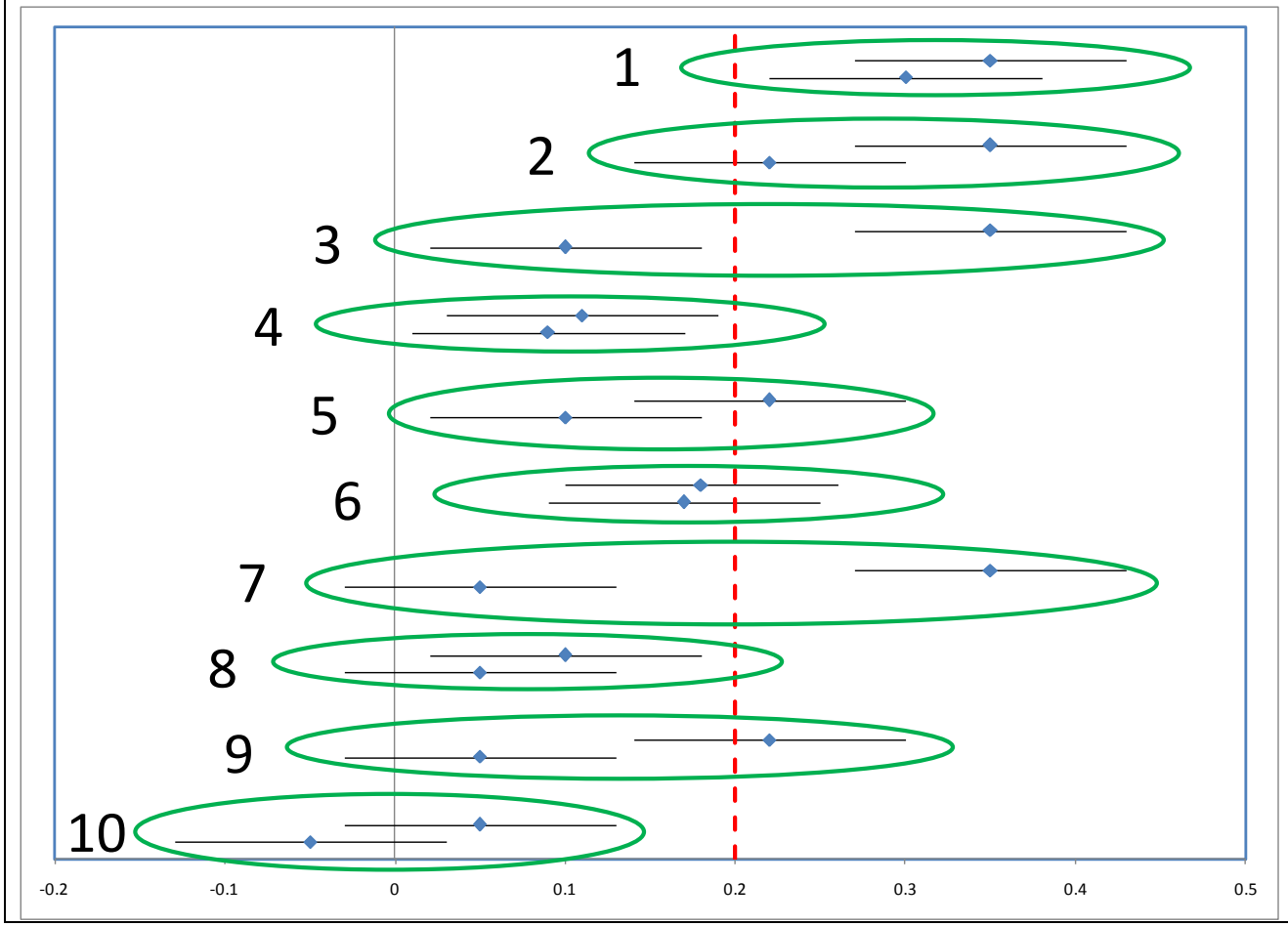
1) Do the two studies both support the conclusion that the effect size is ***greater than 0***?

2) Do the two studies both support the conclusion that the effect size is ***greater than the minimum clinically significant effect size*** (as defined in the graph by an effect size of 0.2)?

3) Do the two studies both support the conclusion that the effect size is ***less than the minimum clinically significant effect size*** (as defined in the graph by an effect size of 0.2)?

Qualitative consistency can be judged separately for these three questions; a pair of studies may be qualitatively consistent in some ways but not others. For each of the situations depicted in the figure, the right portion lists the corresponding determinations of qualitative consistency. Some questions are not applicable to a given pair of results because neither study would support that type of conclusion (e.g., in Situation #1, the 3rd question would not be supported by either study and therefore the issues of qualitative consistency in the 3rd sense would not apply).

Table 21. Qualitative Consistency of Two Studies



ES >0?	ES >0.2?	ES <0.2?
✓	✓	NA
✓	X	NA
✓	X	X
✓	NA	✓
✓	NA	X
✓	NA	NA
X	X	X
X	NA	✓
X	NA	X
NA	NA	✓

NOTES: Each point is the result of a single study with its 95% CI. The dashed line at 0.2 represents the minimum difference considered to be clinically significant. In the right-hand cells, a checkmark ✓ means that the two studies are qualitatively consistent with respect to the question at the top of the column. An X means that the two studies are NOT qualitatively consistent with respect to the question at the top of the column NA means that the question at the top of the column does not apply because neither study supports that conclusion. ES denotes effect size.

Decision Point 13: Was at least one study a multicenter study?

Multicenter trials may increase the strength of a one or two-study evidence base because they demonstrate partial replication of findings; they have shown that different investigators at different centers can obtain similar results using the same protocol. We defined a multicenter trial as any trial that met the following two conditions: 1) ≥3 centers and 2) either ≥100 patients or at least 3 centers enrolled ≥20 patients/center.

Decision Point 14: Is the magnitude of effect extremely large?

When considering the strength of evidence supporting a qualitative conclusion based on only one or two studies, magnitude of effect becomes very important. If a single study finds a large effect with a narrow confidence interval, then new evidence is unlikely to overturn the qualitative conclusion. To resolve this decision point, we consulted the effect size and the 95% confidence interval around the effect size for the study (with two studies, we consulted the interval around the random effects summary statistic).

For any outcome to be considered "large", two criteria must have been met:

- The observed point estimate must have been at least a SMD of 0.8, which corresponds to an odds ratio of 3.74 using the formula recommended by Sanchez.(120) If other units were used (e.g., points on a scale), the observed point estimate must have been four times larger than the minimum clinically significant difference. We used 4x because this was the ratio of large to small effects suggested by Cohen (1988).(79)
- The confidence interval around the estimate must have been fully above a SMD of 0.5, which corresponds to an odds ratio of 2.28 using the formula recommended by Sanchez.(120) If other units were used (e.g., points on a scale), the confidence interval must have been fully above 2.5x the minimum clinically significant difference. We used 2.5x because this was the ratio of moderate to small effects suggested by Cohen (1988).(79)

Figure 7. General Section of Strength-of-Evidence System

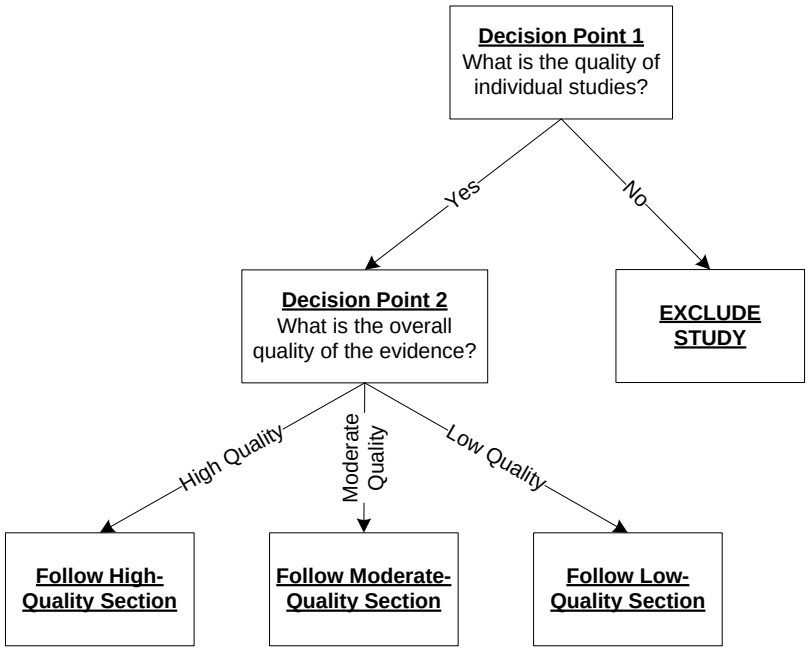


Figure 8. Highest Quality Pathway of Strength-of-Evidence System

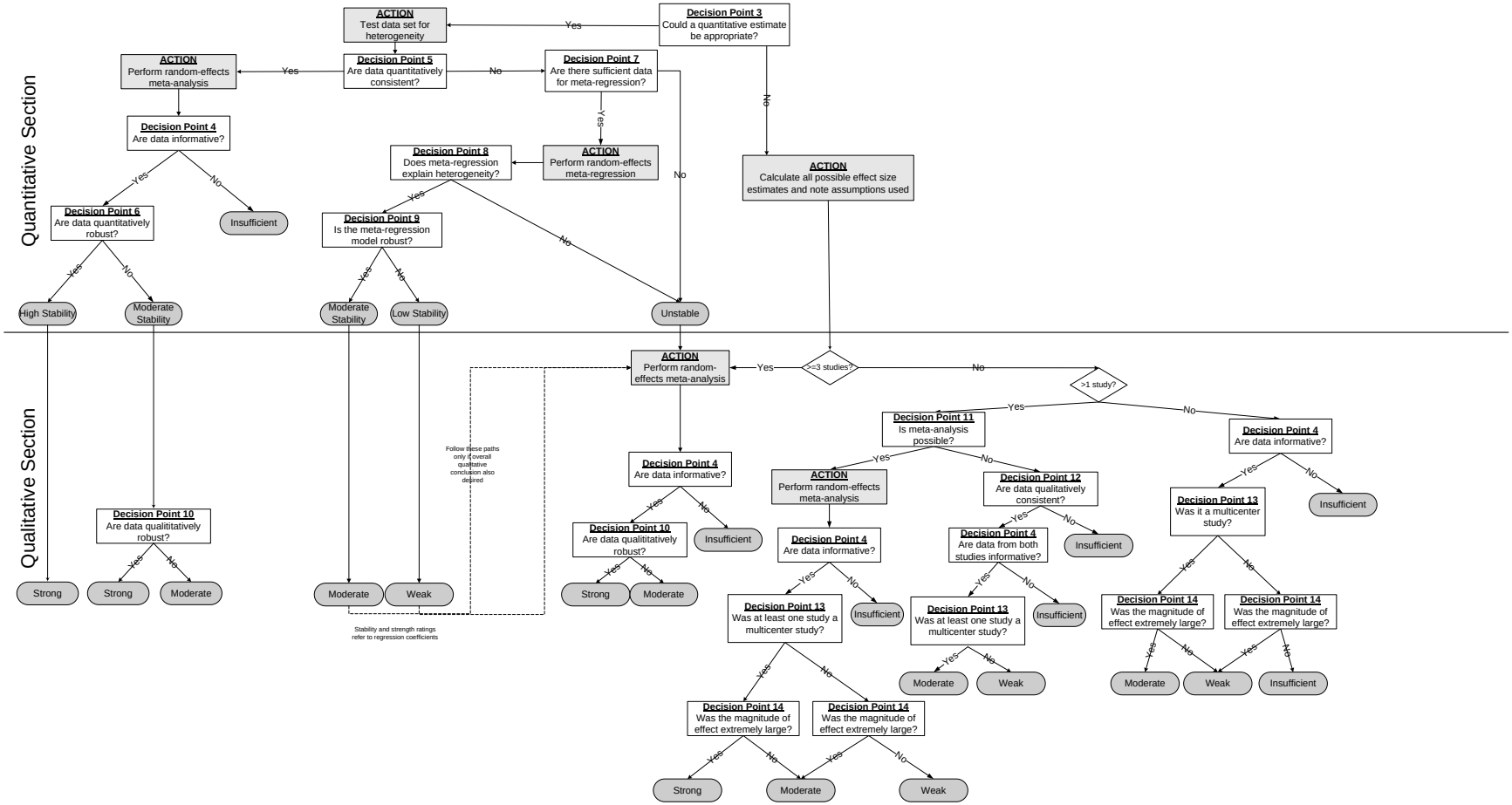


Figure 9. Moderate Quality Pathway of Strength-of-Evidence System

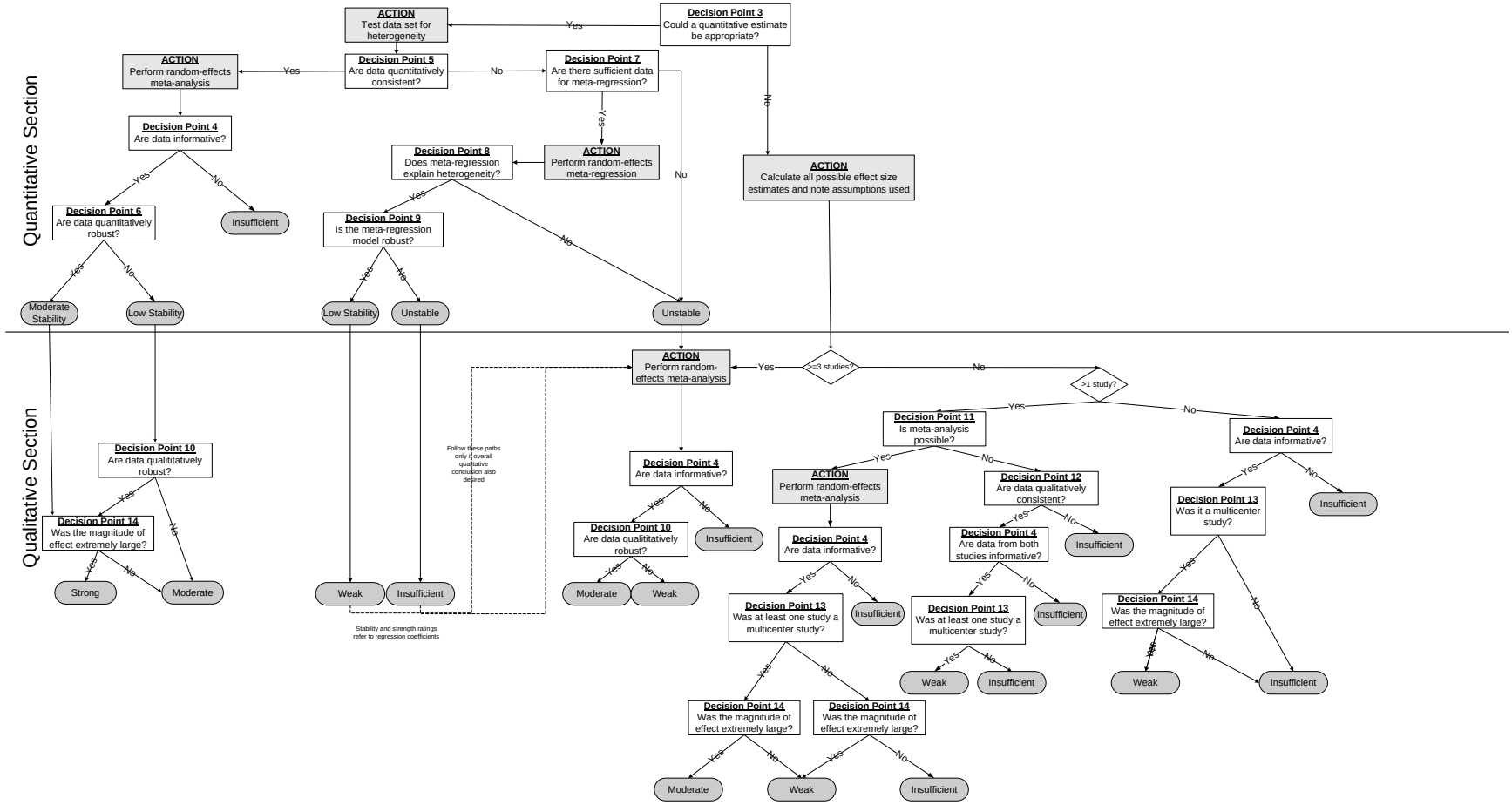
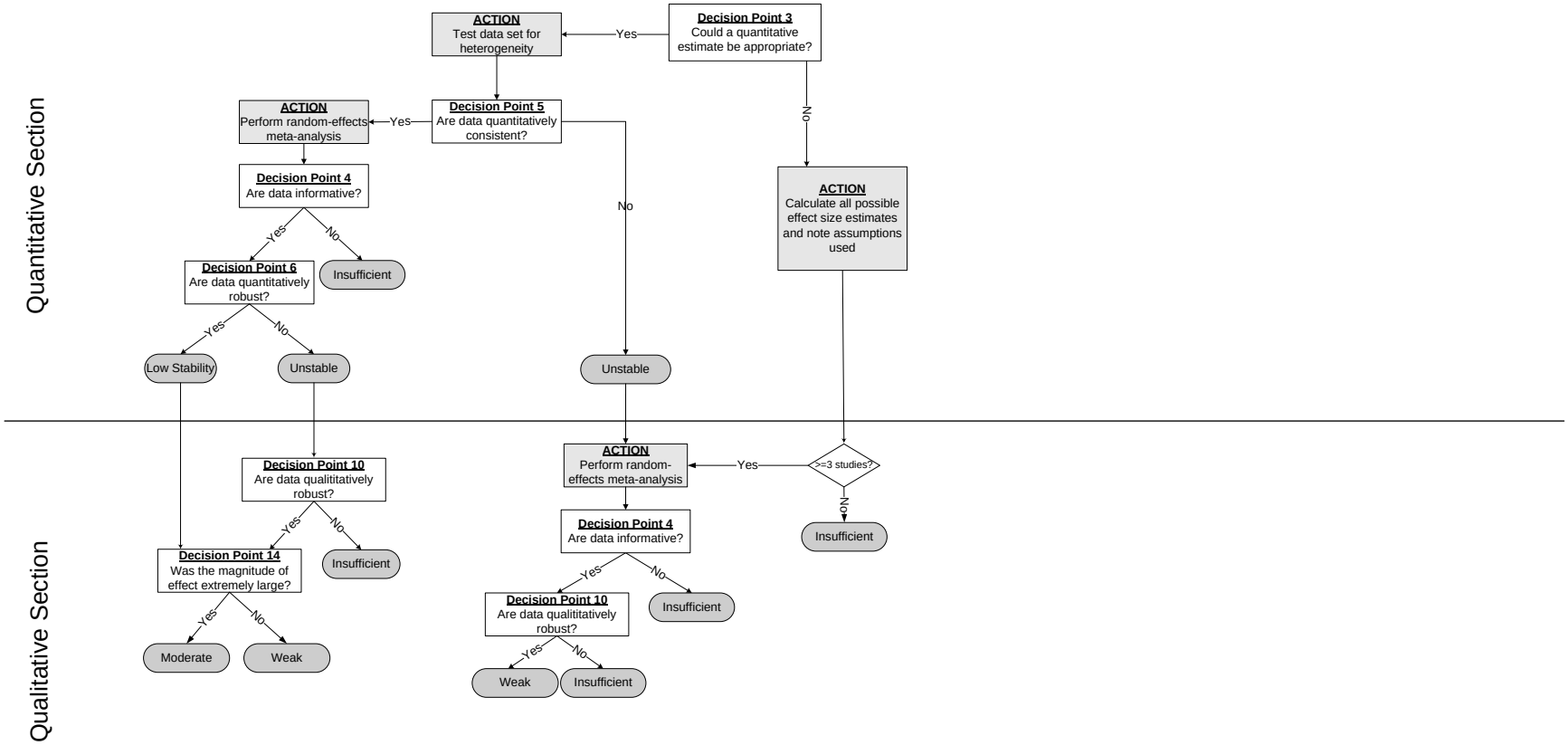


Figure 10. Lowest Quality Pathway of Strength-of-Evidence System



Appendix D. Quality Assessment Scores

Table 22. Quality Assessment for Intellectual Status and Language or Communication Skills

ECRI Institute's Controlled Trial Instrument	Intellectual Status and Language/Communication Skills						
	Remington et al.(66)[1]	Cohen et al.(70)[1]	Zachor et al.(69)[1]	Sallows & Graupner(71)	Howard et al.(65)[1]	Eikeseth et al.(67) & Eikeseth et al.(68)	Smith et al.(72)
1. Were participants randomly assigned to the study's groups?	N	N	N	Y	N	N	Y
2. Did the study use appropriate randomization methods?	N	N	N	Y	N	N	Y
3. Were methods used to make the participants in the study's groups comparable?	Y	Y	Y	Y	Y	NR	Y
4. Were subjects assigned to groups based on factors other than individual, parent or provider preference?	N	N	N	Y	N	N	Y
5. Were characteristics of participants in the different study groups comparable at the time they were assigned to groups?	Y	Y	Y	Y	Y	Y	Y
6. Did participants in the different study groups have similar levels of performance on all of the outcome variables at baseline?	Y	Y	Y	Y	Y	Y	Y
7. Was the comparison of interest prospectively planned?	Y	Y	Y	Y	Y	Y	Y
8. Did ≥85% of participants complete the study?	Y	Y	Y	Y	N	Y	Y
9. Was there a ≤15% difference in completion rates in the study's groups?	Y	Y	Y	Y	Y	Y	Y

ECRI Institute's Controlled Trial Instrument	Intellectual Status and Language/Communication Skills						
	Remington et al.(66)[1]	Cohen et al.(70)[1]	Zachor et al.(69)[1]	Sallows & Graupner(71)	Howard et al.(65)[1]	Eikeseth et al.(67) & Eikeseth et al.(68)	Smith et al.(72)
10. Were all of the study's groups concurrently treated?	Y	Y	Y	Y	Y	Y	Y
11. Was compliance with treatment ≥85% in both of the study's groups?	NR	NR	NR	NR	NR	NR	NR
12. Was there concealment of allocation?	N	N	N	NR	N	N	NR
13. Were outcome assessors blinded to the group to which the participants were assigned?	Y	Y	N	N	N	Y	Y
14. Was the outcome measure of interest objective and was it objectively measured?	N	N	N	N	N	N	N
15. Were the same instruments used to measure the outcomes in all of the study's groups?	Y	Y	N	N	N	Y	Y
16. Was the instrument used to measure the outcome standard?	Y	Y	Y	Y	Y	Y	Y
17. Were the follow-up times in all the study's relevant groups approximately equal?	Y	Y	Y	Y	Y	Y	Y
18. Was the same treatment given to all the participants enrolled in the experimental group?	Y	Y	Y	Y	Y	Y	Y
19. Was the same treatment given to all participants enrolled in the control group?	N	N	Y	Y	Y	N	Y
20. Were all of the study's groups treated at the same center?	Y	Y	Y	Y	Y	Y	Y

ECRI Institute's Controlled Trial Instrument	Intellectual Status and Language/Communication Skills						
	Remington et al.(66)[1]	Cohen et al.(70)[1]	Zachor et al.(69)[1]	Sallows & Graupner(71)	Howard et al.(65)[1]	Eikeseth et al.(67) & Eikeseth et al.(68)	Smith et al.(72)
21. Was the treatment provider's adherence to the intervention protocol (treatment fidelity) assessed?	NR	Y	NR	Y	NR	NR	Y
22. Was the funding for this study derived from a source that does not have a financial or proprietary interest in its results?	Y	Y	Y	Y	Y	Y	Y
Overall Quality Score	**6.8**	**7.0**	**6.4**	**8.2**	**6.4**	**6.6**	**9.1**

[1] Study authors reported between group differences on chronological age. However, subsequent analysis conducted by the authors using age as a covariate indicated that age did not have an effect on any of the outcomes of interest. Thus, we considered the study groups to be comparable.

NR Not reported
N No
Y Yes

Table 23. Quality Assessment for Adaptive Behavior

ECRI Institute's Controlled Trial Instrument	Adaptive Behavior					
	Remington et al.(66)[1]	Cohen et al.(70)[1]	Sallows & Graupner(71)	Howard et al.(65)[1]	Eikeseth et al.(67) & Eikeseth et al.(68)	Smith et al.(72)
1. Were participants randomly assigned to the study's groups?	N	N	Y	N	N	Y
2. Did the study employ stochastic randomization?	N	N	Y	N	N	Y
3. Were methods used to make the participants in the study's groups comparable?	Y	Y	Y	Y	NR	Y
4. Were subjects assigned to groups based on factors other than individual or provider preference?	N	N	Y	N	N	Y
5. Were characteristics of participants in the different study groups comparable at the time they were assigned to groups?	Y	Y	Y	Y	Y	Y
6. Did participants in the different study groups have similar levels of performance on all of the outcome variables at baseline?	Y	Y	Y	Y	Y	Y
7. Was the comparison of interest prospectively planned?	Y	Y	Y	Y	Y	Y
8. Did ≥85% of participants complete the study?	Y	Y	Y	N	Y	Y
9. Was there a ≤15% difference in completion rates in the study's groups?	Y	Y	Y	Y	Y	Y
10. Were all of the study's groups concurrently treated?	Y	Y	Y	Y	Y	Y
11. Was compliance with treatment ≥85% in both of the study's groups?	NR	NR	NR	NR	NR	NR
12. Was there concealment of allocation?	N	N	NR	N	N	NR
13. Were outcome assessors blinded to the group to which the participants were assigned?	Y	Y	N	N	Y	Y
14. Was the outcome measure of interest objective and was it objectively measured?	N	N	N	N	N	N

ECRI Institute's Controlled Trial Instrument	Adaptive Behavior					
	Remington et al.(66)[1]	Cohen et al.(70)[1]	Sallows & Graupner(71)	Howard et al.(65)[1]	Eikeseth et al.(67) & Eikeseth et al.(68)	Smith et al.(72)
15. Were the same instruments used to measure the outcomes in all of the study's groups?	Y	Y	Y	Y	Y	Y
16. Was the instrument used to measure the outcome standard?	Y	Y	Y	Y	Y	Y
17. Were the follow-up times in all the study's relevant groups approximately equal?	Y	Y	Y	Y	Y	Y
18. Was the same treatment given to all the participants enrolled in the experimental group?	Y	Y	Y	Y	Y	Y
19. Was the same treatment given to all participants enrolled in the control group?	N	N	Y	Y	N	Y
20. Were all of the study's groups treated at the same center?	Y	Y	Y	Y	Y	Y
21. Was the treatment provider's adherence to the intervention protocol (treatment fidelity) assessed?	NR	Y	Y	NR	NR	Y
22. Was the funding for this study derived from a source that does not have a financial or proprietary interest in its results?	Y	Y	Y	Y	Y	Y
Overall Quality Score	**6.4**	**7.0**	**8.6**	**6.4**	**6.6**	**9.1**

[1] Study authors reported between group differences on chronological age. However, subsequent analysis conducted by the authors using age as a covariate indicated that age did not have an effect on any of the outcomes of interest. Thus, we considered the study groups to be comparable.

NR Not reported
N No
Y Yes

Table 24. Quality Assessment for Problem Behavior

ECRI Institute's Controlled Trial Instrument	Problem Behavior			
	Remington et al.(66)[1]	Eikeseth et al.(67) & Eikeseth et al.(68)	Sallows & Graupner(71)	Smith et al.(72)
1. Were participants randomly assigned to the study's groups?	N	N	Y	Y
2. Did the study employ stochastic randomization?	N	N	Y	Y
3. Were methods used to make the participants in the study's groups comparable?	Y	NR	Y	Y
4. Were subjects assigned to groups based on factors other than individual or provider preference?	N	N	Y	Y
5. Were characteristics of participants in the different study groups comparable at the time they were assigned to groups?	N	Y	Y	Y
6. Did participants in the different study groups have similar levels of performance on all of the outcome variables at baseline?	Y	Y	Y	Y
7. Was the comparison of interest prospectively planned?	Y	Y	Y	Y
8. Did ≥85% of participants complete the study?	Y	Y	Y	Y
9. Was there a ≤15% difference in completion rates in the study's groups?	Y	Y	Y	Y
10. Were all of the study's groups concurrently treated?	Y	Y	Y	Y
11. Was compliance with treatment ≥85% in both of the study's groups?	NR	NR	NR	NR
12. Was there concealment of allocation?	N	N	NR	NR
13. Were outcome assessors blinded to the group to which the participants were assigned?	Y	Y	N	Y
14. Was the outcome measure of interest objective and was it objectively measured?	N	N	N	N
15. Were the same instruments used to measure the outcomes in all of the study's groups?	Y	Y	Y	Y
16. Was the instrument used to measure the outcome standard?	Y	Y	Y	Y
17. Were the follow-up times in all the study's relevant groups approximately equal?	Y	Y	Y	Y
18. Was the same treatment given to all the participants enrolled in the experimental group?	Y	Y	Y	Y
19. Was the same treatment given to all participants enrolled in the control group?	N	N	Y	Y
20. Were all of the study's groups treated at the same center?	Y	Y	Y	Y

ECRI Institute's Controlled Trial Instrument	Problem Behavior			
	Remington et al.(66)[1]	Eikeseth et al.(67) & Eikeseth et al.(68)	Sallows & Graupner(71)	Smith et al.(72)
21. Was the treatment provider's adherence to the intervention protocol (treatment fidelity) assessed?	NR	NR	Y	Y
22. Was the funding for this study derived from a source that does not have a financial or proprietary interest in its results?	Y	Y	Y	Y
Overall Quality Score	**6.4**	**6.6**	**8.6**	**9.1**

[1] Study authors reported between group differences on chronological age. However, subsequent analysis conducted by the authors using age as a covariate indicated that age did not have an effect on any of the outcomes of interest. Thus, we considered the study groups to be comparable.

NR Not reported
N No
Y Yes

Table 25. Quality Assessment for Achievement/Development

ECRI Institute's Controlled Trial Instrument	Achievement/Development				
	Zachor et al.(69)[1,2]	Cohen, et al.(70)[2,3]	Sallows & Graupner(71)[1]	Smith et al.(72)[1]	Smith et al.(72)[3]
1. Were participants randomly assigned to the study's groups?	N	N	Y	Y	Y
2. Did the study employ stochastic randomization?	N	N	Y	Y	Y
3. Were methods used to make the participants in the study's groups comparable?	Y	Y	Y	Y	Y
4. Were subjects assigned to groups based on factors other than individual or provider preference?	N	N	Y	Y	Y
5. Were characteristics of participants in the different study groups comparable at the time they were assigned to groups?	Y	Y	Y	Y	Y
6. Did participants in the different study groups have similar levels of performance on all of the outcome variables at baseline?	Y	Y	Y	Y	Y
7. Was the comparison of interest prospectively planned?	Y	Y	Y	Y	Y
8. Did ≥85% of participants complete the study?	Y	Y	Y	Y	Y
9. Was there a ≤15% difference in completion rates in the study's groups?	Y	Y	Y	Y	Y
10. Were all of the study's groups concurrently treated?	Y	Y	Y	Y	Y
11. Was compliance with treatment ≥85% in both of the study's groups?	NR	NR	NR	NR	NR
12. Was there concealment of allocation?	N	N	NR	NR	NR
13. Were outcome assessors blinded to the group to which the participants were assigned?	N	Y	N	Y	Y
14. Was the outcome measure of interest objective and was it objectively measured?	N	Y	N	N	Y
15. Were the same instruments used to measure the outcomes in all of the study's groups?	Y	Y	Y	Y	Y
16. Was the instrument used to measure the outcome standard?	Y	Y	Y	Y	Y
17. Were the follow-up times in all the study's relevant groups approximately equal?	Y	Y	Y	Y	Y
18. Was the same treatment given to all the participants enrolled in the experimental group?	Y	Y	Y	Y	Y
19. Was the same treatment given to all participants enrolled in the control group?	Y	N	Y	Y	Y
20. Were all of the study's groups treated at the same center?	Y	Y	Y	Y	Y

ECRI Institute's Controlled Trial Instrument	Achievement/Development				
	Zachor et al.(69)[1,2]	Cohen, et al.(70)[2,3]	Sallows & Graupner(71)[1]	Smith et al.(72)[1]	Smith et al.(72)[3]
21. Was the treatment provider's adherence to the intervention protocol (treatment fidelity) assessed?	NR	Y	Y	Y	Y
22. Was the funding for this study derived from a source that does not have a financial or proprietary interest in its results?	Y	Y	Y	Y	Y
Overall Quality Score	**6.8**	**7.5**	**9.1**	**9.1**	**9.5**

[1] As measured by scores on standardized achievement tests
[2] Study authors reported between group differences on chronological age. However, subsequent analysis conducted by the authors using age as a covariate indicated that age did not have an effect on any of the outcomes of interest. Thus, we considered the study groups to be comparable.
[3] As measured by classroom placement

NR Not reported
N No
Y Yes

Table 26. Quality Assessment for Family Wellbeing

ECRI Institute's Controlled Trial Instrument	Family Wellbeing Remington et al.(66)[1]
1. Were participants randomly assigned to the study's groups?	N
2. Did the study employ stochastic randomization?	N
3. Were methods used to make the participants in the study's groups comparable?	Y
4. Were subjects assigned to groups based on factors other than individual or provider preference?	N
5. Were characteristics of participants in the different study groups comparable at the time they were assigned to groups?	Y
6. Did participants in the different study groups have similar levels of performance on all of the outcome variables at baseline?	Y
7. Was the comparison of interest prospectively planned?	Y
8. Did ≥85% of participants complete the study?	Y
9. Was there a ≤15% difference in completion rates in the study's groups?	Y
10. Were all of the study's groups concurrently treated?	Y
11. Was compliance with treatment ≥85% in both of the study's groups?	NR
12. Was there concealment of allocation?	N
13. Were outcome assessors blinded to the group to which the participants were assigned?	Y
14. Was the outcome measure of interest objective and was it objectively measured?	N
15. Were the same instruments used to measure the outcomes in all of the study's groups?	Y
16. Was the instrument used to measure the outcome standard?	Y
17. Were the follow-up times in all the study's relevant groups approximately equal?	Y
18. Was the same treatment given to all the participants enrolled in the experimental group?	Y
19. Was the same treatment given to all participants enrolled in the control group?	N
20. Were all of the study's groups treated at the same center?	Y
21. Was the treatment provider's adherence to the intervention protocol (treatment fidelity) assessed?	NR
22. Was the funding for this study derived from a source that does not have a financial or proprietary interest in its results?	Y
Overall Quality Score	**6.4**

[1] Study authors reported between group differences on chronological age. However, subsequent analysis conducted by the authors using age as a covariate indicated that age did not have an effect on any of the outcomes of interest. Thus, we considered the study groups to be comparable.

NR Not reported
N No
Y Yes

Table 27. Median Quality of Studies Addressing Key Questions

Study	Cognitive and Language Status	Adaptive Behavior	Problem Behaviors	Classroom Placement	Achievement Status[1]	Family Wellbeing
Key Question 1						
Remington et al. (2007)(103)	6.8	6.4	6.4	NR	NR	6.4
Cohen et al. (2005)(65)	7.0	7.0	NR	7.5	NR	NR
Howard et al. (2005)(65)	6.4	6.4	NR	NR	NR	NR
Median Score	**6.8**	**6.4**	**6.4**	**7.5**	**NR**	**6.4**
Key Question 2						
Eikeseth et al. (2007)(67) & Eikeseth et al. (2002)(68)	6.6	6.6	6.6	NR	NR	NR
Zachor et al. (2006)(70)	6.4	NR	NR	NR	6.8	NR
Howard et al. (2005)(65)	6.4	6.4	NR	NR	NR	NR
Median Score	**6.4**	**6.5**	**6.6**	**NR**	**6.8**	**NR**
Key Question 3						
Sallows & Graupner (2005)(71)	8.2	8.6	8.6	8.6	9.1	NR
Smith et al. (2000)(72)	9.1	9.1	9.1	9.1	9.5	NR
Median Score	**8.6**	**8.8**	**8.8**	**8.8**	**9.3**	**NR**

[1] Measured as scores on standardized achievement tests, change in diagnostic status, or change in severity of symptoms

NR Not reported

Appendix E. Patient Characteristic Tables

Table 28. Participant Eligibility Criteria for Included Studies

Study	Inclusion criteria	Exclusion criteria
Remington et al. 92007)(66)	Criteria included: 1) all children met diagnosis of autistic disorder based on the ADI-R, 2) between 30 and 42 months of age, 3) free of any other chronic or serious medical condition, and 4) lived in the family home.	NR
Eikeseth et al. (2007)(67) & Eikeseth et al. (2002)(68)[1]	Criteria included: 1) diagnosis of childhood autism according to the ICD-10 criteria and ADI-R, 2) chronological age at the time of intake between 4 and 7 years, 3) deviation IQ of 50 or above on the WPPSI-R or ratio IQ of 50 or above on the Bayley Scales of Infant Development-Revised, and 4) absence of major medical conditions other than autism.	NR
Cohen et al. (2006)(70)	Criteria included: 1) primary diagnosis of autistic disorder or PDD-NOS confirmed by the ADI-R, 2) pretreatment IQ above 35 on the Bayley Scales of Infant Development, 3) chronological age between 18 and 42 months at diagnosis and 48 months at treatment onset, 4) no severe medical limitations, 5) residence within 60 km of treatment agency, 6) no more than 400 hours of previous behavioral intervention, and 7) parents agreement to actively participate	NR
Zachor et al. (2006)(69)	All children met the established criteria for autistic disorder/PDD-NOS according to DSM-IV criteria.	Children with identified medical abnormalities (e.g., seizures, hearing deficiencies) were excluded.
Howard et al. (2005)(65)	Criteria included: 1) diagnosis of AD or PDD-NOS according to DSM-IV criteria before age 48 months, 2) entry into an intervention program before 48 months of age, 3) English as the primary language spoken in home, 4) no significant medical conditions other than AD or PDD-NOS, and 5) no prior treatment of more than 100 hours.	NR
Sallows & Graupner (2005)(71)	Criteria included: 1) age at intake between 24 and 42 months, 2) ratio estimate (mental age/chronological age) of the Mental Developmental Index of 35 or higher, 3) neurologically within normal limits 4) a diagnosis of autistic disorder that met DSM-IV and ADI-R criteria.	NR
Smith et al. (2000)(72)	Criteria included: 1) chronological age between 18 and 42 months, 2) residence within a one hour drive of treatment site, 3) IQ ratio between 35 and 75, d) diagnosis of autistic disorder or PPD-NOS, and e) absence of major medical problems.	NR

[1] Same patient population

ADI-R Autism Diagnostic Interview-Revised
ASD Autism spectrum disorder
PDD-NOS Pervasive Developmental Disorder-Not Otherwise Specified
NR Not reported
WPPI-R Wechsler Preschool and Primary Scale of Intelligence-Revised

Table 29. Baseline Characteristics of Children in Included Studies

Study	Group	Number of Children	Pretreatment Mean Age in Months (SD)	Primary Diagnosis (Number of Children)	Gender (Number Boys)	Previous Intervention (Number of Children)	Family Socioeconomic Status (Number of Children)	Parent Education
Remington et al. (2007)(66)[1]	EIBI	23	35.7 (4.0)	AD	NR	NR	NR	Mean (SD) with University education: mothers 10 (43); fathers 10 (50)
	Standard care	21	38.4 (4.4)	AD	NR	NR	NR	Mean (SD) with University education: mothers 4 (19); fathers 9 (45)
Eikeseth et al. (2007)(67) & Eikeseth et al. (2002)(68)[2]	IBT	13	66.31 (11.31)	AD	8 (62%)	NR	NR	NR
	Intensive "eclectic" special education	12	65.00 (10.95)	AD	12 (100%)	NR	NR	NR
Cohen et al. (2006)(70)	EIBI	21	30.2 (5.8)	AD-20 PDD-NOS- 1	18 (86%)	NR	NR	Mother mean years 15.3 (SD 2.9) Father mean years 15.8 (SD 2.9)
	Standard care	21	33.2 (3.7)	AD- 15 PDD-NOS-6	17 (81%)	NR	NR	Mother mean years 13.1 (SD 1.6) Father mean years 11.8 (SD 2.3)
Zachor et al. (2006)(69)[3]	ABA	20	27.7 (range 22 to 34 months)	AD or PDD-NOS	19 (95%)	NR	NR	NR
	Eclectic	19	28.8 (range 23 to 33 months)	AD or PDD-NOS	18 (95%)	NR	NR	NR

Study	Group	Number of Children	Pretreatment Mean Age in Months (SD)	Primary Diagnosis (Number of Children)	Gender (Number Boys)	Previous Intervention (Number of Children)	Family Socioeconomic Status (Number of Children)	Parent Education
Howard et al. (2005)(65)[4]	EIBI	29	30.86 (5.16)	24 AD, 5 PDD-NOS	25 (86%)	NR	NR	Mean years mother 14.10 (SD 2.34), father 14.61 (SD 2.77)
	Intensive "eclectic" intervention	16	37.44 (5.68)	12 AD, 4 PDD-NOS	13 (81%)	NR	NR	Mean years mother 13.00 (SD 1.83), father 13.13 (SD 2.56)
	Non-intensive early intervention program	16	34.56 (6.53)	9 AD, 7 PDD-NOS	16 (100%)	NR	NR	Mean years mother 13.00 (SD 1.41), father 13.00 (SD 1.81)
Sallows and Graupner (2005)(71)	Center-directed UCLA (Lovaas) model	13	33.23 (3.89)	AD	11 (85%)	NR	Median Income: $62,000 (range 35 to 100+)	Number who completed college 9 of 12 mothers and 10 of 12 fathers
	Parent-directed UCLA (Lovaas) model	10	34.20 (5.06)	AD	8 (80%)	NR	Median Income: $59,000 (range 30 to 100+)	Number who completed college 9 of 10 mothers and 6 of 9 fathers

Study	Group	Number of Children	Pretreatment Mean Age in Months (SD)	Primary Diagnosis (Number of Children)	Gender (Number Boys)	Previous Intervention (Number of Children)	Family Socioeconomic Status (Number of Children)	Parent Education
Smith et al. (2000)(72)	Intensive UCLA (Lovaas) Model	15	36.07 (6.0)	7 AD and 8 PDD-NOS	12 (80)	NR	Median income: $40 to $50,000 (range <10 to 100+)	Median years: mother 12 (range 10 to 16 +), father 13 (range <6 to 16+)
	Parent-directed UCLA (Lovaas) Model	13	35.77 (5.37)	7 AD and 6 PDD-NOS	11 (84)	NR	Median income: $40 to $50,000 (range <10 to 100+)	Median years: mother 15 (range 12 to 16 +), father 15 (range 12 to 16+)

[1] The authors of the Remington et al. study indicated that the comparison group was on average three months older than the experimental group. However, when the authors explored age as covariate in the main analysis, they found that this difference had no effect on any of the outcomes.

[2] Same patient population

[3] The authors of the Zachor et al. study indicated that analysis of the background data on the fathers' and mothers' education of participants from both groups did not reveal significant differences.

[4] In Howard et al., children in the EIBI group were statistically younger at intake (pretreatment) than children in the other two groups. However, when the authors explored age as a covariate in a regression analysis, they found that this difference had no effect on any of the outcomes.

ABA	Applied Behavior Analysis
AD	Autistic Disorder
ASD	Autism Spectrum Disorder
DD	Developmental Delay
EIBI	Early Intensive behavioral intervention
EIC	Early Intervention Center
IBT	Intensive behavioral treatment
LD	Language Delay
NR	Not reported
PDD-NOS	Pervasive Developmental Disorder-Not Otherwise Specified
SD	Standard deviation
UCLA	University of California as Los Angles (where the Lovaas Young Children with Autism Project is located)

Appendix F. Treatment Characteristics and Individual Study Results of Studies Addressing Key Question 1

Table 30. Treatment Characteristics of Studies Addressing Key Question 1

Study	Intervention	Theoretical orientation	Primary Setting	Method of Instruction	Hours per Week	Therapists	Supervision	Total Duration of Treatment (or study)	Other/Concomitant Treatment
Remington et al. (2007)(66)	Early Intensive Behavioral Intervention	ABA	Home	One-to-one teaching	Average of 25.6 hours/week of ABA	Multidisciplinary tutor team of 3 to 5 therapist trained in ABA and parents	All therapists including parents supervised by experienced consultant	2 years	At baseline no child was attending school, 15 (65%) speech therapy, 11 (48%) dietary restrictions, 1 (4%) medications, 6 (26%), high dose vitamins, and 5 (22%) homeopathic intervention. At 24 months, 17 (74%) children were attending mainstream school for an average of 13.28 hours/week, 6 (26%) speech therapy, 4 (17%) PEC, 8 (35%) sign language or Makaton Communication System, 12 (52%) dietary restrictions, medications 1 (4%), high dose vitamins 7 (30%), and homeopathic intervention 1 (4%)

Study	Intervention	Theoretical orientation	Primary Setting	Method of Instruction	Hours per Week	Therapists	Supervision	Total Duration of Treatment (or study)	Other/Concomitant Treatment
	Standard care (included a variety of publically funded interventions)	14 (67%) recievied speech therapy, 11 (52%) received TEACCH, 16 (76%) PEC, and 10 (48%) sign language or Makaton communication	Various settings-home, school, and clinics	None of the interventions were delivered one-to-one or intensively	NR	Various professionals and parents-no specifics reported	NR	2 years	At baseline no child was attending school, 12 (57%) speech therapy, 3 (14%) dietary restrictions, 1 (5%) medication, 0 high-dose vitamins, and 5(24%) homeophathic intervention. At 24 months 10 (48%) children in mainstream school for 22.3 hours/week, 11 (52%) special needs school for 13.6 hours/week, 10 (48%) speech therapy, 6 (29%) dietary interventions, 4 (19%) prescription medication, 1 (5%) high dose vitamins, and 1 (5%) homeopathic interventions

Study	Intervention	Theoretical orientation	Primary Setting	Method of Instruction	Hours per Week	Therapists	Supervision	Total Duration of Treatment (or study)	Other/Concomitant Treatment
Cohen et al. (2006)(70)	EIBI	ABA (specifically the Lovaas UCLA treatment model)	Home	In-home one-to-one instruction, peer play training, and regular education classroom inclusion when appropriate	First year: 35 to 40 hours per week of home instruction; second year: 26 to 31 hours of home instruction, 3 to 5 hours of peer play, and 6 to 9 hours of preschool; during third year in-home training hours decreased and preschool hours increased	1 to 3 community recruited tutors trained by staff who participated in a 3 to 4 month internship at UCLA and parents who participated in a 12 to 18 hour training workshop over 2 to 3 days and continued with weekly trainings.	Consultants from UCLA made on-site visits 2 to 4 times per year and monitored treatment adherence. Each child's intervention was supervised by a UCLA-trained individual who held a Master's degree in Psychology and was board certified in ABA	3 years	NR
	Majority of children (n = 17) were enrolled in a public school Special Day Class (SDC). Other children received a mix of services.	Mixed (behavioral and develop-mental)	Classroom for most children	For children enrolled in the SDC, the child teacher ratio varied from 1:1 to 3:1.	Classes operated 3 to 5 days per week for up to 5 hours per day.	Special education teachers	NR	3 years	Most children received other services such as speech, occupational, and behavioral therapy 0 to 5 hours per week.
Howard et al. (2005)(65)	EIBT	ABA	Mixed: home, school, and community	Mostly one-to-one instruction	25 to 30 hours per week for children <3 yrs and 35 to 40 hours per week for children >3 yrs	4 to 5 instructional assistants trained a supervised by staff with a Masters' degrees in psychology or special education and supervisory experience in ABA spent 6 to 9 hours per week with each child. Parents who received training in ABA assisted.	Supervised under the direction of Board Certified Behavior Analyst	1.2 years	NR

Study	Intervention	Theoretical orientation	Primary Setting	Method of Instruction	Hours per Week	Therapists	Supervision	Total Duration of Treatment (or study)	Other/Concomitant Treatment
	Generic education programming (GP)	No specific orientation. Authors indicate that children received developmentally appropriate educational services in local special education classrooms	Classroom	One to six adult: child ratio	15 hours of intervention per week	Special education teachers and/or certified speech and pathologists	NR	1.2 years	13 of 16 children in this group received individual or small group speech and language therapy on to two times per week.

ABA Applied behavior analysis
EIBI Early intensive behavioral intervention
NR Not reported
PEC Picture Exchange Communication System
TEACCH Treatment and Education of Autistic and Related Communication Handicapped Children
UCLA University of California at Los Angles

Table 31. Intellectual/Cognitive Status of Studies Addressing Key Question 1

Study	Instrument	Group	Number of Children	Baseline Score Mean (SD)	First Follow-up Mean (SD)	First Follow-up Effect Size Estimate SMD (95% CI) and p-value[1,2]	Last Follow-up Mean (SD)	Last Follow-up Effect Size Estimate SMD (95% CI) and p-value[1,2]
Remington et al. (2007)(66)	Stanford-Binet & Bayley Scales	EIBI	23	61.43 (16.43)	68.78 (20.49) FU-1 year	0.573 (-0.031 to 1.177), p = 0.063	73.48 (27.28) FU-2 years	0.594 (-0.011 to 1.198), p = 0.054
		SC	21	62.33 (16.64)	58.90 (20.45) FU 1 year		60.14 (27.76) FU-2 years	
Cohen et al. (2006)(70)[3]	Bayley Scales of Infant Development (II)	EIBI	21	61.60 (16.40)	77.00 (21.97) FU- 1 year	0.489 (-0.140 to 1.119), p = 0.128	87.00 (24.17) FU-3 years	0.598 (-0.036 to 1.236), p = 0.065
		SC	19	59.40 (14.70)	66.00 (16.60) FU-1 year		73.00 (19.71) FU-3 years	
Howard et al. (2005)(65)[4]	Bayley Scales of Infant Development (II) and WPPSI-R	EIBI	28	58.54 (18.15)	89.88 (20.87) FU-1.2 years	1.234 (0.568 to 1.901), p <0.001	NR	NR
		SC	16	59.88 (14.85)	68.81 (15.32) FU-1.2 years		NR	

[1] All effect size estimates were calculated by ECRI Institute unless otherwise indicated using the standardized mean difference formula. Positive values indicate better outcomes for experimental group, which is listed first under the group column.

[2] Baseline to follow-up effect size estimates were calculated using a correlation correction estimate of 0.50 to account for time. Positive values indicate better outcomes.

[3] Means and confidence intervals were abstracted from figures presented in the study and may not be exact. Confidence intervals were converted to standard deviation units by ECRI Institute.

[4] The Bayley scale was used at baseline to measure overall IQ for most children (n = 42), and to account for chronological age the WPPSI was administered at follow-up.

CI Confidence interval
EIBI Early Intensive Behavioral Intervention
FU Follow-up
SC Standard care
SD Standard deviation
SMD Standardized mean difference
WPPSI-R Wechsler Primary Preschool Scales of Intelligence-Revised

Table 32. Expressive and Receptive Language Skills of Studies Addressing Key Question 1

Study	Instrument	Group	Number of Children	Baseline Score Mean (SD)	First Follow-up Score Mean (SD)	First Follow-up Effect Size Estimate SMD (95% CI) and p-value[1,2]	Last Follow-up Score Mean (SD)	Last Follow-up Effect Size Estimate SMD (95% CI) and p-value[1,2]
Cohen et al. (2006)(70)[3]	Reynell Receptive	EIBI	21	52.00 (14.28)	65.00 (24.16) FU-1 year	0.769 (0.142 to 1.396), p = 0.016	72.00 (19.77) FU-3 years	0.550 (-0.066 to 1.167), p = 0.080
		SC	21	53.00 (15.38)	52.00 (14.27) FU-1 year		62.00 (25.26) FU-3 years	
	Reynell Expressive	EIBI	21	53.00 (14.23)	64.00 (28.56) FU-1 year	0.634 (0.015 to 1.254), p = 0.002	78.00 (28.56) FU-3 years	0.396 (-0.215 to 1.007), p = 0.204
		SC	21	51.00 (13.18)	49.00 (16.48) FU-1 year		66.00 (29.66) FU-3 years	
Howard et al. (2005)(65)	Reynell Receptive	EIBI	25	52.16 (18.44)	71.31 (22.72) FU-1.2 years	0.989 (0.283 to 1.695) p = 0.006	NR	NR
		SC	13	49.00 (13.61)	49.21 (16.08) FU-1.2 years		NR	
	Reynell Expressive	EIBI	25	51.88 (12.91)	70.46 (22.88) FU-1.2 years	1.162 (0.442 to 1.881), p = 0.002	NR	NR
		SC	13	48.77 (11.61)	46.79 (12.81) FU-1.2 years		NR	

[1] All effect size estimates were calculated by ECRI Institute unless otherwise indicated using the standardized mean difference formula. Positive values indicate better outcomes for experimental group, which is listed first under the group column.

[2] Baseline to follow-up effect size estimates were calculated using a correlation correction estimate of 0.50 to account for time. Positive values indicate better outcomes.

[3] Means and confidence intervals were abstracted from figures presented in the study and may not be exact. Confidence intervals were converted to standard deviation units by ECRI Institute.

CI Confidence interval
EIBI Early Intensive Behavioral Intervention
FU Follow-up
SC Standard care
SD Standard deviation
SMD Standardized mean difference

Table 33. Frequencies of Children by Group Achieving a Score on the Reynell Verbal Comprehensive Achievement and Expressive Language Scale

Study	Instrument	Group	Number of Children	Baseline	1 Year Follow-up Number of Children (%)	2 Year Follow-up Number of Children (%)
Remington et al. (2007)(66)	Reynell Verbal Comprehensive	EIBI	23	4 (17%)	19 (83%)	21 (91%)
		SC	21	3 (14%)	11 (52%)	11 (52%)
	Reynell Expressive Language	EIBI	23	2 (8.7%)	17 (74%)	21 (91%)
		SC	21	1 (4.8)	8 (38%)	10 (48%)

Note: The authors reported that some children were unable to obtain a score on the Reynell Developmental Language Scales, so raw data were incomplete and the authors chose to analyze the frequency of children who obtained a score.

EIBI Early Intensive Behavioral Intervention
SC Standard care

Table 34. Non-verbal Communication Skills of Studies Addressing Key Question 1

Study	Instrument	Group	Number of Children	Baseline Score Mean (SD)	First Follow-up Mean (SD)	First Follow Effect Size Estimate SMD (95% CI) and p-value[1,2]	Last Follow-up Mean (SD)	Last Follow-up Effect Size Estimate SMD (95% CI) and p-value[1,2]
Remington et al. (2007)(66)	Early Social Communication Scales (Initiation)	EIBI	21	3.33 (4.40)	7.71 (7.52) FU-1 year	0.259 (-0.394 to 0.912), p = 0.437	11.76 (9.41) FU-2years	0.086 (-0.564 to 0.737), p = 0.795
		SC	16	3.63 (4.92)	6.19 (8.79) FU-1 year		11.19 (13.86) FU-2years	
	Early Social Communication Scales (Responding)	EIBI	21	5.29 (3.62)	8.95 (4.18) FU-1 year	0.578 (-0.086 to 1.241), p = 0.088	11.29 (3.47) FU-2 years	0.469 (-0.190 to 1.128), p = 0.163
		SC	16	5.94 (3.91)	7.13 (5.21) FU-1 year		10.06 (4.99) FU-2 years	
Cohen et al. (2006)(70)[3]	Merrill-Palmer Scale	EIBI	21	82.40 (16.40)	93.00 (24.16) FU-1 year	0.273 (-0.342 to 0.88), p = 0.384	95.00 (23.07) FU-3 years	0.000 (-0.612 to 0.612), p = 1.000
		SC	20	73.40 (11.90)	79.00 (16.02) FU-1 year		86.00 (24.57) FU-3 years	
Howard et al. (2005)(65)	Merrill-Palmer Scale	EIBI	21	80.14 (11.86)	101.67 (19.14) FU-1.2 years	1.035 (0.301 to 1.770), p = 0.006	NR	NR
		SC	13	77.69 (12.33)	82.53 (16.76) FU-1.2 years		NR	

[1] All effect size estimates were calculated by ECRI Institute unless otherwise indicated using the standardized mean difference formula. Positive values indicate better outcomes for experimental group, which is listed first under the group column.

[2] Baseline to follow-up effect size estimates were calculated using a correlation correction estimate of 0.50 to account for time.

[3] Means and confidence intervals were abstracted from figures presented in the study and may not be exact. Confidence intervals were converted to standard deviation units by ECRI Institute.

CI Confidence interval
EIBI Early Intensive Behavioral Intervention
FU Follow-up
SC Standard care
SD Standard deviation

Table 35. Adaptive Behavior of Studies Addressing Key Question 1

Study	Instrument	Group	Number of Children	Baseline Score Mean (SD)	First Follow-up Mean (SD)	First Follow-up Effect Size Estimate SMD (95% CI) and p-value[1,2]	Last Follow-up Mean (SD)	Last Follow-up Effect Size Estimate SMD (95% CI) and p-value[1,2]
Remington et al. (2007)(66)	VABS-Composite Score	EIBI	23	114.78 (26.89)	169.70 (49.07) FU-1 year	0.549 (-0.053 to 1.152), p = 0.074	202.83 (61.98) FU-2 years	0.357 (-0.239 to 0.954), p = 0.240
		SC	21	113.57 (29.78)	145.76 (45.56) FU-1 year		182.86 (58.89) FU-2 years	
	VABS- Communication	EIBI	23	23.52 (11.35)	42.83 (18.25) FU- 1 year	0.406 (-0.191 to 1.004), p = 0.183	54.74 (24.43) FU-2 years	0.322 (-0.273 to 0.918), p = 0.289
		SC	21	21.62 (10.81)	34.62 (17.17) FU-1 year		46.00 (24.51) FU-2 years	
	VABS-Daily Living	EIBI	23	24.13 (7.49)	39.52 (14.71) FU-1 year	0.414 (-0.184 to 1.012), p = 0.175	50.22 (16.46) FU-2 years	0.471 (-0.129 to 1.070), p = 0.124
		SC	21	25.43 (10.56)	35.52 (14.34) FU- 1 year		44.67 (16.99) FU-2 years	
	VABS-Socialization	EIBI	23	29.57 (6.65)	38.52 (12.57) FU-1 year	0.386 (-0.211 to 0.983), p = 0.205	43.52 (15.94) FU-2 years	0.057 (-0.534 to 0.649), p = 0.850
		SC	21	28.29 (7.48)	33.14 (11.77) FU-1 year		41.48 (14.52) FU-2 years	
	VABS-Motor Skills	EIBI	23	37.57 (6.37)	48.83 (6.84) FU-1 year	0.717 (0.107 to 1.327), p = 0.021	54.35 (9.12) FU-2 years	0.613 (0.008 to 1.218), p = 0.047
		SC	21	38.24 (7.06)	44.48 (7.70) FU-1 year		50.71 (8.21) FU-2 years	

Study	Instrument	Group	Number of Children	Baseline Score Mean (SD)	First Follow-up Mean (SD)	First Follow-up Effect Size Estimate SMD (95% CI) and p-value[1,2]	Last Follow-up Mean (SD)	Last Follow-up Effect Size Estimate SMD (95% CI) and p-value[1,2]
Cohen et al. (2006)(70)[3]	VABS-Composite Score	EIBI	21	69.8 (8.10)	82.00 (12.08) FU-1 year	1.123 (0.472 to 1.773), p = 0.001	80.00 (20.87) FU-3 years	0.737 (0.112 to 1.362), p = 0.021
		SC	21	70.6 (9.60)	71.00 (10.98) FU-1 year		69.00 (15.38) FU-3 years	
	VABS- Communication	EIBI	21	69.4 (11.8)	80.00 (17.57) FU-1 year	0.725 (0.101 to 1.349), p = 0.023	84.00 (23.07) FU-3 years	0.510 (-0.105 to 1.124), p = 0.104
		SC	21	65.0 (6.80)	66.0 (12.08) FU-1 year		69.00 (24.16) FU-3 years	
	VABS-Daily Living	EIBI	21	73.20 (9.20)	79.00 (14.28) FU-1 year	0.542 (-0.074 to 1.158), p = 0.084	79.00 (21.97) FU-3 years	0.653 (0.033 to 1.274), p = 0.039
		SC	21	72.70 (12.5)	72.00 (9.88) FU-1 year		68.00 (12.08) FU-3 years	
	VABS-Socialization	EIBI	21	70.30 (10.9)	85.00 (12.08) FU-1 year	1.011 (0.369 to 1.654), p = 0.002	87.00 (14.28) FU-3 years	0.824 (0.194 to 1.454), p = 0.010
		SC	21	75.10 (13.0)	77.00 (14.28) FU-1 year		77.00 (25.26) FU-3 years	

Study	Instrument	Group	Number of Children	Baseline Score Mean (SD)	First Follow-up Mean (SD)	First Follow-up Effect Size Estimate SMD (95% CI) and p-value[1,2]	Last Follow-up Mean (SD)	Last Follow-up Effect Size Estimate SMD (95% CI) and p-value[1,2]
Howard et al. (2005)(65)	VABS-Composite Score	EIBI	26	70.46 (11.85)	81.32 (11.14) FU-1.2 year	1.282 (0.558 to 2.006), p = 0.001	NR	NR
		SC	13	71.62 (10.47)	68.25 (9.86) FU-1.2 year		NR	
	VABS- Communication	EIBI	28	66.18 (10.02)	85.44 (14.73) FU-1.2 year	1.309 (0.623 to 1.994), p <0.001	NR	NR
		SC	15	66.20 (8.70)	68.69 (14.18) FU-1.2 year		NR	
	VABS-Daily Living	EIBI	28	70.71 (10.14)	76.56 (11.59)	1.335 (0.633 to 2.037), p <0.001	NR	NR
		SC	14	73.43 (10.39)	65.19 (8.84) FU-1.2 year		NR	
	VABS-Socialization	EIBI	28	72.79 (11.26)	82.08 (11.73) FU-1.2 year	1.185 (0.495 to 1.875), p = 0.001	NR	NR
		SC	14	75.07 (12.09)	70.56 (11.77) FU-1.2 year		NR	
	VABS-Motor Skills	EIBI	28	95.11 (11.70)	98.16 (12.01) FU-1.2 year	0.468 (-0.197 to 1.134), p = 0.168	NR	NR
		SC	13	92.08 (13.84)	89.50 (10.06) FU-1.2 year		NR	

Note: Remington et al. reported raw scores whereas Cohen et al. and Howard et al. reported standardized scores.

[1] All effect size estimates were calculated by ECRI Institute unless otherwise indicated using the standardized mean difference formula. Positive values indicate better outcomes for experimental group, which is listed first under the group column.

[2] Baseline to follow-up effect size estimates were calculated using a correlation correction estimate of 0.50 to account for time. Positive values indicate better outcomes.

[3] Means and confidence intervals were abstracted from figures presented in the study and may not be exact. Confidence intervals were converted to standard deviation units by ECRI Institute.

CI Confidence interval
EIBI Early Intensive Behavioral Intervention
FU Follow-up time
SC Standard care
SD Standard deviation
VABS Vinland Adaptive Behavior Scales Raw Scores

Table 36. Number of Children Achieving Post-treatment Scores within Normal Range on Child-Assessed Outcomes

Study	Outcome	Group	Total Number of Children	Number of Children Reaching Scores in the Normal Range (%) All studies only report at last follow-up
Remington et al. (2007)(66)	IQ	EIBI	23	5 (21%)
		SC	21	3 (14%)
Cohen et al. (2005)(70)	IQ	EIBI	21	12 (57%)
		SC	21	7 (33%)
	Language Comprehension	EIBI	21	8 (38%)
		SC	21	4 (19%)
	Expressive Language	EIBI	21	9 (43%)
		SC	21	6 (29%)
	Adaptive Behavior	EIBI	21	8 (38%)
		SC	21	3 (14%)
	School Placement	EIBI	21	6 (29%)
		SC	21	0 (0.0%)
Howard et al. (2005)(65)	IQ	EIBI	28	13 (46%)
		SC	16	3 (19%)

Note: Both Cohen et al. and Howard et al. defined this outcome as scores on measures of IQ that fell within normal range (85 or higher) at follow-up. The definition used in Remington et al. was somewhat different in that the authors defined normal range as scores falling halfway between the mean baseline IQ for the children in the study (62.0) and the typical, population mean (100), which corresponds to an IQ of 82 or higher. For all other outcomes in the Cohen study, except school placement, change was defined as scores falling within normal range at follow-up (85 or above). The definition of normal range for school placement was placement into a regular integrated classroom without any assistance.

EIBI Early Intensive Behavioral Intervention
SC Standard care

Table 37. Problem Behaviors of Studies Addressing Key Question 1

Study	Instrument	Parent	Group	Number of Parents	Baseline Score Mean (SD)	1 Year Follow-up Mean (SD)	1 Year Follow-up Effect Size Estimate SMD (95% CI), p-value[1]	2 Year Follow-up Mean (SD)	2 Year Follow-up Effect Size Estimate SMD (95% CI), p-value[1]
Remington et al. (2007)(66)	Development Behavior Checklist (Total Score)	Mother	EIBI	23	50.26 (22.75)	45.57 (18.79)	0.258 (-0.336 to 0.852), p = 0.395	44.70 (24.20)	0.071 (-0.521 to 0.663), p = 0.814
			SC	21	67.81 (18.77)	57.71 (22.61)		60. 62 (24.72)	
		Father	EIBI	16	46.67 (22.15)	43.67 (16.28)	0.176 (-0.528 to 0.883), p = 0.622	45.19 (20.94)	0.045 (-0.660 to 0.749), p = 0.901
			SC	15	57.57 (15.67)	58.02 (21.05)		55.20 (19.44)	
	Nisonger Child Behavior Rating Form (Positive Social Behavior)	Mother	EIBI	23	10.57 (4.25)	15.22 (4.09)	0.733 (0.122 to 1.344), p = 0.019	15.30 (4.69)	0.490 (-0.110 to 1.090), p = 0.110
			SC	21	9.29 (3.47)	11.00 (4.10)		11.86 (4.84)	
		Father	EIBI	16	8.94 (3.47)	13.06 (3.04)	0.643 (-0.080 to 1.365), p = 0.081	12.69 (4.06)	0.304 (-0.405 to 1.012), p = 0.401
			SC	15	8.73 (3.67)	10.40 (4.75)		11.20 (5.19)	
	Autism Screening Questionnaire	Mother	EIBI	23	19.26 (4.93)	16.43 (5.56)	0.323 (-0.273 to 0.918), p = 0.288	15.96 (5.63)	0.245 (-0.349 to 0.839), p = 0.419
			SC	21	21.14 (5.47)	20.14 (6.55)		19.29 (7.22)	
		Father	EIBI	16	20.88 (4.54)	18.44 (5.54)	0.345 (-0.365 to 1.054), p = 0.341	19.88 (6.16)	0.096 (-0.609 to 0.800), p = 0.790
			SC	15	21.07 (6.41)	20.73 (7.45)		19.47 (7.46)	

[1] All effect size estimates were calculated by ECRI Institute unless otherwise indicated using the standardized mean difference formula. Baseline to follow-up effect size estimates were calculated using a correlation correction estimate of 0.50 to account for time. Positive values indicate better outcomes for experimental group, which is listed first under the group column.

CI Confidence interval
EIBI Early Intensive Behavioral Intervention
SC Standard care
SD Standard deviation

Table 38. Parent/Family Wellbeing of Studies Addressing Key Question 1

Study	Instrument	Parent	Group	Number of Parents	Baseline Score Mean (SD)	1 Year Follow-up Mean (SD)	1 Year Follow-up Effect Size Estimate SMD (95% CI), p-value[1]	2 Year Follow-up Mean (SD)	2 Year Follow-up Effect Size Estimate SMD (95% CI), p-value[1]
Remington et al. (2007)(66)	QRS-F (Stress)	Mother	EIBI	23	6.43 (4.290	7.48 (4.70)	0.417 (-0.180 to 1.015), p = 0.171	8.52 (2.97)	0.269 (-0.325 to 0.863), p = 0.375
			SC	21	7.24 (4.19)	6.48 (4.08)		8.29 (3.62)	
		Father	EIBI	16	6.81 (4.26)	7.88 (4.27)	0.376 (-0.334 to 1.087), p = 0.299	8.94 (3.62)	0.113 (-0.592 to 0.818), p = 0.753
			SC	15	5.87 (3.19)	5.53 (3.00)		7.60 (2.72)	
	HADS (Anxiety Scale)	Mother	EIBI	23	9.35 (4.21)	10.48 (5.12)	0.638 (0.031 to 1.244), p = 0.039	9.13 (4.53)	0.204 (-0.389 to 0.797), p = 0.501
			SC	21	9.76 (4.87)	7.87 (4.60)		8.62 (4.43)	
		Father	EIBI	16	8.89 (4.76)	8.52 (4.72)	0.134 (-0.571 to 0.840), p = 0.709	8.38 (4.08)	0.169 (-0.537 to 0.875), p = 0.639
			SC	15	7.93 (3.67)	7.00 (3.16)		8.13 (4.10)	
	HADS (Depression Scale)	Mother	EIBI	23	8.13 (4.12)	8.04 (5.80)	0.308 (-0.288 to 0.903), p = 0.311	7.09 (4.97)	0.181 (-0.411 to 0.774), p = 0.549
			SC	21	8.71 (3.68)	7.19 (4.26)		6.90 (3.94)	
		Father	EIBI	16	5.69 (4.42)	6.56 (5.25)	0.634 (-0.088 to 1.355), p = 0.085	7.00 (5.34)	0.557 (-0.161 to 1.275), p = 0.128
			SC	15	7.07 (3.61)	5.27 (2.99)		5.93 (3.83)	

Study	Instrument	Parent	Group	Number of Parents	Baseline Score Mean (SD)	1 Year Follow-up Mean (SD)	1 Year Follow-up Effect Size Estimate SMD (95% CI), p-value[1]	2 Year Follow-up Mean (SD)	2 Year Follow-up Effect Size Estimate SMD (95% CI), p-value[1]
	KIPP-PC	Mother	EIBI	23	127.30 (27.00)	127.39 (23.79)	0.011 (-0.581 to 0.602), p = 0.972	128.00 (19.62)	0.063 (-0.529 to 0.655), p = 0.835
			SC	21	133.10 (19.37)	133.43 (18.23)		132.43 (17.94)	
		Father	EIBI	16	120.94 (20.23)	122.56 (19.70)	0.265 (-0.442 to 0.973), p = 0.463	122.81 (22.47)	0.099 (-0.606 to 0.804), p = 0.782
			SC	15	124.73 (19.66)	131.40 (15.68)		128.53 (9.70)	

[1] All effect size estimates were calculated by ECRI Institute unless otherwise indicated using the standardized mean difference formula. Baseline to follow-up effect size estimates were calculated using a correlation correction estimate of 0.50 to account for time. Positive values indicate better outcomes for experimental group, which is listed first under the group column.

CI Confidence interval
EIBI Early Intensive Behavioral Intervention
HADS Hospital Anxiety and Depression Scale
KIPP-PC Kansas Inventory of Parental Perceptions Positive Contributions scale
QRS-F Questionnaire on Resources and Stress Friedrich Short Form
SC Standard care
SD Standard deviation

Appendix G. Treatment Characteristics and Individual Study Results of Studies Addressing Key Question 2

Table 39. Treatment Characteristics of Studies Addressing Key Question 2

Study	Intervention	Theoretical Orientation	Primary Setting	Method of Instruction	Hours per Week	Therapists	Supervision	Total Duration of Treatment (or study)	Other/Concomitant Treatment
Eikeseth et al. (2007)(67) & Eikeseth et al. (2002)(68)[1]	Behavioral based on Lovaas/UCLA method	ABA	Classroom	One-to-one instruction	Mean hours 28.00 (range 20 to 35)	Teachers, aides and parents	10 hours per week from staff who had a minimum of 1,500 of experience implementing UCLA model	3 years	NR
	Eclectic/ developmental approach	Developmental (deriving mostly from TEACCH and the DIR model)	Classroom	One-to-one instruction	Mean hours 29.08 (range 20 to 41)	Multidisciplinary team of school personnel	2 hours per week of consultation from supervisor trained in methods of treatment	3 years	NR
Zachor et al. (2006)(69)	EIBI	ABA	Center-based	One-to-one instruction	35 hours per week	Skilled behavior therapists	All therapist supervised by trained behavior analyst who designed each child's individual treatment program	1 year	Supervised inclusion program in a regular preschool was added for children who had attained sufficient skills to participate and learn from typically developing children.
	Eclectic/ developmental approach	Developmental (deriving mostly from TEACCH and the DIR model)	Center-based	One-to-one and group instruction	Children received the same number of hours of treatment as the ABA group—35 hours per week.	Special education teachers with experience in autism	NR	1 year	Supervised inclusion program in a regular preschool was added for children who had attained sufficient skills to participate and learn from typically developing children.

Study	Intervention	Theoretical Orientation	Primary Setting	Method of Instruction	Hours per Week	Therapists	Supervision	Total Duration of Treatment (or study)	Other/Concomitant Treatment
Howard et al. (2005)(65)	EIBI	ABA	Mixed: home, school, and community	Mostly one-to-one instruction	25 to 30 hours per week for children <3 yrs and 35 to 40 hours per week for children >3 yrs	4 to 5 instructional assistants trained a supervised by staff with a Masters' degrees in psychology or special education and supervisory experience in ABA spent 6 to 9 hours per week with each child. Parents who received training in ABA assisted.	Supervised under the direction of Board Certified Behavior Analyst	1.2 years	NR
	Eclectic/ developmental approach	Developmental (deriving mostly from TEACCH and the DIR model)	Classroom	Mostly one-to-one instruction	25 to 30 hours of intervention per week	Special education teachers	Teachers received consultation with staff with 1 to 2 years of experience in ABA	1.2 years	7 of 16 children received speech therapy one to two times per week

[1] Same patient population

ABA Applied behavior analysis
DIR Developmental Individual-Difference Relationship
EIBI Early intensive behavioral intervention
NR Not reported
TEACCH Treatment and Education of Autistic and Related Communication Handicapped Children
UCLA University of California in Las Angeles

Table 40. Intellectual/Cognitive Status for Key Question 2

Study	Instrument	Group	Number of Children	Baseline Score Mean (SD)	First Follow-up Mean (SD)	First Follow-up Between Group Effect Size Estimate SMD (95% CI) and p-value[2]	Last Follow-up Mean (SD)	Last Follow-up Between Group Effect Size Estimate SMD (95% CI) and p-value[2]
Eikeseth et al. (2007)(67) & Eikeseth et al. (2002)(68)[1]	Bayley Scales of Infant Development (II) and WPPSI-R	EIBI	13	61.92 (11.31)	79.08 (18.09) FU- 1 year	0.784 (-0.030 to 1.598), p = 0.059	86.9 (25.0) FU-3 years	0.789 (-0.025 to 1.604), p = 0.058
		Eclectic	12	65.17 (14.97)	69.50 (18.38) FU- 1 year		71.9 (28.4) FU-3 years	
Zachor et al. (2006)(69)	Stanford-Binet & Bayley Scales	EIBI	20	76.1 (15.2)	NR	NR	NR	NR
		Eclectic	14	79.6 (17.0)	NR		NR	
Howard et al. (2005)(65)[3]	Bayley Scales of Infant Development (II) and WPPSI-R	EIBI	28	58.54 (18.15)	89.88 (20.87) FU-1.2 years	1.213 (0.549 to 1.878), p <0.001	NR	NR
		Eclectic	16	53.69 (13.50)	62.13 (19.63) FU-1.2 years		NR	

[1] Same participant population

[2] All effect size estimates were calculated by ECRI Institute unless otherwise indicated using the standardized mean difference formula. Baseline to follow-up effect size estimates were calculated using a correlation correction estimate of 0.50 to account for time Positive values indicate better outcomes for experimental group, which is listed first under the group column.

[3] The follow-up period for this study was 14 months.

CI Confidence interval
EIBI Early Intensive Behavioral Intervention
FU Follow-up
SD Standard deviation

Table 41. Language/Communication Skills for Key Question 2

Study	Instrument	Group	Number of Children	Baseline Score Mean (SD)	First Follow-up Mean (SD)	First Follow-up Between group effect size estimate SMD (95% CI) and p-value[3]	Last Follow-up	Last Follow-up Between group effect size estimate SMD (95% CI) and p-value[3]
Eikeseth et al. (2007)(67) & Eikeseth et al. (2002)(68)[1,2]	Reynell Scale-Receptive Language	EIBI	12	49.0 (16.42)	58.47 (17.11)	0.740 (-0.153 to 1.632), p = 0.104	3 years[5]	NR
		Eclectic	9	50.38 (15.46)	47.55 (17.25)			
	Reynell Scale-Expressive Language	EIBI	12	45.12 (13.44)	67.39 (17.81)	1.412 (0.448 to 2.376 0, p = 0.004		NR
		Eclectic	9	51.24 (19.24)	49.00 (18.69)			
Howard et al. (2005)(65)[4]	Merrill-Palmer Scale	EIBI	21	80.14 (11.86)	101.67 (19.14)	0.804 (0.128 to 1.479), p = 0.020	NR	NR
		Eclectic	16	67.44 (16.69)	73.56 (24.94)			
	Reynell Scale-Receptive Language	EIBI	25	52.16 (18.44)	71.31 (22.72)	0.739 (0.091 to 1.387), p = 0.025	NR	NR
		Eclectic	16	45.38 (14.97)	49.93 (19.62)			
	Reynell Scale-Expressive Language	EIBI	25	51.88 (12.91)	70.46 (22.88)	0.730 (0.083 to 1.377), p = 0.027	NR	NR
		Eclectic	16	43.88 (6.69)	47.67 (23.39)			

[1] Same participant population.

[2] First follow-up at 1 year.

[3] All effect size estimates were calculated by ECRI Institute unless otherwise indicated using the standardized mean difference formula. Baseline to follow-up effect size estimates were calculated using a correlation correction estimate of 0.50 to account for time Positive values indicate better outcomes for experimental group, which is listed first under the group column.

[4] The follow-up period for this study was 14 months.

[5] The Reynell was not administered at follow-up in this study.

CI Confidence interval
EIBI Early Intensive Behavioral Intervention
SD Standard deviation

Table 42. Adaptive Behavior of Studies Addressing Key Question 2

Study	Instrument	Group	Number of Children	Baseline Score Mean (SD)	First Follow-up Mean (SD)	First Follow-up Between group effect size estimate SMD (95% CI) and p-value[3]	Last Follow-up Mean (SD)	Last Follow-up Between group effect size estimate SMD (95% CI) and p-value[3]
Eikeseth et al. (2007)(67) & Eikeseth et al. (2002)(68)[1,2]	VABS-Composite Score	EIBI	13	55.77 (8.96)	67.00 (16.30)	0.826 (0.009 to 1.644), p = 0.048	67.9 (17.1)	1.614 (0.711 to 2.517), p <0.001
		Eclectic	12	60.0 (13.20)	60.17 (11.69)		49.5 (13.0)	
	VABS-Communication	EIBI	13	58.23 (9.21)	73.93 (16.55)	1.181 (0.331 to 2.031), p = 0.006	78.5 (22.3)	1.614 (0.711 to 2.517), p = 0.001
		Eclectic	12	63.17 (16.11)	61.58 (13.37)		56.0 (16.3)	
	VABS-Daily Living	EIBI	13	56.92 (9.80)	66.15 (16.55)	0.261 (-0.527 to 1.049), p = 0.516	66.1 (18.1)	0.925 (0.100 to 1.750), p = 0.028
		Eclectic	12	57.00 (15.92)	62.50 (10.97)		50.4 (20.2)	
	VABS-Socialization	EIBI	13	59.92 (7.19)	69.92 (17.26)	0.110 (-0.675 to 0.895), p = 0.784	72.2 (14.4)	1.441 (0.560 to 2.321), p = 0.001
		Eclectic	12	62.17 (10.32)	70.67 (13.66)		58.1 (9.6)	
Howard et al. (2005)(65)[4]	VABS-Composite Score	EIBI	26	70.46 (11.85)	81.32 (11.14)	0.980 (0.323 to 1.637), p = 0.003	NR	NR
		Eclectic	16	69.81 (10.48)	69.25 (12.91)			
	VABS-Communication	EIBI	28	66.18 (10.02)	85.44 (14.73)	1.463 (0.777 to 2.149), p <0.001	NR	NR
		Eclectic	16	63.69 (9.68)	64.13 (14.18)			

Study	Instrument	Group	Number of Children	Baseline Score Mean (SD)	First Follow-up Mean (SD)	First Follow-up Between group effect size estimate SMD (95% CI) and p-value[3]	Last Follow-up Mean (SD)	Last Follow-up Between group effect size estimate SMD (95% CI) and p-value[3]
	VABS-Daily Living	EIBI	28	70.71 (10.14)	76.56 (11.59)	0.348 (-0.271 to 0.966), p = 0.270	NR	NR
		Eclectic	16	68.06 (11.61)	70.00 (11.92)			
	VABS-Socialization	EIBI	28	72.79 (11.26)	82.08 (11.73)	0.726 (0.093 to 1.359), p = 0.024	NR	NR
		Eclectic	16	75.50 (14.25)	75.00 (18.01)			

[1] Same participant population

[2] First follow-up at 1 year and last follow-up at 3 years

[2] All effect size estimates were calculated by ECRI Institute unless otherwise indicated using the standardized mean difference formula. Positive values indicate better outcomes for experimental group, which is listed first under the group column.

[3] All effect size estimates were calculated by ECRI Institute unless otherwise indicated using the standardized mean difference formula. Baseline to follow-up effect size estimates were calculated using a correlation correction estimate of 0.50 to account for time Positive values indicate better outcomes for experimental group, which is listed first under the group column

[4] The follow-up period for this study was 14 months.

CI Confidence interval
EIBI Early Intensive Behavioral Intervention
SD Standard deviation
VABS Vinland Adaptive Behavior Scales

Table 43. Problem Behaviors for Key Question 2

Study	Achenbach Child Behavior Checklist- Behavior Category[1]	Group	Number of Children	Follow-up Mean (SD)	Between Group Effect Size Estimate SMD (95% CI)[2]	p-value
Eikeseth et al. (2007)(67) & Eikeseth et al. (2002)(68)[1]	Withdrawn	EIBI	13	59.4 (6.3)	0.342 (-0.448 to 1.133)	0.369
		Eclectic	12	61.4 (5.3)		
	Somatic	EIBI	13	55.0 (7.0)	0.359 (-0.431 to 1.150)	0.373
		Eclectic	12	58.0 (9.6)		
	Anxious/depressed	EIBI	13	57.8 (6.1)	0.106 (-0.679 to 0.891)	0.791
		Eclectic	12	57.1 (7.1)		
	Social	EIBI	13	62.3 (6.3)	0.864 (0.043 to 1.684)	0.039
		Eclectic	12	67.2 (4.9)		
	Thought	EIBI	13	68.1 (9.6)	0.047 (-0.737 to 0.832)	0.906
		Eclectic	12	68.5 (7.0)		
	Attention	EIBI	13	59.0 (5.4)	0.540 (-0.259 to 1.338)	0.186
		Eclectic	12	62.1 (6.1)		
	Delinquent	EIBI	13	56.0 (5.2)	0.654 (-0.151 to 1.460)	0.111
		Eclectic	12	59.0 (3.8)		
	Aggressive	EIBI	13	57.3 (4.5)	1.407 (0.531 to 2.283)	0.002
		Eclectic	12	63.7 (4.6)		

[1] Last follow-up was 3 years. No baseline data were reported for this outcome.

[2] All effect size estimates were calculated by ECRI Institute unless otherwise indicated using the standardized mean difference formula. Positive values indicate better outcomes for experimental group, which is listed first under the group column. Lower values represent improved behavior.

CI Confidence interval
EIBI Early Intensive Behavioral Intervention
SD Standard deviation

Appendix H. Treatment Characteristics and Individual Study Results of Studies Addressing Key Question 3

Table 44. Treatment Characteristics of Studies Addressing Key Question 3

Study	Intervention	Theoretical orientation	Primary Setting	Method of Instruction	Hours per Week	Therapists	Supervision	Total Duration of Treatment (or study)	Other/Concomitant Treatment
Sallows et al. (2005)(71)	Clinic-directed EIBI	ABA (mostly Lovaas method)	Home	One-to-one instruction	37 to 39 hours	Therapists trained in ABA	6 to 10 hours per/week of in-home supervision from senior therapists and weekly consultation by clinic supervisor	2 years	NR
	Parent-directed EIBI	ABA (mostly Lovaas method)	Home	One-to-one instruction	31 to 32 hours	Parents trained in ABA	6 hours per month of in-home supervision from senior therapist and consultation every 2 months	2 years	NR

Study	Intervention	Theoretical orientation	Primary Setting	Method of Instruction	Hours per Week	Therapists	Supervision	Total Duration of Treatment (or study)	Other/Concomitant Treatment
Smith et al. (2000)(72)	Clinic-directed EIBI	ABA (Lovaas method)	Home and classroom once children gained skills to participate in group instruction and naturalistic setting	Mostly one-to-one instruction	30 hours per week for 2 to 3 years	4 to 6 student therapists working under close supervision of study authors and parents provided 5 hours of treatment per week with student therapists for 3 months.	NR	3 years	NR
	Parent-directed EIBI	ABA (Lovaas method)	Home	Mostly one-to-one	Parents received a total of 5 hours of training per week for 3 to 9 months. Between sessions parents provided 5 hours of instruction to their child per week.	Parents and parent trainers	One hour of supervision per week by senior author of study and further supervision if needed.	3 years	Throughout parent training, children were enrolled in public special education classes for 10 to 15 hours per week.

ABA Applied behavior analysis
EIBI Early intensive behavioral intervention
NR Not reported

Table 45. Intellectual/Cognitive Status for Key Question 3

Study	Instrument	Group	Number of Children	Baseline Score Mean (SD)	Post treatment-Mean (SD)	Post treatment between group effect size estimate SMD (95% CI), p-value[1]	Length of Follow-up
Sallows &Graupner (2005)(71)	Bayley Scales of Infant Development	CD	13	50.85 (10.57)	73.08 (33.08)	0.208 (-0.619 to 1.034), p = 0.622	4 years
		PD	10	52.10 (8.98)	79.60 (21.80)		
Smith et al. (2000)(72)	Stanford-Binet Intelligence Scale or Bayley Scales of Infant Development	CD	15	50.53 (11.18)	66.49 (24.08)	0.875 (0.097 to 1.652), p = 0.027	3 years
		PD	13	50.69 (13.88)	49.67 (19.74)		

[1] All effect size estimates were calculated by ECRI Institute unless otherwise indicated using the standardized mean difference formula. Baseline to follow-up effect size estimates were calculated using a correlation correction estimate of 0.50 to account for time. Positive values indicate better outcomes. Positive values indicate better outcomes for experimental group, which is listed first under the group column.

CD Center directed
CI Confidence intervals
PD Parent directed

Table 46. Language and Communication Skills for Key Question 3

Study	Instrument	Group	Number of Children	Baseline Score Mean (SD)	Post treatment Mean (SD)	Post treatment between group effect size estimate SMD (95% CI), p-value[1]	Length of Follow-up
Sallows &Graupner (2005)(71)	Merrill-Palmer Scale	CD	13	70.58 (16.54)	77.58 (25.24)	0.011 (-0.813 to 0.836), p = 0.978	4 years
		PD	10	82.67 (14.94)	89.44 (18.35)		
Smith et al. (2000)(72)	Merrill-Palmer Scale	CD	15	21.60 (4.49)	64.33 (18.74)	0.857 (0.081 to 1.633), p = 0.030	3 years
		PD	13	21.92 (5.50)	49.17 (21.43)		
Sallows &Graupner (2005)(71)	Reynell Receptive Language	CD	13	38.85 (6.09)	55.85 (36.23)	0.338 (-0.493 to 1.168), p = 0.425	4 years
		PD	10	38.78 (6.44)	65.78 (25.81)		
	Reynell Expressive Language	CD	13	47.92 (6.17)	53.38 (31.91)	0.200 (-0.626 to 1.026), p = 0.635	
		PD	10	48.44 (6.96)	59.22 (25.13)		
Smith et al. (2000)(72)	Reynell Receptive Language	CD	15	13.47 (3.60)	42.87 (22.29)	0.547 (-0.209 to 1.304), p = 0.156	3 years
		PD	13	13.69 (3.73)	33.00 (16.86)		
	Reynell Expressive Language	CD	15	15.13 (0.52)	44.53 (23.48)	0.435 (-0.316 to 1.186), p = 0.256	
		PD	13	16.31 (2.69)	36.23 (21.19)		

[1] All effect size estimates were calculated by ECRI Institute unless otherwise indicated using the standardized mean difference formula. Baseline to follow-up effect size estimates were calculated using a correlation correction estimate of 0.50 to account for time. Positive values indicate better outcomes. Positive values indicate better outcomes for experimental group, which is listed first under the group column.

CD Center directed
CI Confidence intervals
PD Parent directed

Table 47. Adaptive Behaviors Key Question 3

Study	Instrument	Group	Number of Children	Baseline Score Mean (SD)	Post treatment Mean (SD)	Post-treatment between group effect size estimate SMD (95% CI), p-value	Length of Follow-up
Sallows &Graupner (2005)(71)	VABS Composite Score	CD	13	59.54 (5.31)	69.00 (28.04)	0.172 (-0.653 to 0.998), p = 0.682	4 years
		PD	10	60.90 (5.94)	66.70 (14.68)		
	VABS-Communication	CD	13	57.46 (4.97)	73.69 (32.32)	0.073 (-0.752 to 0.898), p = 0.862	4 years
		PD	10	63.20 (5.58)	81.40 (24.33)		
	VABS-Daily Living	CD	13	63.92 (5.53)	66.23 (25.95)	0.120 (-0.705 to 0.945), p = 0.776	4 years
		PD	10	64.20 (3.68)	64.20 (12.42)		
	VABS-Socialization	CD	13	58.38 (6.17)	73.92 (23.49)	0.410 (-0.423 to 1.242), p = 0.335	
		PD	10	60.30 (5.76)	68.90 (10.11)		
Smith et al. (2000)(72)	VABS Composite Score	CD	15	63.44 (9.35)	61.19 (29.72)	0.204 (-0.540 to 0.949), p = 0.591	3 years
		PD	13	65.17 (9.44)	58.50 (16.58)		
	VABS-Communication	CD	15	58.20 (5.56)	67.87 (30.08)	0.478 (-0.275 to 1.231), p = 0.213	
		PD	13	62.00 (6.11)	60.77 (17.26)		
	VABS-Daily Living	CD	15	69.93 (8.37)	62.33 (25.76)	0.001 (-0.742 to 0.744), p = 0.998	3 years
		PD	13	70.62 (11.50)	63.00 (16.97)		
	VABS-Socialization	CD	15	62.40 (7.82)	66.33 (24.78)	0.220 (-0.525 to 0.965), p = 0.563	3 years
		PD	13	69.15 (8.75)	68.92 (16.94)		

[1] All effect size estimates were calculated by ECRI Institute unless otherwise indicated using the standardized mean difference formula. Baseline to follow-up effect size estimates were calculated using a correlation correction estimate of 0.50 to account for time. Positive values indicate better outcomes. Positive values indicate better outcomes for experimental group, which is listed first under the group column.

CD Center directed
CI Confidence intervals
PD Parent directed
VABS Vinland Adaptive Behavior Scales

Table 48. Problem Behaviors for Key Question 3

Study	Instrument	Behavior	Group	Number of Children	Post treatment Mean (SD)	Post-treatment between group effect size estimate SMD (95% CI), p-value[1]	Length of Follow-up
Smith et al. (2000)(72)	Parent Version of Achenbach Child Behavior Checklist	Withdrawal	CD	15	59.33 (10.26)	0.091 (-0.652 to 0.834), p = 0.810	3 years
			PD	13	60.17 (7.81)		
		Somatization	CD	15	56.11 (8.16)	0.000 (-0.743 to 0.743), p = 1.00	
			PD	13	56.11 (8.16)		
		Anxiety/depression	CD	15	52.22 (5.24)	0.850 (0.075 to 1.626), p = 0.032	
			PD	13	59.67 (11.59)		
		Social problems	CD	15	60.11 (13.46)	0.337 (-0.411 to 1.085), p = 0.377	
			PD	13	64.33 (11.34)		
		Thought problems	CD	15	67.11 (10.82)	0.225 (-0.520 to 0.970), p = 0.554	
			PD	13	64.47 (12.74)		
		Attention problems	CD	15	64.78 (10.32)	0.336 (-0.412 to 1.084), p = 0.378	
			PD	13	67.50 (4.18)		
		Delinquency	CD	15	54.67 (9.24)	0.537 (-0.219 to 1.293), p = 0.164	
			PD	13	59.00 (6.42)		
		Aggression	CD	15	56.11 (9.10)	0.366 (-0.383 to 1.115), p = 0.338	
			PD	13	59.67 (10.41)		
	Teacher Version of Achenbach Child Behavior Checklist	Withdrawal	CD	15	61.89 (7.04)	1.154 (0.353 to 1.956), p = 0.005	3 years
			PD	13	55.00 (4.40)		
		Somatization	CD	15	52.33 (4.95)	0.372 (-0.377 to 1.121), p = 0.331	
			PD	13	54.86 (8.47)		
		Anxiety/depression	CD	15	54.22 (5.26)	0.074 (-0.669 to 0.817), p = 0.846	
			PD	13	54.57 (4.08)		

Study	Instrument	Behavior	Group	Number of Children	Post treatment Mean (SD)	Post-treatment between group effect size estimate SMD (95% CI), p-value[1]	Length of Follow-up
		Social problems	CD	15	59.78 (9.59)	-0.264 (-1.010 to 0.482), p = 0.488	
			PD	13	57.43 (8.02)		
		Thought problems	CD	15	64.67 (13.62)	-0.187 (-0.931 to 0.557), p = 0.623	
			PD	13	62.57 (7.55)		
		Attention problems	CD	15	64.89 (12.80)	-0.293 (-1.040 to 0.453), p = 0.441	
			PD	13	61.57 (9.29)		
		Delinquency	CD	15	53.44 (6.39)	0.096 (-0.647 to 0.839), p = 0.800	
			PD	13	54.00 (5.13)		
		Aggression	CD	15	60.00 (10.81)	0.489 (-0.265 to 1.242), p = 0.204	
			PD	13	55.71 (5.53)		

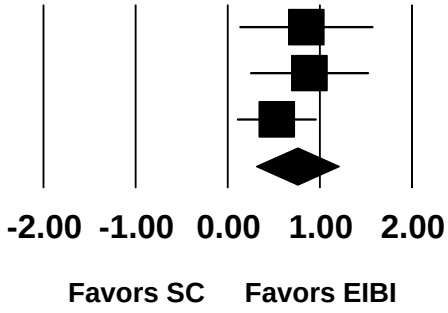
[1] All effect size estimates were calculated by ECRI Institute unless otherwise indicated using the standardized mean difference formula. Positive values indicate better outcomes for experimental group, which is listed first under the group column. Only posttreatment ratings were reported in the study.

CD Center directed
CI Confidence intervals
PD Parent directed

Appendix I. Results of Meta-Analyses and Sensitivity Analysis

Figure 11. Key Question 1: Robustness Test of 1 Year Follow-up IQ Meta-Analysis (One Study Removed)

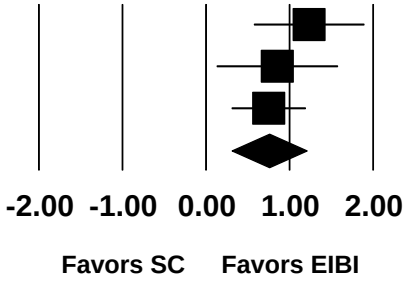
Study name	Statistics with study removed				Std diff in means (95% CI) with study removed
	Point	Lower limit	Upper limit	p-Value	
2007 Remington	0.854	0.124	1.584	0.022	
2006 Cohen	0.888	0.241	1.535	0.007	
2005 Howard	0.533	0.097	0.969	0.017	
Summary ES	0.750	0.302	1.199	0.001	



Random-Effects Meta Analysis

Figure 12. Key Question 1: Robustness Test of 1 Year Follow-up IQ Meta-Analysis (Cumulative by Publication Date)

Study name	Cumulative statistics				Cumulative std diff in means (95% CI)
	Point	Lower limit	Upper limit	p-Value	
2005 Howard	1.234	0.568	1.901	0.000	
2006 Cohen	0.854	0.124	1.584	0.022	
2007 Remington	0.750	0.302	1.199	0.001	
Summary ES	0.750	0.302	1.199	0.001	



Random-Effects Meta Analysis

Figure 13. Key Question 1: Difference in IQ Scores at Last Follow-up

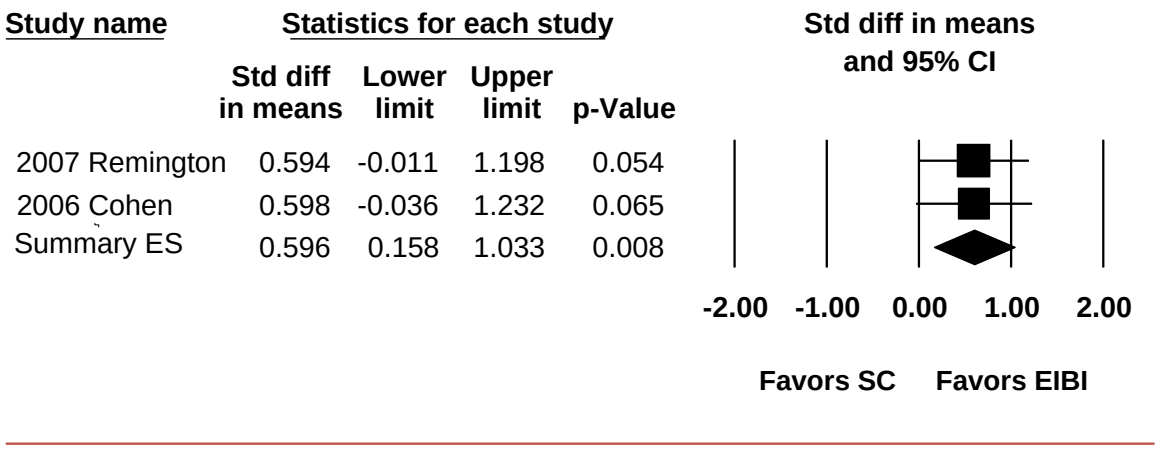


Figure 14. Key Question 1: Robustness Test for of Children Achieving IQ Score of 85 or Above (One Study Removed)

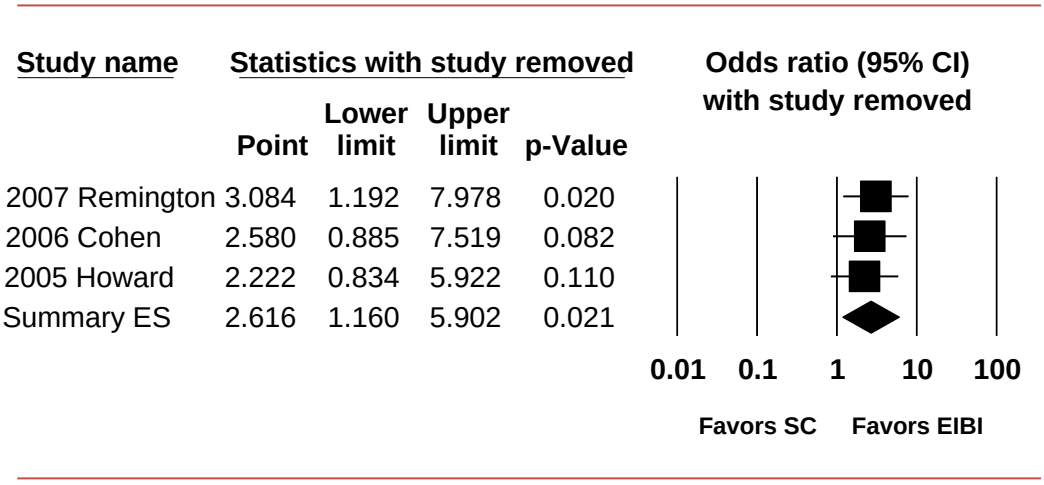
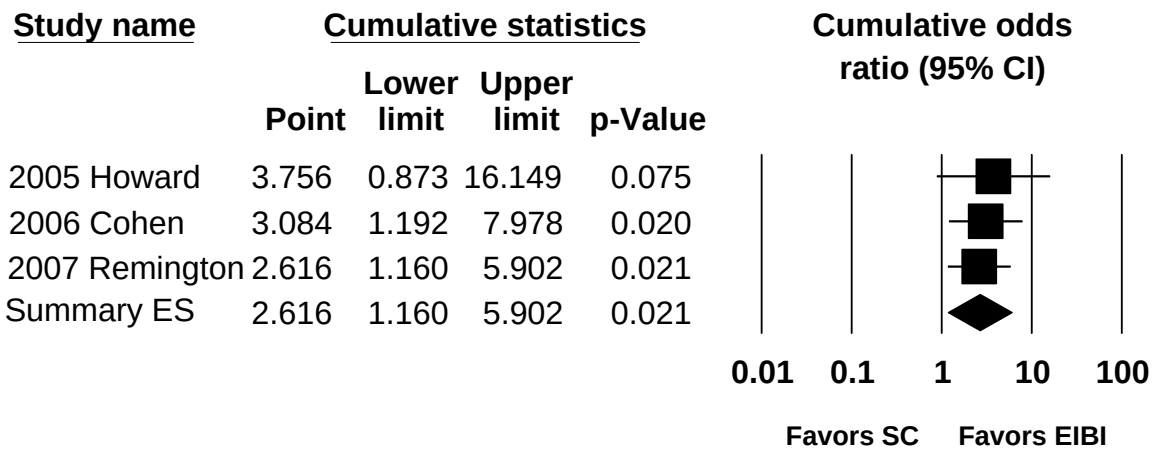
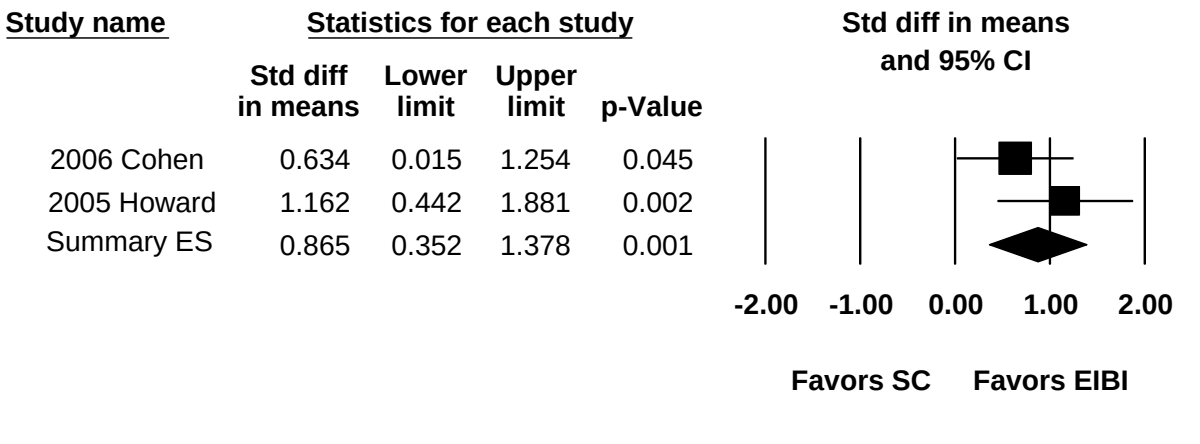


Figure 15. Key Question 1: Robustness Test for of Children Achieving an IQ Score of 85 or Above (Cumulative by Publication Date)



Random-Effects Meta Analysis

Figure 16. Key Question 1: Meta-Analytic Results of Reynell Scales at 1 Year Follow-up for Expressive Language



Random Effects Meta Analysis/I^2=15.58

Figure 17. Key Question 1: Meta-Analytic Results of Reynell Scales at 1 Year for Receptive Language

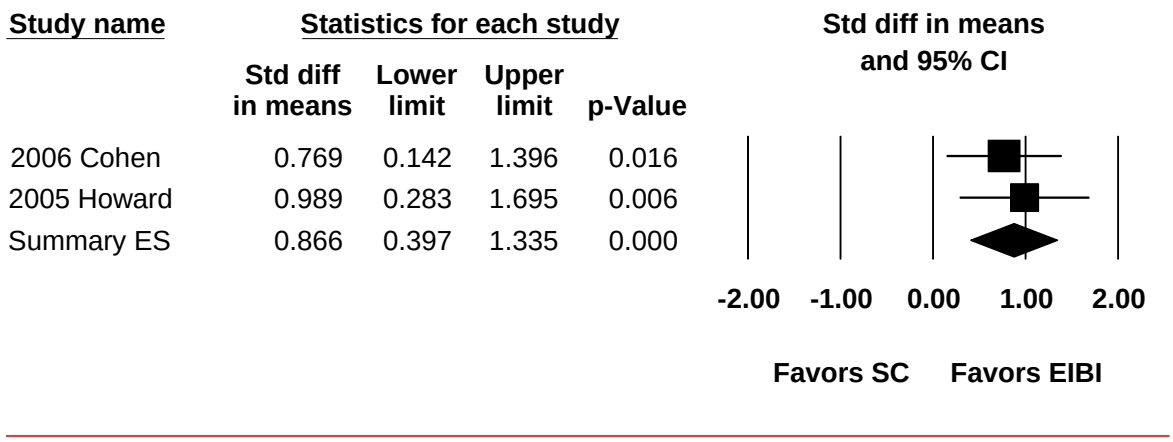


Figure 18. Key Question 1: Meta-Analytic Results of Merrill Palmer Scale of Non-Verbal Communication at 1 Year

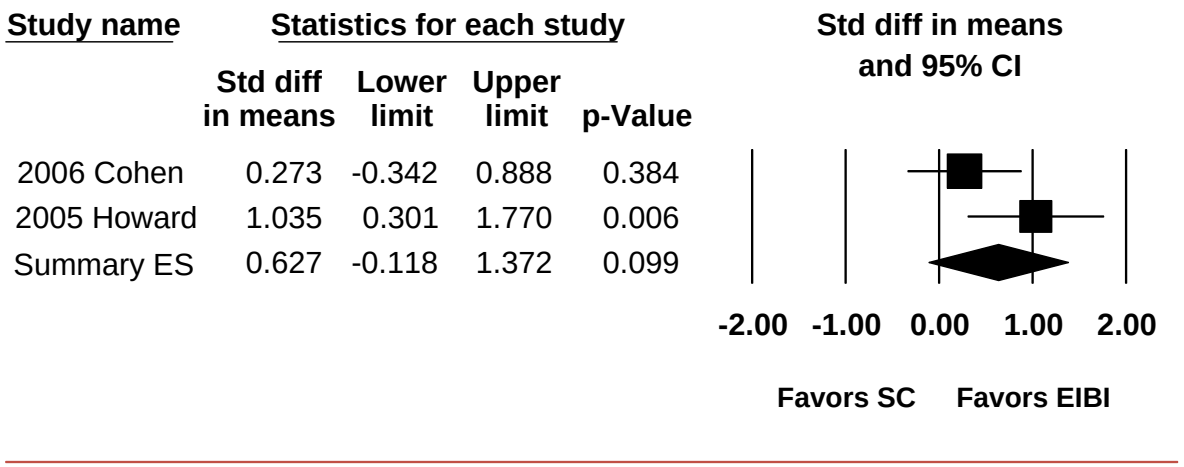
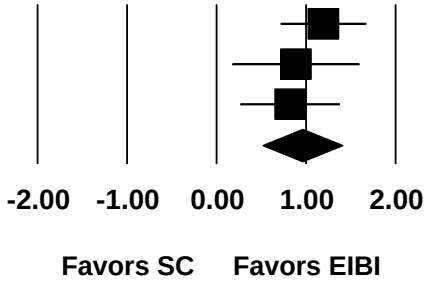


Figure 19. Key Question 1: Robustness Test for Vinland Adaptive Behavior Composite Score at 1 Year Follow-up (One Study Removed)

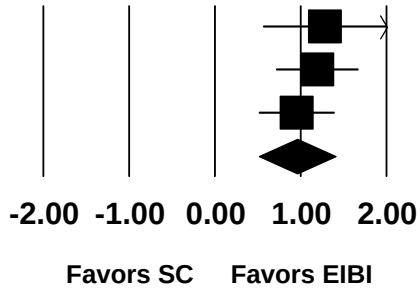
Study name	Statistics with study removed				Std diff in means (95% CI) with study removed
	Point	Lower limit	Upper limit	p-Value	
2007 Remington	1.194	0.710	1.678	0.000	
2006 Cohen	0.887	0.171	1.604	0.015	
2005 Howard	0.822	0.261	1.383	0.004	
Summary ES	0.953	0.507	1.400	0.000	



Random Effects Meta Analysis

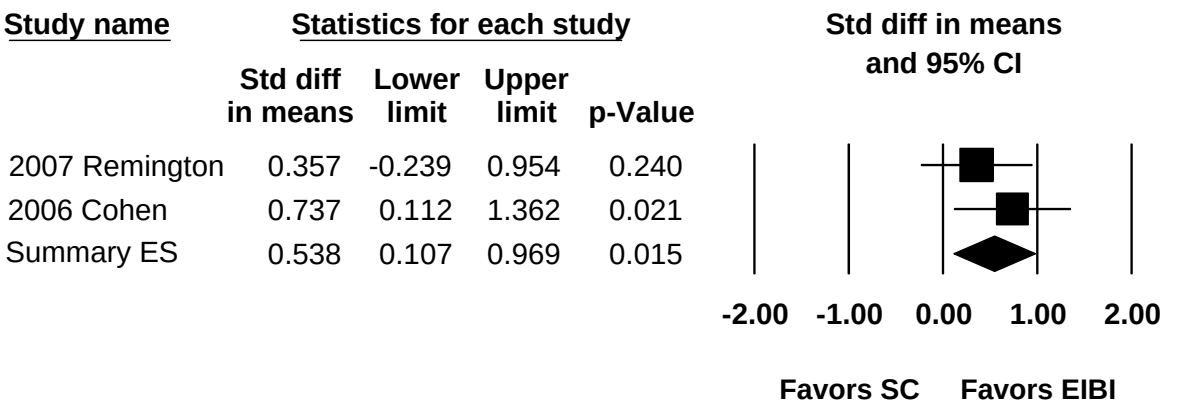
Figure 20. Key Question 1: Robustness Test for Vinland Adaptive Behavior Composite Score at 1 Year Follow-up (Cumulative by Publication Date)

Study name	Cumulative statistics				Cumulative std diff in means (95% CI)
	Point	Lower limit	Upper limit	p-Value	
2005 Howard	1.282	0.558	2.006	0.001	
2006 Cohen	1.194	0.710	1.678	0.000	
2007 Remington	0.953	0.507	1.400	0.000	
Summary ES	0.953	0.507	1.400	0.000	



Random Effects Meta Analysis

Figure 21. Key Question 1: Vinland Adaptive Behavior Composite Score at Last Follow-up



Random Effects Meta Analysis/I^2=0.00

Appendix J. Treatment Guidelines, Information from Professional Groups, and Third Party Payer Coverage Policies

Table 49.Treatment Guidelines for ASDs Identified through National Guideline Clearinghouse (NGC) and Healthcare Standards (HCS)

Reference	Title	Guideline Objective	Treatment interventions considered in report	Summary of Recommendations for Non-pharmacological Treatment Interventions
New York Department of Health Early Intervention Program 2008(81)	*Clinical Practice Guideline: Autism/Pervasive Developmental Disorders* *Assessment and Intervention for Young Children ages 0-3*	To provide recommendations about best practices for assessment and intervention for young children with autism, with a primary focus on children under 3 years of age	Early intervention services (behavioral and education intervention programs); DIR; Sensory Integration; Auditory Integration Training; Facilitated Communication; Music Therapy; Touch Therapy	**Intensive Behavioral and Educational Intervention Programs:** • The recommendations specify that treatment should be initiated when the child is young, include a minimum of approximately 20 hours per week of individualized behavioral intervention using ABA techniques, and that the number of hours should be reviewed and revised when necessary and child's progress monitored. The evidence reviewed for the guidelines was insufficient to predict the optimal number of hours that will be effective for any given child. Specific behavioral strategies that are useful for children with autism include techniques such as: prompting, modeling, fading and reinforcement. **Other treatments:** • There is no research evidence that intervention approaches based on DIR, sensory integration therapy, auditory integration therapy, facilitated communication, music therapy, and touch therapy are effective as intervention for young children with autism. Without evidence from controlled studies using accepted scientific methodology that demonstrates effectiveness, interventions based on these approaches cannot be recommended as primary interventions for young children with autism.

Reference	Title	Guideline Objective	Treatment interventions considered in report	Summary of Recommendations for Non-pharmacological Treatment Interventions
Scottish Intercollegiate Network (SIGN) 2007(82)	*Assessment, diagnosis and clinical interventions for children and young people with autism spectrum disorders. A national clinical guideline*	To provide evidence-based recommendations on the assessment, diagnosis and clinical interventions for children and young people with autism spectrum disorders (ASD)	Support for early communication skills, Interventions for social communication and interaction, Intensive behavioral programs, Behavioral interventions, Pharmacologic therapy (Risperidone, Methylphenidate, Melatonin), and Service provision (training of healthcare personnel, provision of information for parents/carers, education and skills interventions for parents of preschool children with ASD)	**Intensive Behavioral and Educational Intervention Programs:** • The Lovaas program should not be presented as an intervention that will lead to normal functioning. Behavioral Interventions should be considered to address a wide range of specific behaviors in children and young people with ASD, both to reduce symptom frequency and severity and to increase the development of adaptive skills. **Other treatments:** • Interventions that support communication in ASD are indicated, such as the use of visual augmentation (e.g., in the form of pictures of objects). • Interventions to support social communication should be considered for children and young people with ASD, with the most appropriate intervention being assessed on an individual basis. • Auditory integration training is not recommended. • Facilitated communication should not be used as a means to communicate with children and young people with ASD.

Reference	Title	Guideline Objective	Treatment interventions considered in report	Summary of Recommendations for Non-pharmacological Treatment Interventions
Burrows, Canadian Pediatric Society 2004 (reaffirmed 2008)(83,84)	*Early Intervention for children with autism*	To briefly describe the main educational interventions (programs) that are intended to result in global improvement in autism and review the status of the evidence regarding their effectiveness. Behavioral techniques that limit their aim to changing specific areas of functioning in autism were not reviewed	Early Intensive Behavioral Intervention (usually referring to the Lovaas method) and "normalized teaching" Other models for intensive autism treatment (LEAP, Floor Time, and TEACCH) were described but not critiqued because of a paucity of controlled trials	**Intensive Behavioral and Educational Intervention Programs:** • The quality of the existing studies on educational treatment programs was suboptimal but did show a trend toward a positive outcome from intervention. • However, there is no evidence to support adopting a single autism treatment program as the gold standard. • Although evidence of efficacy for educational treatment programs was weak, the studies to date do suggest some guiding principles that may be of use in planning treatment. Given the available information, it appears reasonable to set a target of a minimum of 15 hours a week of structured, individualized teaching; the family should be involved in service provision; and there should be an ongoing program evaluation and adjustment to meet the child's needs. • There is a great need for well-designed and well- implemented studies in this area including identifying the common effective elements of treatment programs; studies involving children across the full spectrum of autism; studies that identify the optimal age and IQ range of children receiving these services, optimal program intensity, duration of treatment and parental involvement; the magnitude of effectiveness of these programs; and direct comparison of the various intensive treatment programs.
Ludwig, Alberta Heritage Foundation for Medical Research 2001(85)	*Intensive Intervention Programs for Children with Autism*	Summarized three systematic/critical reviews done previously by ECRI Institute, British Columbia Office of Health Technology Assessment Report (BCOHTA) and Smith	Intensive intervention programs, including: Lovaas Therapy, The Rutgers Autism Program, The TEACCH Program, The Denver Model, The LEAP Program, The Autism Preschool Program, and Princeton Child Development Program	**Intensive Behavioral and Educational Intervention Programs:** • Because of methodological limitations and weaknesses of the existing research, evidence for the efficacy or effectiveness of one intervention over another remains limited. It does appear that children improve in functioning with intensive intervention programs, but it remains to be determined if one program is more effective than another. • There was insufficient evidence to establish a relationship between amount (intensity and duration) of any intensive intervention program and outcome measures (IQ, language development, adaptive behavior tests).

Reference	Title	Guideline Objective	Treatment interventions considered in report	Summary of Recommendations for Non-pharmacological Treatment Interventions
McGahan, Canadian Coordinating Office for Health Technology Assessment (CCOHTA) 2001(86)	*Behavioral Interventions for Preschool Children with Autism*	To summarize the evidence and expert opinions regarding behavioral therapy, describe Canadian issues and initiatives, analyze the legal case findings, and identify key factors that influence the provision of services to preschoolers with autism in Canada	Behavioral Interventions: Lovaas, Douglass Developmental Disabilities Center Program, LEAP, May Institute, Autism Preschool Program, Princeton Child Development Institute Program, TEACCH, The Denver Model, and Others	**Intensive Behavioral and Educational Intervention Programs:** • The literature on efficacy of behavioral interventions lacks controlled trials and most studies have methodological flaws that make interpretation of their results difficult. • However, the existing evidence suggests that behavioral intervention, including a minimum intensity of approximately 20 hours per week of one-on-one applied behavioral analysis, can improve aspects of functioning, in particular IQ, in autistic children. • Still to be determined is what subset of children derive the most benefit, which components of therapy are integral to a positive outcome, whether similar results would be observed in older children, whether there are definable long term functional benefits, or if gains in IQ translate into happier, better functioning people. Policy makers, program developers and clinical researchers should evaluate progress in therapy to determine if therapy is or continues to of benefit.

Reference	Title	Guideline Objective	Treatment interventions considered in report	Summary of Recommendations for Non-pharmacological Treatment Interventions
British Columbia Office of Health Technology Assessment 2000(87)	*Autism and Lovaas treatment: A systematic review of effectiveness evidence* *Critical appraisal of submitted cost-benefit models of 'Lovaas' early intensive behavioral intervention for autism*	To determine if early, intensive behavioral therapy for preschool-aged children with autism resulted in normal functioning, or essentially a cure To conduct a critical appraisal of two cost-benefit analyses	Lovaas method, TEACCH	**Intensive Behavioral and Educational Intervention Programs:** • While many forms of intensive behavioral therapy clearly benefit children with autism, there is insufficient evidence to establish a causal relationship between a particular program of intensive behavioral treatment and the achievement of normal functioning. • There is insufficient effectiveness evidence to establish a relationship between the amount (per day and total duration) of any form of early comprehensive treatment program ad overall outcome. • Randomized trials of alternative early intensive treatment programs are needed. • There is insufficient evidence to conduct a cost-benefit analysis of early, intensive treatment programs in terms of "normalization" of children with autism. • Regarding the one included TEACCH publication, the authors conclude that auxiliary home interventions increase developmental functioning in young autistic children above and beyond gains due to school-based interventions.

Table 50. Guidelines/Practice Parameters Identified through Other Sources

Reference	Title	Guideline Objective	Treatment interventions considered in report	Summary of Recommendations for Non-pharmacological Treatment Interventions
Ministry of Health, New Zealand 2008(89)	*New Zealand Autism Spectrum Disorder Guideline*	To provide evidence-based guidance on ASD in both children and adults in New Zealand	Comprehensive treatments, educational treatments and psychosocial treatments	**Intensive Behavioral and Educational Intervention Programs:** • Treatment should encourage functional development, skills for independent living to minimize stress on the person with ASD and their family. • Treatment plans should be comprehensive, and include behavioral needs, educational interventions, psychosocial treatments, communication, environmental and systems issues and the suitability (or not) of medication. • Professionals, people with ASD, family, and carers should work together to evaluate treatment approaches before and during implementation. • All behavioral interventions should be of good quality and incorporate the following principles: person-centered planning, functional assessment, positive intervention strategies, multifaceted interventions, focus on environment, meaningful outcomes, focus on ecological validity and systems-level intervention. • When severe behaviors are evident, people with ASD need to be assessed for co-morbid conditions such as seizures, attention deficit hyperactivity disorder (ADHD), anxiety disorders and depression.

Reference	Title	Guideline Objective	Treatment interventions considered in report	Summary of Recommendations for Non-pharmacological Treatment Interventions
Association for Science in Autism Treatment(46)	*Summaries of Scientific Research on Interventions on Autism*	To describe and summarize the existing research on psychological, educational and therapeutic interventions and provide recommendations for each	Applied Behavioral Analysis, Animal Therapy, Art Therapy, Auditory Integration Therapy, Augmentative Communication, Developmentally based Individual difference Relationship based Intervention (DIR), Facilitated Communication, Holding Therapy, Music Therapy, Oral-Motor Training/Therapy, Patterning, Picture Exchange Communication System (PECS), TEACCH, Psychoanalytic and Humanistic Play Therapy, Recreational Sports/Exercise, Relationship Development Intervention, Sensory Integrative Therapy, Socialization related classes, Social Stories, Son Rise, Video Modeling, and Vision Therapy	**Intensive Behavioral and Educational Intervention Programs:** • Applied Behavioral Analysis: ABA is an effective intervention for ASD. This program should be supervised by a qualified behavior analyst. As there is scientific support for this program, professional and families may wish to obtain additional information about this approach. Larger studies with strong scientific designs are needed to assess the long-term outcomes of early, intensive ABA and other comprehensive ABA intervention programs. **Other treatments:** • In general, for all other therapies assessed, the authors of the report concluded that researchers may wish to conduct studies with strong scientific designs to evaluate the therapies, and professionals should present them as untested and encourage families who are considering one of these interventions to evaluate it carefully.

Reference	Title	Guideline Objective	Treatment interventions considered in report	Summary of Recommendations for Non-pharmacological Treatment Interventions
Prior and Roberts 2006(90)	*Early Intervention for Children with Autism Spectrum Disorders: Guidelines for Best Practice*	To summarize the research findings related to early intervention for autism, outline the kinds of programs available in Australia, identify research and evidence based guidelines for best practice in early intervention and provides a list of contacts for programs across Australia	Educational interventions including Applied Behavioral Analysis (ABA), Relationship Development Intervention (RDI), Picture Exchange Communication System (PECS), Auditory Integration Training (AIT), Treatment and Education of Autistic and related Communication handicapped Children (TEACCH), Music Intervention Therapy, and family based interventions such as The Hanen Program	**Intensive Behavioral and Educational Intervention Programs:** • The most systematic evidence available has come from intensive behavioral programs such as Lovaas or Applied Behavior Analysis. • Evaluations on intensive behavioral programs show improved learning and behavioral development in a significant proportion of children. These methods do not suit all children, however, and strict conditions of timing, intensity and quality of therapist training influence the success of these methods. • The following are key elements necessary for effective intervention: an autism specific curriculum focusing on attention, compliance, imitation, language, and social skills; a highly supportive teaching environment which provides predictability and routine and addresses challenging behaviors, obsessions and ritual behaviors; provides support for children in their transition from the preschool classroom; promotes a partnership between parents and treatment professionals; provides services for a minimum of 20 hours a week over a at least a two year period; adapts to meet the individual child's needs by taking account of their strengths and weaknesses and family circumstances. • Other programs have not shown sufficient evidence of short or long term improvement to qualify for unreserved support.
National Early Childhood Technical Assistance Center (NECTAC)(91)	*Elements of Effective Programs*	To present a consensus opinion about what are the most important elements of treatment programs for individuals with an ASD	Seven well-known treatment models families are likely to recognize and frequently request	• The six elements identified as part of all effective treatment programs include: the earliest possible start to intervention, individualization of services to meet unique needs of the child and his/her family, systematic teaching strategy that builds toward meaningful goals, specialized curriculum that focuses on ASD deficits, the amount of time in which the child is being taught or actively learning, and family involvement. • In addition, the three other important elements that were identified as part of some, but not all effective programs include a structured environment, programs guided by information about child development, and interventions that include interactions with typically developing children.

Table 51. Guidelines/Practice Parameters by State

Reference	Title	Guideline Objective	Recommended Non-pharmacological Treatment Interventions may be found at:
California Departments of Education and Developmental Services 1997(121)	*Best Practices for Designing and Delivering Effective Programs for Individuals with ASDs: Recommendations of the Collaborative Work Group on Autistic Spectrum Disorders*	To summarize suggested strategies for addressing common issues in program development, transition planning, provision of effective staff development, and program evaluation	www.isciii.es/htdocs/centros/enfermedadesraras/pdf/aut_gcalif.pdf
Connecticut Birth to Three System 2008(122)	*Service Guideline: Autism Spectrum Disorder Intervention guidance for service providers and families of young children with ASDs*	The purpose of this guideline is to help families and service providers develop and carry out intervention plans for families of children who have characteristics of disorders on the Autism Spectrum, including Pervasive Developmental Disorder (PDD).	http://www.birth23.org/Publications/Autism%202008.pdf
Indiana Institute on Disability and Community 2001(123)	*Early Intervention for Young Children with Autism Spectrum Disorders: Recommendations for Designing Effective Programs*	Written for both family members and professionals, this publication describes the key components of an effective early intervention program for young children with an autism spectrum disorder and provides practical recommendations for implementing these key components.	http://www.iidc.indiana.edu/
Maine Administrators of Services for Children with Disabilities (MADSEC) 2000(124)	*Report of the MADSEC Autism Task Force*	Perform a detailed analysis of methodologies used to educate children with autism, focusing on the scope and quality of the scientific research to determine each method's effectiveness. Based upon the research analysis, this report makes recommendations for the consideration of decision makers.	www.madsec.org

Reference	Title	Guideline Objective	Recommended Non-pharmacological Treatment Interventions may be found at:
New Jersey 2004(125)	*Service Guidelines For Children with Autism Spectrum Disorders*	To enhance the capacity of families to meet the developmental needs of children, birth to age three, who have delays or disabilities, by providing quality services and support to families and their children.	http://www.state.nj.us/health/fhs/documents/autismguidelines.pdf
New Mexico Family Infant Toddler Program 2004(126)	*Autism Spectrum Disorders - Guidance on providing supports and services to young children with autism spectrum disorders and their families*	To provide guidance on providing support and services to young children with ASDs and their families	http://www.health.state.nm.us/ddsd/fit/pdf%5CAutism-Spectrum-Disorders.pdf
The New York State Education Department Office of Vocational and Educational Services for Individuals with Disabilities Special Education Policy, Planning and Partnerships 2004(127)	*The Availability and Effectiveness of Programs for Preschool Children with Autism*	To report on the availability and effectiveness of approved programs providing special education services to preschool children with autism	http://www.vesid.nysed.gov/specialed/autism/preschoolstudy.htm
New York State Department of Health Early Intervention Program 1999(128)	*Clinical Practice Guideline: Report of the Recommendations: Autism/Pervasive Developmental Disorders. Assessment and Intervention for Young Children (Age 0-3 Years)*	This document provides an extraordinarily thoughtful and balanced presentation of the critical issues in assessment and intervention for this group of children. There is no doubt in my mind that readers will find the Guideline to be a valuable resource, as it will allow numerous individuals with different levels of expertise to gain a firm understanding and make highly informed decisions with respect to assessment and intervention for young children with autism and pervasive developmental disorders.	http://www.health.state.ny.us/community/infants_children/early_intervention/autism/index.htm

Reference	Title	Guideline Objective	Recommended Non-pharmacological Treatment Interventions may be found at:
North Dakota Department of Public Instruction 2003(129)	*Guidelines: Identifying, Serving, and Educating Children and Youth with Autism*	To review and discuss the issues relative to the assessment and education of individuals with autism, including best practice strategies, family support and early intervention	www.dpi.state.nd.us/speced/guide/autism.pdf
Ohio Developmental Disabilities Council(130)	*Service Guidelines for Individuals with Autism Spectrum Disorder/Pervasive Developmental Disorder (ASD/PDD) Birth through Twenty-one*	To provide recommendations based on the current knowledge about "best practices" for the assessment of individual needs and the delivery of appropriate services for children and young adults with ASD	http://ddc.ohio.gov/Pub/Child/htm
Washington State Infant Toddler Early Intervention Program(131)	*Successes in Serving Families and Infants and Toddlers with Autism*	To insure services are reasonably calculated to confer developmental benefit, this guideline describes a Individualized Family Service Plan (IFSP) process that includes the family and shares information about the importance of integrated services, methods and approaches	http://www1.dshs.wa.gov/word/adsa/iteip/SLM_Autism.doc.
Department of Health Services State of Wisconsin 2007(132)	*Intensive In Home Service*	To describe intensive in home services and provide guidelines for how they should be implemented	http://dhs.wisconsin.gov/bdds/waivermanual/waiverch04_08.pdf#page=85

Table 52. Third Party Payer Coverage Policies for Services to Individuals with Autism Spectrum Disorder

Third-party Payer	Web site	Coverage Policy	Policy/ Bulletin Number	Treatments Considered to be Experimental and Not Covered
Aetna(133)	http://www.aetna.com/cpb/medical/data/600_699/0648.html	For pervasive developmental disorder (PDD), intensive educational interventions and alternative/augmentative communication aids are covered.	0648	Auditory Integration Training; Chelation Therapy; Cognitive Rehabilitation; Elimination Diets; Facilitated Communication; Holding Therapy; Immune Globulin Infusion: Music therapy and rhythmic entrainment interventions; nutritional supplements; Secretin infusion; Sensory Integration Therapy; Vision therapy.
American Medical Association(134-136)	http://coverageandpayment.medirecgs.com	As of 2008, payment for these services may not be made if the service was provided to either a patient in a hospital outpatient department or to an inpatient of the hospital by an independently practicing Physical/occupational therapist: cognitive skills development and sensory integrative techniques.	NR	NR
Athens area Health Plan Select, Inc., Athens Georgia 2005(137)	http://www.aahps.com/pdfs/PPO_EOC.pd f	Treatment for autism shall be covered on the same basis as other diagnosed neurological disorders.	NR	NR
Blue Cross/Blue Shield of Alabama(138)	http://provider.medica.com/router/default.pdf?doc=/C1/CoveragePolicies/Document%20Library/Chelation Therapy_CP.pdf	NR	NR	Chelation therapy
Blue Cross/Blue Shield of Massachusetts(139,140)	http://www.bluecrossma.com/common/en_US/medical_policies/281%20Early%20Intervention%20Special%20Needs%20Chapter%20766%20prn.pdf	Early intervention is covered if child is 3 or less with an established, biological or environmental risk; has a known disabling physical or mental condition; four or more risk factors.	281,439	Recreational services; orthoptic (vision) training; auditory integration training; facilitated communication; cognitive rehabilitation therapy; sensory integration therapy.

Third-party Payer	Web site	Coverage Policy	Policy/ Bulletin Number	Treatments Considered to be Experimental and Not Covered
Blue Cross/Blue Shield of Tennessee(141)	http://www.bcbst.com/mpmanual/speech_-_language_therapy_occupational_therapy_and_physical_therapy_for_autism.htm	NR	NR	Speech/language therapy, occupational therapy and physical therapy for the treatment of autism are considered investigational except when the Tennessee State Mandate applies. The Tennessee State Mandate applies to individual policies, fully insured accounts, and self-funded accounts not governed by ERISA, and to children with ASDs less than 12 years of age. Specifically, the mandate states: A contract or policy of an insurer that provides benefits for neurological disorders, whether under an individual or group health insurance policy providing coverage on an expense-incurred basis, an individual or group service contract issued by a health maintenance organization, a self-insured group arrangement to the extent not preempted by federal law or a managed health care delivery entity of any type or description shall provide benefits and coverage for the treatment of ASDs that are at least as comprehensive as those provided for other neurological disorders.
Cigna(142)	http://www.cigna.com/health/provider/medical/procedural/coverage_positions/index.html	NR	0447	Sensory integration therapy; auditory integration therapy; facilitated communication; augmentative communication devices; chelation therapy; cognitive behavioral therapy; cognitive rehabilitation; dietary/nutritional interventions; hyperbaric oxygen therapy; intensive intervention programs (e.g., Lovaas, ABA), immune globulin therapy; music therapy, secretin infusion; vision therapy.

Third-party Payer	Web site	Coverage Policy	Policy/ Bulletin Number	Treatments Considered to be Experimental and Not Covered
Health Partners(143)	http://www.healthpartners.com/policies/	Medical policy for PDD currently being revised.	NR	NR
Kaiser Permanente Health Plan, Northern California Region 2004(144)	https://www.kaiserpermanente.org	Mental health services for PDD or autism are covered, including evaluation, crisis intervention, outpatient visits, psychological testing, visits for the purpose of monitoring drug therapy, inpatient psychiatric care, and structured multidisciplinary programs of psychiatric care as an alternative to inpatient psychiatric care.	NR	NR
MAMSI Life and Health Insurance Company State of Maryland(145)	http://www.mamsiunitedhealthcare.com/s/g/md/0726299-0105MD.pdf	Habilitative including speech, occupational and physical therapy) services are limited to 50 visits per year combined per condition. Treatment related to autism or PDD except as it relates to habilitative services for children under the age of 19 is excluded. However, the assessment of these disorders is covered.	NR	Policy also routinely excludes the following treatments which are sometimes used to treat ASD: art therapy; massage therapy; mental health services; therapy for eyes and eye exercises; special education, counseling therapy or care for learning deficiencies or behavioral problems; confinement, treatment, services or supplies related to learning disabilities, mental retardation and/or mental deficiency; educational assessments and vocational training.
Medica(146,147)	http://provider.medica.com/C9/MedicalPolicies/default.aspx	NR	NR	Lovaas therapy/intensive early intervention behavior therapy services/intensive behavioral intervention; Health Research Institute/Pfeiffer Treatment Center Protocols; Sensory Integration Therapy; Auditory Integration Training; Chelation Therapy.
Premera Blue Cross 2008(148)	http://www.asha.org/about/legislation-advocacy/2008/PremeraBlueCross.htm	Speech-generating devices (SGD) and other Augmentative and Alternative Communication (AAC) devices are covered.	NR	NR

Third-party Payer	Web site	Coverage Policy	Policy/ Bulletin Number	Treatments Considered to be Experimental and Not Covered
Regence Blue Cross/ Blue Shield(149)	http://blue.regence.com/trgmedpol/index.html	Augmentative communication devices and systems (ACD), also known as augmentative and alternative communication devices and speech generating devices are covered if recommended by a therapist, individual either unable to communicate or learn to communicate through mean such as writing; willingness to use device; if for a degenerative disease, device is able to meet individual's anticipated needs; if pre-literate but anticipated to learn to read and spell, device should have spelling and text capabilities in addition to symbols.	52	NR
Wellmark Blue Cross/ Blue Shield(150,151)	http://www.wellmark.com/e_business/provider/medical_policies/medical_policies.asp	NR	08.03.04; 08.01.06	Sensory Integration therapy; chelation therapy

Appendix K. Previous Systematic Reviews

Table 53. Previous Systematic Reviews

Citation	Search Strategy	Key Inclusion/ Exclusion Criteria	Evidence Base	Participant Characteristics	Outcomes Assessed	Method of Assessing Study Quality	Type of Review	Results and/or Authors' Conclusions
Case-Smith & Arbesman (2008)(97) *Evidence-Based Review of Interventions for Autism Used in or of Relevance to Occupational Therapy*	The following databases were searched for studies published between 1986 and 2007: Medline, CINAHL, ERIC, PsycINFO, Social Sciences Abstracts, Sociological Abstracts, Lingustics and Language Behavior Abstracts, RehabData, Latin American and Caribbean Health Sciences Literature, and EBSCOHost.	Studies were included if 1) they provided evidence for an intervention approach used with children or adolescents with ASD, had been peer reviewed, were published between 1986 and 2007, and addressed a performance area or intervention approach with the domain of occupational therapy. Studies were excluded if they were descriptive studies (Level IV and V evidence-case series, single subject designs, expert opinion, etc), published before 1986, used qualitative methods to the exclusions of quantitative methods, were not peer reviewed, or contained fatal flaws.	49 studies: 18 RCTs, systematic reviews, or meta-analyses; 17 non-randomized trials such as cohort studies; 14 before and after, one group designs. Overall number of children not reported. This was a very broad review that included studies that assessed focal interventions and comprehensive interventions. The comprehensive interventions included studies on TEACCH, DIR, IBI, and Parent Directed Approaches. Studies/reviews that overlapped with ECRI Institute's current review include: Bassett et al. (2000), Cohen et al. (2006), Diggle et al. (2003), Sallows & Graupner (2005), and Smith et al. (2000).	Young children or adolescents with ASDs.	The outcomes assessed were those reported in the studies that made up the evidence base for this review and ranged from specific behaviors such as imitation and play to broader behaviors such as overall IQ.	Quality was assessed by categorizing studies into Level I evidence (RCTs, systematic reviews, or meta-analyses, k = 18); Level II (non-randomized trials, k = 17), and Level III (pre-post, single group studies).	Narrative	The authors concluded that the research literature offers strong positive evidence for occupational therapists to use comprehensive, individualized analysis of the child's performance to develop intervention strategies. But, no overall conclusions were made about the efficacy of one particular intervention or one intervention compared to another.

Citation	Search Strategy	Key Inclusion/ Exclusion Criteria	Evidence Base	Participant Characteristics	Outcomes Assessed	Method of Assessing Study Quality	Type of Review	Results and/or Authors' Conclusions
Reichow & Wolery (2008)(101) *Comprehensive Synthesis of Early Intensive Behavioral Interventions for Young Children with Autism Based on the UCLA Young Autism Project Model*	The authors conducted an electronic database search (specific databases not reported), reviewed references of other reviews on the same topic, hand searched selected journals, and contacted experts in the field.	Studies were included if a) they specifically assessed the UCLA Young Autism Program; included participants that were diagnosed with autistic disorder, ASD, or PDD-NOS; c) participants were less than 84 months at beginning of treatment; d) mean duration of treatment ≥1 year, e) at least one child outcome was measured; f) the experimental design was a case series, non-randomized or randomized controlled trial, and the publication was in English in a peer-reviewed journal.	The evidence-base consisted of 14 studies: 2 RCTs, 5 non-randomized controlled trials, 2 prospective case series studies, and 5 retrospective case controlled studies. Total number of children 373 The following studies overlapped with ECRI Institute's current review: Eikeseth et al. (2007 & 2002), Cohen et al. (2006), Sallows & Graupner (2005), and Smith et al. (2000). Studies not included in ECRI Institute's current review did not meet our inclusion criteria because they were not prospective controlled trials (k = 7) or they did not meet current diagnostic standards established in the DSM-IV or TR edition (k = 1).	In 50% of the included studies a 100% of children had a diagnosis of autism. The other studies included children with ASD and PDD-NOS. In 43% of studies children were less than 36 months, and in 50 percent of studies the mean pre-treatment IQ was between 40 and 55.	IQ, adaptive behavior, expressive and receptive language, academic placement, psychopathology, and diagnostic recovery	The studies were assessed for experimental rigor using the Evaluative Method for Determining Evidence-Based Practices in Autism (Reichow et al. in press), study design, method for group assignment (e.g., random assignment, therapist availability or parent selection), procedural fidelity. Based on this system, 3 studies received a strong rating, 5 received a adequate or moderate rating, and 5 received a weak rating.	Quantitative review that included a meta-analyses, homogeneity testing, sensitivity analysis (e.g., test for publication bias), and moderator analyses.	The primary analysis in this review focused on changes in IQ scores from pre to post treatment of children who received the UCLA program as most studies reported on this outcome. No comparative analyses were conducted as the studies varied in terms of the comparison condition. The authors did not attempt to group studies according to comparison condition. The results of a random effects meta-analysis of 12 studies indicated that the mean effect size for IQ was 0.69 ($p < 0.001$), which suggests that EIBI is effective in increasing IQ in children with autism. However, the authors concluded that this finding should be interpreted with caution as the studies included in their analysis contained methodological flaws (e.g., single group designs, non-equivalent study groups, etc).

Citation	Search Strategy	Key Inclusion/ Exclusion Criteria	Evidence Base	Participant Characteristics	Outcomes Assessed	Method of Assessing Study Quality	Type of Review	Results and/or Authors' Conclusions
Rogers & Vismara (2008)(98) *Evidence-Based Comprehensive Treatments for Early Autism*	Searched PsycINFO for studies published from 1998 to 2006	Studies involving comprehensive treatment approaches for autistic children aged 5 years or younger were included (e.g., IBI, also included studies on more focal types of treatment such as pivotal response training and medication therapy). Studies that did not report analyses of child progress using general measures of children's language or intellectual development; studies targeting only one domain (i.e., social behavior) and case reports and studies not published in peer-reviewed journals were excluded.	Study criteria were defined as follows: Type 1 studies – RCTs (k = 4); type 2 studies - comparison group design or single-subject (k = 6); type 3 studies – uncontrolled studies (k = 11) Total number of children 421 The following studies overlapped with ECRI Institute's current review: Cohen et al. (2006); Eikeseth et al. (2002); Howard et al. (2005); Sallows & Graupner (2005); and Smith et al. (2000). Other studies were excluded because they did not meet the current reviews inclusion criteria (e.g., uncontrolled studies).	Children aged 26 months to 4 years diagnosed with autism and related disorders	Outcomes measured include language development, IQ, and adaptive behavior	Nathan and Gorman (2002) study criteria: Type 1 evidence RCTs, Type 2 non-randomized controlled trials, and Type 3 evidence non-controlled trials	Narrative	Although based on small study samples, all studies noted improvements in language, communication, IQ, and reduction in severity of autism symptoms. To gain additional insight on the effects of established treatment programs, authors stress the need for larger, well-powered studies, inclusion of multi-site studies to ascertain occurrence of "recovery" in autism, and the importance of education and training of researchers to enhance treatment delivery to culturally diverse populations.

Citation	Search Strategy	Key Inclusion/ Exclusion Criteria	Evidence Base	Participant Characteristics	Outcomes Assessed	Method of Assessing Study Quality	Type of Review	Results and/or Authors' Conclusions
Doughty (2004)(152) *What is the evidence for the effectiveness of behavioral and skill-based early intervention in young children with Autism Spectrum Disorder (ASD)?*	The following databases searched from 2000 to 2003: MEDLINE, EMBASE, CINAHL, PsycINFO, AMED, and ERIC, Current Contents, Web of Science, Evidence-based medicine reviews, Cochrane, DARE, NHS Economic Evaluation Database, and Health Technology Assessment Database.	Studies were excluded if mean/median age ≥8 years, <5 subjects were in treatment/control groups, studies did not use standardized/ validated outcome measures, and non-English language studies.	This review examined intensive or comprehensive behavioral or skill-based interventions that treat or manage symptoms of ASD. 10 studies made up the evidence base: 5 primary studies (1 cohort, 4 CTs [2 RCTs] and 5 secondary studies (SRs). Overall number of patients in primary studies was 171. The following primary studies overlapped with ECRI Institute's current review: Eikeseth et al. (2002) and Smith et al. (2000). The other 3 studies did not meet the current review's inclusion criteria (i.e., did not address intervention of interest [k = 1], not a controlled trial [k = 1], and incomparable study groups [k = 1]).	Patients aged <8 years diagnosed with ASD (75% of sample) as classified by DSM-IV and/or ICD-10	Behavioral change and the development of reciprocal social interaction and/or communication skills	NHMRC Hierarchy of Evidence (1999)	Narrative	According to the authors, behavioral interventions have shown to improve functional status of children with ASD; however, one program does not appear to be more effective than another. In addition, it was unclear that the definition of included programs (including intensive behavioral, parent training or parent-managed behavioral therapy) in the primary studies was consistent. Two primary studies that graded high- level evidence concluded parent-training intervention was more effective than usual care for improving communication, and IBI may be more effective than parent training.

Citation	Search Strategy	Key Inclusion/ Exclusion Criteria	Evidence Base	Participant Characteristics	Outcomes Assessed	Method of Assessing Study Quality	Type of Review	Results and/or Authors' Conclusions
Levy et al. (2006)(99) *Interventions for Young Children With Autism: A Synthesis of the Literature*	Interventional studies published from 1975 to April 2001.	Studies were required to be interventional, involve a treatment/ comparison or single-group study design and have ≥50% population diagnosed with autism. Studies involving medical interventions (i.e., diet, pharmacology) were excluded.	23 studies were examined and separated into 6 categories: parent involvement, intensive behavioral intervention, multi-component intervention, language/ speech treatment, setting, and other interventions. Overall number of subjects was not reported. None of the studies included in this review overlapped with ECRI Institute's review mainly because the studies were not controlled trials or did not address a treatment of interest.	66% of study samples were aged 3-8; similar disability status in treatment/ comparison (or control)	Outcomes ranged from intelligence scores to problem behaviors.	NR	No meta-analyses performed, but the author calculated individual study effect sizes using the standardized mean difference (or Cohen's d)	The author indicates that studies associated with large effect sizes (Cohen's d of ≥0.8 shared three similar characteristics: 1) the interventions were comprehensive in that they focused on a variety of areas, including language, behavior, social skills, etc; 2) the interventions were intensive and lasted for a long time; and 3) the interventions involved the child's parents.

Citation	Search Strategy	Key Inclusion/ Exclusion Criteria	Evidence Base	Participant Characteristics	Outcomes Assessed	Method of Assessing Study Quality	Type of Review	Results and/or Authors' Conclusions
Finch & Raffaele (2003)(153) *Intensive Behavioral Intervention for Children with Autism: A Review of the Evidence*	Searched the following databases MEDLINE, CINAHL, AMED, Cochrane Reviews and occupational therapy specific databases (dates searched not reported in article)	Studies included were those which were published in journals and involved experimental trials (versus descriptive studies) of IBI with children with a diagnosis of autism or PDD.	7 studies-1 RCT, 4 non-randomized controlled trials, and 2 case series (single group study). Overall number of children not reported Only one study included in this review overlapped with ECRI Institute's Review-Smith et al. (2000). Other studies did not meet ECRI Institute's inclusion criteria-case series (k = 2), retrospective case controlled trial (k = 2), or did not meet current diagnostic standards established in the DSM-IV or DSM-IV-TR (k = 3).	Children under eight years of age who were diagnosed with PDD or autism.	Intelligence (IQ), class placement, and social and adaptive functioning	Critical Review Form for Quantitative Studies developed by Law et al. (1998).	Narrative	The authors concluded that "research indicates some positive gains in IQ, class placement and adaptive behavior for most children receiving IBI. However, it is difficult to conclude strongly that IBI is effective for all children with autism based on limitations in the evidence." The authors list the following limitations: not all children in studies had a diagnosis of autism, diagnostic criteria has changed since older studies, outcomes varied between studies, no studies directly compared IBI to other intensive interventions, and limited long-term follow-up.

Citation	Search Strategy	Key Inclusion/ Exclusion Criteria	Evidence Base	Participant Characteristics	Outcomes Assessed	Method of Assessing Study Quality	Type of Review	Results and/or Authors' Conclusions
Diggle et al. (2002)(154) *Parent-mediated early intervention for young children with autism spectrum disorder*	The following databases were searched from 1966 to 2002: ERIC, The Cochrane Controlled Trials Register, MEDLINE, EMBASE, PsycINFO, CINAHL, Dissertation Abstracts International, Social Sciences Abstracts, Sociological Abstracts, Linguistics and Language Behavior Abstracts, National Research Register	Studies focused on parent-implemented early intervention with a comparison group receiving no treatment, a waiting-list group or a comparison intervention. Studies involving subjects with a dual diagnosis (i.e., ADHD, OCD) were included. Studies that included drug treatments, or treatments that aim to have physiologic effects (i.e., dietary intervention), and surgical interventions were excluded.	2 RCTs (n = 63) One of the 2 studies overlapped with ECRI Institute's current review-Smith et al. (2000). The other study by Jocelyn et al. (1998) was excluded as the diagnostic criteria used in this study were not based on current standards established in the DSM-IV or TR edition.	Children aged 1 year to 6 years 11 months diagnosed with ASD	Primary outcomes included child language progress, child positive behavioral change, and parent interaction style. Secondary outcomes included parent confidence, and reduction in levels of parental stress.	Studies evaluated for their risk of potential biases (e.g., selection bias), and graded as low, moderate, or high risk	Qualitative-no quantitative analyses conducted due to heterogeneity of included studies (e.g., differences in interventions delivered and outcomes measured)	According to the authors, two significant results were found in favor of parent training in one study: child language and maternal knowledge of autism. In the second study, intensive intervention (involving parents, but delivered by professionals) was associated with better child outcomes on measures of IQ and non-verbal cognitive ability than what was found in the parent-mediated early intervention.

Citation	Search Strategy	Key Inclusion/ Exclusion Criteria	Evidence Base	Participant Characteristics	Outcomes Assessed	Method of Assessing Study Quality	Type of Review	Results and/or Authors' Conclusions
Chorpita et al. (2002)(100) *Toward Large-Scale Implementation of Empirically Supported Treatments for Children: A Review and Observations by the Hawaii Empirical Basis to Services Task Force*	PsycINFO, studies previously reviewed by Lonigan and Elbert Task Force on Empirically Supported Psychosocial Interventions for Children, the American Academy of Child and Adolescent Psychiatry Practice Parameters, and personal communication with members of the Lonigan and Albert Task Force and other national scholars in effectiveness research	Studies that examined comprehensive treatments (designed to improve overall functioning, address multiple symptoms, and exist over long term) and focal treatments (designed to eliminate undesirable autistic behaviors) that included a pill or placebo control, an alternative treatment condition, or a wait-list.	The evidence considered in Chorpita et al. on comprehensive interventions was based on a previous systematic review by Rogers (1998). Rogers has since published an update of her review - Rogers and Vismara (2008), which is described in this table. Subjects included in comprehensive treatment studies and overall number of subjects was not reported. Focal treatments (15 controlled single-subject experimental designs) specific to FCT/ABA and Caregiver-Based Intervention Programs.	Children aged 2 to 15 diagnosed with autism and related disorders	Primary outcomes included behavioral changes for both children and parents (i.e., termination of self-injury, parent's level of distress, knowledge of autism)	NR	Narrative	Although clinical improvements were frequently observed in autistic children undergoing comprehensive treatments, research failed to rule out alternative explanations for improvement deemed an essential component for efficacy by study authors. Focal treatments, however, demonstrated both efficacy and effectiveness, including one trial demonstrating that a Caregiver-Based Intervention Program was superior to day care alone. The effectiveness of intensive sessions of FCT/ABA for children aged 2 to 15 were noted in as short as 2 weeks time and were often associated with clinically important changes in behavior, including the termination of self-injury.

Citation	Search Strategy	Key Inclusion/ Exclusion Criteria	Evidence Base	Participant Characteristics	Outcomes Assessed	Method of Assessing Study Quality	Type of Review	Results and/or Authors' Conclusions
Ludwig & Harstall (2001)(155) *Intensive Intervention Programs for Children with Autism*	Searched MEDLINE, PreMEDLINE, EMBASE, Best Evidence 2000, HTA, EED, DARE, HealthSTAR, PsycInfo, CINAHL, ERIC, Dissertation abstracts, CMA practice guideline, Cochrane, NHS Centre for Reviews and Dissemination, and other databases from 1985 to 2000	IBI programs for children with ASD	Ludwig et. Harstall evaluated three systematic reviews by: ECRI Institute– 14 studies (6 CTs [1 RCT]) 8 single group, pre-post study design) BCOHTA – 5 CTs, same studies as ECRI Institute with exception of 1 CT Smith – 12 studies (2 not previously discussed in the other reviews) N = 394 Two of the previous reviews assessed by Ludwig & Harstall are described in this report—ECRI Institute's previous review and BCOHTA (Basset et al. 2000). The third review by Smith is not discussed as this review was published prior to search dates for ECRI Institute's current review.	Children diagnosed with autism	Intelligence tests, adaptive behavior and behavior problems, tests of language performance and development, personality assessment, autism rating/ assessment scales, and parent measures	NR	Narrative-summarized findings of previous reviews to come to overall conclusions.	A broad range of programs were analyzed, including the strict operant discrimination learning of the Lovaas program vs. the more developmentally oriented programs of the Denver Model and TEACCH Program. The authors concluded that although children seem to improve in functioning with behavioral intervention programs, it is not readily apparent if one program is more effective than another.

Citation	Search Strategy	Key Inclusion/ Exclusion Criteria	Evidence Base	Participant Characteristics	Outcomes Assessed	Method of Assessing Study Quality	Type of Review	Results and/or Authors' Conclusions
McGahan L. (2001)(86) *Behavioural Interventions for Preschool Children with Autism*	DIALOG, MEDLINE, EMBASE, HealthSTAR, ERIC, PscyINFO from 1995 to 2000; CINAHL*direct* – no date limit; Current Contents Search, Cochrane Library to Issue 2, 2001; University of York NHS Centre for Reviews and Dissemination; CCOHTA	Studies included had an intervention generally accepted as a valid behavioral procedure by professionals in the field of behavior modification or ABA. Studies with methodological flaws were excluded.	The evidence assessed in this review comes from 5 previous systematic reviews, including ECRI Institute's previous review produced in 1999. Two of the other reviews evaluated in McGahn's review are described in this table: Bassett et (2000) and Ludwig and Harstall. The other two reviews, both published prior to 2000, are not presented in the table.	Children with an identifiable diagnosis of autism, a related PDD, or the presence of "autistic-like-behavior"	Functional improvement, time to improvement, social skills, communications skills, academic performance, and cognitive function	New York Assessment: strong evidence is described as being based on two or more studies that met criteria for adequate evidence about efficacy and having at least moderate applicability to the topic, where the evidence consistently and strongly supported the recommen-dation. Methods for assessing evidence were not defined by task groups from either Maine or California.	Narrative	The authors present recommendations of three working groups. The panels agree that the recipients of IBI consistently show significant functional improvement when compared to controls. In addition, although IBI interventions may be implemented between 18 and 40 hours per week and significant improvements may occur in as early as 20 hours per week, the groups stress the importance of periodically reviewing and revising hours of IBI based on the child's progress. The California Collaborative Work Group specifically addresses the importance of the administration of programs by trained professionals and the necessary inclusion of entrance and exit criteria.

Citation	Search Strategy	Key Inclusion/ Exclusion Criteria	Evidence Base	Participant Characteristics	Outcomes Assessed	Method of Assessing Study Quality	Type of Review	Results and/or Authors' Conclusions
Bassett et al. (2000)(87) Autism and Lovaas treatment: A systematic review of effectiveness evidence	MEDLINE (1966 to 1999), HealthStar (1975 to 1999), Embase (1988 to 1999), CINAHL (1982 to 1999), Current Contents (1996 to 1999), and combined Science and Social Sciences CitationIndex (1989 to 1999)	Studies that examined pre-school population, described interventions as early, applied behavioral analysis, behavior therapy, or intensive, home-based program, measured overall function, and included a treatment and control group.	4 CTs plus secondary evidence including previous reviews and critical appraisal debates. Total number of children 96. None of the four CTs overlapped with the evidence base for the current review as the diagnostic criteria used in these studies were not based on current standards established in the DSM-IV or TR edition.	Preschool children with autism.	Primary outcome -Intellectual functioning, specifically normal functioning. Other outcomes include- language, social interaction and play, adaptive or self-care skills, mal-adaptive behavior	The authors provide a critical appraisal of the included studies and suggest that studies, particularly those by Lovaas et al. (1997) and McEachin et al. (1993), are limited by numerous methodologic al flaws (namely lack of random assignment).	Narrative	Overall, the authors conclude that there is insufficient evidence to establish a causal relationship between a particular program of IBI and the achievement of "normal functioning." Further, the benefits in terms of overall functioning found by Lovaas (1987) and McEachin (1993) have not been corroborated by independent researchers.

ABA – Applied behavioral analysis
ADHD – Attention deficit/hyperactivity disorder
APA – American Psychological Association
ASD – Autism spectrum disorder;
BCOHTA – British Columbia Office of Health Technology Assessment Report
CT – Controlled trials
FCT/ABA – Functional Communication Training and Applied Behavior Analysis
IBI – Intensive behavioral intervention
NHMRC – National Health and Medical Research Council
NR – Not reported
NZHTA – New Zealand Health Technology Assessment
OCD – Obsessive compulsive disorder
PDD-NOS – Pervasive Developmental Disorder-Not Other Wise Specified
RCT – Randomized controlled trials
SR – Systematic review

Appendix L. Names of Those Involved in the Preparation of This Report

ECRI Institute Personnel

All ECRI Institute personnel involved in the preparation of this report may be contacted at:

ECRI Institute
5200 Butler Pike
Plymouth Meeting, PA 19462
Telephone: (610) 825-6000
Facsimile: (610) 834-1275

Stacey Uhl, M.S.S.
Lead Research Analyst

Internal Review Committee

Karen Schoelles, M.D., S.M.
Medical Director

Wendy Bruening, Ph.D.
Senior Research Analyst

Joann Fontanarosa, Ph.D.
Research Analyst

Meredith Noble, M.S.
Research Analyst

External Reviewers

Christopher H. Schmid, Ph.D.
Professor, Tufts University School of Medicine
Director, Biostatistics Research Center
Institute for Clinical Research and Health Policy Studies
Tufts Medical Center
800 Washington Street, Box #63
Boston, MA 02111

Mark Wolery, Ph.D.
Department of Special Education
Box 328, Peabody College
Vanderbilt University
Nashville, TN 37203