
Michele Spreckley, MCSP, and Roslyn Boyd, PhD, MSC (Physiotherapy)

Objective To review the effectiveness of applied behavior intervention (ABI) programs for preschool children with autism spectrum disorder (ASD) in their cognitive, adaptive behavior, and language development.

Study design Systematic reviews, randomized or quasirandomized controlled trials (RCT) of ABI delivered to preschool children with ASD were reviewed. Quantitative data on cognitive, language, and behavior outcomes were extracted and pooled for meta-analysis (RevMan 4.2).

Results Thirteen studies met the inclusion criteria. Six of these were randomized comparison trials with adequate methodologic quality (PEDro ≥ 6). Meta-analysis of 4 studies concluded that, compared with standard care, ABI programs did not significantly improve the cognitive outcomes of children in the experimental group who scored a standardized mean difference (SMD) of 0.38 (95%CI −0.09 to 0.84; P = .1). There was no additional benefit over standard care for expressive language; SMD of 0.37 (95%CI −0.09 to 0.84; P = .11), for receptive language; SMD of 0.29 (95%CI −0.17 to 0.74; P = .22) or adaptive behavior; SMD of 0.30 (95%CI −0.16 to 0.77; P = .20).

Conclusions Currently there is inadequate evidence that ABI has better outcomes than standard care for children with autism. Appropriately powered clinical trials with broader outcomes are required. (J Pediatr 2009;154:338-44)

Recently the prevalence of autism spectrum disorder (ASD) was reported as 38.9 per 10 000 children in a population-based sample from southeast England. There have been reports that the prevalence of ASD is increasing. In 2002, Chakrabarti and Fombonne concluded that the rate of pervasive developmental disorders (PDD) is higher than reported 15 years ago. Although this increase in ASD arguably may be due to a change in diagnostic criteria, improved screening and early detection, ASD is a concern for health and educational professionals who are providing intervention programs for these increasing numbers.

Childhood ASD is a pervasive developmental disorder that is characterized by abnormal functioning in 3 main areas of development before the age of 3 years: (1) reciprocal social interaction, (2) communication, and (3) stereotyped repetitive behavior. In addition to these diagnostic features a range of nonspecific problems commonly is identified, such as anxiety, sleeping and eating disturbances, temper tantrums, self- and other-directed aggression. Autism is understood to be part of a spectrum. Children with ASD present with great variability in severity and clinical picture, with some attaining functional language whereas others have no effective communication; some remaining isolated and aloof while others are affectionate to particular people. Many children have a limited play repertoire that can be repetitive and perseverant about certain objects. Others develop stereotypical behaviors such as hand flapping, walking on tiptoes, or body rocking. The diversity of ASD suggests that no one treatment has been effective for all children, and response to treatment may depend on the level of intelligence. Between 50% and 70% of children with ASD also present with an intellectual disability. Measures of cognitive functioning are often used as an outcome after intervention. The original study of applied behavior intervention (ABI) by Lovaas et al in 1987 claimed that children who received intensive behavioral treatment (IBT) made significant gains in IQ scores. Many...
of the subsequent studies evaluating ABI have tried to replicate the original study by Lovaas et al.\(^6\) and have measured similar outcomes to add to the body of evidence for ABI.

ABI is based on the theory of applied behavioral analysis (ABA) and may be known as applied behavior analysis intervention (ABAI) or IBT.\(^7\) It is a method developed by Lovaas et al.\(^8\) of teaching appropriate behaviors by breaking tasks down into small discrete steps and training in a systematic and precise way called discrete trial training. This approach is based on the concept that children with ASD have significant difficulties with learning, being unable to learn through imitation, and listening as their normal peers do. Initially ABI was administered at a high intensity, with 40 hours of direct training each week administered to the child by student therapists under supervision.\(^9\) In the earlier studies of discrete trial learning “aversives” were used to encourage children to participate. Aversives are negative responses by the adult given to the child when noncompliant behavior occurs. In later studies, aversives tend not to be used in ABI programs.\(^10\) The current recommended 40-hour weekly intervention in the home setting is a major burden to the family. More recently the intensity has been reduced to 30 hours in response to families’ other needs.\(^10\) This is still a problematic level of input for most families and service providers.\(^10\) The primary aim of this systematic review was to determine the efficacy of ABI in enhancing cognition, language and adaptive behavior when provided to preschool children with ASD.

**METHODS**

**Search Strategy**


The search strategy comprised the following MeSH headings or Key words: (1) autism or autistic spectrum disorder or Asperger syndrome or (PDD) or child development disorder; and (2) behavior therapy or early childhood intervention or applied behavior analysis or early behavioral treatment; and (3) cognition outcomes or rehabilitation or child health outcomes.

Studies were downloaded into Endnote, version 9 (Thomson Reuters, New York, New York), and duplicates were deleted. Studies were identified by title and abstract and screened by the authors to assess whether they met the selection criteria set out below.

**Selection Criteria**

To be included in the meta-analysis, studies had to meet the following selection criteria.

1. Trials included systematic reviews, randomized controlled trials (RCT), quasirandomized controlled trials, or controlled trials.
2. Participants comprised preschool children with a diagnosis of ASD or PDD.
3. Interventions included those that focused on ABI approaches to behavioral management. These included direct behavior management for the child, parent education and training, and consultation with caregivers in the community.
4. Interventions were delivered to the parents/caregivers and/or directly to the child, by special educators, teachers, speech pathologists, psychologists, or other allied health professional students.
5. Studies occurred while the children were of preschool age between 18 months and 6 years.
6. Outcomes included cognitive, language, or adaptive behavior outcomes.

For studies that were excluded, see Study Flow Diagram (Figure 1).

**Data Extraction**

For qualitative analysis the Physiotherapy Evidence Database (PEDro) Scale of quality assessment was used in evaluating these articles by the 2 authors independently. The reliability of this scale has been established.\(^11\) The PEDro scale consists of 11 qualitative measures scoring 1 or 0 for each item.\(^12\) It was not possible to blind the subjects receiving intervention or the therapists delivering the intervention, therefore the maximum possible PEDro score was 9. Studies scoring 6 or more on the PEDro scale are considered to have adequate internal validity for quantitative meta-analysis (Table 1).
Table I. Methodological quality assessment of included studies—PEDro scale

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>Total score</th>
</tr>
</thead>
<tbody>
<tr>
<td>McEachin</td>
<td>1993</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Koegel</td>
<td>1996</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>6*</td>
<td></td>
</tr>
<tr>
<td>Jocelyn</td>
<td>1998</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sheinkop</td>
<td>1998</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smith</td>
<td>2000</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>7*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eikeseth</td>
<td>2002</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>6*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bernard-Opitz</td>
<td>2004</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sallows</td>
<td>2005</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>7*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Howard</td>
<td>2005</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohen</td>
<td>2006</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eldevik</td>
<td>2006</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Magiati</td>
<td>2007</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eikeseth</td>
<td>2007</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6*</td>
<td></td>
</tr>
</tbody>
</table>

Scale of item score 0 = absent/ unclear, 1 = present, The PEDro scale criteria are: (1) specification of eligibility criteria (2) random allocation (3) concealed allocation (4) prognostic similarity at baseline (5) subject blinding (6) therapist blinding (7) assessor blinding (8) greater than 85% follow up of at least one key outcome (9) intention to treat analysis (10) between group statistical comparison for at least one key outcome (11) point estimates and measures of variability provided for at least one key outcome. Studies included in meta-analysis.

*Studies scoring >5.
†Studies included in meta-analysis.

Data Synthesis

Further quantitative analysis was conducted in Review Manager (RevMan), version 4.2 for Windows, (The Nordic Cochrane Centre, The Cochrane Collaboration, Copenhagen, Denmark). Pooled data for treatment effect were calculated across trials with a fixed effect model. Data were analyzed with effect sizes, standardized mean differences (SMD), and 95% confidence intervals (CI). The effect of heterogeneity ($I^2$) measures the degree of inconsistency across studies that is due to variability rather than chance. The larger percentage values show increasing heterogeneity.\(^{13}\)

RESULTS

Description of Studies

A search of the databases identified 64 abstracts that required further investigation (Figure 1). Twenty-five papers were retrieved for detailed examination, and 13 studies met the inclusion criteria.\(^7,10,14-24\) Of these 13 studies, 6 were RCTs or quasi-RCTs with PEDro score $\geq 6$. Four of these 6 studies had adequate data for meta-analysis.\(^7,10,18,23,24\) Twelve articles did not meet the inclusion criteria with reasons for exclusion tabulated in the Study Flow Diagram (Figure 1). Studies were not included in the final analysis if they did not include “discrete trial training” as part of the intervention.

Study Participants

Of the 4 studies, one\(^7\) used a diagnosis of ASD according to the criteria based on the Diagnostic and Statistical Manual of Mental Disorders DSM-IV.\(^25\) Another study did not use a standardized diagnostic instrument.\(^10\) Two studies\(^18,23\) excluded children who had an IQ score less than 50 and used the criteria from the World Health Organization.\(^26\) All of the studies included children within the age range of 18 months to 6 years.

Types of Intervention

The content and intensity of interventions administered to the treatment groups and control groups are tabulated in Table II. Four of the studies\(^7,10,18,23\) based intervention in the treatment group on the ABA developed by Lovaas et al.\(^8\)

In all the studies the comparison groups also received intervention so there were no true control groups, that is, groups receiving no intervention. Eikeseth et al\(^18,23\) provided an eclectic treatment to their comparison groups that was designed to reflect best practice in services to children with autism. This intervention had elements of TEACCH (Treatment and Education of Autistic and related Communication Handicapped Children)\(^27\) and of ABI training.\(^9\) Each child received a combination of interventions based on recommendations from the child’s multidisciplinary team. Two studies provided ABI intervention to families in the comparison group with less intensity and supervision, 30 hours per week compared with 39 in the intervention group\(^7\) and 5 hours per week compared with 30 hours.\(^10\)

Qualitative Analysis

Of the 13 studies retrieved and tabulated, there were 4 studies with adequate data for meta-analysis. Two were randomized clinical trials,\(^7,10\) and 2 studies by Eikeseth et al were quasirandomized trials, with children being allocated into groups depending on availability of therapists, with a follow-up at 12 months\(^18\) and at 22 months later.\(^23\) These 4 studies were identified as moderate to high quality, scoring 6 or more out of a possible score of 9 on the PEDro scale (Table I). The studies of sufficient quality to be considered for meta-
Three studies reported sufficient data on expressive language outcomes with the Reynell Language Scales. The follow-up study by Eikeseth et al in 2007.

Two other studies scored more than 6 on the PEDRO scale but did not have adequate data for meta-analysis. The remaining 7 studies scored between 4 and 5 points out of a possible score of 9 on the PEDro and were considered to be of poor quality.

**Primary Outcome**

**COGNITIVE.** The tests used to measure cognitive outcome were the Bayley Scales of Infant Development (2nd edition), the Wechsler Intelligence Scales for Children, and the Wechsler Preschool and Primary Scale of Intelligence. Three studies shown in Figure 2(i) reported sufficient data on cognitive outcomes to be pooled for meta-analysis.7,10,23 Children who received ABI (n = 41) scored a standardized mean difference (SMD) of 0.38 (95% CI −0.09 to 0.84 P = .11) compared with children who received standard care (n = 35) when tested after intervention. Heterogeneity was low (I² = 33.1%).

**LANGUAGE.** Three studies used the Reynell Developmental Language Scales for assessing language outcomes at preschool age 1 to 6 years (infant, primary). The Clinical Evaluation of Language Fundamentals was used for assessing outcomes after intervention, 4 years later at 7 years of age. The follow-up study by Eikeseth et al in 2007 did not report on language outcomes with the Reynell Language Scales. Three studies reported sufficient data on expressive language shown in Figure 2 (ii) and receptive language for meta-analysis in Figure 2 (iii).

**Extensive Language Outcomes after Intervention.** Children who received intervention (n = 41) scored a SMD of 0.37 (95% CI −0.09 to 0.84; P = .11) compared with the children who received standard care (n = 35) when tested after intervention. One study demonstrated a difference in favor of the intervention with a SMD of 0.97 (95% CI 0.14 to 1.81) and another study favored the comparison group.7 There was moderate heterogeneity between studies (I² = 47.0%).

**Receptive Language Outcomes after Intervention.** Children who received intervention (n = 41) scored a SMD of 0.29 (95% CI −0.17 to 0.74; P = .22) compared with children receiving standard care (n = 35). One study favored the comparison group with a SMD of −0.30 (95% CI −1.13 to 0.53). There was low heterogeneity between groups (I² = 28.3%).

**ADAPTIVE BEHAVIOR.** There were no scales used for measuring adaptive behavior outcomes below the age of 5 years. The Vineland Adaptive Behavior Scale was used to measure outcomes at 5 years of age and older. Three studies reported sufficient data on adaptive behavior for meta-analysis7,10,23 shown in Figure 2 (iv). Overall children who received ABI (n = 41) scored a SMD of 0.30 (95% CI −0.16 to 0.77; P = .20) compared with the children who received standard care (n = 35) when tested after intervention. There was moderate heterogeneity showing increasing inconsistency across studies (I² = 65.9%).

**DISCUSSION**

Four randomized or quasi-randomized clinical trials met inclusion criteria and had primary outcomes analyzed. The meta-analyses of these studies showed that ABI did not result in significant improvement in cognitive, language, or adaptive

---

**Table II. Structure and content of intervention programs**

<table>
<thead>
<tr>
<th>Study</th>
<th>N = Total sample</th>
<th>Duration (months)</th>
<th>Experimental intervention method</th>
<th>Mean age (MA) at intake</th>
<th>Intensity of intervention in experimental group</th>
<th>Comparison group method</th>
<th>Mean age (MA) at intake</th>
<th>Intensity of intervention in comparison group</th>
<th>Intervention providers</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smith19 n = 28</td>
<td>24-36</td>
<td>ABI (Lovas et al 1981)</td>
<td>Weekly 30 hours over 2-3 years</td>
<td>Parent training ABI at home</td>
<td>Weekly 5 hours in home over 3-9 months</td>
<td>By student therapists</td>
<td>Intellectual functioning</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sallows7 n = 23</td>
<td>48</td>
<td>ABI (Lovas 1987) 1:1 at home</td>
<td>Weekly 39 hours Year 1 &amp; 37 hours Year 2</td>
<td>ABL parent directed group</td>
<td>Weekly 31 hours Year 1 &amp; 30 hours Year 2</td>
<td>By therapists</td>
<td>Language Adaptive Sociomotorial Academic Class placement Parent evaluation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eikeseth18 n = 25</td>
<td>12</td>
<td>ABI (Lovas et al., 1981) in kindergarten</td>
<td>Weekly 28 hours plus 2 hour weekly meetings over 1 year</td>
<td>Eclectic treatment in kindergarten MA = 65 months</td>
<td>Weekly 29 hours plus 2 hour weekly meetings</td>
<td>By teacher 4-6 hours; by aide rest of time</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eikeseth22 n = 25</td>
<td>32</td>
<td>ABI in kindergarten and school setting</td>
<td>Reduced to weekly 18 hours after school start</td>
<td>Eclectic treatment in kindergarten and school MA = 65 months</td>
<td>Reduced to weekly 16 hours after school start</td>
<td>By teachers and aides</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CCC. Childcare center.
behavioral outcomes compared with standard care. This meta-analysis has limitations in that there was high variability in the studies included, difficulty establishing control groups, and no standardization of the comparison intervention, poor homogeneity, limited information on retention in the intervention groups, and lack of strict inclusion and exclusion criteria. Despite the barrage of requests for more comprehensive early intervention for children with ASD, only a small number of efficacy studies have been performed, enrolling only 76 children. Current evidence does not support ABI as a superior intervention for children with ASD. The instruments used to measure change were primarily discriminative and secondarily evaluative; these may not be able to detect low-level changes.

Figure 2. Summary of meta-analyses of the effect of ABI versus standard care.
Most of the participants met the standard criteria for a diagnosis of ASD. One study included children with Asperger syndrome and PDD who may make more substantial gains with intervention because of a lower incidence of cognitive impairment. For secondary analysis, Sallows and Graupner found better cognitive outcomes for “rapid learners” compared with “moderate learners.”

The 4 studies included used randomized or quasirandomized allocation for assignment of experimental and comparison groups; however, all the comparison groups received some form of intervention, in some cases including ABI at reduced intensity compared with intervention groups. This highlights the difficulties researchers have in designing research for young children with marked disability because it may be considered unethical to withhold treatment. On the other hand, this replicates current clinical practice in which parents are accessing multiple programs. Two studies reported contamination of both the treatment and comparison groups with concurrent non-study treatments.

An alternative option, where a no-treatment control group is considered unethical, is to randomly allocate children to receive intervention either immediately or after a controlled delay. It is possible to conduct clinical trials of intervention for children with ASD in the clinical setting with appropriate random allocation compared with standard clinical best practice to assess treatment effects of interventions. A non-systematic review concluded that for best practice a minimum of 20 hours a week of intervention over 2 years—taking account of each child’s strengths and weaknesses and family circumstances—was essential for young children with ASD to make gains.

The results of this review should be interpreted with caution because the theoretical construct and program content of the ABI, and the ages of children involved, varied substantially in all 4 studies. The follow-up study by Eikeseth et al. extended into school years. Two of the studies compared traditional ABI developed by Lovaas et al specifically with other treatments. For the other 2 studies, the content of the intervention was the same for the comparison group, although at reduced intensity (80% and 16%). In 1 study, the comparison groups achieved outcomes equivalent to the intervention group. This may be because parents were supervised and trained in ABI and became very skilled in delivery.

A variety of outcome measures were used. The 4 studies used child-related outcomes only. Analysis of the impact on parent outcomes is beyond the scope of this review, although 1 study has addressed this topic. Assessments need to address outcomes for family functioning, as well as a child’s general development.

ASD is a common disorder and has substantial impact on family functioning, parental stress, and limitations in educational attainment for these children. Further research is needed to investigate the cause and nature of ASD in various subgroups and should be sufficiently powered to evaluate critical periods for intervention, the optimum intensity, and mode of delivery for achieving successful outcomes. Outcomes should be measured that are most closely related to the 3 areas of abnormal functioning that characterize ASD, such as social skills, child internalizing, externalizing behaviors, parent-child interaction, and family well-being. A multicenter, randomized control trial is essential to take into account the changes in children with ASD because of natural history. What is too often forgotten is that the overwhelming majority of children with ASD change over time as part of their development as opposed to change resulting from an intervention.

We thank Professor Margot Prior, Department of Psychology, University of Melbourne for her reviewing of a prior draft of the manuscript for which she received no compensation, Associate Professor Dinah Reddihough for granting of study leave for the semester at La Trobe University in 2006, and Chris Spreckley for his technological support.

REFERENCES


343
In this 1959 issue of *The Journal*, Blattner provides another insightful summary of the evolving knowledge of an infectious disease, glandular toxoplasmosis. The protozoan parasite was discovered in 1908 in a North African rodent, the gondi, and about the same time in the wild rabbit in Brazil. Human infection was not established definitively until 1939. In a historical sequence which is unusual, catastrophic congenital disease was well characterized as the consequences of human infection before the more common and broad spectrum of *Toxoplasma* disease in children and adults was recognized. Standardization of the Sabin-Feldman dye test for *Toxoplasma* antibodies, and establishment of diagnostic titers, led to the recognition of relatively common asymptomatic infection, as well as a spectrum of clinical disease.

Blattner reviews the evidence for a *Toxoplasma* cervical lymphadenitis syndrome, which characteristically includes swelling and pain of cervical lymph nodes (especially posterior nodes), back and neck pain (sometimes simulating dread poliomyelitis, which was pandemic at the time), and fatigue. Characteristically there is neither high fever nor sore throat. In a Danish study of 100 cases of lymphadenopathy of unknown origin reported in 1956, 5% of cases were caused by *Toxoplasma*. These clinical features summarized by Blattner are accurate 50 years later.

The evidence for the cat as the definitive host was not yet established at Blattner's 1959 writing. One suspect transmitter in England was the “budgerigar, a psittacine bird becoming increasingly popular as a pet.” Budgerigar (parakeet) owners take heart. The psittacine is off the perch for this one. Apparently bird droppings frequently end up by coincidence in the same vicinity as cat droppings.

Sarah S. Long, MD
Section of Infectious Diseases
St. Christopher’s Hospital for Children
Philadelphia, Pennsylvania
10.1016/j.jpeds.2008.08.047