Progress and outcomes for children with autism receiving parent-managed intensive interventions☆

Peter Bibby[a],*, Svein Eikeseth[b], Neil T. Martin[c,1],
Oliver C. Mudford[c], David Reeves[d]

[a]*Autism & Developmental Disorders Education Research, 62 Oakhurst Grove,
London, England SE22 9AQ, United Kingdom*
[b]*Akershus College, Norway*
[c]*Keele University, England, United Kingdon*
[d]*University of Manchester, Manchester, England, United Kingdom*

Abstract

Parent-managed behavioral interventions for young children with autism are under-researched. We analyzed data from 66 children served by 25 different early intervention consultants. After a mean of 31.6 months of intervention IQ scores had not changed ($N = 22$). Vineland adaptive behavior scores had increased significantly by 8.9 points ($N = 21$). No children aged > 72 months attained normal functioning, i.e., IQ > 85 and unassisted mainstream school placement ($N = 42$). Progress for 60 children across 12 months was found for mental age (5.4 months), adaptive behavior (9.7 months), and language (5.1 months). The interventions did not reproduce results from clinic-based professionally

directed programs. The effectiveness of the parent-managed intervention model as it has developed and the adequacy of professional services in that model are discussed. © 2001 Elsevier Science Ltd. All rights reserved.

1. Introduction

1.1. Progress and outcomes for children with autism receiving parent-managed intensive interventions

Over the last two decades, researchers have shown that children with autism may benefit from intensive behavioral treatment (Anderson, Avery, DiPietro, Edwards, & Christian, 1987; Birnbrauer & Leach, 1993; Eikeseth, Smith, Jahr & Eldevik, in press; Fenske, Zalenski, Krantz, & McClannahan, 1985; Harris, Handleman, Gordon, Kristoff, & Fuentes, 1991; Harris & Handleman, 2000; Lovaas, 1987; McEachin, Smith, & Lovaas, 1993; Weiss, 1999). Most notably, Lovaas (1987) studied a group of 19 children receiving UCLA-model early intensive behavioral treatment (Lovaas et al., 1981; Lovaas, 1993, Lovaas, 1996; Smith, Donahoe, & Davis, 2000). The interventions were early, children were younger than 46 months at the start of treatment, and intensive, in that trained adults provided one-to-one treatment for 40 hr per week for two or more years. At follow-up when the children were between six and seven years of age IQ score had increased by an average of 20 points to a mean of 83, and nine children (47%) achieved "best outcome," defined as normal intellectual functioning (IQ > 85) and regular educational placement without special assistance (Smith, Groen, & Wynn, 2000). In contrast, a control group that had received up to 10 hr per week of behavioral treatment obtained a mean IQ of 52, and a group that had received no behavioral treatment had a mean IQ of 58. Gains were maintained by eight of the nine "best outcome" children at a second follow-up conducted when the children averaged 13 years of age (McEachin et al., 1993). More than 10 replications of the UCLA model of treatment are underway with assistance from the National Institutes of Mental Health (NIMH). Results from these studies will provide further evidence on the effects of clinic-based professionally directed UCLA model treatment.

These scientific publications (particularly Lovaas, 1987) and parental testimonies reporting outcomes of normal functioning (e.g., Maurice, 1993) have prompted many parents of children with autism to seek similar behavioral treatment. The demand has been international and far exceeds the supply of available competent clinics and professionals who have sufficient education, training, experience, and expertise to direct treatment (Foxx, 2000). A potential solution to the demand/availability gap with respect to specialist clinics has been for parents to initiate and manage their own children's treatment programs.

The differences between various models for providing early behavioral intervention services must be emphasized. Research studies cited so far have been

conducted from specialist centers and clinics. Children's treatment was directed and provided by professionals with considerable expertise in applied behavior analysis and young children with autism. Variations in treatment methods employed by these centers have been described (e.g., Harris & Handleman, 2000), and the UCLA model has been differentiated from others (Smith, 1999; Smith, Donahoe, & Davis, 2000).

The present study investigated the effects of parent-initiated approximations to the UCLA workshop model (Lovaas, 1996), not of the UCLA model, nor of other professionally directed center/clinic-based models. Children's treatment was not provided as part of the present study. A generalized description of the development, methods, and content of children's parent-initiated interventions is needed to place this research in its proper context. The following description is based on direct observation of initial and follow-up workshops and children's treatment and on discussions with therapists, families, and service providers. Quantitative data on some treatment variables are reported later in this paper.

Families obtained initial information about behavioral early intervention from parents' organizations, internet websites, e-mail networks, or professionals. They located and recruited a consultant to provide overall direction for their child's program. Families also hired three to five potential therapists after advertizing in local universities and colleges, newspapers, and unemployment offices. In some cases, therapists included family members and volunteers. Generally, 30 to 40 hr of one-to-one therapist-to-child treatment per week were recommended. We obtained no information about therapist turnover, but development and maintenance of a therapy team was cited as the greatest barrier to UK behavioral early intervention programs by 71% of 141 families surveyed by Johnson and Hastings (2000).

The child's program commenced with a workshop provided by the consultant for usually two six-hour days at the child's home. A mixture of theory (explanation of basic behavioral principles), modeling, and practice (e.g., using discrete trials training) was provided to the family and novice therapists. The consultant was hired for further one-day follow-up workshops with a median frequency of four times per year. The child's progress and curriculum were reviewed at these follow-ups, as was therapist performance.

Children's curriculums differed across consultants and children but could generally be traced to their origins in the UCLA treatment manual, also known as *the ME-book* (Lovaas et al., 1981). More specific examples of such curriculums have been published (e.g., Leaf & McEachin, 1999; Schreck, 2000; Taylor & McDonough, 1996; Weiss, 1999). Teaching methods were prescribed as discrete trials training (Koegel, Russo, & Rincover, 1977) and, for generalizable skills that had been established, incidental teaching (Hart & Risley, 1975). One-to-one discrete trials teaching sessions were conducted in children's homes, most often in a child's bedroom. Self-help skills such as eating, dressing and toileting were taught in task-appropriate settings.

These types of programs have been described elsewhere as "parent-directed programs" (Smith, Buch, & Gamby, 2000). We prefer to describe the interventions we observed as *parent-managed* because the extent to which parents

directed some important aspects (e.g., teaching methods and curriculum) was not consistent or obvious. Programs of this type are also known as "workshop-model programs," from the type of consultancy service provided.

To our knowledge, the earliest parent-managed program in England began in 1994. By mid-1998, when the present study commenced, there were approximately 250 individual programs and, mid-2000, we estimate that as many as 650 children receive some form of this intervention in England. Worldwide, many thousands of children and their families are involved. From a recent paper by Jacobson (2000) and discussions with colleagues in the US and other countries, we know that the general situation that we describe for the UK is common internationally.

Whether parent-managed programs can reproduce the positive results of clinic-based programs is largely unknown. Only two reports from these types of intervention have been published to date. Sheinkopf and Siegel (1998) evaluated the progress of nine children receiving 19 hr per week of treatment for 16 months and reported that children's IQ had risen by a mean of 27 points. Smith, Buch, and Gamby (2000) reported outcomes for six boys with autism two to three years after their parents had initiated intensive behavioral programs following the parent-managed version of the UCLA model. Outcomes were mixed, but two of the children showed increases of > 10 points on standardized measures of intelligence, language, and adaptive behavior. These reports provide some support for the notion that parent-managed programs may enhance functioning in children with autism, though none of the children in these studies achieved normal educational and intellectual functioning. More data on the effects of parent-managed intensive programs are needed to advise parents, professionals, and education providers. These are required urgently, given the rapidly increasing numbers of individual children's programs.

The study was designed and commenced before any of the publications cited later than 1997 were available. At that time there were few published caveats regarding the generalizability of Lovaas's (1987) clinic-based results to parent-managed programs, whether the latter were claiming to be conducted according to current UCLA guidelines, or not. It was not unusual for research findings from clinic-based services to be emphasized by advocates for the treatment with the implicit (or explicit) assertion that 47% best outcome (normal intellectual and educational functioning) could be expected from workshop model professional services. This study was designed to assess the validity of this extrapolation.

The purposes of this study were: 1. To assess whether parent-managed intensive behavioral intervention reproduced Lovaas's (Lovaas, 1987) clinic-based results with respect to outcomes; 2. To measure progress of children who were receiving parent-managed intensive behavioral intervention; 3. To determine variables that may predict children's progress and outcomes. The study extends on that of Sheinkopf and Siegel (1998) by including multiple measures coherent with UCLA/NIMH studies and six times as many subjects.

2. Method

2.1. Participants

Parents who had sought legal advice from the first author concerning public funding for behavioral intervention for their children with autism were invited to participate in the study: 75 of 92 consented. Further characteristics of this group were described in Mudford, Martin, Eikeseth, and Bibby (in press). Sixty-six children (11 girls and 55 boys) participated through to the end of the study. Reasons for dropout or exclusion were: Four children were withdrawn by parents' request; one had ceased to receive behavioral treatment before any assessments were conducted; one was excluded because an occasion for assessment could not be arranged; two had an additional diagnosis of cerebral palsy; and one girl was re-diagnosed with Rett's Disorder. All children were reported by parents to be receiving ongoing intensive behavioral intervention, which had commenced at a mean age of 45.0 months (SD 11.2 months). When the first set of assessments were conducted for this Study 11 children (18% of the sample) were in their first 12 months of behavioral intervention, 31 children (52%) were in their second year of treatment, and 18 (30%) had been receiving behavioral treatment for more than two years. All but two, who had a Pervasive Developmental Disorder diagnosis, had received an independent diagnosis of autism or autistic spectrum disorder. No children who completed the study had secondary diagnoses of additional neurological disorders or epilepsy. Although 65 of 66 programs had been financed initially by families, by the end of this study all but three children's behavioral programs were funded wholly or partially by Local Education Authorities.

2.2. Measures

We used child assessment instruments that had been selected for the UCLA/ NIMH affiliated studies. Thus the data obtained from children in the present study can be compared directly with those of other recent and ongoing studies of early behavioral intervention. Results are reported in the form of standard scores and/or mental ages or age equivalents (derived from raw scores and tables in the test manuals).

2.2.1. Cognitive development
Cognitive development was assessed with Bayley Scales of Infant Development (2nd ed.) (Bayley, 1993), WPPSI-R (Wechsler Preschool and Primary Scale of Intelligence - Revised) (Wechsler, 1990), or WISC-III (Wechsler Intelligence Scale for Children, 3rd ed.) (Wechsler, 1992). The choice of assessment instrument depended on the age and ability of the child. First, WPPSI-R was attempted with children between the ages of 3.5 and 7.25 y, and WISC-III with children aged over 7.25 y. If the child did not obtain a valid score, as defined in the Wechsler manuals, the Bayley was used. Standard scores could not be deter-

mined for Bayley test results because nearly all assessments were conducted when the children were older than 3.5 years. Therefore, Bayley ratio IQ scores were computed: (age equivalent divided by chronological age) times 100. Cognitive performance was also assessed with the Merrill-Palmer Scales of Mental Tests (Stutsman, 1948).

2.2.2. Language skills

The Reynell Developmental Language Scales (3rd UK ed.) (Edwards et al., 1997) provided measures of age equivalence in comprehension and expression separately. Children who scored below the floor level on comprehension or expression were given an age-equivalent of 20 months, one month less than the minimum age equivalent in this edition of the tests.

2.2.3. Adaptive behavior

The Vineland Adaptive Behavior Scales (Survey Form Interview Edition) (Sparrow, Balla, & Cicchetti, 1984) provided data on children's communication, daily living skills, socialization, and problem behaviors as described by their parent(s) in a structured interview with a trained rater. Comparative raw data were obtained from a large ($N = 684$) study of norms obtained with children and adults with autism (Carter et al., 1998). From these we computed Vineland composite standard scores for children in the same age ranges as our group. These means are reported in the Results section.

2.2.4. Ratings of behavior

Research assistants rated each child's speech, behavioral excesses and deficits common in autism as part of the interview conducted with parent(s) after the completion of each child assessment. Copies of the interview are obtainable on request. Speech was rated on a five-point scale, with descriptors at each point derived from the Autism Diagnostic Interview - Revised (Lord, Rutter, & LeCouteur, 1994): 1 = No recognizable (to nonfamily members) speech, 2 = One-word utterances (at least two recognizable words), 3 = Single words (>10 words used in context), 4 = Phrase speech (two or more word phrases including a verb and used in context), 5 = Conversational speech. The frequency of behavioral excesses (aggression to people or property, self-injury, tantrums) was rated on a four-point scale similar to that used in standardized behavior rating forms (e.g., Aman, Tasse, Rojahn, & Hammer, 1996): 1 = no problem, 2 = once a week or less, 3 = several times per week, 4 = at least daily. A *Sum Pathology Score* (Lovaas, 1987), with a maximum score of eight, was derived by awarding one point for presence of: stereotypy, appearance of sensory deficit, and problem behavior; and one point for absence of: speech, appropriate toy play, emotional attachment to parent, play with other children, and toilet training. Ratings were made early in the interview for retrospective preintervention levels and later for current child behaviors.

The reliability of these ratings was not assessed. The interviewer had spent several hours with the children in their homes before questioning parents, i.e.,

knew the child somewhat, so parental report that diverged significantly from the interviewer's observations was unlikely. Also, previous research has indicated that parents' reports on specific behaviors can coincide closely with direct observations (Siegel, Anders, Ciaranello, Bienenstock, & Kraemer, 1986). Therefore we believe that parents gave quite accurate accounts when face-to-face with the interviewer.

2.3. Additional data from interview with parents

The parent interviews included questions that provided information about the child's age at start of intervention, duration of intervention, intensity (hours per week) of one-to-one treatment, funding, psychological assessments external to the study, and school placement.

2.3.1. Treatment personnel

Parents were asked about program consultancy. The majority of consultants named (23 of 25) were nonUK-resident multinational service providers. Consultants were grouped according to qualifications: Group 1, serving 20% of children at the start of the study, fulfilled all the criteria for UCLA Level II (Smith & Lovaas, 1998); Group 2 provided for 29% of the sample and were Ph.D.-level and/or registered clinical psychologist-level, but did not fulfil all the current UCLA criteria; and, Group 3 were paraprofessionals with master's-level or bachelor's-level qualifications at most serving 51% of children. By the end of the Study 10% of families changed consultant so that the percentages in each group changed to 19% for Group 1, 35% in Group 2, and 47% in Group 3. There were 18 Group 3 consultants named. Seven worked for agencies headed by a Group 1 consultant (and consulted to 11 children); seven were employed through agencies headed by a Group 2 level individual (18 children); and four could be described as independent paraprofessionals (5 children).

2.3.2. Supplementary treatments

The interview with parents included discussion of supplementary treatments that the child had received during the previous 12 months. The majority (81%) of children received one or more interventions chosen by their parents as treatment for autism or its symptoms (e.g., diets, vitamins, minerals, secretin infusions, homeopathic preparations, sensory treatments) during the 12 months in which children participated in the study. This high frequency of parent-selected alternative treatments is consistent with another recent report on children receiving early intensive behavioral intervention (Smith & Antolovich, 2000).

2.4. Early data

We did not assess directly children's performance prior to initiation of behavioral intervention. Results from standardized tests conducted by independent professionals before intervention commenced or up to three months after was

classed as "Early" data. The time of testing relative to the start of behavioral intervention is reported for the individual Early data in the Results section. These data were obtained by scrutiny of copies of all professional reports, collated by Education Authorities in the statutory procedure for assessment of children's educational needs. Additional Early assessment data were obtained by writing to psychologists who had assessed the children. Early measures of IQ obtained from independent professionals were based on the following instruments that were considered to produce valid IQ scores: Bayley and Wechsler tests, Griffiths Mental Development Scale (Griffiths, 1970), and British Ability Scales (Elliot, Murray, & Pearson, 1977; Elliot, Smith, & McCullough, 1997). Results from Leiter tests (Leiter, 1969) were not included in statistical analyses because they have been shown to overestimate IQ (Shah & Holmes, 1985). Received Merrill-Palmer IQ scores were not included for the same reason (see Results section). It is worth noting here that mean Early IQ was 50.8 following the rules described, but could be inflated to 65.9 if we included scores from Leiter, Merrill-Palmer, and Vineland (cf. Lord & Schopler, 1989; Lovaas, 1987; Rutter & Lockyer, 1967; Sheinkopf & Siegel, 1998).

2.5. Assessment procedures

Child assessments and parent interviews were conducted twice. First assessment was performed between September 1998 and May 1999. Henceforth the first assessment is labeled "T1." Second assessment (T2) occurred 12 months (± 2 weeks) later for every child.

Assessments and interviews were conducted by seven psychology graduates with a minimum of 12 months' postgraduate experience working directly with children with autism in the same age range as those in the present sample. All had extensive experience of using behavioral teaching procedures with young children, and five worked concurrently as behavioral early intervention therapists. These research assistants were trained to administer the specific instruments by a clinical psychologist who provided assessment services for one of the UCLA/ NIMH replication studies.

Research assistants administered assessments in the children's homes. On each occasion of measurement, all tests were conducted in a single day. Direct child testing occurred over 1.5 to 4 hr depending on the child's performance and behavior. Children were given sufficient breaks and rewards for compliance to facilitate optimal performance (following recommendations in Sattler, 1992). Testing protocols prescribed in the test manuals were not otherwise modified.

Parent interviews for Vineland Scales and concerning independent variables took place after child assessments were complete. The interview lasted approximately one hour.

Steps were taken to negate some potential sources of bias. Research assistants

were blind to the hypotheses and design of the study. They did not conduct assessments for children with whom they had ever worked as therapists. A different researcher conducted assessments on each occasion of measurement to ensure that the second set of measurements was not influenced by the results of the first.

2.6. Reliability

Since IQ scores have received most emphasis in previous reports, a second assessment of IQ was obtained on 17 occasions, i.e., for 13% of IQ scores obtained. The reliability checks were conducted within two weeks either side of the primary assessment. The primary assessor was blind to whether an assessment of reliability had taken place (or was planned), and the reliability assessor was blind to previous scores. Reliability checks were made by a different research assistant (2 occasions), the third author (7 occasions), a trained research assistant recruited and trained mid-way through the study (4 occasions), and the psychologist who had trained the original group of assessors (4 occasions). Scores from different assessors were highly correlated ($r = .91, p < .01$) and did not differ in level: $t(16) = .85$, ns.

Reliability of the Early IQ data from assessments independent of the present study was estimated based on nine occasions when an independent assessment had been conducted within six months of one of the assessments conducted for the study. Correlation between the study score and the independent score was $r = .81$ ($p < .01$), with no significant difference in level: $t(8) = .75$, ns.

2.7. Design

The study plan at the outset had been to assess the performance of all children longitudinally for two years with a third assessment 12 months after the second set (T2) was conducted. However, after the children's first assessment (T1), the data suggested already that parental hopes and advocates' claims for 47% best outcome rate were over-optimistic. We considered whether we had a duty to families of children with autism and professionals who advise them to publicize the results at this point in the study, which would severely jeopardize its continuance. A university ethics committee reviewed these preliminary results and advised us to continue the study for another year. Given results after the second round of assessments, we decided there were sufficient data to publish the results and terminate the study.

The study design thus became longitudinal over 12 months for individual children and cross-sectional with respect to time into treatment because children were first assessed at different times after the start of treatment. The design is a variant of the accelerated multicohort longitudinal design (Kazdin, 1998).

3. Results

3.1. Analyses

Results for IQ and Vineland adaptive behavior composite are reported for Early data (i.e., score from assessment closest to start of behavioral intervention) and on all measures at T1 (i.e., the time of the first assessment for this study) and at T2 (i.e., time of the second study assessment 12 months after T1). All statistical tests on progress were one-tailed as previous reports suggested a general hypothesis predictive of improvement in scores over time. Statistical significance is reported at the $p < .01$ level unless otherwise stated. Regarding individual data, a change of 15 points (generally one standard deviation for the assessment instruments employed) was chosen to indicate a "notable" change in level. Also 15 points is the bandwidth for levels of mental retardation (American Psychiatric Association, 1994).

3.2. Representativeness of early data

Early IQ scores were obtained for 22 children, and Early Vineland adaptive behavior composite standard scores for 21. Nine children had Early scores on both measures. There was no significant difference between the means ($t = .45$) or variances ($F = .00$) of IQ scores at T1 when comparing children for whom Early data were available with those for whom they were not. Similarly, for children with Early Vineland data, means were not different ($t = .82$), and neither were variances ($F = .32$) at T1. Therefore, there was reason to believe that children for whom Early data were found were representative of the group as a whole.

4. Analysis 1: changes from early data to T2 data

IQ data for the 22 children for whom Early data were obtained are shown in Table 1. Mean Early IQ at an average age of 43.4 months was 50.8. The change to mean IQ of 55.0 at T2 when the average age was 75.5 was not significant: $t = 1.37$, ns. Children are ordered in Table 1 with those who made greatest positive changes from Early data to T2 at the top. The first six children listed in Table 1 showed increases of > 15 IQ points. The IQs of the last two children listed decreased notably. The majority (64%) did not change appreciably on this measure.

When Early IQ data from Merrill-Palmer and Leiter tests were included with the data shown in Table 1, the number of children with Early data increased to 29. The mean Early IQ was 62.0. Comparable T1 and T2 means (i.e., Merrill-Palmer scores used when that test or Leiter had been used Early) were 65.6 and 63.8 respectively. There were no significant differences between these Early and

Table 1
Individual data for 22 children with Early IQ scores ordered from greatest to least gain in score across whole duration of program

Child	Early IQ			Follow-up IQ		Duration of Program[3]
	Age	Time[1]	IQ[2]	T1	T2	
1	29	−4	28	53	56	24
2	36	−2	86	119	111	31
3	51	+3	68	79	92	29
4	27	−5	41	54	60	42
5	48	+2	35	43	51	18
6	41	+2	14	38	30	24
7	41	0	46	61	60	71
8	34	0	27	40	40	43
9	32	−2	32	48	42	29
10	48	+2	31	34	38	18
11	34	+1	70	79	72	17
12	45	+3	60	69	62	45
13	81	+2	44	54	46	37
14	45	−10	49	41	50	32
15	52	−2	87	91	86	27
16	53	0	86	78	80	25
17	34	−14	50	50	40	40
18	67	+3	77	52	66	24
19	41	−4	47	32	36	27
20	40	−1	45	39	34	29
21	43	+3	44	34	26	26
22	32	−6	51	40	31	37
mean	43.4		50.8	55.6	55.0	31.6
SD	12.6		20.6	22.0	22.3	11.9

Notes: [1] Time of early assessment in months before (−), coincident with (0), or after (+) child's program commenced; [2] IQ from Wechsler, Bayley, Griffiths, or British Ability Scales; [3] Months of one-to-one treatment by T2.

follow-up means. The inclusion of these other test scores increases the mean absolute IQ level at all occasions of measurement but does not change the finding that group mean IQ did not change across 31.6 months of intervention.

Mental age was derived from Early IQ scores when the test manuals permitted their derivation. Mean Early mental age was 21.3 months ($N = 17$), increasing significantly to 38.1 months at T2 ($t = 6.50$). So, as a group, children performed better on IQ tests from the Early assessments to T2 in that they responded as if they had achieved 16.8 more months of typical development. However, considering that this was gained during 31.6 months of intervention, it can be seen why IQ did not increase, and the group did not make progress to catch up cognitively with typical children.

Vineland Adaptive Behavior Scales composite scores are shown in Table 2 for 21 children for whom Early scores were obtained. There was a significant increase of 8.9 points from the Early mean of 54.5: $t = 1.92$, $p = .03$. The first

Table 2

Individual data for 21 children with early Vineland scores ordered from greatest to least gain in score across whole duration of program

Child	Early VAB[2]			Follow-up VAB		Duration of Program[3]
	Age	Time[1]	VAB	T1	T2	
23	35	+1	58	102	112	28
24	46	+3	59	115	108	28
12	41	0	49	70	75	45
25	34	0	51	73	91	23
7	41	0	39	58	65	71
26	40	−2	54	64	76	48
27	42	0	52	77	69	44
3	51	+3	77	79	86	29
28	45	+2	48	51	54	29
13	82	+3	42	46	46	37
10	49	+2	49	46	49	18
29	35	+3	53	54	53	29
14	56	+1	63	48	62	32
6	41	+2	44	48	43	24
11	34	+1	66	62	68	17
30	30	0	54	60	52	47
31	38	0	50	50	48	27
19	45	0	47	48	44	27
32	66	+1	47	43	36	26
33	66	+1	45	38	32	32
34	39	−1	97	57	62	37
mean	45.5		54.5*	61.4	63.4*	33.2
SD	12.8		13.0	19.4	21.9	12.3

Notes: [1] Time of early assessment in months before (−), coincident with (0), or after (+) child's program commenced; [2] Vineland adaptive behavior composite standard score; [3] Months of one-to-one treatment at T2; * Means significantly different $t(20) = 1.92$, $p = .03$ (one-tailed).

seven children listed in Table 2 increased their Vineland composite score by > 15 points. One child (34) made a notable loss.

We could not compare Vineland scores with the published norms in Carter et al. (1998) because those norms were divided into mute and verbal subgroups by a behavioral definition that differed from those we employed to assess presence or absence of speech. The raw data obtained from the study by Carter et al. (1998) were analyzed by extracting data for groups of children within the age range of children in the present study when Early Vineland data were obtained and when T2 data were collected. Means from the Carter et al. (1998) data were 53.7 ($N = 250$) and 49.7 ($N = 271$), respectively. There was no significant difference between means at the Early data age ($t = .26$, ns) but the difference of 13.7 points at T2 age was significant ($t = 3.94$, $p < .01$).

On this measure of adaptive behavior, the group did make progress toward reducing developmental delay, and at follow-up (T2) they out-performed children with autism who had not received intensive behavioral treatment (Carter et al., 1998). Following the Vineland manual, however, the mean at T2 of 63.4 is

still at the "low" level of adaptive skills, i.e., > 2 SDs below the typical population mean of 100 (Sparrow et al., 1984).

Sum pathology scores decreased significantly by 2.7 points to 2.2 points on average from the time that behavioral intervention started to T2: t(59) = 9.67. That is, the number of behavioral excesses and deficits associated with autistic disorder reduced during the time that children were receiving behavioral treatment. Speech, rated by researchers from parent interviews, improved with mean rating increasing from 1.5 to 3.6: t(59) = 14.94. Before intervention, 37 children (62%) had no recognizable words (i.e., rating of 1), but that figure had dropped to 5 (8%) by T2.

5. Analysis 2: outcomes

Lovaas (1987) described outcomes for children at a minimum age of 72 months after they had received at least 24 months of behavioral intervention. At T2, 42 children in the present study met both criteria. Their mean age was 85.5 months (SD 11.7), and they had received behavioral intervention for an average of 36.8 months (SD 10.3). Mean total hours of one-to-one treatment by T2 was 4743 (SD 1508).

The principal measures used by Lovaas (1987) to describe outcomes were IQ score for intellectual functioning and school placement for educational functioning. Smith, Groen, and Wynn (2000) detailed the criteria for "best outcome" used in Lovaas (1987) and McEachin et al. (1993): IQ > 85, and successful regular (mainstream) school placement with no special assistance.

5.1. IQ scores

Mean IQ score at T2 for our "outcome" group of 42 was 53.1 (SD 20.0, 95% confidence interval = 46.8 to 59.4). This finding was significantly less than Lovaas's (Lovaas, 1987) mean score of 83 for outcome in the intensively treated group, even if the unreported standard deviation had been as high as 40, 7.6 points greater than standard deviation at follow-up (McEachin et al., 1993): t(59) = 3.09, *p* < .01. There was no difference between the obtained T2 mean and either of Lovaas's control groups at outcome: 52 for the low intensity behavioral intervention group, and 58 for the group that received no behavioral treatment. Early IQ scores were available for 12 children in our outcome group. From these Early data, only Child 3 (Table 1) from this subgroup moved into the nonretarded IQ range (i.e., IQ > 69), while Child 18 moved to IQ < 70.

5.2. School placements at T2

Thirty-eight children in the outcome group were continuing to receive home and community-based one-to-one behavioral intervention for a mean of 23.7 (SD

Table 3
Mean scores at first assessment (T1), second assessment (T2) and difference across 12 months
(T2 score − T1 score). The group of children who contributed scores at "outcome" (right
column) is defined in the text. N = 60 except where noted.

Measure	T1	T2	T2 − T1	Outcome[1]
IQ				
deviation or ratio	57.6	58.5	0.9[2]	53.1
mental age (months)	34.0	39.4	5.4[3]**	
Vineland				
Composite				
months	31.8	41.4	9.7**	
standard score	58.8	61.0	2.2	55.2
Communication				
months	35.2	47.5	12.3**	
standard score	66.9	70.3	3.5*	66.3
Daily Living				
months	33.7	41.3	7.6**	
standard score	59.0	56.9	−2.1	50.4
Socialization				
months	26.3	36.2	9.8**	
standard score	62.8	65.9	3.1	63.0
Maladaptive raw score	16.3	15.2	−1.1	
Merrill-Palmer (mental age, months)	48.8	55.67	6.8**	58.3
Reynell (months)				
Comprehension	31.9	37.5	5.7**	38.1
Expressive	31.2	35.7	4.6**	35.6

Notes: [1] n = 42, [2] n = 58, [3] n = 51; ** p < .01, * p < .05, one-tailed paired samples t-tests
comparing T1 with T2, i.e., significant progress.

10.4) hours per week. They attended school additionally for an average of 15.9
(SD 11.0) hours per week. Placements were: 25 (66%) in mainstream (regular)
schools for at least part of the school week, with 23 receiving full-time individual
adult support, and two receiving part-time individual adult assistance; 13 attend-
ing school in special education settings, with nine having specially assigned
one-to-one assistance at least part of their time at school. Additionally, we have
data on school placement for six children in the study who had ceased to receive
EIBI between T1 and T2, including two who had received > 24 months of
behavioral treatment but were aged < 72 months. Five were receiving their
education at special needs schools, and one had no school placement.

    None of the 42 children aged > 72 months were in mainstream school without
individual adult support. That is, there were no "educationally normal function-
ing children" in the outcome subgroup, therefore no "best outcomes."

5.3. Additional description of outcomes

    Table 3 (right column) shows means for IQ, Vineland adaptive behavior
standard scores, Reynell language age equivalents (months), and Merrill-Palmer
mental ages (months) for the outcome group of 42. The principal reason for

presenting these data are for comparison with other researchers' published data (e.g., Smith, Groen, & Wynn, 2000; Smith, Buch, & Gamby, 2000). When considering comparison with results from studies that have included a wide range of IQ assessment instruments (e.g., Sheinkopf & Siegel, 1998), it is noteworthy that the mean Merrill-Palmer ratio IQ was 68.2, 15 points higher than IQ obtained from Bayley and Wechsler tests.

5.4. Partial best outcomes

Altogether, 10 out of 66 children from our study exceeded the IQ > 85 criterion at T2. Half of these children entered treatment with an IQ score above criterion with Early IQ data from a variety of instruments showing, for different children, 117 and 107 on Leiter, 98 on Merrill-Palmer, and 87 and 86 on Griffiths. Another three of the 10 children had Early data indicating preintervention IQ score < 85. These three children made progress to achieve the IQ criterion for best outcome. No Early data were available for the other two children with IQ > 85 at T2.

Further follow-up six to nine months after T2 of the six children who had T2 IQ scores > 85 but were aged < 72 months found that three were succeeding in regular schools without any extra adult support, and one was receiving minimal support (5 hr/week). These four children, who were not included in the "outcome" group of 42 because they were too young at T2, may be described as approximating Lovaas's criteria for "best outcome."

6. Analysis 3: changes from T1 to T2

All children who received behavioral intervention through the 12 months from their first assessment at T1 to their second assessment at T2 were included in these analyses ($N = 60$, except when stated). Group means on all standardized tests and differences from T1 scores to T2 scores are presented in Table 3. A significant positive value for the T2-T1 difference indicates that mean score increased over 12 months, i.e., children made progress during 12 months.

IQ scores did not change significantly from T1 to T2: $t(57) = 0.71$, ns. Group mean mental age increased by 5.4 months in the 12 months from T1 to T2: $t(50) = 4.55$. There was no significant change on Vineland composite standard score over the 12 months from T1 to T2: $t(59) = 1.60$, ns. However, Vineland composite age equivalent increased significantly from T1 to T2 by 9.7 months to 41.4 months: $t(59) = 6.51$. Vineland domain standard scores showed significant progress between assessments in Communication (up 3.5 points to 70.3), but not in Socialization or Daily Living Skills. Children did make progress on all Vineland domain age equivalents in the 12 months from T1 to T2. The domain with most progress was Communication with a 12.3-month increase. Socialization increased by 9.8 months, and Daily Living Skills showed least progress with 7.6 months. Except in Communication, children's progress was less than would

Table 4
Mean, standard deviation, and range of selected child and treatment variables (N = 60)

Measure	Mean	SD	Range
Age at T2 (months)	77.4	15.0	50–121
Age at start of intervention (mo.)	45.0	11.2	30–79
Duration of intervention at T2 (mo.)	32.8	10.6	15–71
Total 1-to-1 hours	4256	1452	1365–9538
Mean overall 1-to-1 hours/week	30.3	5.5	14–40
1-1 hours/week from T1 to T2	27.1	9.8	3–40
School hours/week from T1 to T2	11.6	9.3	0–30

be expected from typically developing children in 12 months. The descriptors from the Vineland manual indicate that, as a group, children were at the "low" adaptive level (i.e., standard score < 70, 2 SDs below the mean) on all domains except Communication at T2. Maladaptive behavior raw scores were stable and continued to show "significant" problems.

Other standardized measures showed progress across the year with increases of 5.7 months and 4.6 months for Reynell Comprehension and Expressive scales respectively, and 6.8 months for Merrill-Palmer. As with Vineland scales, these gains were all less than 12 months of typical children's development and therefore were not sufficient to reduce developmental delays.

6.1. Individual changes from T1 to T2

Fig. 1 shows individual changes in IQ scores and Fig. 2 shows changes in Vineland composite scores across 12 months. Thirteen different children (22%) made notable gains and three (5%) showed losses in one or more areas of development over the year. Four (of 58) children's IQs increased by 15 or more points during the year: up by 15 to 68, 19 to 100[1], 29 to 98[2], and 37 to 135. One child lost > 15 points: down 17 to 47. Seven (of 60) children increased their Vineland Composite standard score by > 15 points, with two progressing from low to adequate levels of adaptive behavior (up 18 to 91, and up 36 to 110[3]). Regarding language gains, five (of 60) children gained > 12 months in age equivalence on both Reynell scales, i.e., their language delay was reduced across the year. The language of two children regressed, i.e., age equivalent in months on Comprehension and Expressive scales was lower at T2 than it had been at T1.

7. Analysis 4: treatment variables and predictors of progress

Table 4 presents summaries of some variables concerning children and their educational interventions. The last four variables reported in Table 4 relate to the

amount of one-to-one home-based treatment received or to the time spent at school. Taken as independent variables, the implication is that these factors may influence a child's rate of performance gain. However, it seems at least equally likely that the influence will be the other way around and that children who perform at a relatively high level will tend to spend more time at school and receive less one-to-one treatment. Therefore, these variables were not analyzed as potential predictors of progress.

7.1. Predictors

The two comprehensive measures of child performance, and those for which there were Early data, were IQ and Vineland adaptive behavior composite. There were strong correlations between Early IQ with IQ at T1 and T2 (for both, $r =$ .78), and between T1 and T2 IQs ($r = .93$). Correlations between Early Vineland composite standard scores and those at T1 and T2 were nonsignificant ($r = .28$, ns; $r = .35$, ns), but score at T1 was highly correlated with score at T2 ($r = .91$).

Three potential predictors of progress were investigated: age at T2; age at start of intervention; and, duration of intervention at T2. Regression analysis was used to examine the impact of each of these on progress from Early data to T1, Early data to T2, and across 12 months from T1 to T2. In view of the relatively small sample sizes, each predictor was investigated individually. To control for differences between children on earlier scores, score at start of time period analyzed was entered into each regression analysis along with each predictor.

Progress on IQ from Early data to T1 was predicted *only* by age at start of behavioral treatment ($r = -0.39$, p(1-tailed) $= .04$). Dividing this group at the median of start age (43 months) showed that children who started before this age progressed an average of 10.8 IQ points and children who started later lost 2.4 points. No predictors of IQ progress from Early data to T2 or from T1 to T2 were found. The presence of any recognizable speech before intervention (as reported by parents) did not affect progress on IQ. Progress on the Vineland was not predicted by any of the variables investigated.

Grouping of consultants by professional background was described in the Method section. There were insufficient numbers of children with Early data to analyze any effects of the three consultant groups on change from Early data to T2. Considering changes in IQ and Vineland from T1 to T2, there was no significant effect of group membership. The analyses may have been confounded by a further variable, which was "frequency of consultant visits during the 12 months": Group 1 had visited a mean of 3.8 times (range 1 to 17); Group 2, 3.2 times (0 to 8); and Group 3, 8.6 times (2 to 26). There was a difference in these means: F(2, 57) = 4.78, $p = .01$ as the frequency of visits by Group 3 consultants was significantly higher than for either of the other groups. No differences between consultant groups were found when the analysis of T1 to T2 changes was repeated with "frequency of visits" as a covariate.

8. Discussion

This study was designed to evaluate progress and outcomes for children with autism receiving parent-managed behavioral interventions. First we analyzed IQ-test data from 22 children and adaptive behavior scores from 21 children across an average of 31.6 months of treatment. Mean IQ scores for 22 children did not change significantly, maintaining around the borderline of mild/moderate mental retardation. Significant increase in age-equivalence scores was found on the IQ measures, but this progress was not sufficient to reduce developmental delays (i.e., progress in 12 months was less than 12 months of typical children's development). Dividing this group at the median of age of starting intervention (43 months) showed that children who started before this age made a mean gain of 10.8 IQ points, and children who started later lost 2.4 points on average. Vineland adaptive behavior scores improved by 8.9 points to a mean of 63.4. Examination of individual data showed that 27% of children made notable IQ gains and 33% made notable gains in adaptive behavior (notable gains defined as increase of > 15 points).

The small and nonsignificant gain obtained in IQ was comparable with data reported by Lord and Schopler (1989) for children who had not received intensive behavioral treatment. When we analyzed raw data from the study by Carter et al. (1998) extracting data for groups of children within the age range of children in the present study, we found that the children's adaptive scores were similar at the age of Early data whereas at follow-up the children from the present study had a mean adaptive score 13.7 points higher than children of same age in the Carter et al. study. These gains in adaptive behavior could indicate that children may be able to cope with less restrictive environments in the future than if they had not received behavioral intervention.

At a mean of 20.8 months into treatment, progress was assessed over a period of 12 months for a larger group of children ($N = 60$). No significant improvement in standard or ratio scores in the areas of intellectual functioning, adaptive functioning, or language was found. However, individual data showed that 22% of the children reduced their developmental delays in one or more broad domains: cognitive ability, adaptive behavior, and language. Fewer than 10% showed marked loss of performance in any of these areas. Significant progress was found on the majority of age-equivalent measures, but those gains were generally not sufficient to reduce developmental delays. Data from a structured parent interview revealed a significant improvement in speech and autistic behaviors. With speech, for example, at the last assessment 92% of children had used at least one-word utterances that were recognizable. For behavior ratings, even though not reflected in Vineland maladaptive scores, parents reported that their children were less autistic in their behaviors. These improvement in skills as reported by the parents and as measured by improvement in age-equivalence scores on the standardized tests may reflect behavioral changes that parents recognize and value, and thus may encourage parents to continue with the intervention.

Outcome was examined for 42 children who were at least 72 months of age and had received two years or more of intervention. None of these children fulfilled Lovaas's (Lovaas, 1987) criteria for "best outcome." Considering all 66 children in the study, ten children (17%) were determined to be in the range of normal intellectual functioning (IQ > 85) at their last assessment, and three (5%) have subsequently achieved educational normal functioning also. Five children with IQ > 85 at follow-up were already in the normal intellectual functioning range before the start of behavioral intervention by some measure. However, IQ scores of 98, 107, and 117 from nonverbal tests (Leiter and Merrill-Palmer) may have overestimated initial IQ for three of these five children (Shah & Holmes, 1985; this study).

The results of the present study are similar to those reported by Smith, Buch, and Gamby (2000) and Smith, Groen, and Wynn (2000). Larger gains were obtained by Sheinkopf and Siegel (1998), but they included Merrill-Palmer scores in their outcome data, which may have inflated IQ scores as suggested from our results. Overall, research to date suggests that although parent-managed programs may bring about gains in language, adaptive, and intellectual functioning, they may not be as large as in center-based programs, the latter reporting 20 IQ-point increases (Harris & Handleman, 2000; Harris et al., 1991; Lovaas, 1987; McEachin et al., 1993) and up to 47% normal functioning (Lovaas, 1987).

Failure to replicate Lovaas's (Lovaas, 1987) results may be due to variations in the group of children before they commenced intervention, their treatment, or a combination of these broad factors. Considering children first, we had estimated that 16% of our sample had IQ < 37, and 57% were older than 40 months if mute or 46 months if speaking at start of treatment (Mudford et al., in press) and therefore would not qualify for intake by Lovaas's criteria. However, the mean IQ at baseline was similar when a similar mix of tests was included. It should be recalled here that initial IQ did not predict progress in the present study. So, while there were differences in children, these seem unlikely to account for all of the difference between 47% and 0% best outcome.

Variations in treatment may account for much of the difference in outcome between parent-managed and center-based studies (Smith, Buch, & Gamby, 2000). With a mean of 30.3 hr per week across the intervention, children received, on average, nearly 10 fewer hours per week of one-to-one intervention than in Lovaas (1987). However children did receive more intensive one-to-one treatment than in other studies that reported more positive results (Anderson et al., 1987; Fenske et al., 1985; Harris et al., 1991). What happened to children during these hours of treatment is likely to be a critical factor but, as in most studies (e.g., Anderson et al., 1987; Fenske et al., 1985; Harris et al., 1991; Lovaas, 1987; Sheinkopf & Siegel, 1998), we did not have direct measures of the quality of the treatment. Lack of such measures is a limitation of the present study. Smith, Buch, and Gamby (2000) did collect data on therapist performance, and found that parent-employed therapists were less accurate at conducting elements of discrete trials training than therapists employed at a UCLA/NIMH

replication site clinic. This may have been a significant factor in the present study.

We do have data that attest to infidelity of all children's programs to the UCLA model, especially to that described for the Lovaas (1987) study, but also to the workshop model (Lovaas, 1996; Smith, Buch, & Gamby, 2000). Nearly 80% of children's programs were supervised by consultants who did not have authorization to provide UCLA-model programs, despite most children's programs being specified as "Lovaas programs." Consultants' supervision of children's treatment was reduced from weekly (Lovaas, 1987) to approximately once every three months. Thus, we can be sure that no children received consultancy services of the quality and quantity that produced 47% best outcomes (Lovaas, 1987). We can not assert that children received competent behavioral services: No consultants conformed to all elements of the standards which had been recommended in 1998 by the Autism Special Interest Group of the Association for Behavior Analysis (http://www.wmich.edu/aba/Austimwebfile.htm). Jacobson (2000) expressed similar concerns about quality control with respect to US services for families of children with autism.

Finally, with regard to procedural differences from Lovaas (1987), 81% of children received one or more supplementary "treatments." The effects of the additional interventions could not be analyzed due to the variety and combinations of treatments, durations, and doses. However, there is little evidence that these interventions have beneficial effect, at least on the type of standardized outcome measures employed in the present study (e.g., Mudford et al., 2000; Smith, 1996; Smith & Antolovich, 2000). Although the use of alternative treatments must limit conclusions about the effects of concurrent behavioral intervention, this limitation is perhaps unavoidable in research on educational interventions for autism when families are encouraged to try these alternatives by the media (including electronic mailing lists) and word-of-mouth from other parents.

Predictors of progress and/or outcome have proved elusive as in earlier studies. Although we did find that children who started intervention before age 43 months made more progress on IQ over their whole time of behavioral treatment, this was at the margin of statistical significance ($p = .04$, from a one-tailed test). The finding provides support to those of Harris and Handleman (2000) and Fenske et al. (1985) in that those children who start behavioral intervention earlier achieve better progress. Outcome on the IQ measure was also predicted by earlier IQ score, similar to previous reports (Harris & Handleman, 2000; Lovaas & Smith, 1988). In other words, the subgroup of children who had a relatively high IQ at the start of treatment generally maintained this superiority, even if they did not make progress on this measure.

The research design included no comparison group, and we compared our findings with previously published data. The study was designed to assess the reproducibility of Lovaas's (Lovaas, 1987) findings by parent-managed behavioral intervention as it had developed in England, in which case Lovaas's outcomes were the appropriate criterion against which to compare our results. Even so, it would be of value to know whether children in the study were helped

more or less than by the range of "standard" provisions for young children with autism in the UK. Research to describe typical provision for young children with autism has yet to be conducted, so it would have been problematic to identify an appropriate comparison group. Relevant UK data are presently lacking (Jordan, Jones, & Murray, 1998). Therefore, we encourage research to describe typical provision and to assess its effects.

English parent-managed intensive interventions for autism have not proved as effective as might be predicted through extrapolation from Lovaas's (Lovaas, 1987) findings. The results of this study do not call into question previous research that has demonstrated that behavioral interventions are helpful for persons with autism (Matson, Benavidez, Compton, Pacawskyj, & Baglio, 1996). Neither do the findings suggest that genuinely competent behavior analytic autism specialists working within a clinic-based model can not achieve substantial benefits for many young children with autism. What we have shown is that the model of parent-managed applied behavior analysis provision that has developed has not proved sufficient to bring about these benefits in most cases.

Notes

1. Also made notable gains in adaptive behavior and language
2. Also made gain in adaptive behavior
3. Also made gain in language

References

Aman, M. G., Tasse, M. J., Rojahn, J., & Hammer, D. (1996). The Nisonger CBRF: A child behavior rating form for children with developmental disabilities. *Research in Developmental Disabilities, 17*, 41–57.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Anderson, S. R., Avery, D. L., DiPietro, E. K., Edwards, G. L., & Christian, W. P. (1987). Intensive home-based early intervention with autistic children. *Education and Treatment of Children, 10*, 352–366.

Bayley, N. (1993). *Manual for the Bayley scales of infant development* (2nd ed.). London: The Psychological Corporation.

Birnbrauer, J. S., & Leach, D. J. (1993). The Murdoch early intervention program after two years. *Behaviour Change, 10*, 63–74.

Carter, A. S., Volkmar, F. R., Sparrow, S. S., Wang, J-J, Lord, C., Dawson, G., Fombonne, E., Loveland, K., Mesibov, G., & Schopler, E. (1998). The vineland adaptive behavior scales: supplementary norms for individuals with autism. *Journal of Autism and Developmental Disorders, 28*, 287–302.

Edwards, S., Fletcher, P., Garman, M., Hughes, A., Letts, C., & Sinka, I. (1997). *Manual for the reynell developmental language scales III: the University of Reading Edition.* Windsor, England: NFER Nelson.

Eikeseth, S., Smith, T., Jahr, E., & Eldevik, S (in press). Intensive behavioral treatment at school for four to seven year old children with autism: A one-year follow-up. *Behavior Modification.*

Elliot, C. D., Murray, D. J., & Pearson, L. S. (1977). *British Ability Scales.* Windsor, England: NFER.

Elliot, C., Smith, P., & McCullough, K. (1997). *British Ability Scales* (2nd ed.). Windsor, England: NFER-Nelson.

Fenske, E. C., Zalenski, S., Krantz, P. J., & McClannahan, L. E. (1985). Age at intervention and treatment outcome for autistic children in a comprehensive intervention program. *Analysis and Intervention in Developmental Disabilities, 5*, 49–58.

Foxx, R. M. (2000)., May). *Clinical, legal, and ethical issues in behavioral treatment for autism* (Panel discussion, R. M. Foxx - Chair). Association for Behavior Analysis Convention, Washington, DC.

Griffiths, R. (1970). *The abilities of young children.* London: Child Development Research Centre.

Harris, S. L., & Handleman, J. S. (2000). Age and IQ at intake as predictors of placement for young children with autism: a four- to six-year follow-up. *Journal of Autism and Developmental Disorders, 30*, 137–142.

Harris, S. L., Handleman, J. S., Gordon, R., Kristoff, B., & Fuentes, F. (1991). Changes in cognitive and language functioning of preschool children with autism. *Journal of Autism and Developmental Disorders, 21*, 281–290.

Hart, B., & Risley, T. R. (1975). Incidental teaching of language in the preschool. *Journal of Applied Behavior Analysis, 8*, 411–420.

Jacobson, J. W. (2000). Early intensive behavioral intervention: Emergence of a consumer-driven service model. *The Behavior Analyst, 23*, 149–171.

Johnson, E., & Hastings, R. P. (2000). *Facilitating factors and barriers to the implementation of intensive home-based behavioral intervention for young children with autism.* Manuscript submitted.

Jordan, R., Jones, G., & Murray, D. (1998). *Educational interventions for children with autism: A literature review of recent and current research* (Final report). School of Education, University of Birmingham, England B15 2TT.

Kazdin, A. E. (1998). *Research design in clinical psychology* (3rd ed.). Boston: Allyn and Bacon.

Koegel, R. L., Russo, D. C., & Rincover, A. (1977). Assessing and training teachers in the generalized use of behavior modification with autistic children. *Journal of Applied Behavior Analysis, 10*, 197–205.

Leaf, R. B., & McEachin, J. J. (1999). *A work in progress: Behavior management strategies and a curriculum for intensive behavioral treatment of autism.* New York: DRL Books.

Leiter, R. G. (1969). *General instructions for the Leiter International Performance Scale.* Chicago: Stoetling.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview - revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*, 659–685.

Lord, C., & Schopler, E. (1989). Stability of assessment results of autistic and non-autistic language-impaired children from pre-school years to early school age. *Journal of Child Psychology and Psychiatry, 30*, 575–590.

Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55*, 3–9.

Lovaas, O. I. (1993). The development of a treatment-research project for developmentally disabled and autistic children. *Journal of Applied Behavior Analysis, 26*, 617–630.

Lovaas, O. I. (1996). The UCLA young autism model of service delivery. In C. Maurice, G. Green, & S. C. Luce (Eds.), *Behavioral intervention for young children with autism: a manual for parents and professionals* (pp. 241–248). Austin, TX: Pro-Ed.

Lovaas, O. I., Ackerman, A., Alexander, D., Firestone, P., Perkins, M., & Young, D. B. (1981). *Teaching developmentally disabled children: the ME book.* Austin, TX: Pro-Ed .

Lovaas, O. I., & Smith, T. (1988). Intensive behavioral treatment for young autistic children. In B. B. Lahey & A. E. Kazdin (Eds.), *Advances in clinical child psychology, Vol. 11* (pp. 285–324). New York: Plenum Press.

Matson, J. L., Benavidez, D. A., Compton, L. S., Pacawskyj, T., & Baglio, C. S. (1996). Behavioral treatment of autistic persons: A review of research from 1980 to the present. *Research in Developmental Disabilities, 17*, 433–465.

Maurice, C. (1993). *Let me hear your voice: a family's triumph over autism.* New York: Fawcett Columbine.

McEachin, J. J., Smith, T., & Lovaas, O. I. (1993). Long-term outcome for children with autism who received early intensive behavioral treatment. *American Journal on Mental Retardation, 97,* 359–372.

Mudford, O. C., Cross, B. A., Breen, S., Cullen, C., Reeves, D., Gould, J., & Douglas, J. (2000). Auditory integration training for children with autism: no behavioral benefits detected. *American Journal on Mental Retardation, 105,* 118–129.

Mudford, O. C., Martin, N. T., Eikeseth, S., & Bibby, P. (in press). Parent-managed behavioral treatment for pre-school children with autism: Some characteristics of UK programs. *Research in Developmental Disabilities.*

Rutter:M., & Lockyer, L. (1967). A five to fifteen year follow-up study of infantile psychosis. *British Journal of Psychiatry, 113,* 1169–1182.

Sattler, J. M. (1992). *Assessment of children* (revised 3rd ed.). San Diego, CA: Author.

Schreck, K. A. (2000). It can be done: An example of a behavioral individualized education program (IEP) for a child with autism. *Behavioral Interventions, 15,* 279–300.

Shah, A., & Holmes, N. (1985). The use of the Leiter International Performance Scale with autistic children. *Journal of Autism and Developmental Disorders, 15,* 195–203.

Sheinkopf, S. J., & Siegel, B. (1998). Home-based behavioral treatment of young children with autism. *Journal of Autism and Developmental Disorders, 28,* 15–23.

Siegel, B., Anders, T. F., Ciaranello, R. D., Bienenstock, B., & Kraemer, H. C. (1986). Empirically derived subclassification of the autistic syndrome. *Journal of Autism and Developmental Disorders, 16,* 275–293.

Smith, T. (1996). Are other treatments effective? In C. Maurice, G. Green, & S. C. Luce (Eds.), *Behavioral intervention for young children with autism: a manual for parents and professionals* (pp. 45–59). Austin, TX: Pro-Ed.

Smith, T. (1999). Outcome of early intervention for children with autism. *Clinical Psychology: Science and Practice, 6,* 33–49.

Smith, T., & Antolovich, M. (2000). Parental perceptions of supplemental interventions received by young children with autism in intensive behavior analytic treatment. *Behavioral Interventions, 15,* 83–97.

Smith, T., Buch, G. A., & Gamby, T. E. (2000). Parent-directed intensive early intervention for children with pervasive developmental disorder. *Research in Developmental Disabilities, 21,* 297–309.

Smith, T., Donahoe, P. A., & Davis, B. J. (2000). The UCLA treatment model. In J. S. Harris & J. S. Handleman (Eds.), *Preschool programs for children with autism, 2nd ed* (pp. 23–39). Austin, TX: Pro-Ed.

Smith, T., Groen, A. D., & Wynn, J. W. (2000). Randomized trial of intensive early intervention for children with pervasive developmental disorder. *American Journal on Mental Retardation, 105,* 269–285.

Smith, T., & Lovaas, O. I. (1998). Intensive and early behavioral intervention with autism: the UCLA young autism project. *Infants and Young Children, 10,* 67–78.

Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). *Vineland adaptive behavior scale - survey form edition.* Circle Pines, MN: American Guidance Service.

Stutsman, R. (1948). *Manual for the Merrill-Palmer Scale of Mental Tests.* Windsor: NFER Nelson.

Taylor, B. A., & McDonough, K. A. (1996). Selecting teaching programs. In C. Maurice, G. Green, & S. C. Luce (Eds.), *Behavioral intervention for young children with autism: a manual for parents and professionals* (pp. 63–177). Austin, TX: Pro-Ed.

Wechsler, D. (1990). *Manual for the Wechsler preschool and primary scale of intelligence* (3rd ed.). London: The Psychological Corporation.

Wechsler, D. (1992). *Manual for the Wechsler intelligence scale for children - revised.* London: The Psychological Corporation.

Weiss, M. J. (1999). Differential rates of skill acquisition and outcomes of early intensive behavioral intervention for autism. *Behavioral Interventions, 14,* 3–22.